

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

JOSIANE SOUZA DE PAULA

**DETERMINANTES DO DESEMPENHO EDUCACIONAL
NO ESTADO DE MINAS GERAIS**

UBERLÂNDIA
2017

JOSIANE SOUZA DE PAULA

**DETERMINANTES DO DESEMPENHO EDUCACIONAL
NO ESTADO DE MINAS GERAIS**

Tese apresentada ao programa de Pós-Graduação em Economia da Universidade Federal de Uberlândia, como requisito parcial à obtenção do título de Doutora em Economia.

Área de concentração: Desenvolvimento Econômico

Orientador: Prof. Dr. Cleomar Gomes da Silva

Co-orientadora: Profa. Dra. Ana Maria de Paiva Franco

UBERLÂNDIA
2017

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da UFU, MG, Brasil.

P324d Paula, Josiane Souza de, 1989-
2017 Determinantes do desempenho educacional no Estado de Minas
Gerais / Josiane Souza de Paula. - 2017.
 157 f. : il.

Orientador: Cleomar Gomes da Silva.
Coorientadora: Ana Maria de Paiva Franco.
Tese (doutorado) - Universidade Federal de Uberlândia, Programa
de Pós-Graduação em Economia.
Inclui bibliografia.

1. Economia - Teses. 2. Educação - Aspectos econômicos - Teses. 3. Educação e Estado - Minas Gerais - Teses. 4. Educação - Aspectos econômicos - Pesquisa - Teses. I. Silva, Cleomar Gomes da. II. Franco, Ana Maria de Paiva. III. Universidade Federal de Uberlândia. Programa de Pós-Graduação em Economia. IV. Título.

CDU: 330

JOSIANE SOUZA DE PAULA

**DETERMINANTES DO DESEMPENHO EDUCACIONAL
NO ESTADO DE MINAS GERAIS**

Tese apresentada ao programa de Pós-Graduação em Economia da Universidade Federal de Uberlândia, como requisito parcial à obtenção do título de Doutora em Economia.

Uberlândia, 24 de março de 2017.

Banca Examinadora:

Prof. Dr. Cleomar Gomes da Silva (orientador)

Profa. Dra. Ana Maria de Paiva Franco (UFU)

Prof. Dr. José Waldemar da Silva (UFU)

Prof. Dr. Ricardo da Silva Freguglia (UFJF)

Profa. Dra. Estela Maris Pereira Bereta (UFSCAR)

“A educação é a arma mais poderosa que você pode usar para mudar o mundo.”

Nelson Mandela

AGRADECIMENTOS

Neste ano, completam-se 10 anos desde que iniciei meus estudos em economia. Durante todos esses anos várias pessoas contribuíram para que eu pudesse concluir o doutorado. Inicialmente preciso agradecer a Deus por estar sempre presente na minha vida, por todas as portas abertas e por me dar forças nos momentos de fraqueza.

Agradeço ao meu esposo, Marcos Batista, pelos momentos de carinho, compreensão e por todos os abraços recebidos enquanto escrevia a tese. Aos meus pais, Rafael e Irvone, por todo capital econômico, social e cultural investidos na minha educação. A minha amada avó Léa (*in memoriam*) que me mostrou a importância dos estudos e que vivia dizendo: “Quem não tem educação não tem nada na vida”.

Aos professores do IE/UFU, por todos os ensinamentos passados e pelo incentivo dado durante essa caminhada. Em especial agradeço ao professor Henrique Neder, principal incentivador do tema desse trabalho, por toda sua generosidade em ensinar e em compartilhar suas ideias. A professora Rosana Ribeiro pelos conselhos, livros e sugestões feitas na fase inicial desse trabalho. Ao professor José Flores, meu orientador na época da iniciação científica, pelos seus ensinamentos e por me incentivar a ingressar na vida acadêmica.

Ao professor Cleomar Gomes, pela orientação, paciência e pelas contribuições na realização desse trabalho. Agradeço imensamente a professora Ana Maria Franco, pela sua amizade, apoio, dedicação e por todos ensinamentos passados durante esse tempo. Sou grata também ao professor José Waldemar pela sua valiosa ajuda com a Teoria da Resposta ao Item, pela disposição em ensinar e por todo auxílio. Aos funcionários do IE/UFU, em especial, a Camila Bazani, por sempre atender minhas solicitações com um sorriso.

Novamente, agradeço aos professores José Waldemar e Ana Maria Franco, juntamente aos professores Ricardo Freguglia e Estela Maris por aceitarem o convite para participar da banca examinadora e por todas as contribuições para esta tese e para os trabalhos dela derivados.

Agradeço aos amigos da turma de doutorado Antônio Marcos, Alzemar Delfino, Filipe Prado, Juliene Barbosa e Maria Inês Miranda, pelos efeitos positivos da convivência, pelos debates em aula, pela amizade e pelos bons momentos compartilhados. Em particular, quero agradecer as amigas desde a graduação, Maria Inês e Alanna Santos, por compartilhar angústias, alegrias, ideias e por todo incentivo na fase final desse trabalho. Agradeço a todos os amigos e colegas que fiz durante o doutorado, em especial, à Camila Hermida, pelo carinho e por sempre ter uma palavra de auxílio e à Fernanda Fernandes, pelas aflições compartilhadas, pela parceria durante o estágio docência e pela sua ajuda. Aos amigos do laboratório, em especial, Ludmila

Luísa, Douglas Xavier e Marcelo Silva, pelos cafés e conversas em que dividimos ansiedades, alegrias e conhecimentos.

Sou grata ainda aos meus amigos e familiares, por todo amor, carinho e compreensão.

Por fim, agradeço a CAPES pelo apoio financeiro durante o doutorado.

RESUMO

O modo pelo qual a educação afeta o desempenho futuro dos indivíduos é ainda pouco compreendido. A pontuação dos estudantes em testes padronizados tem sido a principal fonte de medida encontrada pela literatura para mensurar conhecimentos e habilidades. A pressuposição é que os resultados desses testes são um previsor adequado do desempenho futuro dos alunos nas diversas esferas da vida em sociedade. As pesquisas iniciais nessa área encontraram um alto valor preditivo das variáveis relacionadas à família. No entanto, estudos recentes também têm considerado a importância de variáveis relacionadas à escola. Esta tese procurou analisar os determinantes do desempenho educacional em exames padronizados de estudantes localizados no estado de Minas Gerais, a partir de dados da Prova Brasil de 2013. A metodologia empregada tem como base um modelo hierárquico, mais indicado para lidar com a natureza hierarquizada de dados educacionais, para alunos da 4^a série/5ºano do ensino fundamental a partir das notas obtidas nos testes de Matemática. Consideram-se como fatores explicativos do desempenho estudantil as características do ambiente escolar, características da família e fatores individuais dos estudantes. Os resultados demonstram que tanto fatores familiares, bem como o ambiente escolar, são importantes para o desempenho acadêmico. Destaca-se a relevância de variáveis relacionadas à qualidade dos serviços educacionais, como a infraestrutura ofertada pela escola e a presença de professores com ensino superior em Matemática. Além desses, a escolaridade da mãe e o nível de capital social familiar estão entre os fatores que demonstram o valor que é atribuído à educação pela família do estudante. O trabalho infantil é apontado como um elemento prejudicial ao desempenho acadêmico, dado que ele pode reduzir o impacto positivo do capital social da família. Buscou-se, também, comparar o caso de Minas Gerais com o Estado do Ceará, dada a evolução recente dos indicadores educacionais nesse último estado, tendo sido encontradas diferenças entre os fatores que afetam o desempenho escolar, porém, no geral, também se verificou a influência de variáveis relacionadas a características dos alunos. Diferentemente de Minas Gerais, que tem efeitos diferenciados em termos no nível de capital social e da condição de trabalho dos estudantes, no Ceará, há diferenciação entre as escolas quanto ao desempenho em Matemática, em termos de atraso escolar. Ademais, aplicou-se também um modelo logístico hierárquico para verificar os principais determinantes do atraso escolar em Minas Gerais. Os resultados obtidos mostraram que há influência de características dos estudantes e de fatores escolares sobre a probabilidade de atraso, havendo diferenças regionais nesse aspecto. Além disso, observou-se que o desempenho dos estudantes considerados em atraso escolar, em algumas escolas, é menor do que em outras, conforme o seu nível socioeconômico. Espera-se que os resultados encontrados contribuam para o debate e para a agenda de políticas públicas voltadas para à melhoria da qualidade da educação e para redução das desigualdades escolares.

Palavras-chave: Educação; Desempenho escolar; Prova Brasil; Modelo hierárquico; Atraso escolar; Minas Gerais.

ABSTRACT

The way which education affects people's future performance is still poorly understood. Students' scores on standardized tests have been a major source of measure found in the literature to measure knowledge and skills. The assumption is that the results of these tests are suitable predictors of students' future performance in various spheres of life in society. Initial research in this area found a high predictive value in variable related to the family. However, recent studies have also considered the importance of school-related variables. This dissertation aimed at analyzing the determinants of education performance in standardized tests of students located in the state of Minas Gerais, by means of data concerning Prova Brasil (Brazil Test) of 2013. The employed methodology was based on a hierarchical model, more indicated to deal with the hierarchical nature of educational data, for students of the 4th grade / 5th grade of elementary school, based on the grades obtained in the Mathematics tests. School environment characteristics, family characteristics, and individual factors of the students were considered as performance's explanatory factors. The results demonstrate that both familiar factors and school environment are important for academic performance. We highlight the relevance of variables related to the quality of educational services, such as infrastructure offered by the school, and the presence of teachers presenting higher education in Mathematics. In addition, the mother's schooling and the level of social capital are among the factors that demonstrate the value attributed to education by the student's family. Child labor is seen as a detrimental element to academic performance, since it can reduce the positive impact of family social capital. We also sought to compare the case of Minas Gerais to the state of Ceará, given the recent evolution in educational indicators in this state. Differences in the factors that affect student performance were found but, in general, the influence of students' variables was verified. Different from Minas Gerais that has different effects in terms of the level of social capital and working conditions of students, in Ceará there is differentiation among schools concerning the performance in mathematics in terms of school delay. A hierarchical logistic model was applied in order to verify the main determinants of school delay in Minas Gerais. The results showed that there is influence of student characteristics and school factors on the probability of delay, with regional differences in this aspect. In addition, it was observed that in some schools the performance of delayed students is more harmful than in others, according to their socioeconomic level. It is expected that the results will contribute to the debate and to an agenda of public policies aimed at improving the quality of education and to reduce school inequalities.

Keywords: Education; School performance; Prova Brazil; Hierarchical model; School delay; Minas Gerais.

LISTA DE ILUSTRAÇÕES

Figura 1: Tamanho médio da turma no EF, por municípios - Minas Gerais (2015)	32
Figura 2: Distorção Idade-Série no EF, Minas Gerais (2015)	34
Figura 3: Proficiências Médias por Município, 5º Ano do EF, Brasil (2015)	37
Figura 4: Proficiências médias por Município, 9º Ano do EF, Brasil (2015)	37
Figura 5: Retas de regressão para dez escolas	68

LISTA DE GRÁFICOS

Gráfico 1: Médias de Proficiência dos Resultados do Brasil no SAEB (1995 a 2015).....	36
Gráfico 2: Histograma do NSE da Escola	86
Gráfico 3: <i>Boxplot</i> do NSE da Escola por Localização e Rede de Ensino	87
Gráfico 4: Verificação da Normalidade dos Resíduos com Escores Normais (Modelo 5) ..	105
Gráfico 5: Dispersão dos Resíduos Padronizados e Valores Ajustados (Modelo 5).....	106
Gráfico 6: Normalidade dos Resíduos para Efeitos Aleatórios (Modelo 5).....	106
Gráfico 7: Matriz de Dispersão dos Efeitos Aleatórios E Estimados (Modelo 5).....	107

LISTA DE TABELAS

Tabela 1: Ensino Fundamental em Minas Gerais, 2015 (em mil)	31
Tabela 2: Trajetória dos Estudantes do 5º Ano do Ensino Fundamental, Minas Gerais	33
Tabela 3: Índice de Desenvolvimento da Educação Básica (IDEB), Minas Gerais	35
Tabela 4: INSE, Percentual de Escolas por Faixa, Brasil (2014)	38
Tabela 5: Variáveis Explicativas no Nível do Aluno e Descrição	77
Tabela 6: Variáveis Explicativas no Nível da Escola e Descrição	80
Tabela 7: Estatísticas Descritivas - Alunos, 5º Ano, Minas Gerais.....	83
Tabela 8: Estatísticas Descritivas – Professor, Diretor e Escola, 5º Ano, Minas Gerais.....	85
Tabela 9: Modelo Nulo (sem variáveis explicativas)	88
Tabela 10: Modelos com Variáveis Explicativas Aluno e Escola.....	92
Tabela 11: Modelos com Interações e Coeficientes Aleatórios	99
Tabela 12: Comparação da Dimensão das Amostras entre Minas Gerais e Ceará.....	112
Tabela 13: Estatísticas Descritivas – Alunos, 5º Ano, Ceará	113
Tabela 14: Estatísticas Descritivas – Professor, Diretor e Escola, 5º Ano, Ceará.....	114
Tabela 15: Modelo Nulo (Ceará).....	116
Tabela 16: Modelos com Variáveis Explicativas Aluno e Escola (Ceará).....	117
Tabela 17: Modelos com Coeficientes Aleatórios (Ceará).....	121
Tabela 18: Variáveis explicativas no nível da escola	128
Tabela 19: Estatísticas descritivas no nível do aluno, 5º ano, Minas Gerais.....	129
Tabela 20: Estatísticas Descritivas no Nível do Aluno, 9º Ano, Minas Gerais.....	130
Tabela 21: Estatísticas Descritivas no Nível da Escola, 5º Ano, Minas Gerais	131
Tabela 22: Estatísticas Descritivas no Nível da Escola, 9º Ano, Minas Gerais	132
Tabela 23: Modelo Nulo – Probabilidade de Atraso Escolar, Minas Gerais.....	133
Tabela 24: Modelo para a Probabilidade de Atraso no Ensino Fundamental, MG	134

