UNIVERSIDADE ESTADUAL DE MONTES CLAROS - UNIMONTES
CENTRO DE CIÊNCIAS SOCIAIS APLICADAS - CCSA
DEPARTAMENTO DE CIÊNCIAS ECONÔMICAS
CURSO DE CIÊNCIAS ECONÔMICAS

Educação e desenvolvimento econômico: um estudo para a região Norte de Minas Gerais, com ênfase no ensino fundamental

Acadêmico: Daniel Brito Alves
Orientadora: Professora Doutora Maria Elizete Gonçalves

Montes Claros (MG),
Dezembro de 2017

DANIEL BRITO ALVES

Educação e desenvolvimento econômico: um estudo para a região Norte de Minas Gerais, com ênfase no ensino fundamental

Trabalho de conclusão de curso apresentado pelo acadêmico Daniel Brito Alves ao Departamento de Ciências Econômicas da Universidade Estadual de Montes Claros, como pré-requisito para obtenção do título de Bacharel em Ciências Econômicas.

Orientadora:_____
Professora Doutora Maria Elizete Gonçalves

Banca Examinadora: _____
Professor Doutor Luiz Antônio de Matos Macedo

_____
Professora Doutora Ilva Ruas Abreu

Montes Claros (MG),
Dezembro de 2017

RESUMO

Os benefícios individuais e sociais proporcionados pela educação podem possibilitar que uma sociedade alcance um maior nível de desenvolvimento econômico. Considerando a importância de tal fato, o presente estudo objetivou verificar a existência da relação entre educação e desenvolvimento no Norte de Minas e investigar os determinantes do desempenho escolar nas séries iniciais do ensino fundamental, nesta Região. Utilizou-se como metodologias a análise de correlação e a análise exploratória de dados espaciais para verificar a relação entre educação e desenvolvimento. Para identificar os determinantes do desempenho escolar, empregou-se a Teoria de Resposta ao Item e o modelo de regressão hierárquico. Os dados utilizados foram obtidos através do portal do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) e da Federação das Indústrias do Estado do Rio de Janeiro (FIRJAN). A análise de correlação revelou a existência de relação positiva e significativa entre o indicador de desenvolvimento econômico (IFDM) e o indicador de qualidade da educação (IDEB). Tal fato é relevante ao mostrar que é possível se beneficiar de tal relação, mesmo em regiões menos favorecidas socioeconomicamente como o Norte de Minas. Através da análise espacial constatou-se que em alguns municípios a distribuição conjunta do IFDM e do IDEB não ocorre de forma aleatória, no espaço. Dessa forma, é possível também se favorecer do efeito transbordamento da relação entre educação e desenvolvimento. Quanto ao modelo estimado, verificou-se, entre outros resultados, que a reprovação em séries anteriores, bem como o escore de des(capital) cultural, afetaram negativamente o rendimento escolar dos alunos. Além disso, o nível socioeconômico do aluno, os escores de intervenção, de capital cultural, de dever de casa e a infraestrutura da escola afetaram positivamente a proficiência escolar dos alunos.

Palavras chaves: educação, desenvolvimento econômico, Norte de Minas.

ABSTRACT

The individual and social benefits provided by education can enable a society to achieve a higher level of economic development. Considering the importance of this fact, the present study aimed to verify the existence of the relationship between education and development in the North of Minas Gerais and to investigate the determinants of school performance in the initials séries of elementary education in this Region. Were used as methodologies the correlation analysis and the exploratory analysis of spatial data to verify the relationship between education and development. To identify the determinants of school performance, the Item Response Theory and the hierarchical regression model was used. The data used were obtained through the portal of the National Institute of Educational Studies and Research Anísio Teixeira (INEP) and the Federation of Industries of the State of Rio de Janeiro (FIRJAN). The correlation analysis revealed the existence of a positive and significant relationship between the economic development indicator (IFDM) and the educational quality indicator (IDEB). This fact is relevant in showing that it is possible to benefit from such a relationship, even in socioeconomically disadvantaged regions such as the North of Minas Gerais. Through the spatial analysis it was found that in some municipalities the joint distribution of the IFDM and the IDEB does not occur in a random way, in space. In this way, it is also possible to benefit from the overflow effect of the relationship between education and development. As for the estimated model, it was verified, among other results, that the disapproval in previous series, as well as the score of des(capital) cultural, negatively affected the students' school performance. In addition, the student's socioeconomic status, intervention scores, cultural capital, homework, and school infrastructure positively impacted students' school proficiency.

Key words: education, economic development, North Minas.

SUMÁRIO

Introdução **7**

1. Educação e desenvolvimento econômico: uma leitura a partir da Teoria do Capital Humano **9**

1.1 Educação e economia: a visão de alguns economistas e filósofos 9

1.2 Educação e economia: a Teoria do Capital Humano 11

1.2.1 A formulação de Schultz 11

1.2.2 A formulação de Becker 16

1.3 Revisão de literatura: determinantes do desempenho escolar 19

2. Metodologias de análise: Análise de Correlação, Análise Exploratória de Dados Espaciais, Teoria de Resposta ao Item e Modelo de Regressão Hierárquico **25**

2.1 Análise de Correlação 25

2.1.1 Coeficiente de correlação de Pearson 25

2.1.2 Coeficiente de correlação por postos de Spearman 26

2.2 Análise espacial de dados 27

2.2.1 Análise exploratória de dados espaciais (AEDE) e testes de autocorrelação espacial 27

2.2.1.1 *I* de Moran local bivariado 28

2.2.1.2 Diagrama de dispersão de Moran, mapa de significância e mapa de *clusters* 29

2.3 Teoria de Resposta ao Item 30

2.3.1 Modelo de Resposta Gradual de Samejima 31

2.4 O modelo de regressão hierárquico 34

2.4.1 Especificação do modelo incondicional (modelo nulo) 34

2.4.2 Especificação do modelo condicional 35

2.4.3 Método de estimação 36

2.4.4 Coeficiente intra-correlação 37

2.5 Variáveis e Dados 37

2.5.1 Variáveis relacionadas à análise de correlação e à análise espacial de dados 38

2.5.2 Variáveis inseridas na Teoria de Resposta ao Item e na análise de regressão 38

3. Educação e desenvolvimento no Norte de Minas **41**

3.1 Indicadores socioeconômicos do Norte de Minas 41

3.2 Indicadores educacionais do Norte de Minas 45

3.3 Educação e desenvolvimento no Norte de Minas: resultados e análise 46

3.4 Determinantes do desempenho escolar no Norte de Minas: resultados e análise 50
3.4.1 Estimação dos escores 50
3.4.2 Resultados dos modelos de regressão hierárquica..... 53

Considerações finais 57
Referências 59
Anexo I..... 63

Introdução

Uma sociedade pode alcançar um maior nível de desenvolvimento através de investimentos em educação (CALEIRO, 2010). Várias instituições mundiais tem usado esse fundamento como vertente para políticas e relações entre países. Um país pode se tornar mais competitivo se seu fator trabalho for mais bem educado; isso, porque aumentos na produtividade, geralmente decorrentes do capital humano, podem levar a um maior desenvolvimento econômico.

Portanto, a educação é um catalizador do desenvolvimento econômico, como enfatizado por Caleiro (2010). O autor ainda cita o reconhecimento generalizado da importância da educação. Os benefícios dessa seriam individuais (saúde, produtividade, etc.) e sociais (redução dos efeitos da pobreza, aumento da competitividade econômica, etc.).

Tendo em vista a importância dessa temática, o presente estudo teve como objetivo verificar, em um primeiro momento, a existência da relação entre educação e desenvolvimento no Norte de Minas. Posteriormente, foram investigados os determinantes do desempenho escolar nas series iniciais do ensino fundamental, nesta Região. Um bom resultado escolar, na educação básica, é um fator relevante para a formação do capital humano. Schultz (1973), um dos formuladores da teoria do capital humano, enfatiza a importância da educação primária. O autor supõe que mais e melhor instrução primária consistiria no melhor investimento, em termos de produção e lucros.

Sendo assim, buscou-se responder as seguintes questões: i) No Norte de Minas, é possível constatar a relação entre educação e desenvolvimento?; ii) Em que medida os municípios norte-mineiros estão associados espacialmente, em relação ao indicador de qualidade da educação, bem como ao indicador de desenvolvimento econômico? e; iii) Na educação básica, nos anos iniciais do ensino fundamental, quais os principais fatores relacionados ao desempenho escolar, na Região?

Dessa forma, as hipóteses norteadoras desse estudo foram as seguintes: 1) Há uma relação positiva entre educação e desenvolvimento econômico, na região do Norte de Minas Gerais; 2) Existem interações entre municípios norte-mineiros próximos, no que se refere ao indicador de qualidade da educação e ao indicador de desenvolvimento econômico e; 3) Variáveis relacionadas tanto ao aluno como à escola, são capazes de explicar, de forma significativa, o resultado escolar, nas séries iniciais do ensino fundamental, na região norte-mineira.

Para atingir os objetivos propostos, utilizou-se como metodologias a análise de correlação e a Análise Exploratória de Dados Espaciais (AEDE); para verificar a relação entre educação e desenvolvimento na Região. Indicadores relacionados aos alunos e às escolas foram construídos com base na Teoria da Resposta ao Item (TRI). Para identificar os determinantes do desempenho escolar, empregou-se o modelo de regressão hierárquico.

Os dados utilizados foram obtidos através do portal do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) e da Federação das Indústrias do Estado do Rio de Janeiro (FIRJAN).

Entre os resultados, verificou-se a existência de relação positiva entre o indicador de desenvolvimento econômico (IFDM) e o indicador de qualidade da educação (IDEB-M-AI) para a região do Norte de Minas Gerais. Através da análise espacial foi possível constatar que em alguns municípios a distribuição conjunta do IFDM e do IDEB não ocorre de forma aleatória, no espaço. A análise dos determinantes do desempenho escolar mostrou, entre outras constatações, que a reprovação em série(s) anterior(es), bem como o indicador des(capital) cultural afetaram negativamente o rendimento escolar dos alunos. Em contrapartida, o nível socioeconômico do aluno, os indicadores de intervenção, de capital cultural, de dever de casa e o da infraestrutura da escola afetaram positivamente a proficiência escolar dos alunos.

Este trabalho foi dividido em três capítulos. O primeiro trata-se do referencial teórico e da revisão de literatura acerca dos determinantes do desempenho escolar. O segundo capítulo apresenta as metodologias utilizadas. O terceiro aborda a formação econômica e social da região do Norte de Minas Gerais, com apresentação de alguns indicadores socioeconômicos e educacionais; além dos resultados e discussão das metodologias de análise utilizadas. Por fim, são apresentadas as considerações finais.

1. Educação e desenvolvimento econômico: uma leitura a partir da Teoria do Capital Humano

Este capítulo contempla a base teórica que sustenta o presente estudo. Inicialmente apresentam-se, resumidamente, algumas visões de economistas e filósofos acerca da relação entre educação e economia (desenvolvimento econômico). Posteriormente, a Teoria do Capital Humano (TCH) é apresentada de acordo com as concepções dos seus principais teóricos, sendo abordada em dois tópicos: a formulação de Schultz e a de Becker.

Na sequência, são expostos alguns estudos sobre os determinantes do desempenho escolar no ensino fundamental; tendo em vista a importância de um bom resultado escolar nas séries iniciais para a formação do capital humano de uma sociedade.

1.1 Educação e economia: a visão de alguns economistas e filósofos

Desde Adam Smith, os economistas se indagam sobre as causas da riqueza das nações. A acumulação de capital, o puritanismo protestante, os talentos empresariais, o clima e a raça são exemplos de respostas a essa indagação. Na sua obra "A riqueza das nações", a educação já aparece como um fator associado à referida riqueza.

Adam Smith (1776) faz referência ao efeito do treinamento (educação) sobre a produtividade da mão-de-obra da seguinte forma:

> O esforço natural de cada indivíduo no sentido de melhorar sua própria condição, quando sofrido para exercer-se com liberdade e segurança, é um princípio tão poderoso, que ele é capaz, sozinho e sem qualquer ajuda, não somente de levar a sociedade à riqueza e à prosperidade, mas de superar centenas de obstáculos impertinentes com os quais a insensatez das leis humanas muitas vezes obstacula seus atos. (SMITH, 1776, p.49).

Marx (1867) também destacou a importância da educação, ao considerar necessários os investimentos nessa área para se alterarem as relações sociais. Segundo seus argumentos, para adquirir a capacidade de serem formadores do estado, os trabalhadores deveriam investir em educação:

> Para modificar a natureza humana geral de tal modo que ela alcance habilidade e destreza em determinado ramo de trabalho, tornando-se força de trabalho desenvolvida e específica, é preciso determinada formação ou educação, que, por sua vez, custa uma soma maior ou menor de equivalentes mercantis. (MARX, 1867, p. 289).

Marshall (1890), na sua obra *Principles of Economics,* diz que qualidades e habilidades são capital humano e que a educação seria uma responsabilidade do estado e dos pais. Desta forma, aos pais caberia a missão de conscientizar quanto aos benefícios da educação, incentivar e investir uma parcela de seus ganhos na educação dos filhos. O Estado deveria investir na educação de modo que, se entre um grande número de crianças uma se destacasse e contribuísse para a sociedade, já teria valido o investimento feito em massa.

Na década de 1970, Kenneth Arrow (1973), por meio do seu artigo *Higher Education as a Filter,* formalizou a chamada Teoria do Filtro. De acordo com essa teoria, a produtividade do indivíduo seria capaz de explicar as diferenças salariais. A escola, nesse caso, exerceria a capacidade de "filtrar ou peneirar pessoas que possuem determinados atributos valorizados pelo mercado de trabalho, tais como persistência, inteligência etc." (NAZARENO, 2016, p. 19). A escolaridade, nesse contexto, é considerada como um sinalizador da qualidade do indivíduo, o qual passou pela triagem da instituição de ensino.

Ainda nos anos de 1970, Lester Thurow (1975) divulgou seu trabalho denominado de *Generating Inequality: Mechanisms of Distribution in the U.S Economy.* Na ocasião, surgiu a Teoria da fila, na qual, a educação é vista "como uma forma de garantir um bom lugar relativo na fila de competição por postos de trabalho." (NAZARENO, 2016, p. 21). Dessa forma, a escolaridade não aumentaria a produtividade do trabalho. Thurow (1975) embasou seu raciocínio na análise de que, para os anos entre 1950 e 1960, os salários de homens brancos adultos não acompanhou a evolução dos seus níveis educacionais. À luz da Teoria da Fila, tal fato acontece porque "não são as características da oferta (trabalhadores) que determinam a produtividade, mas as características da demanda, ou seja, dos postos de trabalho." (NAZARENO, 2016, p. 20). Esse raciocínio ainda ajuda a explicar como um mesmo indivíduo pode não apresentar a mesma produtividade entre diferentes postos de trabalho.

A Teoria do Filtro e a Teoria da Fila, citadas nos parágrafos anteriores, podem ser consideradas contrapontos da Teoria do Capital Humano (TCH). A TCH aponta para a importância da educação no processo de desenvolvimento econômico. Em consonância com a TCH, recentemente, há uma tendência em considerar a educação como elemento essencial no progresso. De acordo com Castro (1976), "poder-se-ia até mesmo falar numa nova ortodoxia do desenvolvimento econômico, na qual, o fator estratégico de progresso seria a educação". (CASTRO, 1976, p. 3).

As formulações da TCH manifestam-se em diversas correntes econômicas, como por exemplo, em Milton Friedman, que atesta ser o investimento em educação "uma forma de investimento em capital humano precisamente análoga ao investimento em maquinaria,

instalações ou outra forma qualquer de capital não humano. Sua função é aumentar a produtividade econômica do ser humano.” (FRIEDMAN,1977, p. 95).

Tendo em vista os pensamentos de Smith (1776), Marx (1867) e Marshall (1890), a tendência apontada por Castro (1976) e o posicionamento de Milton Friedman (1977), o tópico seguinte refere-se à teoria do capital humano.

1.2 Educação e economia: a Teoria do Capital Humano

1.2.1 A formulação de Schultz

Após 1950, a Teoria do Capital Humano foi formalizada. Theodore Schultz foi considerado o formulador e em 1979 foi destacado pelo prêmio Nobel em Economia. Ele propôs o conceito de educação como capital humano.

Schultz (1973) explicou sua teoria buscando mostrar que os países em desenvolvimento, como o Japão, que foram destruídos na Segunda Guerra Mundial, conseguiram reconstruir suas economias em pouco tempo. A partir de então, passou a se defender a ideia que o investimento na capacidade humana pode implicar no desenvolvimento na área econômica.

Enquanto trabalhava, Schultz (1973) se sentiu meio perplexo pelo fato dos seus conceitos, utilizados para analisar capital e trabalho, estarem se revelando desapropriados para explicar os acréscimos que estavam ocorrendo na produção:

> Durante o ano de minha permanência no Centro, comecei a perceber que os fatores essenciais da produção, que eu identificava como capital e trabalho, não eram imutáveis: sofriam um processo de aperfeiçoamento, o que não era devidamente apreciado, segundo a minha conceituação de Capital e Trabalho. (SCHULTZ, 1973, p.10).

Já não estava sendo possível explicar o crescimento somente através dos fatores tradicionais da produção. Para Schultz (1973), a resposta estaria na melhora da qualidade dos fatores, humanos e mecânicos. Considerava que parte significante da eficácia da produção seria resultado da educação formal e ampliava-se na experiência do trabalho. Os maiores salários seriam pagos aos trabalhadores com maior produtividade e, portanto, que possuíam maior nível de escolaridade. Para o autor, a educação é um investimento em capital humano, podendo ser avaliado pelas suas taxas de retorno positivas. Sendo assim, o investimento em capital humano poderia esclarecer confusões em torno do crescimento econômico e as

alterações nas estruturas de pagamentos e salários. Isso, porque as habilidades adquiridas influenciam no montante dos ganhos do trabalho.

Ele percebeu também, que muitas pessoas estavam investindo em si mesmas. E que esse investimento humano estava constituindo em uma forte influência no crescimento econômico. Foi então que definiu: “o investimento básico no capital humano é a educação.” (SCHULTZ, 1973, p.10).

Em seu estudo, menciona que o conceito de educação está intimamente ligado à cultura de cada sociedade. Mas, há características desse conceito que são constantes entre as culturas, que são o ensino e o aprendizado:

> Assim, educar significa, etimologicamente, revelar ou extrair de uma pessoa algo potencial e latente: significa aperfeiçoar uma pessoa, moral e mentalmente, de maneira a torná-la suscetível de escolhas individuais e sociais, e capaz de agir em consonância; significa prepará-la para uma profissão, por meio de instrução sistemática; e por fim, significa exercitar, disciplinar ou formar habilidades, como, por exemplo, aperfeiçoar o gosto de uma pessoa. (SCHULTZ, 1973, p.18).

Além disso, segundo esse autor, a educação seria um dos setores estáveis da economia, podendo exercer certa influência reguladora. Por exemplo, em meio a crises, em que o desemprego aumenta, ela seria capaz de absorver alguns recursos suplementares como estudantes adultos; em momentos de recuperação, poderia liberar recursos.

Assim como as bases da política do livre comércio e da moderna política fiscal e monetária, os estudiosos do investimento da educação introduzem as bases de uma política de crescimento econômico, a qual confere uma maior importância ao processo de aprendizagem.