LISTA DE QUADROS

Quadro 1: Estudos Multinível de Fatores Associados ao Desempenho Escolar 54

LISTA DE SIGLAS E ABREVIATURAS

ABEP	Associação Brasileira de Empresas de Pesquisa
ANA	Avaliação Nacional da Alfabetização
ANEB	Avaliação Nacional da Educação Básica
ANRESC	Avaliação Nacional do Rendimento Escolar
CPV	Coeficiente de Partição da Variância
CV	Coeficiente de Variação
ECIEL	<i>Programa de Estudios Conjuntos de Educacion en América Latina</i>
EF	Ensino Fundamental
EM	Ensino Médio
ENADE	Exame Nacional de Desempenho de Estudantes
ENEM	Exame Nacional do Ensino Médio
IBGE	Instituto Brasileiro de Geografia e Estatística
IC	<i>Information Criteria</i>
ICC	Índice de Correlação Intraclasse
ICE	Índice de Conservação da Escola
ICMS	Imposto sobre Circulação de Mercadorias e Serviços
IDEB	Índice de Desenvolvimento da Educação Básica
IDHM	Índice de Desenvolvimento Humano por Município
IEA	<i>International Association for the Evaluation of Educational Achievement</i>
IEE	Indicador do Efeito da Escola
INEP	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
INSE	Indicador de Nível Socioeconômico
IPE	Índice de Posse de Equipamentos
IPEA	Instituto de Pesquisa Econômica Aplicada
MQO	Mínimos Quadrados Generalizados
NAEP	<i>National Assessment of Educational Progress</i>
NCS	Nível de Capital Social
NELS	<i>National Education Longitudinal Study</i>
NSE	Nível Socioeconômico
OCDE	Organização de Cooperação e de Desenvolvimento Econômico
OR	<i>Odds Ratio</i>
PB	Prova Brasil
PDE	Plano de Desenvolvimento da Educação
PISA	<i>Programme for International Student Assessment</i>
PNAD	Pesquisa Nacional por Amostra a Domicílio
PNUD	Programa das Nações Unidas para o Desenvolvimento
PPV	Pesquisa sobre Padrões de Vida
PROEB	Programa de Avaliação da Rede Pública da Educação Básica
SAEB	Sistema Nacional de Avaliação da Educação Básica
SARESP	Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo
SIMAVE	Sistema Mineiro de Avaliação da Educação Pública
SINAES	Sistema Nacional de Avaliação da Educação Superior
SPAEC	Sistema Permanente de Avaliação da Educação Básica do Ceará

TPE

Todos Pela Educação

TRI

Teoria da Resposta ao Item

UNESCO

United Nation Educational, Scientific and Cultural Organization

SUMÁRIO

INTRODUÇÃO.....	17
CAPÍTULO 1 – Um Breve Histórico dos Sistemas de Avaliação	21
1.1 O Contexto dos Sistemas de Avaliação da Educação.....	22
1.2 Notas Sobre o Processo de Avaliação Educacional.....	28
1.3 Caracterização do Estado.....	29
1.4 Análise de Dados Educacionais em Minas Gerais.....	31
CAPÍTULO 2 – A Economia da Educação: Para Além da Teoria do Capital Humano	39
2.1 Revisão da Literatura Empírica	46
2.2 Críticas e Desafios aos Métodos de Avaliação Padronizada no Brasil	54
2.3 Fatores Associados à Trajetória Escolar.....	57
CAPÍTULO 3 – Metodologia Econométrica.....	64
3.1 O Modelo Hierárquico Linear	66
3.2 O modelo Logístico Hierárquico	71
CAPÍTULO 4 – Determinantes da Proficiência em Minas Gerais.....	74
4.1 Variáveis Utilizadas.....	75
4.2 Estatísticas Descritivas	82
4.3 Resultado da Análise Econométrica	87
4.3.1 <i>Avaliação de Diagnóstico do Modelo</i>	104
4.3.2 <i>Considerações Gerais sobre os Resultados do Modelo</i>	107
4.4 Uma Análise Comparativa ao Caso do Ceará	110
4.4.1 <i>Estatísticas Descritivas</i>	111
4.4.2 <i>Resultados do Modelo</i>	115
4.4.3 <i>Considerações Gerais Sobre os Resultados do Modelo</i>	123
CAPÍTULO 5 – Os Fatores que Determinam o Atraso Escolar.....	125
5.1 Variáveis Utilizadas.....	127
5.2 Estatísticas Descritivas	129
5.3 Modelo de Regressão Hierárquica Linear para o Atraso Escolar.....	132
5.4 Considerações Gerais Sobre os Resultados do Modelo.....	139
CONSIDERAÇÕES FINAIS	141
REFERÊNCIAS	144
Apêndice: O modelo de Teoria de Resposta ao Item	155

INTRODUÇÃO

A importância da relação entre educação e desenvolvimento diz respeito à estreita ligação que há entre o processo de ampliação educacional e a sustentabilidade do desenvolvimento socioeconômico. No final do Século XX, surgiram intensas discussões e recomendações no campo educacional, considerando a importância da educação para o crescimento e desenvolvimento das nações. Uma referência foi a “Declaração Mundial sobre Educação para Todos”, originada na Conferência Mundial sobre Educação para Todos promovida pela UNESCO em 1990.

No Brasil, a aprovação da Constituição Federal de 1988 representou uma grande conquista para o país, principalmente, pela incorporação de direitos sociais para as camadas mais oprimidas da sociedade. Assim, no âmbito educacional, abriu-se a possibilidade de discutir a ideia de “Educação para Todos”, com um enorme avanço em termos dos direitos da população de acesso à educação. Os Artigos 205 e 206 da Constituição Federal de 1988, por exemplo, deixam claro que a educação é um direito de todos, dever da família e do Estado, devendo ser promovida e incentivada por meio da colaboração da sociedade, tendo em vista o completo desenvolvimento do indivíduo, estando o ensino baseado nos princípios da igualdade de acesso e permanência na escola. Isso significa que o desenvolvimento da educação não é somente responsabilidade das escolas de domínio estatal, mas, também, da família e da sociedade (BRASIL, 1988).

Nas últimas décadas, observou-se uma melhoria no acesso à educação básica no Brasil, dado o aumento do número de crianças no ensino fundamental. Entretanto, essa observação por si só não indica que a educação no Brasil tenha melhorado substancialmente. Por isso, faz-se necessário investigar quais fatores têm influenciado o desempenho escolar dos estudantes brasileiros. Ainda que, atualmente, os trabalhos que exploram a discussão da temática da educação tenham ganhado força, ainda não há evidências robustas sobre quais são os fatores determinantes do aprendizado. Análises semelhantes, que partiram do exame das características que afetam a proficiência dos estudantes, já foram empregadas para diversos países, incluindo o Brasil, alguns estados, cidades e escolas específicas.

O desempenho educacional é determinado por vários fatores que podem influenciar bons resultados ou não. Na maioria das vezes, os determinantes do desempenho escolar são separados em três conjuntos de fatores: aqueles ligados ao estudante e a características do meio familiar, aqueles relacionados à escola e aos professores, além daqueles fatores relativos à

direção e organização do sistema de ensino. Assim sendo, a proficiência obtida pelos estudantes é resultado da interação desses diversos fatores. As pesquisas iniciais nessa área encontraram um maior valor preditivo das variáveis relacionadas à família se comparadas às variáveis relacionadas à escola. Por isso, a pesquisa empírica tem tentado especialmente quantificar a parcela explicada da proficiência de cada um desses fatores.

Diante disso, este estudo procura analisar os determinantes do desempenho educacional em exames padronizados de estudantes localizados no Estado de Minas Gerais, considerando-se como fatores explicativos do desempenho as características do ambiente escolar, características dos pais e fatores sociais relacionados aos estudantes. Além disso, devido à grande magnitude encontrada para o coeficiente do atraso escolar buscou-se compreender os fatores que determinam a probabilidade de o aluno cursar a série na idade adequada. A escolha desse estado como *lócus* da pesquisa está baseada, principalmente, na sua representatividade nacional, pois, assim como o Brasil, Minas Gerais possui extensa população, com regiões heterogêneas em diferentes estágios de desenvolvimento, diversas culturas e atividades econômicas, o que, com certeza, leva a resultados diversos sobre a trajetória escolar de seus estudantes. Em relação ao ano de 2013, a sua escolha se deu por ser as informações mais recentes da Prova Brasil no momento em se iniciou a pesquisa.

A justificativa para a realização deste trabalho é a sua relevância, considerando que estudos sobre o tema ainda não chegaram a conclusões consistentes. Pouco se conhece, por exemplo, sobre a importância de variáveis relacionadas à escola na determinação dos resultados. Muitas vezes as variáveis escolares são as mais passíveis de intervenção no âmbito de políticas públicas. Portanto, é de fundamental importância que sejam desenvolvidos estudos e pesquisas sobre o tema, devido à representação significativa da relação entre desenvolvimento humano e o crescimento dos países. Além disso, pesquisas como esta colaboraram para tornar mais claro o processo de produção e geração de desigualdades, pois essas evidenciam que as diferenças de desempenho podem ser parte de um processo histórico de reprodução de desigualdades sociais arraigadas na nossa sociedade. Desse modo, os resultados desta pesquisa poderão ser relevantes para Minas Gerais, em particular, uma vez que a metodologia utilizada ao considerar a natureza hierárquica dos dados educacionais é capaz de indicar quais fatores contribuem para a proficiência dos estudantes. Ademais, esses resultados poderão servir para um maior esclarecimento e compreensão das diversas informações obtidas nos exames, além de subsidiarem programas e políticas educacionais.

Para tanto, a hipótese preliminar é que as condições socioeconômicas e culturais do aluno importam, mas, também, os fatores relacionados à escola podem ser importantes para explicar o resultado obtido pelos alunos em testes padronizados. Em resumo, o presente trabalho tem como objetivo geral estudar os fatores que afetam o desempenho educacional de estudantes localizados no Estado de Minas Gerais a partir de dados da Prova Brasil de 2013. A intenção é identificar se fatores relacionados ao ambiente familiar, social e escolares influenciam no resultado acadêmico e na trajetória escolar dos estudantes. Desse modo, a finalidade deste trabalho é apresentar uma metodologia alternativa para atender a esse propósito, considerando o seu âmbito e limitações. Cabe salientar que o interesse não é o de esgotar a discussão desses temas e, sim, caracterizá-los, visando a obter um maior entendimento dos resultados que serão obtidos.

A metodologia empregada se baseia em um modelo hierárquico para alunos da 4^a série/5^º ano do ensino fundamental, considerando as notas obtidas no teste de Matemática. E, em um segundo momento, utiliza-se um modelo logístico hierárquico para estudantes do 5^º e 9^º ano do ensino fundamental, tendo como determinante a situação do aluno em relação à idade considerada adequada para a série. Ambos os modelos consideram como fatores explicativos do desempenho as características do ambiente escolar, características da família e fatores sociais relacionados aos estudantes.

Assim sendo, a tese está organizada em cinco capítulos, além desta introdução e das considerações finais. O Capítulo 1 apresenta um breve histórico dos sistemas de avaliação da educação, além de descrever as principais informações sobre a evolução e características do sistema de ensino público em Minas Gerais. A primeira seção apresenta o contexto em que se deu a formação dos principais sistemas de avaliação da educação básica no Brasil e no exterior. Na segunda seção, é discutida a importância do processo de avaliação institucional e suas implicações para ações que visem a melhorias da educação ofertada nas escolas. A terceira seção traz uma breve apresentação de informações sociodemográficas de Minas Gerais e a última seção descreve a evolução e características dos sistemas de ensino no Brasil e em Minas Gerais.