A relação entre educação e desenvolvimento econômico estaria baseada nas habilidades para o trabalho adquiridas no ensino, que refletiria na produtividade. Então, o ponto central dessa teoria seria que a educação altera as habilidades para o trabalho, aumentando a produtividade e, consequentemente, evoluindo a renda e o crescimento econômico:

> Sempre que a instrução elevar as futuras rendas dos estudantes, teremos um investimento. É um investimento no capital humano, sob a forma de habilidades adquiridas na escola. Existem numerosos investimentos no capital humano e as cifras tornam-se elevadas. Pode-se dizer, na verdade, que a capacidade produtiva do trabalho é, predominantemente, um meio de produção produzido. (SCHULTZ, 1973, p. 25).

Se a educação fosse gratuita, as pessoas investiriam nela até o ponto em que o seu rendimento se tornasse nulo. Mas Schultz (1973) menciona que a inversão em educação

implica em despesas. E que, esses custos são mais onerosos do que se imagina: “contudo, o que não é conhecido é que a instrução se apresenta muito mais onerosa do que consta nas despesas escolares, considerando-se que os salários não recebidos pelos estudantes adultos não são consignados como despesas.” (SCHULTZ, 1973, p. 36).

Ao empenhar-se às atividades escolares, o aluno deixa de exercer certa atividade remunerada. Schultz (1973) chama esse custo de oportunidade de “salários não-recebidos”. Esse custo da educação tem sido ignorado por muito tempo. Uma possível explicação para esse fato seria que, no passado, a principal função das universidades seria de capacitar um modesto grupo de estudantes para cargos de elite. A maioria desses estudantes já pertencia a famílias com bons padrões de vida.

Para analisar economicamente o valor da educação, Schultz (1973) estabelece uma distinção entre valores consuntivos e produtivos. As contribuições consuntivas do ensino são divididas segundo aquelas que atendem o consumo presente e as que atendem o consumo futuro; estas últimas são consideradas como investimento. Os valores produtivos são investimentos em futuras capacidades de criar e receber rendimentos.

Se os produtos do ensino fossem destinados somente ao consumo presente, a educação não contribuiria para o crescimento econômico. O resultado seria a elevação da procura por bens de consumo, dadas as modificações na renda. Por mais que essas modificações sejam benéficas para o consumidor, não seriam calculadas na Renda Nacional.

Schultz (1973) acrescenta que se exige um grande esforço para que um país atinja um elevado nível de educação. Ademais, se um nível muito baixo de instrução não pode ser aumentado, a educação não pode ser considerada um determinante do crescimento econômico. Mas, “partindo-se de um nível baixo, existem grandes possibilidades de se poder elevá-lo.” (SCHULTZ, 1973, p.59).

A educação pode ser considerada fator de crescimento econômico apenas quando ela aumenta a produtividade e os lucros futuros. Sendo assim, o autor considera necessário elencar as principais atribuições das instituições de ensino. São elas:

- A pesquisa é uma das funções tradicionais.
- Descobrir e cultivar o talento potencial: os benefícios do investimento na exploração de petróleo e no aprimoramento das técnicas extrativas são evidentes. De maneira similar, são os proveitos de um complexo educativo com a finalidade de revelar talento humano, assim como aprimorar essa habilidade.

- Desenvolver a capacidade de adaptação das pessoas às oportunidades de emprego, relacionadas ao crescimento econômico: as alterações nos meios de produção e nos níveis econômicos também podem modificar as estruturas do mercado de trabalho. A educação, nesse contexto, pode contribuir flexibilizando os reajustamentos ocupacionais.
- Em uma economia de crescimento considerado rápido, as instituições de ensino teriam a atribuição de propiciar a mão-de-obra qualificada.

Para Schultz (1973), entre os anos de 1909 e 1929, a educação ocupou um papel menos importante no crescimento, do que após esse período. Já nos últimos anos, a educação tem sido uma fonte de crescimento maior do que o capital imobilizado. E a perspectiva é que ela continue a ser um importante elemento de crescimento econômico nas próximas décadas.

Dessa forma, depreende-se que o ensino compõe uma poderosa fonte de crescimento econômico. É claro que não constitui em fonte natural, mas produzida pelo próprio homem, envolvendo investimentos. A análise do total desses investimentos e a sua taxa de rendimento seriam suficientes para compreender os seus efeitos sobre o crescimento da economia. Porém, há características do sistema econômico que requer a definição de capital humano.

A obstinação no conceito de capital humano baseia-se, principalmente, em alguns problemas no comportamento da economia. O conceito de capital, limitado a máquinas, equipamentos, terra e patrimônio é insuficiente para explicar o crescimento econômico e as obtenções do bem-estar geradas por esse crescimento:

> Para o estudo do crescimento econômico, abrangendo largos períodos, e entre comunidades tão diferentes, o conceito de capital e de formação de capital deveria ser ampliado, de forma a incluir investimento para saúde, educação e adestramento da própria população, isto é, investimento em seres humanos. (KUZNETS, 1961 apud SCHULTZ, 1973, p.64).

Schultz (1973) expõe outro argumento a favor da importância da definição de capital humano. Ele cita que o investimento em bens de produção (estruturas, equipamentos, máquinas) vem se reduzindo em comparação ao realizado nas habilidades humanas. Portanto, um conceito de capital restrito à ideia de imobilizado pode reter atenção em concepções que não são fundamentais para entender o crescimento econômico. Além disso, tal como as estruturas, os equipamentos e as máquinas, o capital humano deprecia e exige reinvestimento.

Os proprietários do capital humano seriam os trabalhadores, que não podem se separar de suas habilidades. Quanto mais conhecimento adquirirem, melhor será a produtividade.

Schultz (1973) afirma que os investimentos em educação devem ser de boa qualidade e feitos de forma correta para que ela resulte em crescimento econômico.

Algumas circunstâncias, segundo o autor, apoiam o conceito de educação como investimento. Uma delas é o caso do crescimento econômico. A outra diz respeito às modificações na estrutura de pagamentos e salários. A terceira, em especial no caso dos Estados Unidos, é o declínio da desigualdade na distribuição da renda pessoal.

A primeira circunstância, que tem relação com o crescimento econômico, foi descrita com certo detalhe nos parágrafos anteriores que, em síntese, é: se a renda altera, enquanto a proporção do capital permanece substancialmente constante, depreende-se que essa alteração é provocada por outras formas de capital, ainda não contempladas, composta principalmente de capital humano.

A segunda circunstância, sobre as modificações na estrutura de pagamentos e salários, relaciona-se com a produtividade advinda da educação. O ensino proporciona habilidades para o trabalho, que se transformam em produtividade. Uma maior renda seria paga aos indivíduos com maiores produtividades, alterando assim, a estrutura de pagamentos e salários.

Quanto ao papel da educação na distribuição de renda, Schultz (1973) apresenta um pressuposto. O investimento em capital humano, como foi exposto, é capaz de aumentar a renda. Uma distribuição mais equitativa desse investimento pode nivelar os salários. Considerando essas proposições, as alterações do investimento no capital humano consistem em um princípio relevante na redução das desigualdades na distribuição de renda.

Schultz (1973) reafirma o papel da educação nesses contextos: “Em cada circunstância – crescimento, estrutura de pagamento e distribuição de renda – as alterações dos níveis de instrução transformam-se em ponderável fator determinante.” (SCHULTZ, 1973, p. 69).

Nos seus estudos, Schultz (1973) cita que se encontrava, nos Estados Unidos, consistentes provas de que as mais importantes fontes de crescimento consistiam na instrução e pesquisa universitária. Ele faz referência ainda ao trabalho de Horvat (1958), que aponta o conhecimento e a habilidade como investimentos cruciais na determinação do crescimento econômico em países “subdesenvolvidos”.

Além de abordar a educação de uma forma mais ampla, Schultz (1973) faz uma referência à educação primária, considerada profundamente compensatória pela contribuição no crescimento do consumo do estudante e de sua família. Ela amplia os benefícios da produção e as vantagens do dispêndio; e possibilita elevar os futuros salários e gerar oportunidades para incrementar treinamentos práticos. Complementando: “Não é de se

estranhar, portanto, que tenhamos, há longo tempo, adotado a generalização da educação primária.” (SCHULTZ, 1973, p. 75).

Segundo o autor, existe uma contundente suposição que mais e melhor instrução primária consistiria no melhor investimento, em termos de produção e lucros, para os países avaliados como extremamente pobres. Ele usou estimativas da Venezuela para confirmar essa suposição, onde o custo total pôde ser considerado como investimento na capacidade de auferir rendas.

Após essa abordagem sobre o capital humano, na ótica de Schultz, é apresentada, de forma sucinta, a formulação dessa teoria, segundo Becker (1993).

1.2.2 A formulação de Becker

Gary S. Becker foi outro teórico do capital humano e prêmio Nobel em Economia de 1992. Ele discute o tema através de uma visão mais sistemática, analisando o custo-benefício das atividades que aumentam a produtividade. Os seus estudos tiveram início na estimativa do rendimento da educação universitária e secundária nos Estados Unidos.

De acordo com Becker (1993), as tentativas de investimento associadas à instrução seriam semelhantes a outros investimentos pessoais. Esses investimentos possuiriam propriedades comuns e seriam dependentes da teoria moldada para investimentos em estruturas e equipamentos. Ele defende que essa teoria deveria ser revisada.

Na sua visão, a teoria do investimento carece de uma reformulação, na qual, o conceito de investimento humano seja incorporado. Ele ainda cita que o investimento em educação colabora com o crescimento econômico, beneficia os rendimentos individuais e a saúde das famílias. Uma das implicações de sua teoria seria a relação entre os investimentos em capital humano e a redução da pobreza.

A educação poderia ser um determinante da saúde e da expectativa de vida de uma pessoa, não somente em razão de aumentar sua renda, mas também por permitir escolhas mais saudáveis. As taxas de retorno dos investimentos em capital humano seriam maiores, quanto mais tempo de trabalho e saúde tivesse o indivíduo. Para Becker (1993), os investimentos em capital humano constituíam em uma das maneiras mais efetivas de aumentar os níveis de renda dos pobres bem como de sua saúde.

Ainda segundo Becker (1993), é preciso que o governo amplie a educação para todos, ofertando-a com qualidade. A atividade educacional seria componente da produção; em razão disso, merece atenção especial dos planejadores de políticas educacionais.

Quando se pensa em capital, a maioria das pessoas tem em mente conceitos tangíveis. Porém, segundo esse teórico, contas bancárias, ações e edificações não seriam as únicas formas de capital, no sentido de produzirem renda. Para ele:

> Escolaridade, um curso de treinamento de computador, despesas com assistência médica e palestras sobre as virtudes da pontualidade e honestidade são também capital. Isso, porque aumentam os lucros, melhoram a saúde ou adicionam bons hábitos de uma pessoa sobre grande parte da sua vida. (BECKER, 1993, p. 15, tradução nossa).

Educação, treinamento e saúde são os investimentos mais importantes em capital humano, segundo Becker (1993). Seu livro *"Human capital: a theoretical and empirical analysis, with special reference to education"* e outros estudos mostraram que a escolaridade tem aumentado significativamente a renda das pessoas, mesmo após a exclusão dos seus custos diretos e indiretos. Considerando dados de vários países, com culturas e sistemas econômicos diferentes, os rendimentos das pessoas mais instruídas são quase sempre acima da média.

Becker (1993) cita uma preocupação que surgiu após alguns estudos sobre educação nos Estados Unidos. A vantagem dos salários de graduados do ensino médio sobre aqueles que haviam abandonado a escola aumentara consideravelmente. Esse fato levou a discussão sobre se os Estados Unidos forneciam qualidade e quantidade da educação de forma adequada.

Outro fato relevante citado por esse autor é o aumento e a diminuição, ao longo do tempo, da fração de diplomados do ensino médio que vão para o ensino universitário. Refletir sobre o ensino superior como capital humano facilita o entendimento dessas oscilações. Na década de 1970, os benefícios de um diploma universitários reduziram-se. A fração de diplomados do ensino médio que começou a faculdade, nesse período, também caiu. Na década de 1980, esperava-se que o número de matrículas continuaria a declinar. Essa expectativa, principalmente dos educadores, baseava-se na queda do número de jovens de dezoito anos e no aumento da mensalidade universitária. Suposição que não sobressaiu, pois, na década, a fração de formados no ensino médio que entrou na faculdade aumentou. Para o autor, isso fazia sentido. Os benefícios da educação universitária aumentaram entre 1980 e 1990.

Segundo Becker (1993), mensalidades e taxas, para a maioria dos estudantes, não seriam o maior custo de se ingressar na faculdade. Grande parte dos custos de se dedicar à

educação superior seria o salário que o estudante universitário deixa de receber por não trabalhar.

A economia do capital humano também proporcionou mudanças na educação das mulheres. Becker (1993) cita o aumento do ingresso de mulheres na faculdade e a redução da diferença dos salários entre mulheres negras e brancas. Além disso, a educação formal não seria a única forma de investir em capital humano. Os trabalhadores também aprendem e são treinados fora da escola, por exemplo, em programas de capacitação. O mencionado autor faz referência ao estudo de Mincer (1988), o qual indica que o treinamento no trabalho seria uma importante fonte de aumentos nos salários, uma vez que os trabalhadores ganham mais experiência.

Becker (1993) aponta a influência da família nos conhecimentos, habilidades, saúde, valores e hábitos dos seus descendentes. Os pais afetam várias áreas da vida de seus filhos, inclusive, salário, educação e ocupação. A relação positiva entre os rendimentos dos pais e filhos seria menos forte comparada à relação entre os anos de escolaridade dos pais e seus filhos.

O crescimento da renda *per capita* de muitos países durante os séculos XIX e XX é, em parte, devido à expansão do conhecimento científico e técnico que aumenta a produtividade do trabalho e de outros insumos na produção, dita Becker (1993). Segundo ele, a indústria seria dependente de conhecimento sofisticado. Essa crescente dependência reforça significativamente o valor da educação, ensino técnico, programas de treinamento/capacitação e de diversos outros tipos de investimentos em capital humano.

Em países que têm pouca mão de obra qualificada, os novos avanços tecnológicos possuem seu valor reduzido, visto que, são raros os trabalhadores que sabem utilizá-los. A coesão entre inovações tecnológicas e investimento em capital humano é muito importante. O desenvolvimento econômico depende, significativamente, dessa associação. Para Becker (1993), vários países alcançaram um desenvolvimento econômico considerável devido aos avanços no conhecimento tecnológico serem acompanhados de educação e formação da população.

Fica evidenciado, portanto, que o aumento do investimento no capital humano é um importante elemento para o desenvolvimento econômico. Além disso, Becker destaca os ganhos sociais e privados dos investimentos educacionais.

O estudo sobre o capital humano influenciou os estudiosos de tal maneira que, segundo Cunha (1985), na década de 60, mais especificamente no Brasil, houve certa pretensão em privilegiar mais o investimento em capital humano do que em recursos físicos.

Através do que foi exposto, percebe-se que os autores, além de reafirmarem os diversos benefícios da educação e sua relação com o desenvolvimento econômico, também enfatizam a importância da qualidade dessa educação.

No próximo tópico são apresentados alguns estudos sobre os determinantes do desempenho escolar; pois este está diretamente relacionado à formação do capital humano, o qual é de suma importância para o desenvolvimento econômico, como exposto.

1.3 Revisão de literatura: determinantes do desempenho escolar

O objetivo do trabalho de Barbosa e Fernandes (2001) foi investigar os efeitos da escola na proficiência em matemática no ensino fundamental, especificamente, para alunos da 4ª série. Para tanto, os autores controlaram variáveis a nível de aluno e de turma; sendo utilizados dados do Sistema de Avaliação da Educação Básica (SAEB) de 1997 para a região Sudeste. Quanto à metodologia, foram utilizados os modelos multiníveis (hierárquicos). Os autores constataram que a escolaridade do pai, a infraestrutura escolar e a escolaridade do professor impactaram positivamente e significativamente a proficiência do aluno. A defasagem escolar, que segundo os autores reflete a repetência e a entrada tardia na escola, apresentou relação negativa com o rendimento escolar, ou seja, quanto maior for o nível de defasagem, menor é o rendimento do aluno. Outro resultado verificado foi que os alunos que conversam sobre a escola em casa apresentaram, em média, desempenho melhor do que aqueles que não conversam.

Em seu estudo, Soares e Mendonça (2003) objetivaram identificar a relação entre o desempenho escolar, as características da escola e o perfil socioeconômico dos alunos da 4ª série do Ensino Fundamental. Utilizaram dados do Sistema Mineiro de Avaliação da Educação Pública (SIMAVE). Empregou-se como metodologia a regressão hierárquica, que mostrou a existência de porcentagem da variação do desempenho escolar que é devido a diferenças entre as turmas e outra porcentagem da variação do rendimento que é devido a diferenças entre as escolas. Apresentaram impacto positivo no desempenho escolar o nível socioeconômico - medido em nível de escola - e a escolaridade da mãe. A defasagem escolar apresentou impacto negativo na proficiência média do aluno. Os autores destacaram a significância da variável que mede a frequência do aluno a cultos religiosos, que afetou positivamente o rendimento escolar. Outra variável que impactou positivamente o desempenho escolar é a frequência com que os professores da escola passam lições de casa, variável de interesse para a atuação dos órgãos públicos. Os autores ainda constataram que o

nível socioeconômico da escola potencializou o efeito negativo da defasagem escolar sobre o rendimento do aluno.

O estudo de Macedo (2004) teve como objetivo analisar fatores associados ao rendimento escolar dos alunos da 5ª série do ensino fundamental. Além disso, a autora analisou os efeitos sobre esses fatores da inclusão do valor adicionado dos alunos da 5ª série em relação aos alunos que tiveram sucesso na 4ª série de 1999. O rendimento escolar foi representado pelas notas obtidas nos exames padronizados, constantes na base de dados "Avaliação de desempenho: fatores associados" - CEDEPLAR/INEP. Como metodologia, foram empregados modelos de regressão linear. Um dos resultados foi de que, com a inclusão da nota do teste aplicado na 4ª série, a variável sexo perde poder explicativo e a variável raça perde sua significância. Segundo a autora, fazer lição de casa se mostrou um fator importante para a obtenção de boas notas em matemática e português. O fato de o aluno trabalhar acarretou rendimentos escolares mais baixos do que daqueles que não trabalhavam, comprometendo mais fortemente as notas de português do que as notas de matemática. Alunos que leem "histórias em quadrinho" ou "jornais e revistas" apresentaram, em média, rendimentos melhores em relação aos alunos sem esse hábito de leitura. Outra constatação foi de que a repetência refletiu em menores proficiências. A autora não encontrou significância para a escolaridade do pai, porém, a escolaridade da mãe foi capaz de influenciar positivamente o rendimento do aluno. Vale ressaltar que o valor adicionado eliminou a significância da escolaridade da mãe. Ademais, constatou-se, como esperado, o efeito positivo na proficiência do aluno de uma escola melhor estruturada e equipada.

O objetivo do estudo de Soares (2004) foi o de apresentar os fatores escolares associados ao desempenho estudantil, controlando a influência de variáveis relacionadas aos alunos. O autor utilizou dados do SAEB, referentes ao teste de matemática da 8ª série do ensino fundamental. A metodologia utilizada foi o modelo de regressão hierárquico com dois níveis, sendo inseridas variáveis referentes ao aluno no nível 1 e variáveis da escola no nível 2. Entre os resultados, verificou-se que os fatores escolares podem explicar 12,3% da variância total do desempenho dos alunos. O autor cita que apesar da maior parte da variação da proficiência ser atribuída às características do próprio aluno, é possível melhorar o desempenho dos alunos modificando as estruturas escolares. Outra constatação foi de que alunos mais favorecidos aproveitaram de forma mais eficiente a melhoria das condições escolares.