O Capítulo 2 apresenta as principais abordagens teóricas sobre economia da educação e suas principais descobertas empíricas. Assim, a primeira seção aborda as questões fundamentais tratadas pela teoria do capital humano, além do contraponto proposto por Bourdieu. A segunda seção apresenta as críticas e os desafios enfrentados pelas avaliações padronizadas. Na terceira

seção, busca-se apresentar uma revisão da literatura empírica que examinou os fatores associados à trajetória escolar dos estudantes.

O terceiro capítulo expõe a metodologia econométrica que foi utilizada na análise dos determinantes de desempenho escolar em Minas Gerais, no ano de 2013. Ainda, apresenta-se o modelo matemático que fundamenta o modelo hierárquico, bem como os passos necessários para a sua estimação.

No Capítulo 4, tem-se a descrição da base de dados utilizada que fundamentou a escolha das variáveis, tratamentos feitos e as estatísticas descritivas, os resultados da análise econométrica e as principais conclusões encontradas acerca dos determinantes do desempenho educacional. Por exemplo, entre os resultados no nível do aluno destacam-se que: a maior escolaridade materna e o maior interesse dos pais na educação de seus filhos influenciam positivamente o desempenho estudantil; já o trabalho na infância e o atraso escolar afetam negativamente as notas. No nível da escola, observou-se que: uma melhor infraestrutura escolar e um nível socioeconômico maior das escolas levam a resultados positivos sobre o desempenho estudantil, assim como diretores mais experientes e professores com licenciatura em matemática também afetam positivamente as notas.

Já no Capítulo 5 da tese, são apresentados os resultados do modelo logístico hierárquico para o atraso escolar. Entre os resultados encontrados destacam-se que têm maiores chances de estar em atraso escolar estudantes do sexo masculino, não brancos, que trabalham fora de casa, entre outros. Já para os alunos que moram com os pais, têm mães com maior escolaridade, fazem o dever de casa, têm pais que apoiam os estudos, estudam em escolas que possuem notas maiores no teste de Matemática e Português as chances de atraso escolar são reduzidas. Por fim, apresentam-se as Considerações Finais da tese.

CAPÍTULO 1 – Um Breve Histórico dos Sistemas de Avaliação

O desenvolvimento de sistemas de avaliação tornou-se o aspecto central nos processos de reformas educacionais em vários países, fundamentalmente, a partir dos anos 1980. O objetivo principal era servir de base para ações de melhoria da qualidade do ensino, tentando dar mais evidência e clareza às características centrais do processo de aprendizagem. Em particular, esses sistemas buscam entender o que os alunos estão aprendendo, quais fatores melhor explicam os resultados da educação, qual a influência do salário, carreira e formação dos docentes sobre o aprendizado e de que modo desigualdades sociais, econômicas e culturais de um local afetam as oportunidades de aprendizagem (CASTRO, 2009).

Inicialmente, o estudo da educação era feito por meio de indicadores de acesso à escola, número de anos de estudo e fluxo escolar¹, os quais são capazes de representar as condições básicas de um sistema de ensino. Contudo, ao longo dos anos, constatou-se que esses indicadores não deveriam ser tomados como referência de qualidade.

Na década de 1990, foram elencadas quais deveriam ser as prioridades da educação em diversos países, principalmente, naqueles com maiores deficiências no sistema de ensino. Considerando que a educação é um direito fundamental de todos, foram estabelecidas metas relacionadas à universalização do acesso à educação e promoção da equidade; foco no esforço da aprendizagem; oferecimento de um ambiente adequado à aprendizagem; ampliação da educação básica, entre outros (UNESCO, 1990).

No Brasil, medidas legais, como a Lei de Diretrizes e Bases da Educação (1996), instituíram critérios para a avaliação da educação, os quais são considerados para a distribuição de recursos financeiros para as escolas (SOARES, 2005). No entanto, levando-se em conta que a preocupação com a qualidade da educação não está ligada somente ao estabelecimento de um conteúdo base, as formas de avaliação foram ampliadas para serem capazes de subsidiar o monitoramento da oferta educacional.

Enquanto a educação brasileira conquistava a universalização do ensino fundamental, além do crescimento dos níveis médio e superior, as taxas de conclusão continuavam sendo um problema, tendo em vista os índices de evasão e repetência verificados. Somente no final da década de 1990, pode-se observar uma melhora nas estatísticas educacionais, principalmente, por meio de um esforço em ampliar essas informações e torná-las mais acessíveis. Com isso,

¹ O fluxo escolar é representado por indicadores tais como: frequência à escola, abandono, conclusão da educação básica, número de matrículas, atraso escolar, entre outros.

foi possível aos governos obter maiores informações para definição de suas ações em termos de política educacional.

Assim, o processo de avaliação da educação ganhou destaque por sua ideia de construção de um projeto educacional no sentido de poder regular o desenvolvimento do aprendizado, além de representar um instrumento útil de reflexão acerca das ações dos agentes envolvidos. Levando isso em consideração, o objetivo deste capítulo será apenas de apresentar brevemente em que contexto se deu a construção dos sistemas de avaliação da educação em termos internacionais e nacionais. Além de apresentar informações sobre as características do sistema de ensino em Minas Gerais.

1.1 O Contexto dos Sistemas de Avaliação da Educação

Destaca-se que a possibilidade de medir a qualidade da educação é contemporânea. Na década de 1960, o Relatório Coleman consagrou os sistemas de avaliação em grande escala nos Estados Unidos (HEYNEMAN, 2005). Ao originar muitas pesquisas e estudos sobre o problema educacional, esse relatório contribuiu para que, em 1969, o congresso americano criasse a Avaliação Nacional do Progresso em Educação (*National Assessment of Educational Progress* – NAEP). O NAEP visa a supervisionar, nas diversas regiões e estados americanos, a evolução da qualidade educacional em várias áreas do conhecimento. Atualmente, os papéis da avaliação vão além da certificação de que o aluno conseguirá ingressar em uma universidade para obter uma boa colocação profissional, tendo em vista que se procura analisar as instituições, o desempenho dos professores, programas estatais e práticas de ensino ao longo do tempo, buscando compreender os condicionantes sociais do resultado obtido pelos estudantes e identificar meios de melhoria (SCHWARTZMAN, 2005). Ressalta-se que esse sistema serviu de base para a construção das provas do sistema nacional de avaliação da educação no Brasil.

Na década de 1970, um projeto de pesquisas da América Latina, conhecido como Programa de Estudos Conjuntos de Integração Econômica Latino-americana (*Programa de Estudios Conjuntos de Educacion en América Latina* – ECIEL), aplicou testes de rendimento com o objetivo de analisar os determinantes do desempenho escolar para uma amostra de cinco países, incluindo o Brasil (CASTRO, 2005). O principal fator encontrado foi o ambiente familiar do aluno. No entanto, os esforços desses testes foram isolados, não tendo sido os

resultados tão satisfatórios, pois os países da América Latina apresentavam-se, em termos educacionais, muito aquém dos países das outras regiões (WOLFF, 1998).

Além das análises internas nos países, também são realizados exames em nível internacional, como o Programa Internacional de Avaliação de Estudantes (*Programme for International Student Assessment – PISA*), desenvolvido e coordenado por uma parceria entre a OCDE e a UNESCO². O PISA compreende uma avaliação comparada aplicada a estudantes de 15 anos – idade em que se espera ter concluído a escolaridade básica obrigatória na maioria dos países. A avaliação é feita a cada três anos, tendo ocorrido sua última edição em 2015. Nesse caso, avalia-se a proficiência dos estudantes em Leitura, Matemática e Ciências, além da aplicação de questionários sobre variáveis demográficas, socioeconômicas e educacionais, para que seja possível, posteriormente, relacionar o desempenho dos alunos com essas variáveis (INEP, 2014). Os resultados são importantes por se revelarem, muitas vezes, diferentes do que se espera de alguns países, conforme o seu nível de renda e desenvolvimento, indicando que a qualidade da educação não está sujeita somente à disponibilidade de recursos financeiros, mas a condições institucionais, pedagógicas e culturais das nações (SCHWARTZMAN, 2005).

Nas últimas décadas, o Brasil construiu seu sistema de avaliação da educação por meio de um conjunto de testes em larga escala. A ideia-base é apresentar um exame detalhado do sistema de ensino nacional e dos principais fatores associados ao desempenho estudantil, provendo indícios sobre a qualidade da educação. Nesse sentido, o resultado permite subsidiar a criação, revisão e o acompanhamento das políticas educacionais no âmbito municipal, estadual e federal.

O Sistema Nacional de Avaliação da Educação Básica³ (SAEB), desenvolvido pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), surgiu na década de 1990, inspirado no sistema norte-americano de avaliação, que foi a base para a construção das provas de Língua Portuguesa e Matemática. Em termos sintéticos, essa avaliação tem como principal objetivo fornecer informações aos gestores educacionais.

Nas suas edições iniciais, o SAEB oferecia somente informações por estados, sem apresentar maiores detalhes de informações por municípios, escolas e alunos. Por meio da divulgação dessas informações, houve uma pressão política sobre as autoridades educacionais para obtenção de resultados mais satisfatórios. Atualmente, as avaliações são realizadas a cada

² Nos moldes do PISA há também outros exames internacionais como o IEA (*International Association for the Evaluation of Educational Achievement*) e o Timms (*The Third International Mathematics and Science Study*), onde o aluno é avaliado para com vistas a identificar a qualidade de ensino na escola e também o grau de aprendizado do estado ou país (CASTRO, 2005).

³ A educação básica compreende a educação infantil, ensino fundamental e ensino médio.

dois anos, tendo caráter amostral, e seus resultados permitem aferir o desempenho dos sistemas educacionais públicos ou privados em suas diferentes esferas administrativas.

A edição inicial ocorreu em 1990, configurando-se como a primeira realização de exames de proficiência e do levantamento de informações estatísticas em termos nacionais. A amostra foi composta por escolas públicas que ofereciam 1^a, 3^a, 5^a e 7^a séries do Ensino Fundamental na área urbana, tendo sido empregados testes de Língua Portuguesa, Matemática e Ciências, além de redação para os alunos da 5^a e 7^a séries. A segunda prova aconteceu em 1993, mantendo-se o formato da edição anterior, com base em três linhas de análise: rendimento estudantil; perfil e práticas de professores; perfil de diretores e formas de gestão.

Na aplicação seguinte, em 1995, houve a inclusão do ensino médio e de escolas particulares, bem como a adoção de técnicas mais sofisticadas de mensuração do desempenho⁴ e o uso de novas ferramentas de análise de dados sobre características socioeconômicas, culturais e hábitos de estudo. Além disso, foi feita uma redefinição das séries analisadas, tendo como base os anos conclusivos de um determinado ciclo escolar, isto é, 4^a e 8^a séries do ensino fundamental (atuais 5º e 9º ano) e 3^a série do ensino médio. Nessa edição, não foi realizada avaliação de Ciências.

Na sua quarta aplicação é que se deu a elaboração das chamadas Matrizes de Referência⁵, no ano de 1997. Em 1999, seguiu-se o modelo adotado anteriormente, agregando-se os conteúdos às competências cognitivas usadas no modo de criação do conhecimento (INEP, 2001). Nessas duas últimas edições, os alunos do ensino fundamental responderam questões de Língua Portuguesa, Matemática e Ciências, e, para os estudantes do ensino médio, acrescentaram-se as disciplinas de História e Geografia.

Na avaliação de 2001, o SAEB apresentou novas alterações. Em resposta às modificações na Lei de Diretrizes e Bases da Educação, que empregou a regulamentação curricular nacional, modernizaram-se as matrizes de referência do SAEB. Isso foi feito com base em uma consulta aos estados a fim de verificar se havia compatibilidade entre as matrizes e o currículo recomendado pelos sistemas estaduais para as disciplinas abordadas na prova. A partir desse ano, os testes passaram a avaliar somente as disciplinas de Língua Portuguesa e Matemática, o que se mantém até os dias atuais.

⁴ Nesse ano, começou-se a utilizar a Teoria da Resposta ao Item (TRI) na construção do teste e análise de resultados, com o objetivo de tornar as avaliações comparáveis ao longo dos anos.

⁵ As Matrizes de Referência descrevem qual deve ser o recorte representativo do conteúdo a ser avaliado em cada disciplina e série, com base nos currículos vigentes no país. Esse recorte deve proporcionar às ferramentas do SAEB a capacidade de mensurar determinadas competências e habilidades esperadas dos alunos ao final de cada ciclo de ensino.

Entre 1990 e 2003, a amostra de escolas era obtida por meio de sorteio, com resultados para estados, regiões e o país. No ano de 2005, o sistema sofreu uma reestruturação, sendo composto por duas avaliações: Avaliação Nacional da Educação Básica (ANEB) e Avaliação Nacional do Rendimento Escolar (ANRESC), que são avaliações para diagnóstico em larga escala.

Outras mudanças ocorreram em 2009, quando foram incluídas escolas públicas rurais do ensino fundamental, desde que atendessem, legalmente, ao mínimo de alunos matriculados. Já em 2013, além da inclusão experimental da disciplina de Ciências para os alunos do ensino fundamental e médio, foi feita a incorporação de mais uma avaliação externa em larga escala ao SAEB: a Avaliação Nacional da Alfabetização (ANA). Essa é uma avaliação censitária, que engloba alunos do 3º ano do Ensino Fundamental provenientes de escolas públicas e tem como objetivo avaliar os níveis de aprendizado nas disciplinas de Língua Portuguesa e Matemática. A diferença dessa avaliação para a ANEB e a ANRESC/Prova Brasil é a sua periodicidade, que é anual (INEP, 2014a).

A ANEB é uma avaliação amostral, com o mínimo de 10 alunos por classe, nas redes públicas e privadas, com foco na educação básica. Ao conservar os mesmos objetivos, características e procedimentos da avaliação da educação básica realizada nas edições anteriores, a ANEB recebe o nome do SAEB em suas divulgações. Já a ANRESC, também conhecida como Prova Brasil, tem caráter censitário e seu objetivo é mensurar a qualidade do ensino ministrado nas escolas públicas (estaduais, municipais e federais), em áreas urbanas e rurais. Na realização da prova, os estudantes respondem questões de Matemática (com foco na resolução de problemas) e Português (com foco em leitura), além de um questionário com itens sobre o contexto socioeconômico familiar, hábitos de estudo e de leitura, motivação, trajetória escolar, entre outros. Ainda, os professores e diretores preenchem questionários sobre dados demográficos, perfil profissional e condições de trabalho. Um avaliador externo aplica a prova e preenche um questionário com informações da escola relacionadas à infraestrutura, ambiente físico, recursos, entre outros. Com esse exame, ampliaram-se as possibilidades de análises da situação da educação brasileira.

Em 2013, para o 5º ano do ensino fundamental, a escala de nota em Língua Portuguesa (que vai de 0 a 350) foi composta por nove níveis, tendo o primeiro deles agrupado os estudantes com proficiência inferior a 150, e os demais níveis foram organizados em intervalos que variam de 25 em 25 pontos (150-175; 175-200; 200-225; e, assim, sucessivamente). Para a disciplina de Matemática, a escala foi formada por 10 níveis, de 125 a 375 pontos, com intervalos de 25

pontos. A proficiência dos estudantes é apresentada em forma numérica, indicando apenas uma posição na escala e, a partir dessa posição, são realizadas interpretações pedagógicas dos resultados, pois cada nível representa um conjunto de habilidades que os alunos demonstram terem adquirido ao longo da vida estudantil.

Cabe destacar que tanto a ANEB, bem como a Prova Brasil, não pretendem avaliar cada aluno individualmente. Na elaboração das provas, o INEP utiliza uma metodologia chamada de Blocos Incompletos Balanceados, cujo objetivo é aplicar um número elevado de itens a um conjunto de estudantes, sem que cada aluno necessite responder a todas as questões. Assim, são elaborados 21 tipos de cadernos de prova a partir da junção de 2 blocos para cada área do conhecimento, contendo 44 itens de múltipla escolha.

A Prova Brasil parte do princípio de que a educação, conforme prevê a Constituição, é um dever da família e do Estado, sendo ainda de responsabilidade de toda a sociedade, por isso, a necessidade de qualificar a educação oferecida em todo o território nacional. Essa prova é realizada a cada dois anos, na 4^a (5º ano) e 8^a (9º ano) séries do ensino fundamental que tenham o número mínimo legal de alunos matriculados. A Prova Brasil responde a uma demanda da sociedade por maiores informações acerca da educação em níveis municipais e de cada escola, em particular, principalmente, para orientar no repasse de recursos e no estabelecimento de metas e ações para melhorar a educação (INEP, 2014a).

Um ponto de destaque está na comparabilidade dos dados, o que permite às escolas identificarem seus pontos fortes e fracos em relação a outras escolas de sua cidade, do estado ou do país. Mesmo assim, muitas dessas instituições de ensino não conseguiram descobrir a melhor forma de melhorar seus resultados, havendo ainda dificuldades na elaboração de práticas pedagógicas apropriadas, por exemplo, ou seja, o desafio está na ausência de capacidade institucional e de competência técnica para efetivar os objetivos da Prova Brasil (CASTRO, 2009).

Na avaliação de 2005, foram realizados testes nas escolas com o mínimo de 30 alunos por turma, seguindo a metodologia similar à do SAEB. Nesse ano, foram avaliadas 42 mil escolas públicas urbanas, compreendendo um total de 3,3 milhões de alunos e, em 2007, foram 4,5 milhões de alunos avaliados. Nesse ano de 2007, o mínimo de alunos passou a ser 20 por sala, sendo a prova realizada juntamente com a aplicação da ANEB e com os mesmos instrumentos de análise. Além disso, nesse ano, a Prova Brasil passou a fornecer informações sobre o Índice de Desenvolvimento da Educação Básica (IDEB).

O IDEB tem como base as informações do Censo Escolar sobre taxa de aprovação, bem como as médias do SAEB (Estados e país) e da Prova Brasil (municípios). A importância desse indicador é que ele impede que haja um direcionamento das ações para reprovação ou aprovação em massa, no sentido de se evitar aprovação sem aprendizagem e a reprovação como forma de exclusão dos alunos com baixo desempenho a fim de aumentar a nota no exame (CASTRO, 2009).

Desse modo, o IDEB tem permitido a identificação das instituições de ensino que não estão cumprindo totalmente com seus objetivos, além de contribuir para a aplicação de metas mais objetivas. Em suma, optou-se pela inclusão, na rotina escolar, da discussão dos seus resultados e de como analisar o sistema de ensino nacional. No entanto, entre as fragilidades desse índice, está a sua dependência em relação ao desempenho médio dos alunos da escola, ou seja, o IDEB é sensível a alterações na média (SOARES, 2009).

No Brasil, a avaliação da educação superior é feita pelo Sistema Nacional de Avaliação da Educação Superior (SINAES), o qual é composto por avaliações das instituições, cursos e desempenho dos estudantes em termos de ensino, pesquisa, extensão, gestão, responsabilidade social, infraestrutura, entre outros. A avaliação do desempenho de estudantes é feita pelo Exame Nacional de Desempenho de Estudantes (ENADE) desde 2004. Esse exame, de caráter obrigatório, é realizado por alunos que estão em fase de conclusão de cursos de graduação, a cada três anos, para cada área do conhecimento. A prova leva em conta conteúdos programáticos, habilidades e as competências obtidas (INEP, 2016).

A cultura da avaliação se aprofundou no país, começando alguns estados a realizar pesquisas semelhantes ao SAEB para suprir a necessidade de um maior conhecimento sobre os principais condicionantes de ordem social ou institucional que afetam o desempenho de cada escola e de seus estudantes (SCHWARTZMAN, 2005). Como exemplo dessas iniciativas, podemos citar: o estado do Paraná (Boletim da Escola – 2001), São Paulo (Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo – SARESP), Minas Gerais (Sistema Mineiro de Avaliação da Educação Pública – SIMAVE) e Ceará (Sistema Permanente de Avaliação da Educação Básica do Ceará – SPAECE).

1.2 Notas Sobre o Processo de Avaliação Educacional

A avaliação institucional compreende um plano de avaliação que produz a demonstração dos rumos da instituição no esforço pela qualidade. Assim, esse é um termo utilizado para a avaliação de políticas públicas, em especial, políticas setoriais relacionadas à educação, representando um serviço prestado para a sociedade, visto que os participantes da instituição podem reconsiderar sua missão e funções, maneiras de atuação e consequências de suas práticas institucionais (LEITE, 2006).

Apesar disso, Freitas *et al.* (2009) alegam que a avaliação representa uma categoria pedagógica controversa pela sua capacidade de afetar a vida das pessoas. No ambiente escolar, a avaliação pode exercer dois papéis: representar uma ameaça ou ser fonte de desenvolvimento. A avaliação atinge todos os atores no processo, visto que, em determinados momentos, têm-se avaliadores que, posteriormente, podem ser objetos de avaliação. Os autores consideram que o processo avaliativo é cheio de trocas entre momentos formais e informais que podem afetar o direcionamento futuro do estudante. Diante disso, o processo avaliativo não deve ser analisado somente pelo desempenho (positivo ou não) do aluno, mas, também, pelos seus aspectos comportamentais.

Assim, o processo de avaliação toma forma com os mecanismos que levam à sustentação ou exclusão de alguns estudantes no interior da escola. Por isso, os resultados da avaliação da aprendizagem necessitam ser articulados com outras categorias que compõem a esfera da avaliação, pela possibilidade de não se conseguir compreender a complexidade que envolve esse tema. Desse modo, propõe-se um modo de avaliação educacional para além daquele dominante, sugerindo-se a análise com base em três níveis integrados: avaliação em larga escala, avaliação institucional e avaliação em sala de aula. Além de seu lado formal, a avaliação tem um importante lado informal. Quanto mais inicial for o grau de ensino, mais continuada e difusa é a presença da avaliação, visto que professores das séries iniciais do ensino fundamental dirigem o processo de avaliação também por meio da convivência diária. Nesse sentido, o processo avaliativo tem caráter contínuo e processual, o que lhe confere uma certa complexidade (FREITAS *et al.*, 2009).

Para Perrenoud (1999), o processo de avaliação é gerador de hierarquias de excelência que definem o caminho a ser seguido. Assim, o autor sugere que o caminho mais fértil para orientar a pesquisa sobre avaliação é aquele em que o processo avaliativo auxilia o aluno a aprender e a se desenvolver, isto é, um processo que compreende a regulação do aprendizado e

o desenvolvimento estudantil no sentido de um projeto educativo. Portanto, o processo de avaliação do aprendizado é uma ferramenta útil à reflexão dos professores sobre as ações escolares.

Nesse sentido, deve-se considerar que a relação entre o desempenho escolar e a qualidade dos resultados obtidos pela escola tem especificação complexa, uma vez que ela está ligada a questões não observáveis ou que não se podem medir por meio dos métodos tradicionais. O resultado obtido por escola é composto pela interação de diversos fatores inter-relacionados, como, por exemplo, características dos indivíduos e da família, oportunidades experimentadas pelos alunos e a qualidade do ensino (GREMAUD; FELÍCIO; BIONDI, 2007). Por isso, faz-se necessário investigar aspectos sociais, econômicos e culturais, além dos resultados obtidos no exame.

1.3 Caracterização do Estado

Com 586.552,4 km de extensão, o Estado de Minas Gerais é aquele que se apresenta como o maior estado em extensão territorial da Região Sudeste e o quarto maior estado do Brasil (IPEA, 2016). Segundo o IBGE, o estado teria, em 2015, uma população estimada em 20.869.101 habitantes, sendo o segundo estado mais populoso do Brasil, ficando atrás apenas de São Paulo (população estimada de 44.396.484 habitantes para o mesmo ano). Com 10,2% da população total do país, Minas Gerais possui 853 municípios com grande diversidade entre as regiões. A densidade demográfica do estado, em 2010, era de 33,41 habitantes/km (IBGE, 2016b), sendo a taxa de urbanização do estado bastante elevada, com, aproximadamente, 85% em 2010.

O estado conta com um Índice de Desenvolvimento Humano por Município (IDHM), segundo os dados do Programa das Nações Unidas para o Desenvolvimento (PNUD), de 0,731, no ano de 2010, ficando em 9º lugar dentre os 27 estados da Federação. Em 1º lugar, encontra-se o estado do Distrito Federal, com 0,824 de IDHM. O valor de IDHM de Minas é considerado médio, pois se projeta no intervalo de 0,5 e 0,8 (PNUD, 2016).

Com relação à educação, os dados do Estado são significativos, especialmente, aqueles relativos à população jovem. A frequência escolar no ano de 2010, segundo dados do IBGE (2016b), na faixa etária dos 7 aos 9 anos, era de 98,6% dessa população. Na faixa etária dos 10 aos 14 anos, a frequência era de 97,1%, enquanto que, na faixa dos 15 aos 19 anos, ficava em

torno de 67,2%. Outro dado relevante é a taxa de analfabetismo da população acima dos 15 anos, que passou de 12%, no ano de 2000, para 8,3%, em 2010. No entanto, a taxa de analfabetismo na faixa etária dos 15 aos 24 era de 3,2%, em 2000, caindo para 1,4%, em 2010.

De acordo com a Pesquisa Nacional por Amostra a Domicílio (PNAD) de 2013, aproximadamente, 43% da população com mais de 25 anos de idade, tinham ensino fundamental incompleto, e somente 23% possuíam ensino médio completo. O percentual de pessoas nessa faixa etária que não possuíam qualquer instrução é de 9,5%, sendo mais da metade desse total (55%) pessoas com idade superior a 60 anos, no referido ano. No Estado de Minas Gerais, a participação das pessoas que possuem curso superior completo e incompleto é de 11%.