Na dissertação de Pinheiro (2005), objetivou-se avaliar o desempenho escolar a partir de variáveis relacionadas aos alunos e às escolas. A autora utilizou dados do Sistema de

Avaliação Educacional de Pernambuco (SAEPE), referentes aos alunos da 4ª e 8ª séries do ensino fundamental e para os alunos da 3ª série do ensino médio. Quanto à metodologia, utilizou-se o modelo de regressão hierárquico. Entre os resultados, constatou-se que alunos com defasagem idade-série (alunos que não cursam a série na idade correta) possuíram, em média, notas menores do que os alunos que cursaram a série na idade adequada. A autora também verificou que alunos do sexo masculino tenderam a apresentar notas maiores em matemática em comparação aos alunos do sexo feminino. Na disciplina de Língua Portuguesa ocorreu o contrário, os alunos do sexo feminino tenderam apresentar rendimentos melhores do que os alunos do sexo masculino. Outro resultado foi de que o tipo de escola (municipal ou estadual) apresentou influência no rendimento dos alunos. Na 4ª série do ensino fundamental, os alunos de escolas municipais apresentaram, em média, desempenho inferior em relação aos alunos de escolas estaduais. Na 8ª série do ensino fundamental ocorreu o oposto; alunos de escolas estaduais tenderam a apresentar desempenho inferior aos alunos de escolas municipais. O estudo revelou ainda que alunos que estudavam em escolas com bom estado de conservação apresentaram melhores desempenhos em relação àqueles cujo estado de conservação do prédio escolar não era bom. (PINHEIRO, 2005, p. 69). Além disso, a autora relatou que escolas localizadas em municípios com maiores Índices de Desenvolvimento Humano (IDH) tenderam a apresentar melhor desempenho.

O objetivo do estudo de Luz (2006) foi realizar uma análise dos determinantes da proficiência de alunos da rede pública e urbana de ensino nas regiões norte, nordeste e centro-oeste do Brasil. Utilizou-se dados do Instituto Nacional de Estudos e Pesquisas Anísio Teixeira (INEP), do período 2000-2001. Como metodologia, foi utilizada a análise multivariada de correspondência múltipla, especificamente, a análise HOMALS (*Homogeneity Analysis by Means of Least Squares*). Uma das constatações da autora foi de que os fatores familiares e individuais apresentaram uma maior efetividade em relação às características escolares. Segundo o referido estudo, alunos do sexo masculino apresentaram melhor desempenho nas provas de matemática do que alunos do sexo feminino; enquanto o fato do aluno trabalhar fora de casa foi um fator que contribuiu negativamente na proficiência escolar. A repetência estaria negativamente relacionada ao desempenho do aluno, sendo que os repetentes podem ter sua proficiência prejudicada devido à perda de motivação e/ou pelo atraso causado pela reprovação. (LUZ, 2006, p. 16). Os resultados desse trabalho apontaram, também, que o tempo percorrido da casa até a escola prejudica o rendimento do aluno. Conforme apontou a autora, a escolaridade da mãe influencia positivamente o desempenho do aluno; além disso, para que essa influência se mantenha na medida em que o aluno evolui de

série, uma maior escolaridade da mãe também é requerida. Constatou-se que o nível socioeconômico do aluno também influencia seu rendimento e; alunos da rede municipal tendem a apresentar resultados inferiores em relação aos alunos da rede estadual. Além disso, foi verificado que o desempenho do aluno aumenta à medida que a infraestrutura escolar melhora.

No estudo de Biondi e Felício (2007), o objetivo foi identificar variáveis escolares que poderiam ser alvo de políticas educacionais com o propósito de melhoria do desempenho escolar. Além disso, as autoras empregaram variáveis de controle referentes aos alunos. Utilizaram dados do SAEB e do Censo Escolar. A metodologia empregada foi o modelo de regressão estimado pelo método de Mínimos Quadrados Ordinários e de Efeitos Fixos. Foi constatado que as escolas com maior proporção de alunos que trabalhavam fora, apresentaram, em média, proficiências menores. Outro resultado foi de que quanto maior a proporção de alunos que possuem computador em casa, melhor o desempenho da escola. As autoras concluíram, também, que o fato da mãe possuir ensino fundamental completo é capaz de influenciar positivamente o rendimento do aluno. Da mesma forma, alunos que moravam com o pai e a mãe, em média, apresentaram desempenho melhor, em relação àqueles que não residiam com seus pais. Professores que possuíam no mínimo ensino superior completo foram capazes de influenciar melhores rendimentos dos alunos. A experiência do diretor da escola não foi significativa para influenciar o desempenho dos alunos, ao passo que menor rotatividade dos professores proporciona, em média, maiores rendimentos escolares. As escolas com biblioteca e laboratório de ciências apresentaram desempenho melhor em relação àquelas desprovidas desses recursos. Vale ressaltar que Biondi e Felício encontraram um sinal negativo para o coeficiente da variável laboratório de informática. Por isso, as autoras advertem que os laboratórios de informática podem estar sendo usados inadequadamente, alocando o tempo dos estudantes de forma equivocada. (BIONDI e FELICIO, 2007, p. 16). Outro fato interessante no referido estudo é que o número de alunos na turma não foi significativo no sentido de explicar o rendimento escolar.

O trabalho de Menezes-Filho (2007) teve como objetivo examinar o desempenho dos alunos da 4ª e 8ª séries do ensino fundamental e da 3ª série do ensino médio nos testes de proficiência em Matemática, bem como identificar fatores associados a um melhor desempenho escolar. Utilizou-se dados do SAEB de 2003 e aplicou-se regressões econométricas. Entre os resultados, o autor concluiu que a escola explica entre 10% e 30% das diferenças das notas dos alunos e o restante da variação é devida às características do estudante e de sua família. Dessa forma, as variáveis que mais explicaram o rendimento

escolar foram aquelas relacionadas às características familiares e do aluno. (MENEZES-FILHO, 2007, p.20). Nas escolas públicas, o estado de conservação da escola reflete em um melhor desempenho do aluno e o tamanho da turma não foi significativo no sentido de explicar o desempenho escolar. Outra constatação do autor foi de que a presença de computadores para os alunos, para os diretores e para o pessoal administrativo tem pouco impacto sobre o desempenho dos alunos. Comparando escolas estaduais e municipais, os alunos da rede municipal tendem a apresentar resultados melhores. Outro resultado importante foi o de que escolas com maiores porcentagens de alunos que recebem bolsa família tendem a apresentar desempenho reduzido. Além disso, a idade do professor está positivamente relacionada com o desempenho dos alunos, enquanto a escolaridade do docente apresentou pouco efeito sobre o rendimento escolar. O salário do professor e o fato de lecionar em mais de uma escola não tiveram impacto sobre a proficiência do estudante. Alunos que fizeram pré-escola apresentaram, em média, desempenho melhor do que os que não fizeram, e a repetência em séries anteriores são capazes de provocar reduções na proficiência atual. Ademais, as seguintes variáveis apresentaram relação positiva com o rendimento escolar: escolaridade da mãe, morar com os pais (ou com pelo menos um deles), o fato do aluno ler livros ou jornais e possuir um ou mais computadores em casa. No caso do aluno trabalhar dentro de casa mais do que 4horas ou trabalhar fora de casa, seu rendimento escolar, em média, reduz.

No trabalho de Schiefelbein e Simmons (2013), foram analisadas várias pesquisas sobre os determinantes do desempenho cognitivo dos estudantes nos países em desenvolvimento. O estudo foi recomendado pelo *Educational Review and Advisory Group*. Nas pesquisas analisadas, foram utilizados métodos antropológicos, psicológicos, sociológicos, econômicos, linguísticos e estatísticos. Para os referidos autores, os planejadores e educadores possuem especial interesse nesses estudos, pois estes são capazes, muitas vezes, de aferir o efeito relativo dos investimentos que, teoricamente, influenciam a vida do estudante. (SCHIEFELBEIN e SIMMONS, 2013, p. 54). Como resultado, averiguaram que em cinco das oito pesquisas que analisaram a relação do custo por estudante e o desempenho escolar, maiores despesas por aluno não estavam associadas a melhores desempenhos. De 14 estudos, nove constataram que classes maiores estavam associadas a melhor desempenho ou o tamanho da classe não o afetava; porém, os autores advertem sobre esse resultado. Entre 10 pesquisas, sete constataram relação positiva entre a disponibilidade de compêndios e o desempenho do aluno. Um resultado interessante encontrado pelos pesquisadores foi de que o "dever de casa" foi um dos determinantes mais significativos do desempenho escolar,

afetando-o positivamente; conforme seis entre oito estudos. Outro resultado foi de que entre 32 estudos, 19 não identificaram diferenças no rendimento dos alunos que possuíam professores graduados daqueles alunos cujos professores não eram diplomados. Na maioria das pesquisas analisadas por Schiefelbein e Simmons, o nível sócio-econômico do aluno teve um impacto positivo e significativo no rendimento escolar; bem como o nível de instrução dos pais e o acesso a livros e a brinquedos. Segundo os estudos, maiores níveis de repetência influenciam negativamente o resultado do estudante; e o acesso a jardins de infância influenciam significativamente o seu desempenho.

O próximo capítulo aborda as metodologias de análise, bem como os dados utilizados no presente estudo.

2. Metodologias de análise: Análise de Correlação, Análise Exploratória de Dados Espaciais, Teoria de Resposta ao Item e Modelo de Regressão Hierárquico

Segundo Gil (2010), classificar as pesquisas é uma atividade importante. Ao classificar seu estudo, "o pesquisador passa a dispor de mais elementos para decidir acerca de sua aplicabilidade na solução dos problemas propostos para investigação." (GIL, 2010, p. 24). Quanto ao ambiente de pesquisa e à abordagem teórica, pode se classificar esse trabalho como bibliográfico, uma vez que Gil (2010) destaca que será fonte bibliográfica quando o material for obtido em bibliotecas ou bases de dados.

Consoante com seus objetivos mais gerais e com as técnicas de coleta e análise de dados, essa pesquisa pode ser classificada como descritiva, considerando que "também são pesquisas descritivas aquelas que visam descobrir a existência de associações entre variáveis, como, por exemplo, as pesquisas eleitorais que indicam a relação entre preferência político-partidária e nível de rendimentos ou escolaridade." (GIL, 2010, p. 28).

Conforme a finalidade, o presente trabalho pode ser classificado como uma pesquisa aplicada. De acordo com Gil (2010), esse tipo de pesquisa é voltado à aquisição de conhecimentos a fim de aplicação numa situação específica.

2.1 Análise de Correlação

2.1.1 Coeficiente de correlação de Pearson

O coeficiente de correlação de Pearson é uma medida estatística desenvolvida para medir o grau de associação ou de relacionamento linear entre duas variáveis quantitativas. Essa medida não especifica a direção de causalidade, mas designa o quão próximo duas variáveis aleatórias variam juntas.

Este coeficiente pode ser calculado da seguinte maneira:

$$r = \frac{1}{n-1}\sum\left(\frac{x_i - \bar{X}}{S_x}\right)\left(\frac{y_i - \bar{Y}}{S_y}\right)$$

Onde:

n = número de observações;

x_i = i-ésima observação de x; y_i = i-ésima observação de y;

$\bar{X}$= média de x; $\bar{Y}$ = média de y;

S_x = desvio padrão de x; S_y = desvio padrão de y;

$$S_x = \sqrt{\frac{1}{n-1}\sum(x_i - \hat{X})^2}$$

$$S_y = \sqrt{\frac{1}{n-1}\sum(y_i - \hat{Y})^2}$$

O coeficiente de correlação, r, é um número que situa entre -1 e 1, ou seja, $-1 \leq r \leq 1$. O valor indica a força da relação linear entre as variáveis e o sinal expressa a direção (positiva ou negativa) do relacionamento. Quando r assume valores positivos ($r > 0$), significa que aumentos nos valores de uma variável estão associados a aumentos nos valores da outra variável. Ao passo que, se r assume valores negativos ($r < 0$), significa que aumentos nos valores de uma variável estão associados à redução nos valores da outra variável.

Atualmente, os *softwares* estatísticos apresentam, junto com o valor do coeficiente de correlação, o nível de significância com o qual se rejeita, ou não, a hipótese nula de ausência de correlação linear entre as variáveis[1].

2.1.2 Coeficiente de correlação por postos de Spearman

O coeficiente de correlação por postos de Spearman é adequado quando as variáveis são mensuradas em nível ordinal (Lira e Neto, 2006, p. 47). Seu cálculo é derivado do estimador do coeficiente de correlação linear de Pearson.

Supondo duas variáveis, x e y, o coeficiente de correlação pode ser expresso por:

$$\hat{p}_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)}$$

Onde:

$\hat{p}_s$ = coeficiente de correlação de Spearman;

[1]Hipóteses: H0 (hipótese nula) indica ausência de correlação linear entre as variáveis; H1 (hipótese alternativa) indica existência de correlação linear entre as variáveis. Geralmente, utilizam-se os níveis de significância 1%, 5% e 10%. Para valores maiores que 10% não se rejeita a hipótese nula.

d_i (diferença entre as ordenações) = (posto de x_i dentre os valores de x) - (posto de y_i dentre os valores de y); e
n = número de pares de ordenações (x_i, y_i).

Assim como o coeficiente de correlação de Pearson, o coeficiente de correlação por postos de Spearman varia entre -1 e 1, ou seja, $-1 \leq \hat{p}_s \leq 1$. A magnitude da correlação de Spearman aumenta conforme se preserva ou se inverte a relação de ordem entre as variáveis. O sinal positivo do coeficiente indica que os valores de uma variável tendem a aumentar quando os valores da outra variável aumentam. Analogamente, o sinal negativo indica que os valores de uma variável tendem a diminuir quando os valores da outra variável aumentam.

2.2 Análise espacial de dados

A análise espacial de dados pode ser caracterizada como um conjunto de técnicas que permitem visualizar distribuições espaciais, bem como identificar associações espaciais e *clusters*. Por meio da análise exploratória de dados espaciais (AEDE), uma das técnicas de análise espacial, é possível conhecer mais precisamente a natureza da interação do processo espacial.

2.2.1 Análise exploratória de dados espaciais (AEDE) e testes de autocorrelação espacial

Segundo Almeida (2004), a AEDE baseia-se em estatísticas de informação geográfica, capazes de descobrir padrões espaciais nos dados:

> A AEDE é capaz de prover claras dicas e indicações sobre a existência de padrões de associação espacial – tanto em âmbito global quanto local – ou sobre a presença de *clusters* nos dados, ou, ainda, sobre a influência de observações discrepantes (*outliers*). (ALMEIDA, 2004, p. 34).

A AEDE é composta das seguintes técnicas de análise: i) associação espacial global univariada; ii) associação espacial global multivariada; iii) associação espacial local univariada; iv) associação espacial local multivariada; v) análise de *clusters* espaciais e vi) detecção de *outliers* espaciais.

Aliado a essas técnicas de análise estão os testes formais de autocorrelação espacial. Entre os testes estatísticos de associação espacial global estão o *I* de Moran Global e o C de Geary. Para os padrões de associação espacial local usa-se as estatísticas LISA - Índice Local de Associação Espacial, sendo o *I* de Moran Local a principal.

A existência de padrões globais de associação espacial pode estar em conformidade com padrões locais, mas é possível tal situação não ocorrer. Segundo Almeida (2004), o indício de ausência de autocorrelação global pode ocultar padrões de associação local. “O caso oposto ocorre quando uma forte indicação de autocorrelação global pode camuflar padrões locais de associação.” (ALMEIDA, 2004, p. 42).

Com o intuito de superar a limitação dos padrões globais de associação, descrita no parágrafo anterior, desenvolveu-se indicadores de padrões locais de associação espacial. Diante disso, no presente estudo, utilizar-se-á o *I* de Moran local bivariado.

2.2.1.1 *I* de Moran local bivariado

A ideia por trás do *I* de Moran local bivariado é a de verificar se os valores de uma determinada variável de certa unidade espacial guarda alguma relação com outra variável de unidades espaciais vizinhas. O *I* de Moran local bivariado decompõe o indicador global de associação espacial de uma região através da contribuição de cada unidade dessa região em estudo. O *I* de Moran local bivariado pode ser expresso por:

$$I_{kl}^{i} = z_k^i \sum_j w_{ij}\, z_l^i$$

Onde:

I_{kl}^{i} é o indicador de associação espacial local Moran da unidade espacial i em relação às variáveis de interesse z_k e z_l;

z_k^i é a variável de interesse k da unidade espacial i;

w_{ij} é o peso espacial entre as unidades espaciais i e j, mede o grau de interação entre elas; e

o somatório sobre j garante que entrem no cálculo, somente os valores das unidades espaciais vizinhas da observação espacial i.

O *I* de Moran local distribui cada unidade espacial em quatro categorias, correspondendo a um quadrante no diagrama de dispersão de Moran. Essas quatro categorias, bem como o diagrama de dispersão de Moran, serão descritas no próximo tópico.

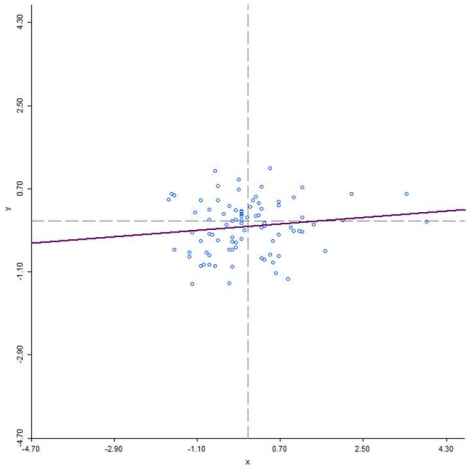
2.2.1.2 Diagrama de dispersão de Moran, mapa de significância e mapa de *clusters*

O diagrama de dispersão de Moran é uma ferramenta de visualização do comportamento de associação espacial. De acordo com Almeida (2004):

> O diagrama de dispersão de Moran multivariado plota, no eixo das abscissas, os valores de uma variável, observados numa determinada região, com o valor médio de uma outra variável, observado nas regiões vizinhas, no eixo das ordenadas. (ALMEIDA, 2004, p. 47).

Ainda segundo Almeida (2004), no diagrama de dispersão de Moran multivariado, os valores das variáveis são apresentados de forma padronizada. A figura 1 mostra um exemplo deste tipo de diagrama.

Figura 1: exemplo de diagrama de dispersão de Moran multivariado – variáveis x e y.



Fonte: elaboração própria a partir do *software* GeoDa.

No primeiro quadrante do diagrama de dispersão de Moran, estão as unidades espaciais do tipo alto-alto. Essas unidades apresentam valores altos de uma variável de interesse e tendem a possuir vizinhos com valores altos de uma outra variável de interesse.

O terceiro quadrante agrupa as unidades do tipo baixo-baixo. Tais unidades apresentam valores baixos de uma variável de interesse e tendem a estar rodeadas de vizinhos com valores também baixos de uma outra variável de interesse.

As unidades do agrupamento do tipo alto-baixo apresentam valores altos de uma variável de interesse e tendem a ter vizinhos com valores baixos de uma outra variável de interesse. São representadas no quarto quadrante.

O tipo de associação baixo-alto apresenta unidades com valores baixos de uma variável de interesse que tendem a estar rodeadas por vizinhos com valores altos de uma outra variável de interesse. Tal tipo de associação é representada no segundo quadrante.

O diagrama de dispersão de Moran não leva em conta a significância do padrão de associação das unidades espaciais em relação às variáveis em estudo. Dessa forma, o diagrama apresenta todas as unidades espaciais em questão, tanto significativas, como não significativas.

Diante dessa característica, uma forma mais eficiente de apresentar os resultados é mapear as estatísticas *I* de Moran, gerando o mapa de significância. (ALMEIDA, 2004, p. 48). Este mapa representa cada unidade espacial da região em estudo, enfatizando aquelas unidades que apresentaram o *I* de Moran local significativo.

A combinação das informações do diagrama de dispersão de Moran com as informações do mapa de significância origina o mapa de *clusters*. "Ele ilustra a classificação em quatro categorias de associação espacial que são estatisticamente significantes." (ALMEIDA, 2004, p. 50). Portanto, o mapa de *clusters* apresenta os agrupamentos alto-alto, baixo-baixo, alto-baixo e baixo-alto que foram estatisticamente significativos.

2.3 Teoria de Resposta ao Item

A Teoria de Resposta ao Item (TRI) apresenta-se como uma alternativa à Teoria Clássica de Medida (TCM). A TRI permite elaborar e analisar escalas, avaliadas através dos instrumentos de medição, como por exemplo, questionários. Composta de modelos probabilísticos para variáveis que não podem ser medidas diretamente (traços latentes), a TRI é amplamente utilizada em avaliações educacionais. (SILVA, 2014).

Os modelos que compõem a TRI representam a possibilidade de determinada resposta a um item ser escolhida em função dos parâmetros que diferenciam este item e do nível do respondente quanto à sua característica (traço latente) que está sendo medida. Quanto aos parâmetros, os modelos podem ser ditos de um, dois e três parâmetros.

Os modelos de um parâmetro avaliam a dificuldade do item, os modelos de dois parâmetros consideram a dificuldade e a discriminação do item e os modelos de três parâmetros consideram, além da dificuldade e da discriminação, a probabilidade de escolha casual.

Considerando o tipo de item (pergunta) do questionário, o modelo TRI mais adequado pode ser o dicotômico, quando os itens apresentam apenas duas opções de resposta, ou então o politômico, adequado quando os itens do teste apresentam mais de duas alternativas. No caso dos modelos politômicos, pode-se considerar apenas o parâmetro de dificuldade do item ou os parâmetros dificuldade e discriminação do item.

Tendo em vista sua adequação com o tipo de questionário utilizado no presente estudo e com a finalidade pretendida na aplicação da TRI, foi aplicado o Modelo de Resposta Gradual de Samejima, apresentado no tópico a seguir.

2.3.1 Modelo de Resposta Gradual de Samejima

O Modelo de Resposta Gradual de Samejima, considerado politômico, é uma generalização do modelo logístico de dois parâmetros; sendo adequado quando os itens do questionário apresentam respostas passíveis de serem ordenadas. Além disso, não é necessário que os itens apresentem o mesmo número de alternativas de respostas. (SILVA, 2014). Outro ponto que favorece a escolha desse modelo no presente estudo é de que não seria necessário um modelo com o terceiro parâmetro, correspondente à probabilidade de escolha casual.

No Modelo de Resposta Gradual de Samejima, a probabilidade de um indivíduo j escolher uma alternativa de resposta, k, de um item i pode ser expressa por:

$$P_{i,k}^{+}(\theta_j) = \frac{1}{1 + e^{-Da_i(\theta_j - b_{i,k})}}$$

A probabilidade do indivíduo $j$obter o escore $k$no item i pode ser expresso por:

$$P_{i,k}(\theta j) = \frac{1}{1 + e^{-Da_i(\theta_j - b_{i,k})}} - \frac{1}{1 + e^{-Da_i(\theta_j - b_{i,k+1})}}$$

Com $i = 1,2,\dots,I,\ j = 1,2,\dots,n$ e $k = 0,1,\dots,m_i$

Onde:

$b_{i,k}$ representa o parâmetro de dificuldade da alternativa k do item i;

a_i representa o parâmetro de discriminação do item i;

θ_j é o traço latente do j-ésimo indivíduo, representando sua habilidade; e

D é o fator de escala constante e igual a 1.

O traço latente, teoricamente, pode assumir qualquer valor real entre -∞ e +∞. Porém, em um teste (ou em uma pesquisa), o escore que mede o traço latente assume valores entre 0 e i (número de itens). Nas palavras de Bortolotti (2003):

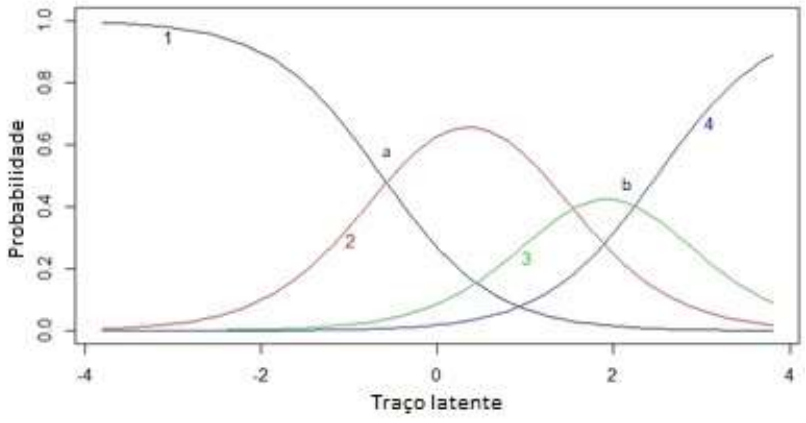
> Desta forma, é necessário estabelecer antes do processo de estimação uma origem que representará a média das habilidades e uma unidade de medida que representará o desvio padrão das habilidades dos indivíduos da população em estudo. (BORTOLOTTI, 2003, p.11).

Segundo Andrade, Tavares e Valle (2000), a métrica usualmente definida para o traço latente é (μ,σ) com μ=0 e σ=1 (média = 0 e desvio padrão = 1). Os programas computacionais, comumente, utilizam essa métrica para a aplicação da TRI. De acordo com Araujo et al (2009), "Embora a utilização desta métrica seja frequente, podem-se fazer transformações lineares de modo a apresentar os resultados em outra métrica qualquer." (ARAUJO, ANDRADE e BORTOLOTTI, 2009, p. 1006). Com o intuito de facilitar sua interpretação, no presente estudo, os escores serão padronizados para apresentarem média 50 e desvio padrão 10; procedimento adotado por Neto et al (2013).

O parâmetro de dificuldade de um item mede o nível de traço latente necessário para que o indivíduo concorde com o item. Através dos parâmetros de dificuldade, é possível indicar os limites da escala do traço latente no qual a alternativa de resposta k passa a ser mais provável que a alternativa anterior, $k-1$. (SILVA, 2014). Dados tais limites de probabilidade, é possível construir a curva categoria resposta do item.

A figura 2 representa a curva categoria resposta de um item hipotético com quatro alternativas de resposta.

Figura 2: Curva categoria de resposta do Item – hipotético



Fonte: Elaboração própria a partir do *software* R.

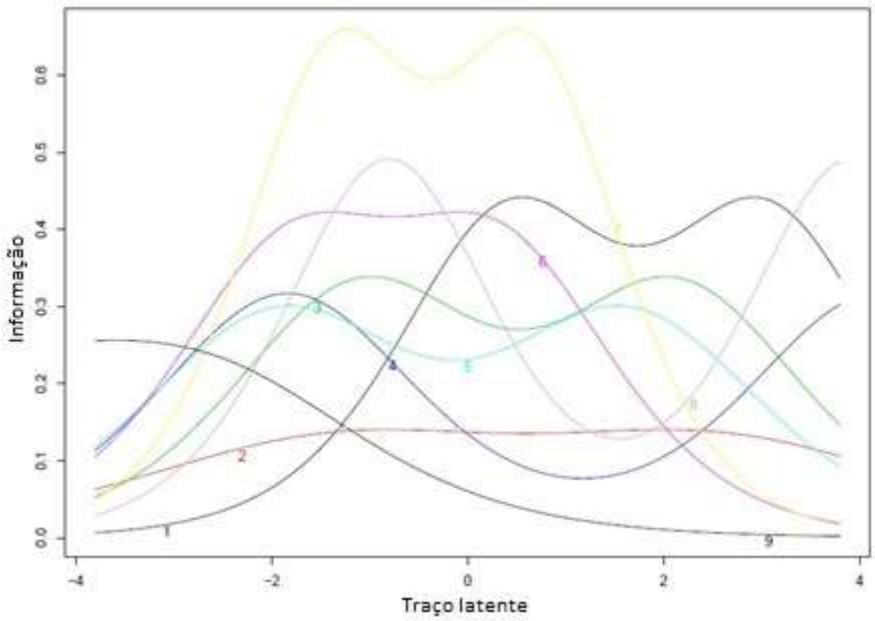
Observa-se, na figura 2, que os indivíduos de traço latente entre -4 e -0.633 (ponto a) tem maior probabilidade de selecionar a primeira categoria de resposta (1). Os indivíduos de traço latente entre 2.49 (ponto b) e 4, apresentam maior probabilidade de selecionar a categoria 4.

Quanto ao parâmetro de discriminação, Quintana (2010) cita ser uma importante característica que reflete a qualidade de um item, capaz de discriminar sujeitos com diferentes níveis de traço latente. Albuquerque e Tróccoli (2004) citam as seguintes categorias de discriminação: nenhuma discriminação (a=0), discriminação muito baixa (0,01≤ a ≤0,34), discriminação baixa (0,35≤ a ≤0,64), discriminação moderada (0,65≤ a ≤1,34), discriminação alta (1,35≤ a ≤1,69) e discriminação muito alta (a ≥ 1,70). Ainda segundo a literatura, itens com discriminação muito baixa são retirados da análise. (NETO et al, 2013).

Outro importante conceito da TRI, é a curva de informação do item, capaz de mensurar a discriminação do item em relação ao traço latente. Os itens que apresentam alto grau de informação são aqueles em que, geralmente, as curvas estão bem distribuídas ao longo do eixo de habilidade, que representa o traço latente. (SILVA, 2014).

A figura 3 representa a curva de informação de itens hipotéticos com quatro alternativas de resposta.

Figura 3: Curva de Informação do Item – hipotético



Fonte: Elaboração própria a partir do *software* R.

Nesse modelo hipotético, dos nove itens representados na figura 3, o item sete apresenta maior informação em relação ao traço latente em comparação aos demais itens. Isso por que tal item apresenta a melhor distribuição ao longo do eixo de traço latente quando comparado com os demais.

Em resumo, com base nos parâmetros de discriminação é possível construir a curva de informação do item e verificar se tal modelo apresenta informações consideradas suficientes; ou seja, se o modelo é capaz de diferenciar os indivíduos e mensurar o traço latente. Através

dos parâmetros de dificuldade é possível construir a curva categoria resposta e, a partir de então, construir os escores, objeto de interesse da aplicação da TRI no presente estudo. Esses escores foram construídos a partir de variáveis referentes aos alunos e às escolas, conforme procedimento adotado por Silva (2015), através do *software* R.

2.4 O modelo de regressão hierárquico

Em um modelo de regressão hierárquico ou multinível, considera-se que existe uma estrutura hierárquica nos dados. Os níveis dessa estrutura hierárquica são representados por submodelos. Cada submodelo expressa o relacionamento entre as variáveis dentro de um determinado nível e também especifica como as variáveis em um nível influenciam os relacionamentos que ocorrem em outros níveis.

No modelo de regressão clássico, os parâmetros (coeficientes de regressão) são fixos; já no modelo hierárquico, o intercepto e o coeficiente angular são considerados aleatórios. Isso significa que podem ser considerados dependentes da influência do nível hierárquico mais elevado.

As principais vantagens dessa metodologia podem ser elencadas da seguinte forma:

i) como o modelo considera a estrutura de agrupamento dos dados, pode-se formular e testar hipóteses a respeito dos efeitos entre os níveis;

ii) a flexibilidade da variável resposta pode ser repartida nos diferentes níveis e;

iii) pode se analisar cada grupo individualmente, isso, devido à estimação de uma equação para cada unidade.

2.4.1 Especificação do modelo incondicional (modelo nulo)

Para a análise dos dados, serão estimados modelos hierárquicos de dois níveis. No nível 1 serão incluídas as variáveis relacionadas ao aluno. No nível 2 estarão as variáveis relacionadas à escola.

Para verificar a proporção da variabilidade entre os dois níveis é estimado, inicialmente, o modelo nulo, isto é, sem variáveis. Então, o modelo de nível 1 pode ser especificado da seguinte maneira:

$$Y_{ij} = \beta_{0j} \qquad (1)$$

Por sua vez, o modelo de nível 2 é assim apresentado:

$$\beta_{0j} = u_{0j} + e_{ij} \quad (2)$$

Substituindo-se (2) em (1), tem-se a equação a seguir:

$$Y_{ij} = u_{0j} + e_{ij} \quad (3)$$

Sendo:

i = 1, 2, ... , I unidades de nível 1;

j = 1, 2, ... , J unidades de nível 2;

Y_{ij} a proficiência do i-ésimo aluno da j-ésima escola;

β_{0j} o intercepto do modelo;

u_{0j} o parâmetro da parte fixa do modelo;

e_{ij} o efeito aleatório referente ao nível 2.

Depois de especificado o modelo incondicional, estima-se o modelo condicional, ou seja, com a inclusão de variáveis.

2.4.2 Especificação do modelo condicional

O modelo hierárquico condicional pode ser especificado da seguinte maneira:

$$y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + e_{ij} \quad (4)$$

$$\beta_{0j} = y_{00} + y_{01}w_j + u_{0j} \quad (5)$$

$$\beta_{1j} = y_{10} + y_{11}w_j + u_{1j} \quad (6)$$

Substituindo-se as equações (6) e (5) na equação (4), tem-se:

$$y_{ij} = y_{00} + y_{10}x_{ij} + y_{01}w_j + y_{11}w_jx_{ij} + u_{1j}x_{ij} + u_{0j} + e_{ij} \quad (7)$$

Onde:

y_{ij} é a proficiência do i-ésimo aluno da j-ésima escola;

β_{0j} corresponde ao intercepto geral do modelo;

β_{1j} corresponde ao coeficiente angular relacionado à variável x (de nível 1), e mede o efeito da variável na proficiência do aluno;

$y_{00}, y_{01}, y_{10}, y_{11}$ são os parâmetros fixos do modelo;

u_{0j} corresponde ao efeito individual da escola, componente de erro aleatório referente ao nível 2,associado ao intercepto;

u_{1j} corresponde ao componente de erro aleatório referente ao nível 2, associado ao coeficiente angular;

e_{ij} representa o componente de erro aleatório relacionado à proficiência do aluno. Esse termo corresponde ao resíduo associado à medida da proficiência do aluno, não explicado pelo modelo. Supõe-se que o erro e_{ij} seja independente dos erros de nível 2 (u_{0j} e u_{1j}).

w_j representa a variável explicativa de nível 2 (da escola) e;

$w_j x_{ij}$ corresponde à interação entre as variáveis explicativas, relacionadas ao aluno e à escola.

Na equação de regressão (7), poderão ser incluídas outras variáveis explicativas de nível de aluno e também de nível de escola. Essa inclusão permite obter o impacto das novas variáveis no rendimento escolar, bem como obter outros termos de interação, alcançando uma maior diminuição da variabilidade total e, consequente aumento da capacidade de explicação da variável dependente pelo modelo resultante.

2.4.3 Método de estimação

No modelo de regressão hierárquico, os parâmetros de efeitos fixos são, geralmente, estimados pelo método da máxima verossimilhança (MLE). Os componentes de variância dos termos de erro podem ser estimados tanto pela MLE como pela máxima verossimilhança restrita (REML). Nas palavras de Fávero e Belfiore (2017):

> Ambas exigem a otimização de determinada função-objeto, que geralmente parte de valores iniciais dos parâmetros e usa uma sequência de interações para encontrar os parâmetros que maximizam a função de verossimilhança previamente definida. (FÁVERO E BELFIORE, 2017, p. 863).

Supondo uma regressão apenas com a constante, seja n o número de observações da variável dependente Y, μ a sua média e σ_Y^2 a sua variância. A estimação por MLE de σ_Y^2 é obtida considerando os n termos $Y_i - \mu$. A estimação de σ_Y^2 por REML é obtida a partir dos (n-1) primeiros termos de $Y_i - \bar{Y}_i$. Portanto, o método REML "gera uma estimação não viesada de σ_Y^2, por esta ser a própria variância amostral obtida pela divisão dos elementos por (n -1)" (Fávero e Belfiore, 2017, p. 863). Dessa forma, utilizar-se-á o método REML no presente estudo, pelo fato desse método "gerar estimações não viesadas das variâncias dos termos de erro em modelos multinível" (FÁVERO E BELFIORE, 2017, p. 864).

2.4.4 Coeficiente intra-correlação

O coeficiente de "intra-correlação" indica o grau de agrupamento da população em estudo. No modelo multinível especificado, esse coeficiente mede a proporção da variância total do resultado dos alunos que é devida às características das escolas. O coeficiente de "intra-correlação" pode ser obtido da seguinte maneira: constrói-se um modelo multinível incondicional (modelo nulo), apresentado na subseção 2.4.1, sob a equação $Y_{ij} = u_{0j} + e_{ij}$. A variância total para a variável dependente nesse modelo é dada por $\sigma_{u0}^2 + \sigma_e^2$. O coeficiente de "intra-correlação" (ρ), então, é calculado pela fórmula:

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

O valor de (ρ) situa-se entre 0 e 1, de forma que quanto maior o seu valor, maior a proporção da variância que é devida ao segundo nível. Este coeficiente é usado para legitimar a aplicação de um modelo multinível ao invés de um modelo de regressão clássico.

2.5 Variáveis e Dados

O Sistema Nacional de Avaliação da Educação Básica (SAEB) tem como público-alvo os alunos do 5º e 9º ano do ensino fundamental e da 3ª série do ensino médio. Tendo em vista que a análise desse estudo possui como foco os anos iniciais do ensino fundamental, propõe-se utilizar as informações referentes aos alunos do 5º ano.

Os microdados referentes ao aluno, à escola e ao índice de qualidade da educação foram obtidos no SAEB e no Censo Escolar, através do portal do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). Os dados relacionados ao índice

de desenvolvimento econômico municipal (IFDM) foram coletados através do portal da Federação das Indústrias do Estado do Rio de Janeiro (FIRJAN).

A coleta dos indicadores do nível socioeconômico da região Norte de Minas e do Estado de Minas Gerais (PIB, PIB *per capita*, renda *per capita*, índice de Gini, Índice Mineiro de Responsabilidade Social – Saneamento, habitação e meio ambiente) foi feita por meio dos sítios do Instituto Brasileiro de Geografia e Estatística (IBGE) e do Índice Mineiro de Responsabilidade Social (IMRS).

Os quadros de 1 a 3 apresentam as variáveis utilizadas.

2.5.1 Variáveis relacionadas à análise de correlação e à análise espacial de dados

O Quadro 1 apresenta as variáveis utilizadas como índice de desenvolvimento econômico e índice de qualidade da educação dos municípios.

Quadro 1: variáveis relacionadas ao índice de desenvolvimento econômico e ao índice de qualidade da educação dos municípios norte-mineiros – 2013.