O nível de escolaridade média da população de Minas Gerais vem também apresentando um crescimento ao longo do tempo. Para a faixa etária da população com mais de 10 anos, é de cerca de 7,2 anos de estudo, em 2013, sendo o nível médio de escolaridade feminino superior ao masculino, 7,4 anos contra 7,0 anos em média, respectivamente. Para a mesma faixa etária, a mesorregião com maior taxa de escolaridade média foi a Região Metropolitana de Belo Horizonte, a qual apresentou oito anos de escolaridade, em média, em 2013, enquanto a região do Jequitinhonha exibiu a menor escolaridade média, com seis anos de estudo.

No que se refere ao nível de renda, na PNAD de 2014, o estado de Minas Gerais registrou o valor de R\$1.769 como o rendimento médio mensal real das pessoas com mais de 10 anos de idade. A população feminina ocupada obteve um rendimento inferior à masculina, R\$1.477 contra R\$1.981, respectivamente. O nível de ocupação total foi de 56,8%, em média, sendo 67,2% entre os homens e 47,1%, entre as mulheres.

No que diz respeito aos indicadores sociais, o índice de Gini da distribuição do rendimento médio mensal do pessoal com mais de 10 anos em Minas Gerais demonstra que o estado é menos concentrado em termos de distribuição de renda do que o país como um todo, visto que o índice, em 2006, era de 0,525, enquanto que, para o Brasil, era de 0,547. O estado que apresentou o maior índice de Gini, em 2006, é o Piauí, com 0,607, e o menor é o Amapá, com 0,448. Em 2010, por volta de 2,56% dos domicílios do estado de Minas Gerais apresentavam rendimentos inferiores a meio salário mínimo, o qual é considerado como o rendimento básico. Sendo assim, os rendimentos inferiores indicam que esses indivíduos estão abaixo da linha da pobreza.

1.4 Análise de Dados Educacionais em Minas Gerais

A educação básica compreende a educação infantil, o ensino fundamental e o ensino médio. Ao terminar o ano letivo, os alunos podem ser aprovados, reprovados ou terem abandonado a escola. Juntas, essas três situações representam o rendimento escolar.

Na última edição do SAEB, em 2015, participaram um total de 57.744 mil escolas, entre as quais, 38.155 eram escolas que ofereciam o 5º ano do Ensino Fundamental e 29.620, o 9º ano. A pesquisa engloba 3.986.190 milhões de estudantes, dos quais, aproximadamente, 52% são estudantes do 5º ano, 46%, estudantes do 9º ano, e o restante são alunos da 3ª série do Ensino Médio⁶.

No Brasil, há 3,17 milhões de alunos matriculados no 5º ano do ensino fundamental, com turmas médias de 18,0 estudantes nas 105 mil escolas que oferecem esse curso. Já para o 9º ano do ensino fundamental, o país possui 2,81 milhões de alunos matriculados, distribuídos em 56,1 mil escolas, com a média de 25 alunos por turma (DATAVIVA, 2016).

Tabela 1: Ensino Fundamental em Minas Gerais, 2015 (em mil)

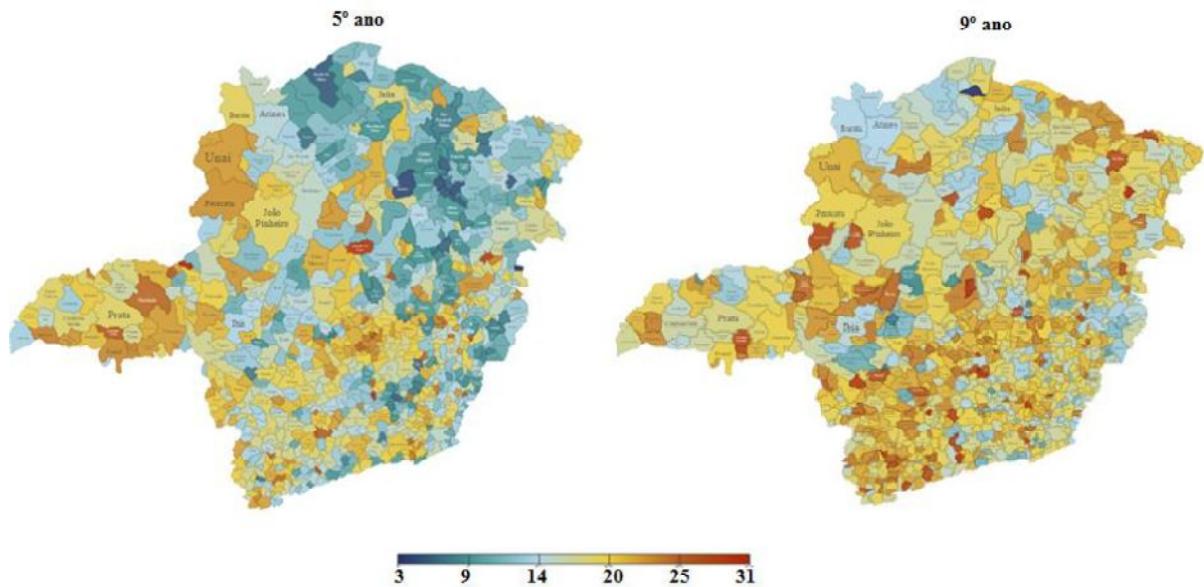
	5º Ano	9º Ano
Número de matrículas	276	308
Número de turmas	15,5	11,2
Número de escolas	9,01	5,22
Tamanho médio das turmas	17,0	27,0
Idade média	10,8	14,9

Fonte: INEP/MEC, 2015. Elaborado por: DataViva

As informações da Tabela 1 mostram que há em Minas Gerais 276 mil alunos registrados no 5º ano do ensino fundamental e 308 mil alunos, no 9º ano. Em média, as turmas são compostas por 17,0 (5º ano) e 27 alunos (9º ano). Conforme a Figura 1, abaixo, nota-se que o tamanho médio das turmas é bem heterogêneo nos municípios do estado. Observando-se a idade média dos estudantes mineiros, nota-se que ela está bem próxima da idade esperada para a série, ou seja, 10 anos para o 5º ano e 14 anos para o 9º ano.

⁶ Vale destacar que os microdados do SAEB de 2015 foram disponibilizados somente em novembro de 2016.

Figura 1: Tamanho médio da turma no EF, por municípios - Minas Gerais (2015)



Fonte: DataViva, 2016 (<http://www.dataviva.info/pt/>)

Considerando uma trajetória regular, o aluno deve iniciar o ensino fundamental aos 6 anos de idade e completá-lo aos 14 anos, ingressando no ensino médio aos 15 anos de idade e concluindo aos 17 anos. O aluno que apresenta dois anos ou mais de diferença da idade recomendada para a série que cursa está em atraso escolar, contribuindo para a origem de imperfeições no fluxo escolar. Com base nos dados da PNAD, a proporção dos estudantes com atraso no ensino fundamental foi mais elevada entre os estudantes da rede de pública (28,7% contra 7,9% da rede particular), homens (31,6%), residentes em área rural (40%) e negros (31,4%) (IBGE, 2016a).

Analizando a trajetória dos estudantes do 5º ano do ensino fundamental (Tabela 2)⁷, em Minas Gerais, vemos que, na rede estadual, em 2013, 99% dos estudantes matriculados foram aprovados, 0,4% foram reprovados e 0,6% deixaram de frequentar a escola durante o ano letivo. Já para a rede municipal, as taxas foram inferiores, tendo 95% dos estudantes matriculados obtido aprovação, 3,6% foram reprovados e 0,5% abandonaram a escola. No ano de 2014, os resultados foram bastante similares, com 99% e 96,1% dos estudantes aprovados, na rede estadual e municipal, respectivamente. Esses valores podem ser justificados pela adoção de sistemas de progressão continuada nos anos iniciais do ensino fundamental em Minas Gerais.

⁷ Na ferramenta Painel Educacional do INEP há apenas informações para os Anos Iniciais do Ensino Fundamental e resultados da ANA 2014. Havia previsão da apresentação dos dados dos Anos Finais do Ensino Fundamental e do Ensino Médio e os resultados da Prova Brasil 2015. No entanto, até o presente momento, essas informações não foram apresentadas.

A taxa de distorção idade-série (uma medida de atraso escolar) capta se os estudantes estão com dois anos ou mais acima da idade recomendada para a série. Em Minas Gerais, percebe-se que essa taxa é de 7% para os alunos do 5º ano da rede estadual e 15,6%, para a rede municipal⁸.

Tabela 2: Trajetória dos Estudantes do 5º Ano do Ensino Fundamental, Minas Gerais

Taxa	2013		2014	
	Estadual	Municipal	Estadual	Municipal
Aprovação	99,00	95,90	99,00	96,10
Reprovação	0,40	3,60	0,50	3,40
Abandono	0,60	0,50	0,50	0,50
Distorção Idade-série	9,50	18,90	7,00	15,60

Fonte: INEPDATA, 2016 (<http://inepdata.inep.gov.br/>)

A demora na matrícula, abandono ou reprovação estão entre as principais causas do atraso escolar. Não ter a idade esperada para a série causa desvantagens aos estudantes, como, por exemplo, afeta a qualidade do aprendizado e compromete a autoestima (HONDA, 2006). Assim, mesmo diante das melhorias observadas, as taxas de conclusão na idade adequada não são tão favoráveis. Conforme o IPEA (2014), menos da metade dos jovens com 15 anos havia concluído o ensino fundamental em 2013, e 20% das pessoas com 18 anos ainda não tinham o diploma de ensino fundamental.

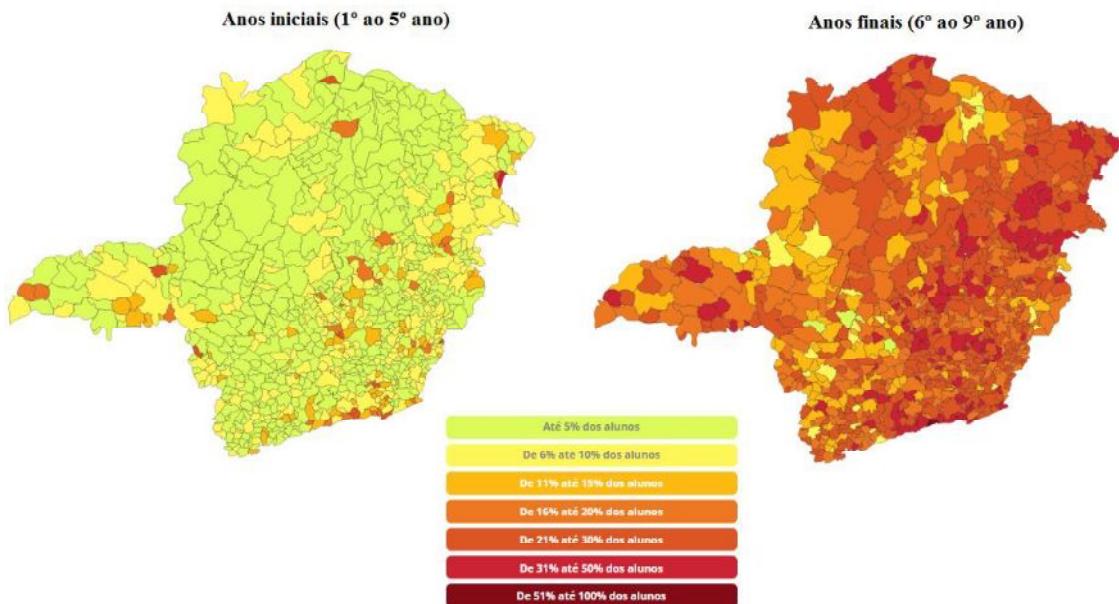
A taxa de frequência escolar bruta no país, em 2015, atingiu 98,6% entre os jovens de 6 a 14 anos de idade, um reflexo da intensificação da expansão educacional da educação básica. No entanto, quanto às pessoas de 15 a 17 anos, essa taxa ficou em 85%, progredindo lentamente (crescimento de apenas 3,4 pontos percentuais comparado com dados de 2005) devido à dificuldade de garantir educação básica obrigatória para essa faixa etária. Embora tenha praticado políticas de correção de fluxo, o país ainda apresenta elevadas taxas de repetência quando comparado aos países da América Latina com níveis elevados para os jovens de 15 anos de idade, que são referência para a avaliação do PISA. Esse resultado leva o país a um desempenho inferior quando comparado a outras nações (IBGE, 2016a).

No entanto, destaca-se que somente a frequência escolar não indica que os alunos estejam prosperando na escola, visto que são várias as razões que afetam negativamente o desenvolvimento do aluno e que podem causar atraso durante o percurso escolar.

⁸ Por trás desse conceito estão diferentes aspectos da educação como a idade de ingresso na escola, taxa de reprovação e abandono.

A Figura 2 mostra dados mais recentes da distorção idade-série para Minas Gerais. Observa-se que, para o ano de 2015, nos anos iniciais do ensino fundamental, grande parte dos alunos estava com a idade adequada para a série, tendo a taxa de distorção idade-série em Minas Gerais, incluindo escolas públicas, privadas, da área urbana e rural, do 1º ao 5º ano, ficado em torno de 5%. No entanto, para os anos finais do ensino fundamental (6º ao 9º ano), essa taxa foi de 21%. Como se pode observar na Figura 2, são pouquíssimos os municípios (dentre os 853 existentes no estado) que estão com a taxa em até 5%, havendo casos em que mais da metade dos estudantes encontra-se em atraso escolar.