Variável	**Especificação**
Índice FIRJAN de Desenvolvimento Municipal	IFDM
Índice de Desenvolvimento da Educação Básica - Municipal – Anos Iniciais	IDEB-M-AI

Fonte: elaboração própria a partir dos dados da FIRJAN, do SAEB e do Censo Escolar/INEP.

O tópico seguinte trata das variáveis utilizadas na Teoria de Resposta ao Item e na análise de regressão.

2.5.2 Variáveis inseridas na Teoria de Resposta ao Item e na análise de regressão

O Quadro 2 apresenta a variável dependente do modelo de regressão hierárquico e as informações referentes ao aluno que foram utilizadas na construção dos escores através da Teoria de Resposta ao Item. Esses escores foram inseridos na regressão hierárquica como variáveis independentes. São apresentadas também duas variáveis explicativas binárias (sexo e cor).

Quadro 2: variáveis relacionadas ao aluno - 5º ano do ensino fundamental – 2013.

Variável	**Especificação**
Média da proficiência do aluno nas disciplinas de Língua Portuguesa e Matemática	Variável dependente
Sexo do aluno	Variável independente
Cor (raça)	Variável independente
Aluno já foi reprovado?	Variável independente
Na sua casa tem televisão em cores?	Escore socioeconômico do aluno
Na sua casa tem aparelho de rádio?	
Na sua casa tem videocassete e/ou DVD?	
Na sua casa tem geladeira?	
Na sua casa tem máquina de lavar roupa (o tanquinho não deve ser considerado)?	
Na sua casa tem carro?	
Na sua casa tem computador?	
Na sua casa tem banheiro?	
Em sua casa trabalha empregado(a) doméstico(a) pelo menos cinco dias por semana?	
Você mora com sua mãe?	
Até que série sua mãe, ou a mulher responsável por você, estudou?	
Sua mãe, ou a mulher responsável por você, sabe ler e escrever?	
Você mora com seu pai?	
Até que série seu pai, ou o homem responsável por você, estudou?	
Seu pai, ou homem responsável por você, sabe ler e escrever?	
Seus pais ou responsáveis incentivam você a estudar?	Escore de intervenção
Seus pais ou responsáveis incentivam você a fazer o dever de casa e/ou os trabalhos da escola?	
Seus pais ou responsáveis incentivam você a ler?	
Seus pais ou responsáveis incentivam você a ir à escola e/ou não faltar às aulas?	
Seus pais ou responsáveis conversam com você sobre o que acontece na escola?	
Com qual frequência você lê: Jornais?	Escore de capital cultural
Com qual frequência você lê: Livros?	
Com qual frequência você lê: Revistas em geral?	
Com qual frequência você lê: Notícias na internet (ex.: blog, notícia)?	
Com qual frequência você costuma ir à/ao: Biblioteca?	
Com qual frequência você costuma ir à/ao: Espetáculo ou exposição (teatro, museu, dança ou música)?	
Em dia de aula, quanto tempo você gasta assistindo à TV, navegando na internet ou jogando jogos eletrônicos?	Escore de (des)capital cultural
Em dias de aula, quanto tempo você gasta fazendo trabalhos domésticos (ex.: lavando louça, limpando o quintal etc.)?	
Atualmente você trabalha fora de casa (recebendo ou não um salário)?	
Você faz o dever de casa de Língua Portuguesa?	Escore dever de casa
Você faz o dever de casa de Matemática?	

Fonte: elaboração própria a partir dos dados do SAEB e do Censo Escolar/INEP.

O Quadro 3 especifica as variáveis relacionadas à escola.

Quadro 3: variáveis relacionadas à escola – 2013.

Variável	**Especificação**
Dependência administrativa da escola (municipal ou estadual)	Variável independente
Nível socioeconômico médio da escola.	Variável independente
Proficiência média da escola.	Variável independente
Sala de diretoria.	Escore de infraestrutura da Escola
Sala de professores.	
Laboratório de informática.	
Laboratório de ciências.	
Sala de atendimento especial.	
Quadra de esportes.	
Biblioteca.	
Banheiro dentro do prédio.	
Banheiro adequado a alunos com deficiência ou mobilidade reduzida.	
Dependências e vias adequadas a alunos com deficiência ou mobilidade reduzida.	
Aparelho de televisão.	
Aparelho de DVD.	
Copiadora.	
Impressora.	
Computador.	
Internet	

Fonte: elaboração própria a partir dos dados do SAEB e do Censo Escolar/INEP.

O próximo capítulo apresenta alguns indicadores socioeconômicos e educacionais; além dos resultados e discussão das metodologias aplicadas ao presente estudo.

3. Educação e desenvolvimento no Norte de Minas

Este capítulo apresenta, inicialmente, alguns indicadores socioeconômicos e educacionais para a região norte-mineira; realizando, sempre que possível, um comparativo com os indicadores do estado de Minas Gerais. Posteriormente, procede-se à apresentação dos resultados e análises das metodologias aplicadas no presente estudo.

3.1 Indicadores socioeconômicos do Norte de Minas

Segundo o Índice Mineiro de Responsabilidade Social (IMRS), o Norte de Minas possui aproximadamente 128 mil quilômetros quadrados, correspondente a cerca de 20% do território de Minas Gerais. De acordo com o Instituto Brasileiro de Geografia e Estatística (IBGE), os 89 municípios do Norte de Minas, em 2000, apresentavam uma população total de 1,49 milhões de pessoas; em 2010, esse número passou para 1,6 milhões de habitantes. Ainda de acordo com estimativas do IBGE, a população norte-mineira total em 2014 seria de 1,69 milhões de pessoas.

A tabela 1 apresenta alguns indicadores socioeconômicos da Região, bem como para o estado de Minas Gerais. Nela, constam o Índice Mineiro de Responsabilidade Social - Saneamento, Habitação e Meio Ambiente (IMRS-SHM) e o Índice FIRJAN de Desenvolvimento Municipal (IFDM).

O IMRS leva em consideração os seguintes indicadores:

- Indicador de déficit habitacional relativo;
- Percentual da população residente em domicílios com banheiro e água encanada;
- Percentual da população urbana residente em domicílios com coleta direta de lixo;
- Disposição final do lixo coletado;
- Existência de Conselho Municipal de Saneamento ou afim;
- Percentual de gastos com o saneamento básico e habitação;
- Percentual da população afetada por doenças relacionadas ao saneamento ambiental inadequado;
- Percentual de cobertura vegetal por flora nativa ou reflorestamento;
- Percentual de área de proteção integral;
- Percentual de área de uso sustentável; e
- Esforço Orçamentário em meio ambiente.

Os valores do IMRS-SHM situam-se entre 0 e 1, e quanto mais próximo de 1 mais favorável é a situação do município.

Quanto ao IFDM, trata-se de um indicador calculado pela Federação das Indústrias do Estado do Rio de Janeiro. O IFDM acompanha o desenvolvimento socioeconômico dos municípios brasileiros em três áreas: emprego/renda, educação e saúde. É baseado em estatísticas oficiais dos ministérios do Trabalho, Educação e Saúde. Seus valores situam-se entre 0 e 1, de forma que quanto mais próximo de 1, maior o nível de desenvolvimento do município.

Tabela 1: Indicadores socioeconômicos do Norte de Minas e de Minas Gerais – 2010.

Indicador	Norte de Minas	Minas Gerais	% NM/MG
PIB (R$ mil)	R$ 13.297.002	R$ 351.133.915	3,79
PIB *per capita*	R$ 8.256	R$ 17.919	46,97
Renda *per capita*	R$ 304,93*	R$ 733,24	41,59
Índice de Gini	0,49*	0,56	-
IMRS-SHM	0,35*	0,44*	-
IFDM (ano 2013)	0,61*	0,68*	-

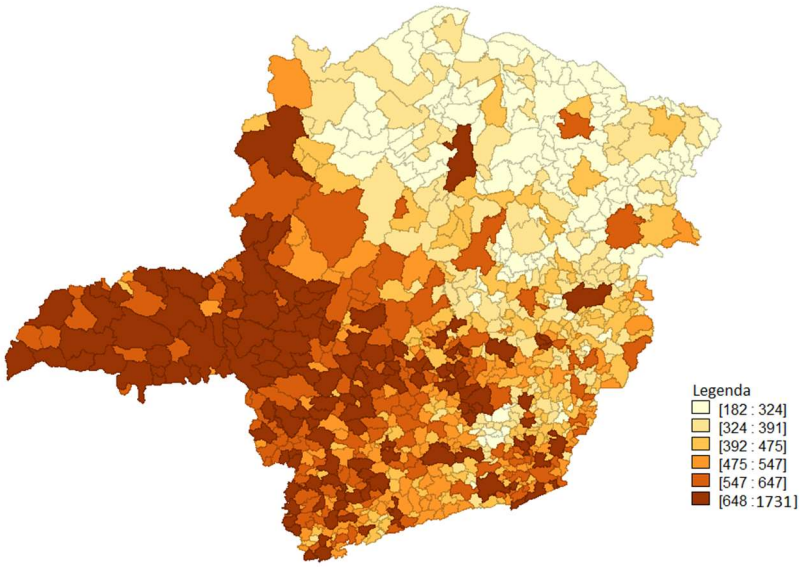
Fonte: elaboração própria a partir dos dados do IBGE, IMRS e da FIRJAN.
*Valores correspondentes à média dos municípios.

No ano de 2010, o PIB dos municípios norte-mineiros representou 3,79% do PIB do estado de Minas Gerais, representando a sexta maior participação do PIB, entre as doze mesorregiões (Tabela 6, do anexo I).

Em termos de PIB *per capita*, o Norte de Minas superou somente o Vale do Mucuri e o Jequitinhonha (Gráfico 1, do anexo I). O valor do indicador, na Região, correspondeu a apenas 46,97% do valor registrado para o Estado.

A média da renda *per capita* dos municípios norte-mineiros foi bem menor que o valor registrado para o Estado. A figura 4 apresenta dados desse indicador por município de Minas Gerais, referente ao ano de 2010. Quanto maior a intensidade da cor, maior é a renda per capita municipal.

Figura 4: Renda *per capita* dos municípios de Minas Gerais, 2010.

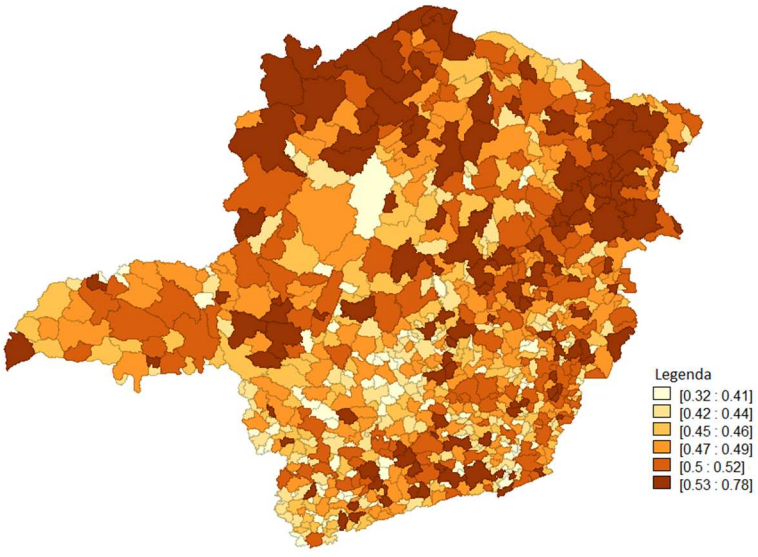


Fonte: elaboração própria a partir dos dados do IMRS.

Através da figura 4, pode-se observar que os municípios que se situam nas mesorregiões como o Noroeste de Minas, Norte de Minas, Jequitinhonha e Vale do Mucuri tenderam a apresentar baixos valores para a renda *per capita*, em relação aos valores dos municípios situados nas regiões Central Mineira, Sul/Sudoeste e no Triângulo Mineiro.

Quanto ao índice de Gini, a média calculada para os municípios do Norte de Minas foi menor que o valor verificado para o Estado. A figura 5 apresenta a distribuição do indicador para os municípios mineiros, relativo ao ano de 2010. Quanto mais intensa a cor do município, maior é o índice de Gini correspondente; logo, maior é a concentração de renda no município em questão.

Figura 5: Índice de Gini dos municípios de Minas Gerais, 2010.



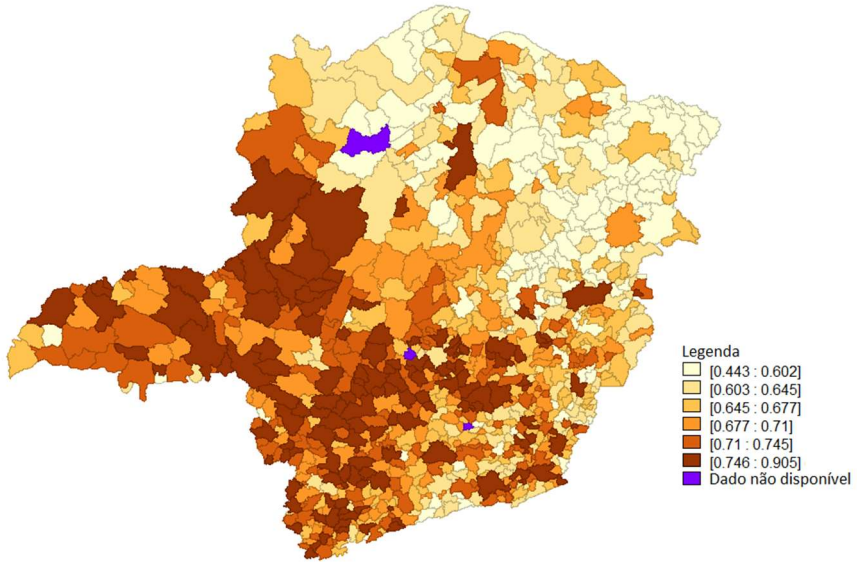
Fonte: elaboração própria a partir dos dados do Atlas do Desenvolvimento Humano no Brasil.

Por meio da Figura, pode-se observar que nas regiões do Norte de Minas, Noroeste, Jequitinhonha e Vale do Mucuri os municípios tenderam a apresentar maiores concentrações de renda em relação aos municípios das regiões Central Mineira, Sul/Sudoeste e do Triângulo Mineiro.

Os municípios do Norte de Minas, quanto à média do IMRS-SHM de 2010, ficaram em décimo primeiro lugar no *ranking* de mesorregiões, ficando acima somente do Vale do Mucuri (Tabela 7, do anexo I). O valor do indicador, para a Região, encontra-se abaixo da média estadual, em 0,09 pontos percentuais.

A média do IFDM para os municípios norte-mineiros, relativa ao ano de 2013, foi menor que a média estadual. A figura 6 retrata a distribuição do IFDM no estado de Minas Gerais. Cores mais intensas retratam maiores valores para o indicador, ou seja, municípios socioeconomicamente mais desenvolvidos.

Figura 6: IFDM dos municípios de Minas Gerais, 2013.



Fonte: elaboração própria a partir dos dados da FIRJAN.

Ao analisar a Figura, observa-se que os municípios localizados na parte superior do Estado apresentam cores mais claras que aqueles da parte inferior. Dessa forma, os municípios do Norte de Minas Gerais tenderam a apresentar valores para o IFDM menores do que os dos municípios de mesorregiões como o Triângulo Mineiro, Central Mineira e Sul/Sudoeste. Os municípios norte-mineiros apresentaram a terceira menor média do IFDM de 2013, entre as mesorregiões de Minas Gerais, ocupando a décima colocação no *ranking* estadual (conforme tabela 8 do anexo I).

Na sequência são apresentados dois indicadores educacionais para a Região: a taxa de analfabetismo e o IDEB, relacionado à qualidade da educação.

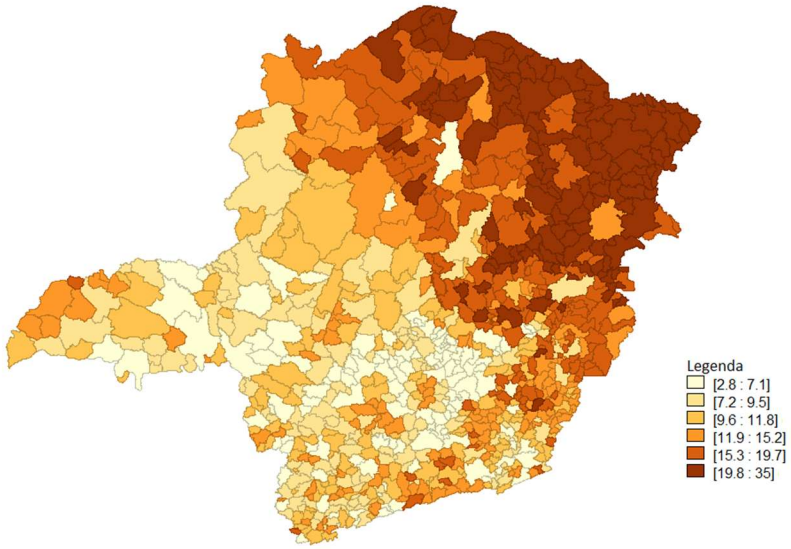
3.2 Indicadores educacionais do Norte de Minas

Entre os vários indicadores educacionais, dois são apresentados neste tópico: a taxa de analfabetismo e o Índice de Desenvolvimento da Educação Básica (IDEB).

O IMRS define a taxa de analfabetismo para pessoas com 15 anos ou mais de idade como a razão entre o número de pessoas de 15 anos ou mais de idade analfabetas e a população total nessa faixa etária, multiplicada por 100. Considera-se analfabetas as pessoas que não são capazes de ler e escrever um bilhete simples em seu idioma. Segundo FERRARO e KREIDLOW (2004), o analfabetismo seria uma forma extrema de exclusão social. Através do IDEB é possível monitorar o nível de qualidade educacional, o qual influencia a formação do capital humano.

A figura 7 apresenta a taxa de analfabetismo para pessoas com 15 anos ou mais de idade, por municípios de Minas Gerais, referente ao ano de 2010. Quanto mais intensa a cor, maior é a taxa de analfabetismo correspondente.

Figura 7: Taxa de analfabetismo para pessoas com 15 anos ou mais de idade, municípios de Minas Gerais, 2010.



Fonte: elaboração própria a partir dos dados do IMRS.

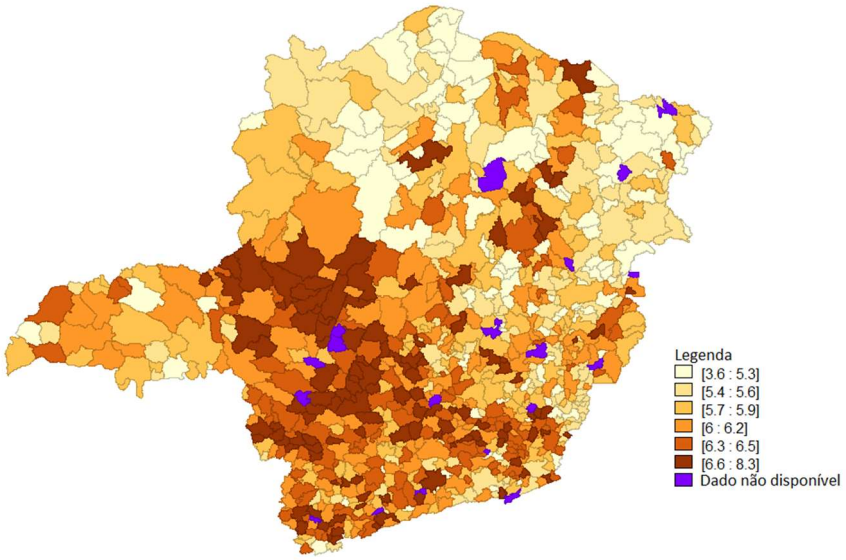
Municípios da parte superior do Estado de Minas Gerais apresentaram cores mais intensas, indicando maiores taxas de analfabetismo nessas localidades. A Região do Norte de Minas, quanto à média da taxa de analfabetismo (das pessoas com 15 anos ou mais de idade),

foi a terceira com maior valor entre as 12 mesorregiões. Além disso, superou a média estadual (conforme tabela 9, do anexo I).