Figura 2: Distorção Idade-Série no EF, Minas Gerais (2015)



Fonte: INEP, 2015. Organizado por QEDU (<http://www.qedu.org.br/>)

Assim, quanto mais elevada a idade, maior é o custo de oportunidade de se estar na escola (VASCONCELLOS, 2005). Os dados apresentados revelam altas taxas de reprovação e abandono escolar de estudantes que ainda estão em idade em que a educação básica é obrigatória, principalmente, para aqueles com menores condições econômicas, elevando-se, assim, as desigualdades educacionais.

A avaliação da qualidade da educação básica no Brasil tem sido feita, principalmente, por meio da análise do IDEB (Índice de Desenvolvimento da Educação Básica). Esse índice combina dados do resultado obtido pelos estudantes nos testes do SAEB e da Prova Brasil, ao final de cada ciclo de ensino, com informações de rendimento escolar.

A criação desse índice retomou o debate sobre a qualidade do ensino ofertado nas escolas, colocando em pauta um desafio a mais para as escolas: garantir que os alunos prossigam, tendo aprendido o conteúdo adequado para a sua idade (ORTIGÃO; AGUIAR, 2013).

De acordo com os dados da Tabela 3, a qualidade da educação apresentou progresso tanto em escolas da rede estadual, bem como da municipal. Evidencia-se que os alunos das escolas públicas apresentaram índice de desenvolvimento de 4,6 para 6,1, entre anos de 2005 e 2015, nos anos iniciais do ensino fundamental, e de 3,6 para 4,6, nos anos finais, para o mesmo período. Observa-se ainda a significativa diferença em todos os anos do valor encontrado para o IDEB entre as escolas públicas e privadas. O estado mineiro teve um desempenho melhor do que o nacional, dado que o IDEB para o Brasil, em 2005, foi igual a 3,6 nas escolas públicas, nos anos iniciais do ensino fundamental, e de 3,2 nos anos finais. Já em 2015, o índice nacional foi de 5,3 no 5º ano e 4,2 para o 9º ano.

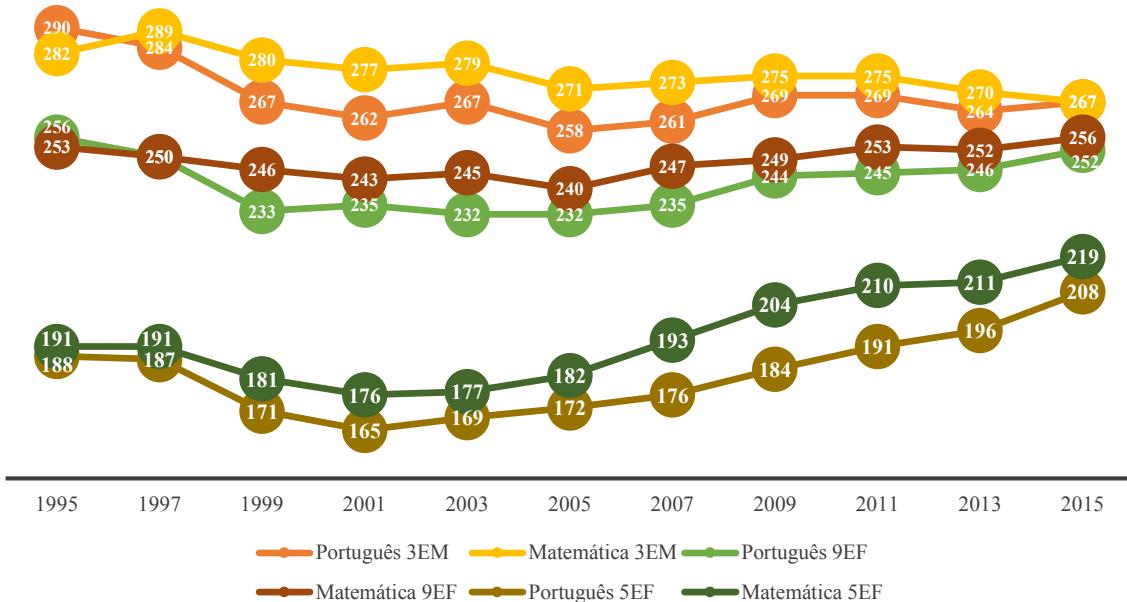
Tabela 3: Índice de Desenvolvimento da Educação Básica (IDEB), Minas Gerais

Ano	Público		Privado	
	Ensino Fundamental		Ensino Fundamental	
	Anos Iniciais	Anos Finais	Anos Iniciais	Anos Finais
2005	4,6	3,6	6,9	6,4
2007	4,6	3,8	6,7	6,4
2009	5,5	4,1	7,1	6,7
2011	5,8	4,4	7,4	6,5
2013	5,9	4,6	7,6	6,3
2015	6,1	4,6	7,8	6,5

Fonte: INEP, 2016

A partir da análise da evolução das notas médias obtidas pelos estudantes no SAEB no país, observa-se que, como esperado, as notas são maiores quanto mais são os anos de estudo, ou seja, o resultado do estudante do ensino médio tende a ser maior do que o do aluno do 5º ano do ensino fundamental (Gráfico 1). No geral, observa-se que as notas estão sempre bem próximas de seus valores médios. Ao longo do período de 1995 a 2015, por exemplo, a média da proficiência obtida em Língua Portuguesa no ensino médio foi de 269, e, em 2015, a nota média nessa disciplina foi de 267. Ainda, destaca-se, considerando a proficiência média, entre 1995 e 2015, que a nota obtida em Matemática foi superior àquela obtida em Português, em todas as faixas de ensino.

Gráfico 1: Médias de Proficiência dos Resultados do Brasil no SAEB (1995 a 2015)



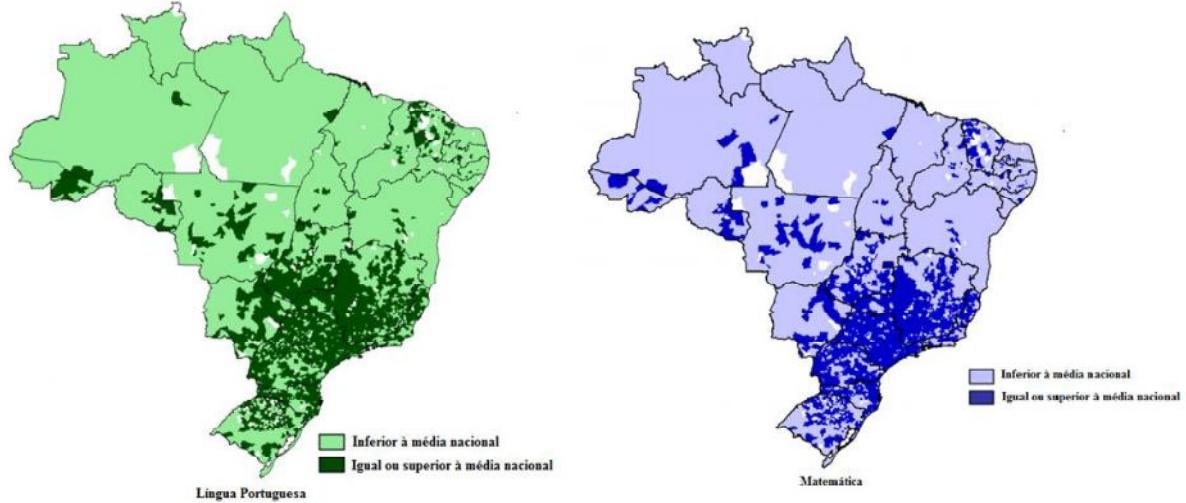
Fonte: Diretoria de Avaliação da Educação Básica – DAEB/INEP, 2016

A proficiência média nacional, em 2015, obtida no 5º ano foi de 219,3 em Matemática e 207,6 em Português. O resultado mineiro foi superior ao brasileiro, sendo de 232,4 em Matemática e 220,7 em Português. Já para o 9º ano do ensino fundamental, a proficiência média no Brasil encontrada foi de 255,8 em Matemática e 251,5 em Língua Portuguesa. Em Minas, a média encontrada no último ano do ensino fundamental foi de 264,6 em Matemática e 258,6 em Língua Portuguesa.

As Figuras 3 e 4 ilustram o exemplo para os municípios brasileiros do 5º ano e 9º ano, tanto em Língua Portuguesa como em Matemática. Observa-se que, em grande parte da região Sudeste, a proficiência encontrada é igual ou superior à média nacional.

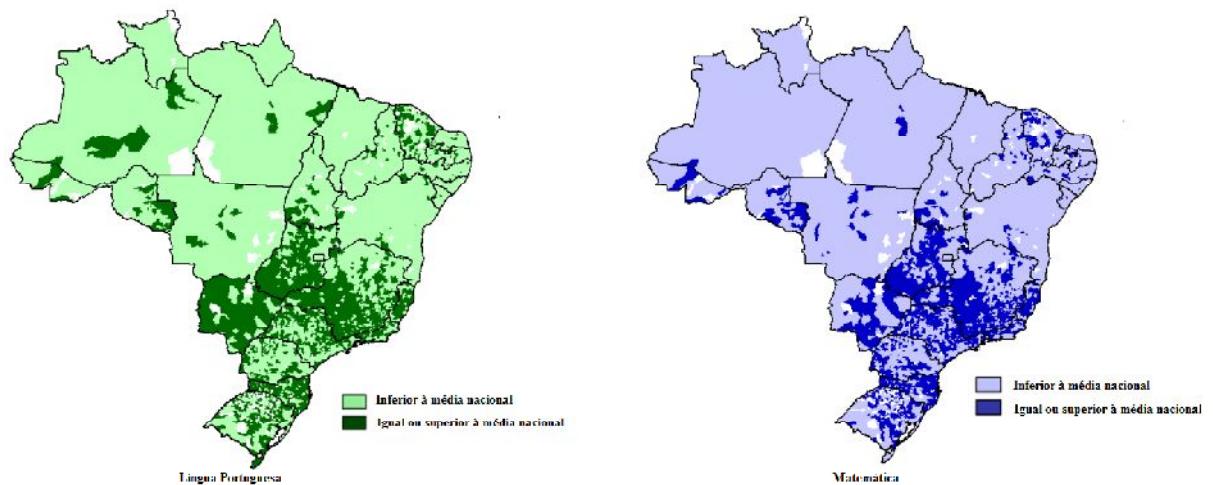
No ano de 2014, o INEP passou a contextualizar as medidas de aprendizado, expondo informações sobre o Indicador de Nível Socioeconômico (INSE) dos alunos. O INSE foi calculado por meio de dados da Prova Brasil e do ENEM (Exame Nacional do Ensino Médio) para os anos de 2011 e 2013, tendo como base informações sobre a escolaridade dos pais, a posse de bens, renda e contratação de serviços pela família dos alunos. A metodologia utilizada baseia-se em um modelo de resposta gradual da Teoria de Resposta ao Item.

Figura 3: Proficiências Médias por Município, 5º Ano do EF, Brasil (2015)



Fonte: Diretoria de Avaliação da Educação Básica – DAEB/INEP, 2016

Figura 4: Proficiências médias por Município, 9º Ano do EF, Brasil (2015)



Fonte: Diretoria de Avaliação da Educação Básica – DAEB/INEP, 2016

Para melhor entendimento, construiu-se uma escala para esse indicador com a finalidade de descrever o padrão de vida dos estudantes e, por meio da análise de cluster, foi possível classificar os estudantes em sete grupos, em uma escala de muito baixo a muito alto. Os resultados apresentados se referem às escolas e municípios que obtiveram, pelo menos, 50% de participação de estudantes do 5º ou no 9º ano do Ensino Fundamental. Esse indicador é útil por fornecer um posicionamento do padrão de vida dos estudantes que compõem uma determinada escola (INEP, 2014b). Na Tabela 4, observa-se que, em 2014, 38,85% das escolas estavam

situadas em um nível socioeconômico médio alto, 26,84%, no nível médio, e 16,36%, no nível médio baixo, indicando que existem mais alunos situados nos níveis médios da escala.

Tabela 4: INSE, Percentual de Escolas por Faixa, Brasil (2014)

Níveis	%	Quantidade
Muito alto	0,18	3
Alto	10,01	170
Médio alto	38,85	660
Médio	26,84	456
Médio baixo	16,36	278
Baixo	5,00	85
Muito baixo	0,24	4

Fonte: INEPDATA, 2015

Franco, Alves e Bonamino (2007) destacam que, para que se alcancem as metas do Plano de Desenvolvimento da Educação (PDE), é preciso que se regularize o fluxo escolar, no sentido de tornarem raras as reprovações e o abandono, e que se melhore a performance escolar de modo que o desempenho da escola mediana no futuro seja equivalente ao de uma escola de alto desempenho no passado.

CAPÍTULO 2 – A Economia da Educação: Para Além da Teoria do Capital Humano

Os estudos sobre economia da educação datam a década de 1950, quando alguns estudiosos se inspiraram na teoria econômica neoclássica para explicar ganhos de produtividade. A ideia básica era tentar medir o valor econômico da educação. Os resultados mostraram que o investimento em “capital humano” era o principal elemento para aumentar os salários em relação aos outros componentes da produção.

Assim sendo, a acumulação do conhecimento era a base da transformação tecnológica da economia mundial. O acúmulo de conhecimento e sua transmissão às novas gerações representaria uma parte cada vez mais importante da atividade econômica, pois um sistema de ensino moderno é fundamental para que uma economia avançada funcione, visto que isso depende de habilidades gerais e específicas que somente a educação é capaz de distribuir (VAIZEY, 1962).

Nos anos 1960, a oferta de pessoas qualificadas era vista como uma forma de limitação ao crescimento. Arthur Lewis (1962) foi um dos autores a tratar essa questão, ressaltando que houve uma confusão da aplicação nos países em desenvolvimento de conclusões estatísticas que tentam medir o rendimento da educação nos países ricos. O autor sugere que a capacidade humana é melhorada pela educação, saúde pública, pesquisa, invenção, mudança institucional e uma melhor organização dos recursos humanos, seja nos negócios, na vida privada ou na vida pública. Portanto, atribuir todas as melhorias de produtividade à educação seria um pouco ingênuo.

Freeman (1995) destaca que, após a Segunda Guerra Mundial, com a mudança nas ideias de economistas sobre o desenvolvimento, o Banco Mundial concluiu que o investimento imaterial no acúmulo de conhecimento era decisivo, ao invés do investimento em capital físico, tal como foi uma vez acreditado, ou seja, sugere-se que o capital humano (conhecimento e as habilidades incorporados nas pessoas), ao invés do capital físico, é vital para a prosperidade econômica do país.

Para Coleman (1988), possivelmente, o desdobramento mais significativo na economia da educação tenha sido a ampliação da noção de que o conceito de capital físico possa incluir o capital humano. De modo semelhante ao capital físico, que é gerado por alterações nos insumos que facilitam a produção, o capital humano é gerado pelas mudanças nas habilidades e capacidades que levam os indivíduos a atuar de novas maneiras.

Formalmente, o conceito de capital humano foi construído no final dos anos 1950 e no início da década seguinte. Mincer (1958) foi um dos primeiros a apresentar a noção de capital humano, tendo como preocupação entender a dinâmica da distribuição de renda entre os indivíduos. Para o autor, dava-se ênfase, anteriormente, a outros fatores para a explicação da distribuição de renda, relegando um papel residual à capacidade dos indivíduos. Em resumo, o autor concluiu que as diferenças nos níveis de treinamento resultavam em diferenças nos níveis de ocupação, bem como nos ganhos obtidos entre as ocupações.

Esse conceito recebeu maior destaque depois que uma publicação da UNESCO (1968), intitulada *Readings in the Economics of Education*, divulgou uma coletânea de textos clássicos sobre a economia da educação. Entre esses, destaca-se o trabalho “*Investment in man: an economist's view*” de Theodore W. Schultz, que foi um marco importante⁹, uma vez que o autor faz referência a várias formas do capital humano, sempre relacionando-o com o processo de crescimento econômico (SCHULTZ, 1961).

Na literatura, as abordagens mais comuns para a mensuração do capital humano compreendem a abordagem baseada no custo, a abordagem com base no rendimento e a abordagem baseada em educação. Os estudos sobre capital humano, inicialmente, foram direcionados para o cálculo do custo da educação, com base nos custos operacionais ou de investimento ao longo de vários anos, tendo em conta o aumento previsto do número de matrículas no sistema educacional (DEBEAUVAIS, 1962).

Nesse sentido, para Amorim (2009), a teoria do capital humano diz respeito à demanda por educação que enfatiza os aspectos de investimentos que circundam a tomada de decisão. Essa teoria surgiu a partir da constatação de que o crescimento da economia não dependia somente do capital e da produtividade, mas, necessita, também, da participação do fator humano.

Já a abordagem baseada na educação estima o capital humano por meio da medição de indicadores de resultados da educação (*proxies*), tais como, taxas de alfabetização, taxas de matrícula, taxas de evasão, repetência, entre outros. Esses estudos, que buscaram apreender a relação entre as características escolares e o desempenho dos alunos, foram estimulados pela pesquisa realizada pelo sociólogo James Coleman *et al.* (1966). Essa pesquisa ficou conhecida como Relatório Coleman e ganhou destaque por apresentar resultados contrários ao senso comum, isto é, que a qualidade da escola não teria papel fundamental na aprendizagem. Os autores encontraram que o ambiente familiar e o meio social em que vivem os estudantes seriam

⁹ Principalmente, após o autor receber o prêmio Nobel de 1979.

o determinante fundamental para o desempenho escolar. Com base nesses resultados, vários autores buscaram explicar quais fatores são determinantes para a qualidade e o aprendizado escolar. A publicação desse relatório resultou em um grave pessimismo pedagógico, ao se concluir que a escola teria pouco impacto no desempenho dos seus estudantes.

Na literatura, ainda não é consensual quais fatores são determinantes para a qualidade e o aprendizado escolar. Andrade e Soares (2008) explicam que há diversos determinantes, além do efeito escola, sobre o resultado obtido pelos alunos. Conforme os autores, a ideia do “efeito escola” é mensurar, por meio do projeto pedagógico e das políticas internas, a capacidade da escola em influenciar o desempenho cognitivo dos alunos.

Destaca-se que, mesmo que o desempenho dos alunos seja explicado por fatores internos e externos à escola (tais como, o aluno, a família, a sociedade, a escola e o sistema de ensino a qual está ligada), esses fatores, por si só, não são capazes de explicar e influenciar o desempenho obtido.

Entretanto, conforme Soares (2009), há dificuldades de se isolar o efeito da escola, uma vez que, normalmente, há escassez e/ou problemas na coleta de dados. Segundo o autor, a identificação de fatores que interferem no desempenho cognitivo dos estudantes em leitura é problemática, visto que essa qualidade do aluno resulta de questões familiares, escolares (como ações que ocorrem nas salas de aula) ou do próprio modelo pedagógico adotado pela escola.

Portanto, o importante a salientar é que o resultado obtido pelos alunos não é de inteira responsabilidade da escola. Teddlie e Reynolds (2000) mostram que há influência de aspectos históricos e empíricos sobre o efeito das escolas. No geral, a literatura especializada tem encontrado que, mesmo que o sucesso ou fracasso nos testes possam ser explicados por fatores extraescolares, existe uma variedade de resultados para instituições que possuem alunos com características socioeconômicas similares.

Para Debeauvais (1962), a dificuldade em definir o conceito de capital humano está no fato de esse associar duas ideias de diferentes naturezas e que são estudadas por disciplinas diferentes. Segundo o autor, o termo “capital” abrange uma categoria de bens que são tidos como “produtivos”, porque esses têm a propriedade de produzir outros bens; já o homem, como o produtor desses bens, é visto como fator “mão de obra”, que deve ser associado ao capital a fim de iniciar o processo pelo qual a riqueza é produzida. Nessa abordagem, as duas categorias são variáveis independentes.

No entanto, em outras áreas do conhecimento, o homem é analisado por seu comportamento individual ou social, e não em sua função de produção de riqueza. Para os

especialistas em educação, a finalidade essencial da educação é obter o maior desenvolvimento das aptidões intelectuais e morais do indivíduo (DEBEAUVAIS, 1962).

A teoria do capital humano, ao tentar compreender os parâmetros que orientam as decisões individuais, se baseia nos postulados idênticos à teoria econômica convencional, isto é, trata os indivíduos como agentes racionais, maximizadores de recursos e com informação perfeita. Nessa visão, a teoria do capital humano somente tem aplicabilidade em uma sociedade com igualdade de oportunidades no acesso e na permanência na escola (RIBEIRO; CACCIAMALI, 2012).

Assim, não se pode esquecer das contribuições do sociólogo Pierre Bourdieu para esse debate e como a teoria do reproduтивismo surge como uma alternativa teórica. Conforme essa doutrina, a escola tem o papel de selecionar indivíduos por meio de valores, atuando no sentido de garantir a reprodução das estruturas sociais, sobretudo, a hierarquia das classes.

Entre as diferenças de Bourdieu e Coleman para explicar os fenômenos sociais, destaca-se que, Bourdieu é crítico das teorias da escolha racional, as quais Coleman adota (na explicação de fenômenos coletivos, o autor optou pelo conceito de individualismo metodológico). Coleman defende a ideia de maximização da utilidade, entendendo a racionalidade de uma forma ampla. De outro lado, Bourdieu acredita que os atores sociais não avaliam racionalmente as suas ações por meio das chances de elas serem bem-sucedidas, mas, sim, conforme a assimilação de estruturas sociais e de situações passadas (o *habitus*) (AQUINO, 2000).

A importância de apresentar os escritos de Bourdieu se deve ao fato de sua obra oferecer respostas originais, bem como apresentar novos questionamentos sobre o pensamento sociológico acerca das funções e do funcionamento social dos sistemas de ensino nas sociedades contemporâneas e das relações que mantêm os diferentes grupos sociais com a escola. Os conceitos e categorias por ele criados têm alto poder explicativo para as análises que incluem os produtos escolares.

O autor é contrário à teoria do capital humano, visto que ele defende que o *habitus* racional é fruto de uma condição econômica singular, determinada pela detenção do capital econômico e cultural indispensável para se compreenderem as oportunidades potenciais disponíveis. Entretanto, a capacidade de estimar e perceber as oportunidades apenas é alcançada em determinadas classes sociais (RIBEIRO; CACCIAMALI, 2012), ou seja, as classes sociais mais fragilizadas não possuem os elementos essenciais para a manifestação da racionalidade econômica.

Em termos gerais, Bourdieu rompe com explicações fundamentadas em capacidades naturais e individuais, lançando uma crítica ao mito do “dom” e mostrando como condições culturais e sociais permitiram o desenvolvimento desse mito. O autor apresenta ainda a crítica aos mecanismos por meio dos quais o sistema de ensino altera as diferenças iniciais (advindas da passagem da herança cultural familiar) para desigualdades de destino escolar (NOGUEIRA; CATANI, 1998). Desse modo, segundo Bourdieu, a escola não contribui na superação do *déficit* cultural enfrentado pelos jovens estudantes, mas, sim, tende a reproduzir essas diferenças.

Para Bourdieu, a presença de uma hierarquia nos campos de produção simbólica compõe uma das mediações por meio das quais se estabelece a censura específica de um determinado campo. Assim, é “a hierarquia dos objetos que, conscientemente ou inconscientemente, orienta os investimentos intelectuais dos agentes, medidos pela estrutura de oportunidades de lucro material e simbólico” (NOGUEIRA; CATANI, 1998, p. 8).

Em um estudo sobre a Faculdade de Letras da França¹⁰, na década de 1960, Bourdieu observou que alunos advindos de classes ricas herdam um ambiente cultural da família, isto é, possuem um privilégio cultural proveniente do maior contato com teatros, cinemas, leitura de livros e do próprio convívio com a família e a sociedade, sendo essa diferença distintiva para o desempenho universitário dos estudantes.

Bourdieu (1998) estabeleceu o conceito de capital cultural para tratar da desigualdade de desempenho escolar de jovens provenientes de classes sociais distintas, buscando relacionar o bom desempenho escolar com a distribuição desse capital entre as classes ou frações dela. Conforme o autor, cada família passa aos seus filhos um capital cultural e um *ethos*, assim como regras de valores tácitos e intensamente interiorizados. Essa transferência ocorre mais por meio indireto do que direto, o que colabora para a definição do capital cultural e das atitudes em relação à escola. O êxito e as diferenças entre as crianças diante da experiência escolar ocorrem em virtude da herança cultural, a qual é diferente entre as classes sociais.

Assim, ao adotar essa posição, o autor rompe com os pressupostos, tanto do senso comum, que credita o sucesso ou fracasso escolar como resultado de aptidões naturais, bem como com a teoria do capital humano. Além disso, ressalva-se que o nível de instrução das pessoas que compõem a família é apenas um indicador que possibilita posicionar o nível de cultura da família, não revelando informações a respeito do conteúdo transmitido por meio da herança cultural das famílias, nem a respeito de seus modos de transmissão (BOURDIEU, 1998).

¹⁰ Bourdieu, P. L'école conservatrice. Les inégalités devant l'école et la culture. *Revue française de sociologie*, 7(3), 1966, p. 325-347.

Diante disso, percebe-se que há uma forte ligação entre o capital cultural e atitudes familiares. A herança cultural é parcela significativa do sucesso escolar, já que envolve hábitos de leitura e estudo de línguas, o que leva à facilidade de expressão verbal, gosto pela cultura, dentre outros. Ainda, famílias com alta herança cultural são capazes de transmitir informações que são úteis na orientação de seus filhos ao longo da trajetória escolar.