A figura 8 mostra o Índice de Desenvolvimento da Educação Básica dos anos iniciais (5º ano do ensino fundamental) - IDEB-M-AI, do ano de 2013, para os municípios de Minas Gerais. Quanto mais intensa a cor, maior é o IDEB do respectivo município em comparação aos de cores menos intensas.

Ao observar a figura, nota-se que os municípios situados na região Norte do estado apresentaram valores para o IDEB-M-AI menores que os dos municípios das regiões como Central Mineira e Sul/Sudoeste.

Figura 8: IDEB-M-AI dos municípios de Minas Gerais, 2013.



Fonte: elaboração própria a partir dos dados do SAEB/INEP.

A média do IDEB-M-AI do Norte de Minas foi menor do que a média da maioria das mesorregiões, no ano analisado (de acordo com a tabela 10, do anexo I). A região norte-mineira apresentou resultado superior para o indicador somente em relação ao Vale do Mucuri. Além disso, a média do Norte de Minas foi menor que a média do Estado.

O próximo tópico trata-se da apresentação dos resultados e análise da relação entre educação e desenvolvimento no Norte de Minas.

3.3 Educação e desenvolvimento no Norte de Minas: resultados e análise

Através da análise gráfica, não é clara a existência de relação linear entre o IFDM e o IDEB-M-AI (Gráfico 2 do anexo I). Dessa forma, a tabela 2 apresenta tanto o coeficiente de correlação de Pearson, como o coeficiente de Spearman, entre o IFDM e o IDEB-M-AI dos

municípios do Norte de Minas. Os valores dos coeficientes foram bem próximos, sendo que ao nível de significância de 1%, foi possível rejeitar a hipótese nula de ausência de correlação entre as variáveis.

Tabela 2: Coeficiente de correlação de Pearson e de Spearman- IFDM e IDEB-M-AI - Norte de Minas Gerais - 2013.

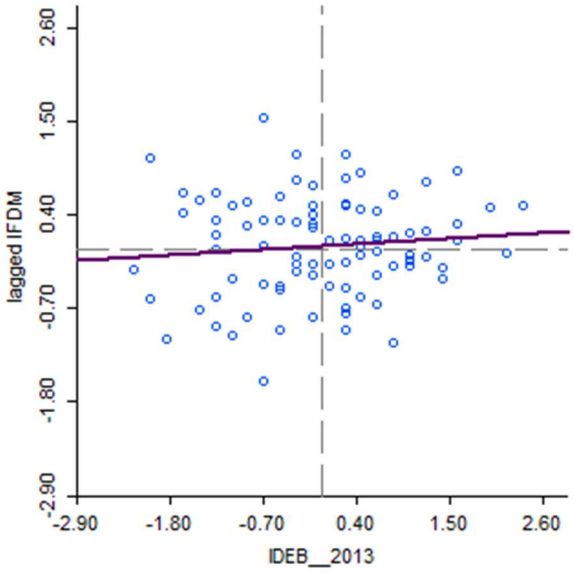
Pearson	0,2805 (0,0085)
Spearman	0,2779 (0,0092)

Nota: nível de significância entre parênteses.
Fonte: elaboração própria a partir dos dados da FIRJAN e do SAEB/INEP.

O sinal do coeficiente estimado foi positivo, indicando que municípios norte-mineiros que apresentaram maiores (menores) valores do IFDM tenderam a apresentar maiores (menores) valores para o IDEB-M-AI. Desta forma, quanto maior o nível de desenvolvimento municipal, maior a qualidade da educação (e vice-versa).

A figura 9 apresenta o diagrama de dispersão de Moran multivariado para o IDEB-M-AI e o IFDM, para os municípios do Norte de Minas.

Figura 9: Diagrama de dispersão de Moran multivariado - IDEB-M-AI e IFDM – Norte de Minas, 2013.

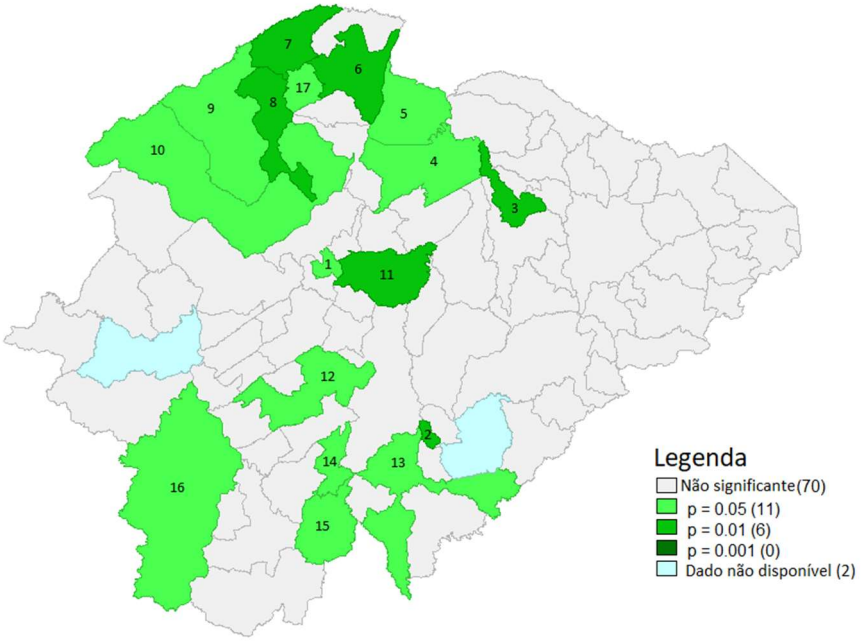


Fonte: Elaboração própria, a partir dos dados da FIRJAN e do SAEB/INEP.

O diagrama de dispersão de Moran apresenta pontos bem distribuídos em seus quatro quadrantes. Esse resultado revela quatro tipos de associação espacial entre os municípios norte-mineiros, em relação às variáveis em estudo: alto-alto, baixo-baixo, alto-baixo, baixo-alto. Entretanto, através do diagrama não é possível verificar a significância de cada caso.

Dessa forma, para verificar se existem casos específicos estatisticamente significativos, utilizou-se o *I* de Moran Local. Apresentar os resultados do *I* de Moran local para cada município, nesse contexto, pode ser inadequado, do ponto de vista didático. Dessa forma, é útil apresentar o mapa LISA de significância (Figura 10).

Figura 10: Mapa LISA de significância - IDEB-M-AI e IFDM – Norte de Minas, 2013.

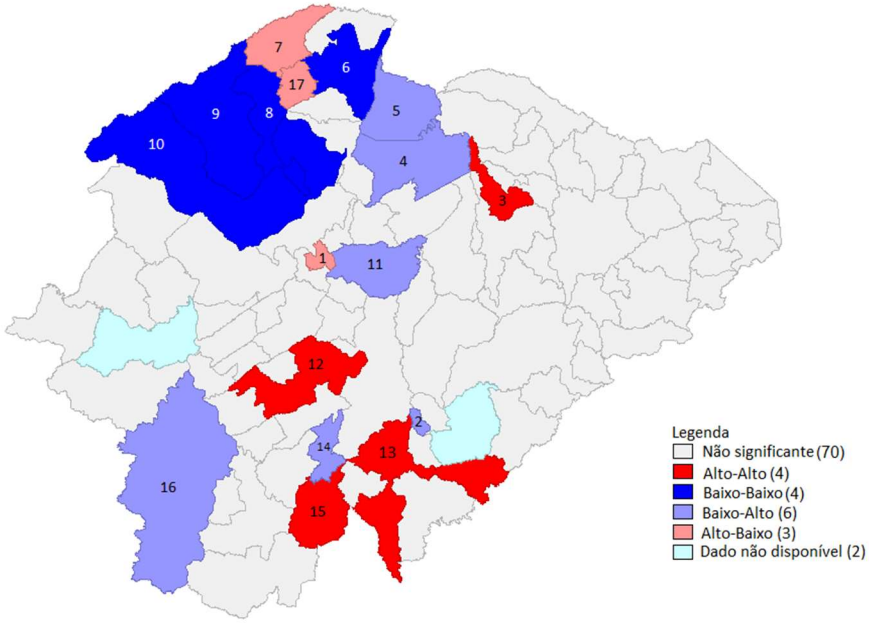


Fonte: Elaboração própria a partir dos dados da FIRJAN e do SAEB/INEP.

Na figura, quanto mais intensa a cor, menor é o nível de significância com o qual pode se rejeitar a hipótese nula de aleatoriedade espacial dos dados. O *I* de Moran local foi significativo para os seguintes municípios: Lontra (1), Glaucilândia (2), Pai Pedro (3), Jaíba (4), Matias Cardoso (5), Manga (6), Montalvânia (7), Cônego Marinho (8), Bonito de Minas (9), Januária (10), São João da Ponte (11), Coração de Jesus (12), Bocaiúva (13), Claro dos Poções (14), Francisco Dumont (15), Buritizeiro (16) e Miravânia (17). Para esses 17 municípios, a distribuição conjunta do IFDM e do IDEB não ocorre de forma aleatória, no espaço.

Outro instrumento importante a ser verificado é o mapa de *clusters*, que, em sua construção, baseia-se no diagrama de dispersão de Moran e no mapa de significância. Através do mapa de *clusters* é possível analisar as formas de interação entre os municípios com *I* de Moran estatisticamente significativo, em relação às variáveis em estudo. O referido mapa para o IDEB-M-AI e o IFDM dos municípios do Norte de Minas, para o ano de 2013, é representado a seguir.

Figura 11: Mapa de *clusters* - IDEB-M-AI e IFDM – Norte de Minas, 2013.



Fonte: Elaboração própria a partir dos dados da FIRJAN e do SAEB/INEP.

O mapa de *clusters* revela que quatro municípios apresentaram interações do tipo alto-alto: Pai Pedro (3), Coração de Jesus (12), Bocaiúva (13) e Francisco Dumont (15). Esses municípios se destacaram por possuir um elevado IDEB-M-AI e tenderam a apresentar vizinhos próximos com valores do IFDM elevados.

Igual número de municípios, sendo eles Manga (6), Cônego Marinho (8), Bonito de Minas (9) e Januária (10), revelou interações do tipo baixo-baixo, ou seja, apresentaram valores de IDEB-M-AI baixo e vizinhos próximos com valores do IFDM baixos.

Observa-se que, também, há municípios com interações do tipo baixo-alto: Glaucilândia (2), Jaíba (4), Matias Cardoso (5), São João da Ponte (11), Claro dos Poções (14) e Buritizeiro (16). Esses municípios registraram baixos valores para o IDEB, sendo cercados por vizinhos próximos com IFDM alto.

Por fim, para os municípios de Lontra (1), Montalvânia (7) e Miravânia (17) constatou-se interações do tipo alto-baixo; de forma que apresentaram um IDEB elevado, tendo vizinhos próximos com baixos valores do IFDM.

Esses resultados apontaram para a formação de alguns clusters de municípios na Região, a partir dos níveis de qualidade da educação e de desenvolvimento regional. Conforme o mapa, para a maioria dos municípios da Região não houve interação espacial significativa entre o IFDM e o IDEB, no ano de 2013.

Avaliada a relação entre educação e desenvolvimento na região norte-mineira, para o ano de 2013, parte-se para o estudo dos determinantes do desempenho escolar no Norte de Minas.

3.4 Determinantes do desempenho escolar no Norte de Minas: resultados e análise

Primeiramente, procede-se à apresentação dos resultados da estimação dos escores referentes ao aluno e à escola, por meio da TRI. Logo após, é realizada uma análise descritiva das demais variáveis inseridas na regressão e, por fim, são expostas as estimações e respectivas análises da regressão hierárquica.

3.4.1 Estimação dos escores

Os quadros de 4 a 9 apresentam os parâmetros de discriminação dos itens para cada escore. Quanto ao escore socioeconômico, o item "Na sua casa tem computador?" foi o mais discriminativo, ou seja, tem maior capacidade de diferenciar os alunos submetidos ao teste. O item "Você mora com sua mãe?" foi o menos discriminativo.

Quadro 4: Parâmetros de discriminação dos itens do escore socioeconômico.

Escore	Nº Item	Item	Discriminação
Socioeconômico	1	Na sua casa tem televisão em cores?	1,01
	2	Na sua casa tem aparelho de rádio?	0,69
	3	Na sua casa tem videocassete e/ou DVD?	0,93
	4	Na sua casa tem geladeira?	0,91
	5	Na sua casa tem máquina de lavar roupa (o tanquinho não deve ser considerado)?	0,71
	6	Na sua casa tem carro?	1,15
	7	Na sua casa tem computador?	1,60
	8	Na sua casa tem banheiro?	1,16
	9	Em sua casa trabalha empregado(a) doméstico(a) pelo menos cinco dias por semana?	0,85
	10	Você mora com sua mãe?	0,62
	11	Até que série sua mãe, ou a mulher responsável por você, estudou?	1,25
	12	Sua mãe, ou a mulher responsável por você, sabe ler e escrever?	1,04
	13	Você mora com seu pai?	0,69
	14	Até que série seu pai, ou o homem responsável por você, estudou?	1,23
	15	Seu pai, ou homem responsável por você, sabe ler e escrever?	1,00

Fonte: Elaboração própria através dos dados do SAEB/INEP.

No escore de intervenção, o item mais discriminativo foi “Seus pais ou responsáveis incentivam você a ler?”, enquanto o item “Seus pais ou responsáveis conversam com você sobre o que acontece na escola?” foi o menos discriminativo.

Quadro 5: Parâmetros de discriminação dos itens do escore de intervenção.

Escore	Nº Item	Item	Discriminação
Intervenção	16	Seus pais ou responsáveis incentivam você a estudar?	3,05
	17	Seus pais ou responsáveis incentivam você a fazer o dever de casa e/ou os trabalhos da escola?	2,66
	18	Seus pais ou responsáveis incentivam você a ler?	3,13
	19	Seus pais ou responsáveis incentivam você a ir à escola e/ou não faltar às aulas?	1,82
	20	Seus pais ou responsáveis conversam com você sobre o que acontece na escola?	1,02

Fonte: Elaboração própria através dos dados do SAEB/INEP.

Quanto ao escore de Capital Cultural, “Com qual frequência você lê: Jornais?” foi o item mais discriminativo. Os itens “Com qual frequência você lê: Notícias na internet (ex.: blog, notícia)?” e “Com qual frequência você costuma ir à/ao: Espetáculo ou exposição (teatro, museu, dança ou música)?” foram os menos discriminativos.

Quadro 6: Parâmetros de discriminação dos itens do escore de Capital Cultural.

Escore	Nº Item	Item	Discriminação
Capital Cultural	21	Com qual frequência você lê: Jornais?	1,17
	22	Com qual frequência você lê: Livros?	1,11
	23	Com qual frequência você lê: Revistas em geral?	1,15
	24	Com qual frequência você lê: Notícias na internet (ex.: blog, notícia)?	0,89
	25	Com qual frequência você costuma ir à/ao: Biblioteca?	0,93
	26	Com qual frequência você costuma ir à/ao: Espetáculo ou exposição (teatro, museu, dança ou música)?	0,89

Fonte: Elaboração própria através dos dados do SAEB/INEP.

No escore de (Des)capital Cultural, tiveram o maior e menor parâmetro de discriminação, respectivamente, os itens “Em dia de aula, quanto tempo você gasta assistindo à TV, navegando na internet ou jogando jogos eletrônicos?” e “Em dias de aula, quanto tempo você gasta fazendo trabalhos domésticos (ex.: lavando louça, limpando o quintal etc.)?”.

Quadro 7: Parâmetros de discriminação dos itens do escore(Des)capital Cultural.

Escore	Nº Item	Item	Discriminação
(Des)capital Cultural	27	Em dia de aula, quanto tempo você gasta assistindo à TV, navegando na internet ou jogando jogos eletrônicos?	0,70
	28	Em dias de aula, quanto tempo você gasta fazendo trabalhos domésticos (ex.: lavando louça, limpando o quintal etc.)?	0,66
	29	Atualmente você trabalha fora de casa (recebendo ou não um salário)?	0,68

Fonte: Elaboração própria através dos dados do SAEB/INEP.

No escore Dever de Casa, a variável "Você faz o dever de casa de Matemática?" teve parâmetro mais discriminativo do que a variável referente ao dever de Língua Portuguesa.

Quadro 8: Parâmetros de discriminação dos itens do escore Dever de Casa.

Escore	Nº Item	Item	Discriminação
Dever de Casa	30	Você faz o dever de casa de Língua Portuguesa?	3,56
	31	Você faz o dever de casa de Matemática?	3,71

Fonte: Elaboração própria através dos dados do SAEB/INEP.

Entre os itens do escore da escola, o mais discriminativo foi "Computador", e o menos discriminativo foi "Sala de atendimento especial".

Quadro 9: Parâmetros de discriminação dos itens do escore da escola.

Escore	Nº Item	Item	Discriminação
Infraestrutura Escolar	32	Sala de diretoria.	0,99
	33	Sala de professores.	1,11
	34	Laboratório de informática.	1,33
	35	Laboratório de ciências.	0,99
	36	Sala de atendimento especial.	0,65
	37	Quadra de esportes.	0,81
	38	Biblioteca.	1,05
	39	Banheiro dentro do prédio.	1,27
	40	Banheiro adequado a alunos com deficiência ou mobilidade reduzida.	0,82
	41	Dependências e vias adequadas a alunos com deficiência ou mobilidade reduzida.	0,66
	42	Aparelho de televisão.	1,57
	43	Aparelho de DVD.	1,68
	44	Copiadora.	0,96
	45	Impressora.	1,80
	46	Computador.	2,85
	47	Internet	2,17

Fonte: Elaboração própria através dos dados do SAEB/INEP.

A tabela 3 apresenta a classificação de cada item quanto ao parâmetro de discriminação, de acordo com as categorias estabelecidas por Albuquerque e Tróccoli (2004).

Tabela 3: Classificação dos itens quanto ao parâmetro de discriminação.

Discriminação	Intervalo	Item	% de Itens
Nenhuma	a=0	-	-
Muito baixa	0,01≤ a ≤0,34	-	-
Baixa	0,35≤ a ≤0,64	10	2,13%
Moderada	0,65≤ a ≤1,34	1, 2, 3, 4, 5, 6, 8, 9, 11, 12, 13, 14, 15, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 46	72,34%
Alta	1,35≤ a ≤1,69	7, 42, 45	6,39%
Muito alta	a ≥ 1,70	16, 17, 18, 19, 30, 31, 43, 44, 47	19,14%
TOTAL	-	47	100%

Fonte: Elaboração própria através dos dados do SAEB/INEP.

Acompanhando a literatura (NETO et al, 2013), optou-se por não excluir itens, uma vez que nenhum apresentou discriminação muito baixa ou categoria inferior. Posteriormente, os escores foram padronizados para assumirem média 50 e desvio padrão 10 e inseridos no modelo de regressão hierárquico. Assim, na especificação dos modelos considerou-se os seguintes escores: socioeconômico, intervenção, capital e des(capital) cultural, dever de casa e infraestrutura escolar.

A Tabela 4 apresenta a análise descritiva das demais variáveis inseridas na regressão.

Tabela 4: Análise descritiva das variáveis inseridas na regressão (que não entraram na análise TRI).

Variável	Obs	Média	Desv. Pad.	Min	Max
Proficiência média	14.297	214,38	46,97	91,29	334,73
Feminino	14.297	0,5118	0,4999	0	1
Branca	14.297	0,1939	0,3954	0	1
Reprovado	14.297	0,1697	0,3754	0	1
Municipal	14.297	0,4600	0,4984	0	1

Fonte: Elaboração própria através dos dados do SAEB/INEP.

O valor médio das variáveis binárias corresponde à proporção de casos para a respectiva categoria. Portanto, pode-se verificar que 51,18% dos alunos analisados são do sexo feminino. Do total de alunos, 19,39% se declararam brancos. Aproximadamente 17% dos alunos já foram reprovados em série(s) anterior(es). Quanto à dependência administrativa, 46% dos alunos analisados estudavam em escola municipal. A variável dependente (proficiência) esteve entre os valores de 91,29 e 334,73, com valor médio de 214,38.

3.4.2 Resultados dos modelos de regressão hierárquica

A tabela 5 apresenta os resultados dos modelos hierárquicos estimados. Na parte dos efeitos fixos, encontram-se as relações entre a variável dependente as variáveis explicativas. Na parte dos efeitos aleatórios podem ser encontradas as "combinações de variáveis explicativas e termos aleatórios" (Fávero e Belfiore, 2017, p. 856). Foram testados efeitos aleatórios nas inclinações, bem como interações das variáveis dos diferentes níveis; entretanto, não tiveram significância estatística. Desse modo, nos efeitos aleatórios dos modelos estimados serão apresentadas somente as estimativas da variância do termo de erro aleatório referente ao nível 2, associado ao intercepto - var(cons); e da variância do termo de erro associado à medida da proficiência do aluno, não explicado pelo modelo - var(residual).

Através das estimativas do modelo nulo foi possível calcular o coeficiente de intra-correlação (ICC), representando a proporção da variância na proficiência dos alunos que é devida às características das escolas. O valor do ICC para o nível da escola (conforme figura 13 do anexo 1) foi de 14,73%, resultado semelhante ao obtido por Gonçalves (2008). Dessa forma, legitima-se a utilização do modelo hierárquico em vez do modelo clássico de regressão.

Tabela 5: Resultados dos modelos hierárquicos.

Variável Dependente: Proficiência média	**Coeficientes**			
Efeitos fixos	**Modelo Nulo**	**Modelo 1**	**Modelo 2**	**Modelo 3**
Nível 1 - Aluno				
Feminino	-	7,2437**	6,9449**	7,1805**
Branca	-	0,1609	-0,0069	0,2056
Reprovado	-	-25,91**	-24,49**	-25,88**
Escore socioeconômico do aluno	-	0,7804**	0,6211**	0,7117**
Escore de intervenção	-	0,5652**	0,5391**	0,5655**
Escore de capital cultural	-	0,5437**	0,5172**	0,5457**
Escore de (des)capital cultural	-	-0,1227**	-0,1270**	-0,1229**
Escore dever de casa	-	0,8688**	0,8119**	0,8651**
Nível 2 - Escola				
Municipal	-	-5,6611*	0,8807	-4,112*
Escore de infraestrutura da escola	-	0,2471*	0,0201	0,080
Proficiência média da escola	-	-	0,7460**	-
Nível socioeconômico médio da escola	-	-	-	1,4112**
Constante	209,61**	70,14**	-65.15**	11,70
Efeitos Aleatórios				
Escola				
var(cons)	317,30**	152,95**	4.74e-12	115,08**
var(residual)	1837,36**	1.499,89**	1.477.85**	1.500,18**
Obs.	14.297			
Grupos (nível 2)	386			

Nota: *p<0,01; ** p<0,001.
Fonte: Elaboração própria, a partir dos dados do SAEB e do Censo Escolar/INEP.

Na análise do efeito de cada variável explicativa sobre a proficiência média do aluno considera-se as demais variáveis constantes. No modelo 1, todas as variáveis tiveram o sinal esperado para os coeficientes estimados. Pode-se observar neste modelo, e também nos demais, que somente a variável indicativa de cor não teve seu coeficiente significativo, fato semelhante ao observado no estudo de Macedo (2004). Verifica-se que os alunos do sexo feminino das escolas do Norte de Minas apresentaram uma proficiência escolar maior que a dos alunos do sexo masculino, em cerca de sete (7) pontos.

O fato de o aluno ter sido reprovado em alguma das séries anteriores tendeu a impactar forte e negativamente seu rendimento escolar para a série analisada (4ª série/5º ano), corroborando com os estudos de Luz (2006). De acordo com Palermo, Silva e Novellino (2014), o valor máximo estimado para o coeficiente de reprovação em estudos já realizados foi de 22; assim, o coeficiente de reprovação estimado para a região do Norte de Minas supera um pouco esse valor. O nível socioeconômico do aluno impactou positivamente o seu rendimento, resultado semelhante ao verificado por Schiefelbein e Simmons (2013), entre outros.

Observa-se ainda que quanto maior o escore de intervenção, maior tende a ser a média da proficiência dos alunos em questão, fato também verificado nos estudos de Barbosa e Fernandes (2001). O coeficiente do escore de capital cultural foi significativo, influenciando positivamente o rendimento escolar dos alunos norte-mineiros, resultado também verificado em Menezes-Filho (2007). Quanto ao escore de (des)capital cultural, este influenciou negativamente a proficiência, assim como nos estudos de Luz (2006) e Menezes-Filho (2007). *Coeteris paribus*, quanto maior o escore de dever de casa maior tendeu a ser a média da proficiência escolar, similar aos estudos de Macedo (2004) e Schiefelbein e Simmons (2013).

Na série analisada, os alunos da rede municipal apresentaram uma redução de 5,66 pontos na proficiência escolar em relação aos alunos da rede estadual. Essa redução foi encontrada também por Luz (2006) e Pinheiro (2005). O escore da escola (infraestrutura) afeta positivamente o rendimento escolar dos alunos em questão, conforme visto em Barbosa e Fernandes (2001) e Soares e Mendonça (2003).

Dois exercícios adicionais foram feitos, tendo por objetivo verificar qual o efeito da proficiência média da escola e do nível socioeconômico médio da escola sobre a proficiência do aluno (modelos 2 e 3 respectivamente).

A inclusão da proficiência média da escola teve como propósito verificar se a proficiência dos colegas, chamada de efeito dos pares, influencia o desempenho escolar de um aluno específico; procedimento semelhante ao adotado por Gonçalves, Rios-Neto e César (2011). Tal variável teve coeficiente significativo, com sinal positivo. Vale ressaltar, contudo, que os coeficientes das variáveis relacionadas à dependência administrativa da escola e ao escore de infraestrutura escolar perderam suas significâncias. A esse respeito, foi constatado que a proficiência média da escola e o escore de infraestrutura escolar apresentam correlação significativa (Figura 12 do anexo 1), ou seja, quanto maior a proficiência dos colegas, melhor a infraestrutura da escola (controlando-se para os demais fatores), fato que pode ter colaborado para a perda de significância do coeficiente dessa última variável. Além disso, a

variância do termo de erro aleatório referente ao nível da escola (associado ao intercepto) perdeu significância.

No modelo 3 foi adicionada a variável nível socioeconômico médio da escola com o propósito de verificar se o nível socioeconômico dos colegas influencia o desempenho escolar de um aluno específico. Tal variável teve coeficiente significativo, e apresentou sinal positivo. Entretanto, cabe certa ressalva em sua interpretação. Através da figura 12 do anexo 1, percebe-se que o escore da escola e o nível socioeconômico médio da escola estão positivamente correlacionados. Fato que pode ter levado a perda de significância do escore da escola; indicando também que os alunos que possuíam melhores níveis socioeconômicos tendiam a estudar nas escolas mais bem estruturadas. Diferente do que ocorreu no modelo 2, os efeitos aleatórios associados à escola permaneceram significativos.

Apesar desses aspectos referentes aos modelos 2 e 3, em geral, houve constância nos valores estimados para os parâmetros dos três modelos (variáveis de nível 1).

Considerações finais

O presente estudo compreendeu duas análises principais e inter-relacionadas: verificar a existência da relação entre educação e desenvolvimento no Norte de Minas e investigar os determinantes do desempenho escolar nas séries iniciais do ensino fundamental, nesta Região. Através da descrição da formação econômica e social do Norte de Minas, foi possível compreender, em certa medida, os seus indicadores econômicos e educacionais, e a forma como estão distribuídos no espaço; sendo constatado que essa região encontra-se entre as menos desenvolvidas do Estado.

A análise de correlação revelou a existência da relação positiva entre o indicador de desenvolvimento econômico (IFDM) e o indicador de qualidade da educação (IDEB-M-AI). Por meio desse resultado, confirma-se a hipótese 1. Tal fato é relevante ao mostrar que é possível se beneficiar de tal relação, mesmo em regiões menos favorecidas socioeconomicamente como o Norte de Minas.

Através da análise espacial foi possível verificar que em alguns municípios norte-mineiros a distribuição conjunta do IFDM e do IDEB não ocorre de forma aleatória, no espaço, confirmando a hipótese 2. Dessa forma, é possível também se favorecer do efeito transbordamento da relação entre educação e desenvolvimento (com a ressalva de verificar os casos em que municípios apresentaram o IDEB elevado e tenderam a apresentar vizinhos próximos com baixos valores do IFDM).

A respeito da análise dos determinantes do desempenho escolar nas series iniciais (5º ano) do ensino fundamental, na região norte-mineira, foi possível constatar que a reprovação em séries anteriores afeta forte e negativamente o rendimento escolar dos alunos em questão. Esse fato deixa evidente que a trajetória escolar passada não está desvinculada do desempenho corrente do aluno, talvez indicando que as dificuldades obtidas em séries anteriores continuam existindo. Contudo, os resultados mostraram que o aluno é capaz de se beneficiar pelo fato de estudarem com colegas de maior proficiência e melhor nível socioeconômico. Esse fato indica que a distribuição dos alunos por turmas não deve ocorrer de forma arbitrária.

Não se constatou diferenças entre proficiências de alunos brancos e não brancos. Além disso, alunos da rede municipal tenderam a apresentar redução na proficiência em relação aos alunos da rede estadual. Comprovou-se ainda que os alunos do sexo feminino tenderam a apresentar proficiências maiores que alunos do sexo masculino. Ademais, o nível socioeconômico do aluno, os escores de intervenção, de capital cultural, de dever de casa e o

da escola afetaram positivamente o rendimento escolar dos alunos; ao passo que o des(capital) cultural o afetou negativamente. Os resultados obtidos para os alunos matriculados nas escolas norte-mineiras são consoantes aos observados na literatura sobre o tema e confirmaram a hipótese 3.

Por fim, através das análises realizadas, foi possível verificar que a região do Norte de Minas pode ser considerada uma das menos desenvolvidas do Estado, supostamente em função da sua formação econômica. Além disso, constatou-se uma relação positiva entre os indicadores de desenvolvimento econômico e da qualidade da educação. Desta forma, é importante se promover políticas educacionais com o propósito de elevar o nível de desenvolvimento econômico. Como visto no capítulo I, com maiores investimentos em educação, propicia-se maior formação de capital humano e, por consequência, maior desenvolvimento econômico. Os resultados da análise dos determinantes do desempenho escolar podem ser utilizados como forma de orientação para tais políticas, na Região.

Referências

ALBUQUERQUE, Anelise Salazar; TRÓCCOLI, Bartholomeu Tôrres. Desenvolvimento de uma escala de bem-estar subjetivo. **Psicologia: teoria e pesquisa**, v. 20, n. 2, p. 153-164, 2004.

ALMEIDA, Eduardo Simões de. Curso de Econometria Espacial Aplicada (Apostila). **Universidade de São Paulo (USP)–Escola Superior de Agricultura "Luiz de Queiroz" (ESALQ). Piracicaba**, 2004.

ANDRADE, Dalton Francisco de; TAVARES, Heliton Ribeiro; VALLE, Raquel da Cunha. Teoria da Resposta ao Item: conceitos e aplicações. **ABE, São Paulo**, 2000.

ARAUJO, Eutalia Aparecida Candido de; ANDRADE, Dalton Francisco de; BORTOLOTTI, Silvana Ligia Vincenzi. Teoria da resposta ao item. **Revista da Escola de Enfermagem da USP**, v. 43, n. spe, p. 1000-1008, 2009.

ATLAS DO DESENVOLVIMENTO HUMANO NO BRASIL. **Perfil dos Municípios de Minas Gerais** – Disponível em: <http://atlasbrasil.org.br/2013/pt/consulta/ >. Acesso em: 20 ago. 2013.

BARBOSA, Maria EF; FERNANDES, Cristiano. A escola brasileira faz diferença? Uma investigação dos efeitos da escola na proficiência em Matemática dos alunos da 4ª série. **Promoção, ciclos e avaliação educacional**. Porto Alegre: ArtMed, p. 155-172, 2001.

BECKER, G.S."**HumanCapital**. "The Concise Encyclopedia of Economics. 2008. Library of Economics and Liberty. <http://www.econlib.org/library/Enc/HumanCapital.html>.Acessadoem 31 de maio de 2016.

BECKER, G. S. **Human capital: a theoretical and empirical analysis, with special reference to education.** Third Edition. Chicago: The University of Chicago Press, 1993.

BIONDI, Roberta Loboda; FELÍCIO, Fabiana de. Atributos escolares e o desempenho dos estudantes: uma análise em painel dos dados do Saeb. **BRASILIA: INEP/MEC**, 2007.

BORTOLOTTI, Silvana Ligia Vincenzi. Aplicação de um modelo de desdobramento graduado generalizado da teoria da resposta ao item-TRI. [Dissertação de Mestrado]. **Universidade Federal de Santa Catarina. Florianópolis, 2003.**

BRAGA, Bruno Marx de Aquino. **Teoria da resposta ao item: o uso do modelo de Samejima como proposta de correção para itens discursivos**. [Dissertação de Mestrado]. Universidade de Brasília. Brasília, 2015.

CALEIRO, Antônio. **Educação e Desenvolvimento: que tipo de relação existe?** Departamento de Economia – Universidade de Évora, Portugal. 2010.

CUNHA, Luiz Antônio. **Educação e desenvolvimento social no Brasil**. Rio de Janeiro: Francisco Alves, 1985.

CASTRO, Claudio de Moura. **Desenvolvimento Econômico, Educação e Educabilidade.** 2 ed. Rio de Janeiro, Tempo Brasileiro, FENAME, 1976.

FÁVERO, Luiz Paulo; BELFIORE, Patrícia. **Manual de análise de dados: estatística e modelagem multivariada com Excel®, SPSS® e Stata®.** 1ª ed. Rio de Janeiro: Elsevier, 2017.

FEDERAÇÃO DAS INDÚSTRIAS DO ESTADO DO RIO DE JANEIRO – FIRJAN. **IFDM por municípios brasileiros**. Rio de Janeiro: Sistema FIRJAN - pesquisas e estudos socioeconômicos. Disponível em: <http://www.firjan.com.br/ifdm/>. Acesso em: 21 jul. 2017.

FERNANDES, Danielle Cireno; HELAL, Diogo Henrique; NEVES, Jorge Alexandre. **Educação, Trabalho e Desigualdade Social.** Belo Horizonte, MG, Argvmentvm, 2009.

FERRARO, Alceu Ravanello; KREIDLOW, Daniel. Analfabetismo no Brasil: configuração e gênese das desigualdades regionais. **Educação & Realidade**, v. 29, n. 2, 2004.

FIGUEIREDO, D.B.F. Filho; SILVA J.A. Jr. **Desvendando os Mistérios do Coeficiente de Correlação de Pearson (r)**. Revista Política Hoje. 2009.

FRIEDMAN, Milton. **Capitalismo e Liberdade**. Rio de Janeiro: Arte Nova, 1977.

GIL, Antônio Carlos. **Como elaborar projetos de pesquisa.** 5. ed. São Paulo: Atlas, 2010.

GONÇALVES, Maria Elizete. Análise de sobrevivência e modelos hierárquicos logísticos longitudinais: uma aplicação à análise da trajetória escolar (4ª a 8ª série-ensino fundamental).[Tese de Doutorado]. **Universidade Federal de Minas Gerais, Belo Horizonte, 2008.**

GONÇALVES, Maria Elizete. **Avaliação do Desempenho Escolar no Ensino Fundamental de Minas Gerais: Uma Análise Baseada no Programa Bolsa Família e no Ideb.** Departamento de Economia – Universidade Estadual de Montes Claros – Unimontes, Brasil. 2014.

GONÇALVES, Maria Elizete; RIOS-NETO, Eduardo Luiz Gonçalves; CÉSAR, Cibele Comini. **Determinantes Socioeconômicos, Demográficos e Institucionais da Repetência no Ensino Fundamental Brasileiro.** Code, Anais do I Circuito de Debates Acadêmicos. 2011.

GUJARATI, Damodar N., and PORTER, Dawn C. **Econometria Básica**. 5ª ed. AMGH Editora, 2011.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Produto Interno Bruto a preços correntes e Produto Interno Bruto per capita segundo as Grandes Regiões, as Unidades da Federação e os Municípios – 2010**. Rio de Janeiro: IBGE. Disponível em: <https://www.ibge.gov.br/home/estatistica/economia/pibmunicipios/2010/>. Acesso em: 30 jan. 2017.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Microdados da Aneb e da Anresc 2015**. Brasília: Inep, 2017. Disponível em: <http://portal.inep.gov.br/basica-levantamentos-acessar>. Acesso em: 30 jan. 2017.

KELNIAR, Vanessa Carla; LOPES, Janete Leige; PONTILI, Rosangela Maria. **A Teoria Do Capital Humano: Revisitando Conceitos.** EPCT – VIII Encontro de Produção Científica e Tecnológica, 2013.

LIRA, Sachiko Araki; NETO, Anselmo Chaves. Coeficientes de correlação para variáveis ordinais e dicotômicas derivados do coeficiente linear de Pearson. **Ciência & Engenharia**, v. 15, n. 1/2, p. 45-53, 2006.

LUZ, Luciana Soares. Os determinantes do desempenho escolar: a estratificação educacional e o efeito valor adicionado. **Anais: XV Encontro Nacional de Estudos Populacionais, ABEP**, 2006.

MACEDO, Gláucia Alves. Fatores Associados ao Rendimento Escolar de Alunos da 5a série (2000) - uma abordagem do valor adicionado. **Anais: XIV Encontro Nacional de Estudos Populacionais**, 2004.

MARSHALL, Alfred. **Princípios de economia: tratado introdutório**. Vol I. Tradução revista de Rômulo Almeida e Ottolmy Strauch. São Paulo: Abril Cultural, 1982.

MARX, Karl; **O Capital, Volume I E II.** Apresentação de Jacob Gorender; Coordenação e revisãode Paul Singer; Tradução de Regis Barbosa e Flávio R. Kothe; Editora Nova Cultural Ltda.; 1996,Círculo do Livro Ltda. Títulosoriginais: Value, Price and Profit; Das Kapital - Kritik der Politisch en konomie. Apresentação de autoria de Winston Fritsch.

MCKEOWN, Rosalyn (2002). **Education for Sustainable Development Toolkit**, versão 2, Julho. Disponível em<http://www.esdtoolkit.org>.

MENEZES-FILHO, Naércio Aquino. **Os determinantes do desempenho escolar do Brasil**. Instituto Futuro Brasil, Ibmec-SP e FEA-USP, 2007.

NAZARENO, Luísa de Azevedo. Desigualdades regionais de renda no Brasil: potencial de queda por meio da educação [Dissertação de Mestrado]. **Brasília: Centro de Estudos Avançados Multidisciplinares, 2016.**

NETO, Joaquim José Soares; Jesus, Girlene Ribeiro de; KARINO, Camila Akemi; ANDRADE, Dalton Francisco de. Uma escala para medir a infraestrutura escolar. **Estudos em Avaliação Educacional**, v. 24, n. 54, p. 78-99, 2013.

OLIVEIRA, Marcos Fábio Martins de; RODRIGUES, Luciene; CARDOSO, José Maria Alves; BOTELHO, Tarcísio Rodrigues. **Formação social e econômica do Norte de Minas.** Marcos Fábio Martins de Oliveira, Luciene Rodrigues (Organizadores). Montes Claros, MG: Editora Unimontes, 2000.

PALERMO, G.A.; SILVA, D.B.N.; NOVELLINO, M.S.F. Fatores associados ao desempenho escolar: uma análise da proficiência em matemática dos alunos do 5º ano do ensino

fundamental da rede municipal do Rio de Janeiro. **Revista Brasileira de Estudos de População**, v. 31, n. 2, p. 367-394, 2014.

PASQUALI, Luiz; PRIMI, Ricardo. **Fundamentos da teoria da resposta ao item**: **TRI**. Avaliação Psicológica, v. 2, n. 2, p. 99-110, 2003.

PINHEIRO, S. M. C. Modelo linear hierárquico: um método alternativo para análise de desempenho escolar. [Dissertação de Mestrado]. **Recife: Departamento de Estatística, Universidade Federal de Pernambuco**, 2005.

QUINTANA, Tatiana Rodrigues. Teoria de resposta do item (TRI) na análise de pesquisa de desempenho na gestão da qualidade em uma siderúrgica. [Monografia] Universidade Federal do Rio Grande do Sul, 2010.

SCHIEFELBEIN, Ernesto; SIMMONS, John. Os determinantes do desempenho escolar: uma revisão de pesquisas nos países em desenvolvimento. **Cadernos de pesquisa**, n. 35, p. 53-71, 2013.

SCHULTZ, T. W. **O Capital Humano: investimentos em educação e pesquisa.** Tradução de MarcoAurélio de Moura Matos. Rio de Janeiro: Zahar, 1973.

SCHULTZ, T. W. **O valor econômico da educação.** 2ª ed. Trad.de P.S. Werneck. Rev. Técnica de C.A. Pajuaba. Rio de Janeiro: Zahar, 1973.

SILVA, Etienne Pereira da. **Teoria De Resposta Ao Item: Análise de Atitude dos Graduandos em Relação às Disciplinas de Estatística.** Monografia. Universidade de Brasília. Brasília, 2014.

SMITH, Adam; **A Riqueza Das Nações - Investigação Sobre Sua Natureza E Suas Causas**.Volume I e II; Ed. Nova Cultural, 1996, Círculo do Livro Ltda.; Título original: An Inquiry into the Nature and Causes of the Wealth of Nations.

SOARES, Tufi Machado; MENDONÇA, Márcia Cristina Meneghin. Construção de um modelo de regressão hierárquico para os dados do SIMAVE-2000. **Pesquisa Operacional**, v. 23, n. 3, p. 421-441, 2003.

THOMAS, J.J. **Introdução à análise estatística**. Tradução de Antônio Amaro Cirurgião. Rio de Janeiro: Zahar Editores, 1978.

WONNACOTT, Ronald J; WONNACOTT, Thomas H. **Econometria**. 2ª ed. Tradução de Maria C. Silva. Rio de Janeiro: Livros Técnicos e Científicos, 1978.

WOOLDRIDGE, Jeffrey. **Introductory econometrics: A modern approach**. 5ª ed. Nelson Education, 2011.

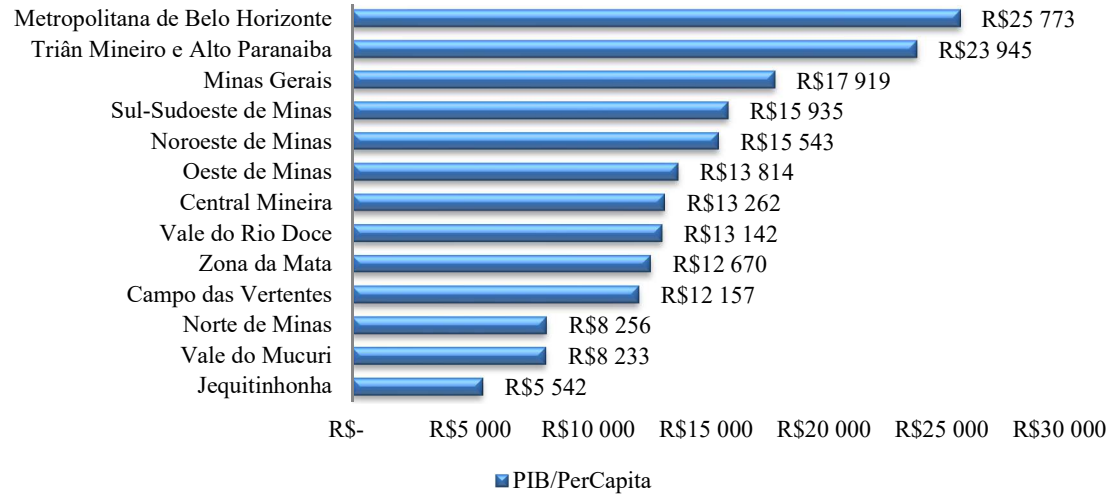
Anexo I

Tabela 6: PIB – mesorregiões de Minas Gerais – 2010.

Mesorregião	**PIB (R$ mil)**	**% PIB**
Metropolitana de Belo Horizonte	R$ 160 689 357	45,76%
Triângulo Mineiro e Alto Paranaíba	R$ 51 269 912	14,60%
Sul-Sudoeste de Minas	R$ 38 867 345	11,07%
Zona da Mata	R$ 27 559 730	7,85%
Vale do Rio Doce	R$ 21 300 089	6,07%
Norte de Minas	**R$ 13 297 002**	**3,79%**
Oeste de Minas	R$ 13 193 233	3,76%
Campo das Vertentes	R$ 6 739 643	1,92%
Noroeste de Minas	R$ 5 694 636	1,62%
Central Mineira	R$ 5 473 060	1,56%
Jequitinhonha	R$ 3 876 872	1,10%
Vale do Mucuri	R$ 3 173 036	0,90%
Minas Gerais	**R$ 351 133 915**	**100%**

Fonte: elaboração própria a partir dos dados do IBGE.

Gráfico 1: PIB *per capita* – mesorregiões de Minas Gerais – 2010.



Fonte: elaboração própria a partir dos dados do IBGE e do IMRS.

Tabela 7: Média do IMRS-SHM por mesorregião de Minas Gerais – 2010.

Mesorregião	**Média do IMRS-SHM**
Metropolitana de Belo Horizonte	0,51
Sul-Sudoeste de Minas	0,48
Oeste de Minas	0,48
Zona da Mata	0,47
Triângulo Mineiro e Alto Paranaíba	0,46
Campo das Vertentes	0,45
Noroeste de Minas	0,44
Central Mineira	0,42
Vale do Rio Doce	0,41
Jequitinhonha	0,36
Norte de Minas	**0,35**
Vale do Mucuri	0,31
Minas Gerais	**0,44**

Fonte: elaboração própria a partir dos dados do IMRS.

Tabela 8: Média do IFDM por mesorregião – Minas Gerais – 2013.

Mesorregião	**Média de IFDM**
Oeste de Minas	0,74
Triângulo Mineiro e Alto Paranaíba	0,72
Sul-Sudoeste de Minas	0,72
Central Mineira	0,71
Metropolitana de Belo Horizonte	0,70
Noroeste de Minas	0,70
Campo das Vertentes	0,68
Zona da Mata	0,67
Vale do Rio Doce	0,64
Norte de Minas	**0,61**
Jequitinhonha	0,59
Vale do Mucuri	0,57
Média dos municípios	**0,68**

Fonte: elaboração própria a partir dos dados da FIRJAN.

Tabela 9: Média da taxa de analfabetismo dos municípios mineiros por mesorregião – pessoas com 15 anos ou mais de idade – 2010.

Mesorregião	**Analfabetismo**
Oeste de Minas	8,7
Campo das Vertentes	8,7
Triângulo Mineiro e Alto Paranaíba	9,3
Sul-Sudoeste de Minas	9,6
Metropolitana de Belo Horizonte	9,8
Central Mineira	11,1
Noroeste de Minas	12,3
Zona da Mata	12,3
Vale do Rio Doce	17,6
Norte de Minas	**20,8**
Jequitinhonha	22,3
Vale do Mucuri	26,1
Minas Gerais	**8,31**

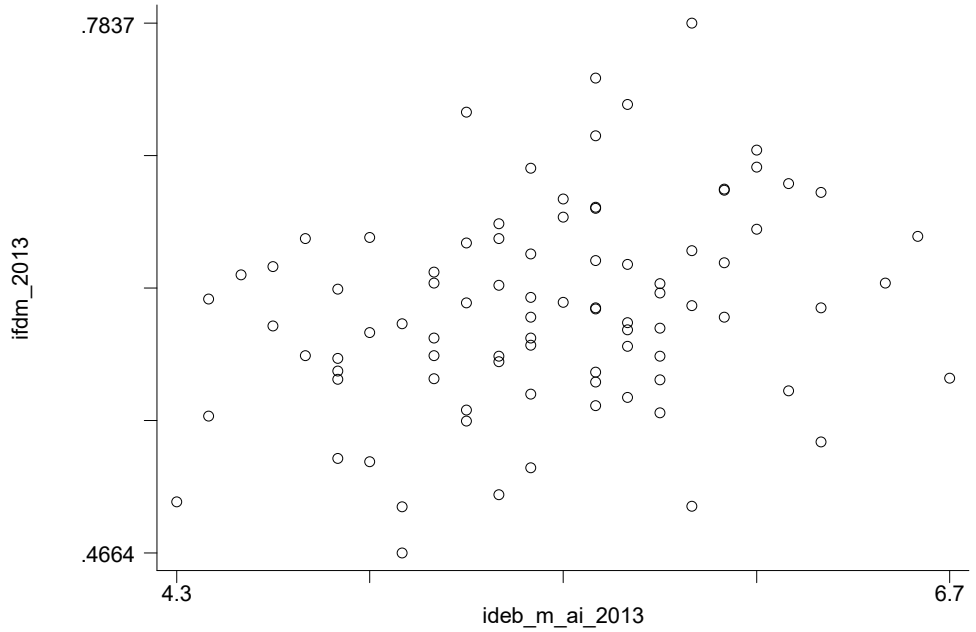
Fonte: elaboração própria a partir dos dados do IMRS.

Tabela 10: Média do IDEB-M-AI por mesorregião – Minas Gerais – 2013.

Mesorregião	**Média do IDEB-M-AI - 2013**
Oeste de Minas	6,43
Central Mineira	6,39
Sul-Sudoeste de Minas	6,25
Triângulo Mineiro e Alto Paranaíba	6,14
Campo das Vertentes	6,11
Zona da Mata	5,96
Metropolitana de Belo Horizonte	5,92
Noroeste de Minas	5,91
Jequitinhonha	5,76
Vale do Rio Doce	5,69
Norte de Minas	**5,45**
Vale do Mucuri	5,43
Média dos municípios	**5,95**

Fonte: elaboração própria a partir dos dados do SAEB/INEP.

Gráfico 2: IFDM e IDEB-M-AI - municípios norte-mineiros, 2013.



Fonte: resultados da pesquisa a partir dos dados do IFDM e do SAEB/INEP.

Figura 12: matriz de correlação das variáveis utilizadas no modelo de regressão.

	med_prof	esc_sco	esc_int	esc_cult	esc_des	esc_dev	esc_inf
med_prof	1.0000						
esc_sco	0.2939	1.0000					
esc_int	0.2151	0.1227	1.0000				
esc_cult	0.0709	-0.0792	-0.1043	1.0000			
esc_des	-0.0393	0.0395	-0.0623	0.0473	1.0000		
esc_dev	0.2911	0.1244	0.2326	-0.1129	-0.0593	1.0000	
esc_inf	0.1278	0.1449	0.0153	-0.0301	-0.0008	0.0246	1.0000
sco_esc	0.3143	0.4909	0.0867	-0.0369	0.0171	0.1212	0.2951
prof_esc	0.4352	0.3545	0.0994	-0.0066	0.0159	0.1503	0.2937

	sco_esc	prof_esc
sco_esc	1.0000	
prof_esc	0.7221	1.0000

Fonte: resultados da pesquisa.

Figura 13: modelo nulo.

```
Mixed-effects REML regression                   Number of obs     =     14,297
Group variable: id_escola                       Number of groups  =        386

                                                Obs per group:
                                                              min =          5
                                                              avg =       37.0
                                                              max =        149

                                                Wald chi2(0)      =          .
Log restricted-likelihood = -74373.409          Prob > chi2       =          .
```

media_proficiencia	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_cons	209.6063	.9987196	209.88	0.000	207.6489	211.5638

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
id_escola: Identity				
var(_cons)	317.2971	27.27866	268.0937	375.5308
var(Residual)	1837.326	22.01994	1794.67	1880.995

LR test vs. linear model: chibar2(01) = 1897.91 Prob >= chibar2 = 0.0000

Fonte: resultados da pesquisa.

Figura 14: Coeficiente de intra-correlação.

Level	ICC	Std. Err.	[95% Conf.	Interval]
id_escola	.1472634	.0109361	.1270924	.1700123

Fonte: resultados da pesquisa.

Figura 15: modelo hierárquico 1.

```
Mixed-effects REML regression                   Number of obs     =     14,297
Group variable: id_escola                       Number of groups  =        386

                                                Obs per group:
                                                              min =          5
                                                              avg =       37.0
                                                              max =        149

                                                Wald chi2(10)     =    3470.03
Log restricted-likelihood = -72848.911          Prob > chi2       =     0.0000
```

media_proficiencia	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
feminino	7.243692	.6648816	10.89	0.000	5.940548	8.546836
branca	.1609164	.8343113	0.19	0.847	-1.474304	1.796137
reprovado	-25.91255	.9156292	-28.30	0.000	-27.70715	-24.11795
escore_sco	.7803603	.0372984	20.92	0.000	.7072567	.8534639
escore_interve	.5651648	.0365897	15.45	0.000	.4934504	.6368793
escore_cap_cult	.5437576	.0331921	16.38	0.000	.4787024	.6088129
escore_desc	-.1226962	.0325637	-3.77	0.000	-.1865199	-.0588725
escore_dever_de_casa	.868779	.0355916	24.41	0.000	.7990207	.9385372
municipal_dep_adm_escola	-5.661077	1.689614	-3.35	0.001	-8.97266	-2.349495
esc_infr_esc	.2471383	.080063	3.09	0.002	.0902178	.4040589
_cons	70.1459	5.832404	12.03	0.000	58.7146	81.5772

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
id_escola: Identity				
var(_cons)	152.9502	14.96488	126.2604	185.2819
var(Residual)	1499.895	17.99025	1465.046	1535.573

LR test vs. linear model: chibar2(01) = 798.60 Prob >= chibar2 = 0.0000

Fonte: resultados da pesquisa.

Figura 16: modelo hierárquico 2.

```
Mixed-effects REML regression                   Number of obs     =     14,297
Group variable: id_escola                       Number of groups  =        386

                                                Obs per group:
                                                              min =          5
                                                              avg =       37.0
                                                              max =        149

                                                Wald chi2(11)     =    7056.07
Log restricted-likelihood = -72468.449          Prob > chi2       =     0.0000
```

media_proficiencia	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
feminino	6.944901	.6529656	10.64	0.000	5.665112	8.22469
branca	-.0068608	.8152304	-0.01	0.993	-1.604683	1.590961
reprovado	-24.49469	.8910421	-27.49	0.000	-26.2411	-22.74828
escore_sco	.6211003	.0353728	17.56	0.000	.551771	.6904297
escore_interve	.5391211	.0358675	15.03	0.000	.4688221	.6094201
escore_cap_cult	.517249	.0322248	16.05	0.000	.4540895	.5804084
escore_desc	-.1270404	.0318437	-3.99	0.000	-.1894529	-.0646278
escore_dever_de_casa	.8118741	.0346664	23.42	0.000	.7439292	.8798189
municipal_dep_adm_escola	.8807038	.7518431	1.17	0.241	-.5928816	2.354289
esc_infr_esc	.0201142	.0363638	0.55	0.580	-.0511576	.0913859
prof_esc	.7460247	.0183404	40.68	0.000	.7100782	.7819713
_cons	-65.15343	5.089477	-12.80	0.000	-75.12862	-55.17824

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
id_escola: Identity				
var(_cons)	4.74e-12	1.01e-11	7.16e-14	3.14e-10
var(Residual)	1477.85	17.48662	1443.972	1512.524

```
LR test vs. linear model: chibar2(01) = 0.00          Prob >= chibar2 = 1.0000
```

Fonte: resultados da pesquisa.

Figura 17: modelo hierárquico 3.

```
Mixed-effects REML regression                   Number of obs     =     14,297
Group variable: id_escola                       Number of groups  =        386

                                                Obs per group:
                                                              min =          5
                                                              avg =       37.0
                                                              max =        149

                                                Wald chi2(11)     =    3633.73
Log restricted-likelihood = -72811.826          Prob > chi2       =     0.0000
```

media_proficiencia	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
feminino	7.180492	.6644448	10.81	0.000	5.878204	8.48278
branca	.2055719	.8334651	0.25	0.805	-1.42799	1.839133
reprovado	-25.88095	.9143587	-28.31	0.000	-27.67306	-24.08884
escore_sco	.7116995	.0381998	18.63	0.000	.6368293	.7865698
escore_interve	.5655087	.0365556	15.47	0.000	.493861	.6371564
escore_cap_cult	.5456634	.0331411	16.46	0.000	.4807081	.6106188
escore_desc	-.1229303	.0325335	-3.78	0.000	-.1866948	-.0591657
escore_dever_de_casa	.8650934	.0355463	24.34	0.000	.7954239	.9347628
municipal_dep_adm_escola	-4.111832	1.529224	-2.69	0.007	-7.109055	-1.114609
esc_infr_esc	.0803843	.0743881	1.08	0.280	-.0654136	.2261822
sco_escola	1.411242	.1534679	9.20	0.000	1.11045	1.712034
_cons	11.70148	8.347413	1.40	0.161	-4.65915	28.06211

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
id_escola: Identity				
var(_cons)	115.0801	12.20531	93.48077	141.67
var(Residual)	1500.181	17.99521	1465.323	1535.869

```
LR test vs. linear model: chibar2(01) = 520.49        Prob >= chibar2 = 0.0000
```

Fonte: resultados da pesquisa.