Em suma, as desigualdades educacionais adviriam da herança cultural dos indivíduos. Assim, o êxito escolar, a permanência na escola, a eliminação de defasagens, entre outros fatores, estariam ligados ao capital cultural existente. Isso porque a reduzida herança cultural evita que as famílias de classes populares possuam informações necessárias sobre o montante de investimentos necessários despendidos durante o ciclo escolar de seus filhos, contrariando a teoria do capital humano.

Entre os elementos relevantes do pensamento do autor, tem-se, por exemplo, a evidência atribuída à relação com o saber. A escola, ao estimar o desempenho de seus estudantes, leva em consideração, especialmente, a maneira de obtenção e uso do saber, isto é, a relação que o aluno possui com o saber. As obras de cultura veiculadas pela escola vão exibir uma relação que tende a ser interessada, trabalhosa, diligente para os estudantes advindos de famílias desprovidas de capital cultural, enquanto que, para os alunos provenientes de meios culturalmente privilegiados, essa relação está marcada pela estima, desenvoltura, elegância e desembaraço verbal ‘natural’.

Na teoria bourdieusiana, o capital cultural existe sob três formas. A primeira é chamada de o estado incorporado, em que a acumulação ocorre sob a forma de arranjos duráveis do organismo, demandando a incorporação tempo e trabalho pessoalmente investido pelo indivíduo receptor. O segundo modo, o estado objetivado, é representado por meio de bens culturais (livros, ferramentas, máquinas) transmissíveis de modo relativamente momentâneo quanto à propriedade jurídica. Por fim, tem-se o estado institucionalizado, representado por títulos e diplomas escolares, o qual possui relativa independência em relação ao seu portador. É nessa última forma de capital cultural que seria possível aplicar o tema das funções sociais do sistema de ensino e entender as relações existentes com o sistema econômico. Em geral, são utilizados como indicadores de capital cultural familiar o nível educacional dos pais e os bens econômicos, embora um, não necessariamente, assegure o outro.

Ribeiro e Cacciamali (2012) chamam atenção que, para Bourdieu, a assimetria de informação se deve somente à desigualdade na distribuição da herança cultural, entendendo

esse autor a assimetria de informação como um dos determinantes do fracasso ou abandono da vida escolar.

Além do conceito de capital cultural, Bourdieu apresenta também a noção de capital social. Depois de realizar várias análises de diversos grupos sociais, o autor constatou que o capital social corresponde ao conjunto de recursos presentes e potenciais que estão relacionados à detenção de uma rede estável de relações mais ou menos institucionalizadas de interconhecimento e inter-reconhecimento. Em outras palavras, capital social diz respeito ao pertencimento a um grupo cujos indivíduos possuem características comuns, sendo eles também unidos por vínculos duráveis e favoráveis. Assim, enquanto não existirem instituições que possibilitem a concentração de todo o capital social do grupo nas mãos de um indivíduo (por exemplo, a família), cada pessoa deve participar do capital coletivo, na proporção direta de sua contribuição. E, na presença de capital social, o aproveitamento do capital econômico e do capital cultural é maior (BOURDIEU, 1998).

Coleman também trabalhou com o conceito de capital social. Na proposta desse autor, a noção de capital social esclarece as relações de confiança que existem entre os membros de uma comunidade. Nesse sentido, o capital social seria o conjunto de relações sociais nas quais os indivíduos estão envolvidos e que os permitem alcançar objetivos por fazerem parte dessas relações. A diferença entre eles é que, para Bourdieu, o capital social é algo que os indivíduos possuem, enquanto que, para Coleman, são as relações sociais que o possuem (AQUINO, 2000).

Em resumo, as obras de Bourdieu destacaram aspectos fundamentais entre a estrutura da sociedade e o sistema de ensino ao revelar como o êxito na trajetória escolar é condicionado, em grande parte, pela proximidade entre a cultura da escola e a da família. Para esse autor, o capital econômico e o capital cultural são os princípios que oferecem maior contribuição para a hierarquização dos grupos e indivíduos na sociedade. Bourdieu revelou como a origem social dos alunos, a distribuição irregular do capital cultural entre as famílias e a tendência da escola em considerar igualmente os estudantes com distintos níveis desse capital tendem a beneficiar os alunos pertencentes aos estratos sociais mais favorecidos, metamorfoseando as desigualdades sociais em desigualdades escolares (INEP, 2014b).

Os principais escritos de Bourdieu em relação à escola contribuem para a análise das estratégias de reprodução, levando-nos a compreender como se configuram as relações entre a escola e a realidade social. Ainda, contribuem para a discussão teórica por revelar os meios por meio dos quais as desigualdades sociais se decompõem em desigualdades escolares.

Analizando as relações entre educação e desigualdade na sociedade brasileira, Barros, Henriques e Mendonça (2000) observaram que o processo de desenvolvimento nacional nas décadas recentes tende a reforçar a heterogeneidade educacional do país. E, ainda, o Brasil apresenta fraco desempenho educacional e atrasos na educação de quase uma década, se comparado à experiência internacional, considerando um país com características de desenvolvimento similares. Diante dessa conjuntura, a avaliação dos fatores que influenciam o desempenho educacional a partir de uma abordagem ampla se torna necessária, uma vez que essa pode auxiliar na articulação de estratégias para diminuição de desigualdades socioeconômicas e culturais.

Conforme destacado anteriormente, as análises sobre qualidade da educação têm enfatizado o impacto de fatores no processo de aprendizado, além dos escolares. A associação entre o resultado de exames padronizados e as informações contidas nos questionários socioeconômicos apresenta a importante função dos fatores econômicos, sociais e culturais das famílias e dos alunos no processo de aprendizagem e na sua trajetória escolar. Portanto, variáveis como nível de renda, acesso a bens, anos de escolaridade dos pais e participação na vida escolar do discente, e ambiente familiar estão propensos a influenciar na performance do estudante.

2.1 Revisão da Literatura Empírica

A escolaridade é uma variável central na determinação do progresso econômico, tanto individual, bem como nacional. De um lado, a expansão da educação aumenta a produtividade do trabalho, contribuindo para o crescimento econômico e a redução da pobreza. Do outro, por se tratar de um ativo não transferível e pela sua facilidade de distribuição, se comparada à maioria dos ativos físicos, a escolaridade conduz a uma maior igualdade e mobilidade social. Portanto, esses fatores esclarecem a importância da expansão educacional para a promoção do crescimento econômico e redução da pobreza e da desigualdade (BARROS; HENRIQUES; MENDONÇA, 2000).

Barbosa-Filho e Pessoa (2010) apresentam uma resenha literária sobre o efeito da educação na economia. Segundo os autores, os resultados que consideram a qualidade da educação, medida a partir do desempenho dos estudantes em avaliações de ciências e

Matemática, têm grande relevância para o crescimento e, provavelmente, é a via mais promissora de investigação.

A eficácia dos insumos escolares sobre o desempenho educacional tem sido frequentemente estudada na literatura, mas ainda não há consenso sobre seu real impacto. Assim, acredita-se que o modo pelo qual a educação afeta o desempenho futuro dos indivíduos ainda não está totalmente claro. Por isso, a literatura empírica tem se direcionado, principalmente, para a tentativa de quantificar qual é a parcela de contribuição de cada um desses fatores (CURI; SOUZA, 2015).

O Relatório Coleman serviu de inspiração para vários trabalhos nessa área por concluir que grande parte do desempenho era explicada por fatores da família e do capital social dela originado, isto é, a escola exercia pouca influência (COLEMAN *et al.*, 1966).

Tradicionalmente, a literatura faz uma analogia à função de produção (COLEMAN *et al.*, 1966; LAU, JAMISON, LOUAT, 1991; HANUSHEK, 2002; RIVKIN, HANUSHEK, KAIN, 2005), relacionando o desempenho dos alunos em exames padronizados (produto) a fatores que descrevam as características da família e dos estudantes, bem como às características da escola e a fatores sociais (insumos). Neste sentido, a função de produção da educação pode ser definida (FRANCO, 2008), em termos sintéticos, como:

$$Y_{ij} = \alpha_0 + A_{ij} \beta_1 + F_{ij} \beta_2 + P_{ij} \beta_3 + D_{ij} \beta_4 + E_{ij} \beta_5 + u_{ij} \quad (1)$$

em que: Y_{ij} representa alguma medida de produto escolar, como por exemplo, a nota obtida pelo aluno i na escola j em um teste padronizado; α_0 é uma constante; A_{ij} corresponde a um vetor de características do aluno i na escola j tais como cor, sexo, defasagem idade-série; F_{ij} é um vetor de variáveis que descrevem características do *background* familiar do aluno i na escola j , tais como escolaridade da mãe, nível socioeconômico, trabalho fora ou dentro de casa; P_{ij} refere-se a um vetor de características do professor do aluno i na escola j , tais como escolaridade do professor, salário, experiência; D_{ij} refere-se a um vetor de características dos diretores do aluno i na escola j , tais como escolaridade, experiência, práticas de gestão; E_{ij} é um vetor de características da escola do aluno i tais como infraestrutura, localização, tamanho das turmas; u_{ij} representa um termo de erro aleatório.

Em geral, é encontrada uma baixa influência de características escolares se comparadas às características dos próprios alunos e das famílias. Segundo Biondi e Felício (2007), esses

resultados são pouco robustos, devido à variabilidade da significância estatística e do efeito encontrado sobre o desempenho escolar das variáveis estudadas.

Schielbein e Simmons (1980) fizeram uma análise dos resultados de 26 estudos sobre os determinantes do desempenho escolar de estudantes pertencentes a países em desenvolvimento, descrevendo suas implicações para a política educacional. Os autores dividem esses determinantes em três níveis: recursos e processos escolares; atributos do professor; e características do aluno. Segundo os autores, é importante destacar que esses podem ser determinantes na teoria, mas, não necessariamente, o são para determinada realidade. Os resultados indicaram que, em mais da metade das observações, variáveis ligadas a características dos estudantes parecem ter maior consistência do que as características da escola ou do professor. No geral, observou-se que entre as principais influências possíveis para o desempenho cognitivo dos alunos em testes escolares estão: a quantidade desejada de alunos por turma, considerando a suas implicações com os custos; a realização de trabalhos extraclasse, visto que esses levam a melhores resultados; entre as características dos professores, diplomação, o tempo de experiência e nível de instrução foram as mais analisadas; entre as características dos estudantes, o nível socioeconômico dos pais, repetência, subnutrição, saúde e educação pré-primária (momentos iniciais da trajetória escolar) foram as variáveis com resultados mais congruentes entre as pesquisas analisadas.

Considerando as diferenças encontradas na literatura, Hill e Rowe (1996) buscaram identificar a importância relativa dos efeitos das turmas e escolas no contexto australiano. Os resultados sobre a eficácia escolar e docente indicaram que há uma variação substancial entre as escolas primárias no desempenho dos alunos em Inglês e Matemática, em especial, no que diz respeito ao nível socioeconômico dos estudantes.

Um estudo sobre escolas nos Estados Unidos investigou de que modo o aprendizado dos estudantes é influenciado pelas diferenças existentes entre escolas, focando na mudança do desempenho dos alunos do ensino médio. Com base em um modelo hierárquico linear, Lee (2000) concluiu que os estudantes são intensamente afetados pelas escolas que frequentam, em particular, pelo tamanho das escolas e pela sua composição social.

Outro trabalho sobre escolas norte-americanas aponta que os estudos dos efeitos escolares nos EUA negligenciaram o papel da sala de aula, considerado por Wenglinsky (2001) como o principal local de aprendizagem. Diante disso, o autor explorou a relação entre as práticas de ensino e a performance estudantil com dados do NAEP da prova de Matemática de 1996. A partir do modelo multinível, concluiu-se que as técnicas utilizadas nas salas de aula, quando

somadas a características dos professores, contribuem para o desempenho da mesma forma que as características dos estudantes.

Já Hanushek (2002) descobriu que as evidências que levaram às conclusões do Relatório Coleman indicam diversos problemas relacionados aos recursos governamentais. Em seu estudo econométrico, o autor encontrou que há grandes diferenças entre professores e escolas, no caso das escolas públicas norte-americanas. Entretanto, essas diferenças não são facilmente descritas pelos recursos empregados, pois elas descrevem, na verdade, somente padrões de resultados e a forma de organização dos incentivos nas escolas.

A pesquisa de Suryadarma *et al.* (2004), considerando estudantes da 4^a série na Indonésia, no ano de 2003, encontrou que tanto variáveis individuais, bem como variáveis escolares, afetam o desempenho. O método utilizado foi a regressão de MQO com os erros-padrão ajustados para heteroscedasticidade, utilizando-se clusters escolares a partir de uma função de produção.

O trabalho de Rivkin, Hanushek e Kain (2005) aponta para a importância do efeito do docente para o aprendizado dos alunos. Segundo os autores, a pesquisa acadêmica sobre a política educacional tem sido fortemente influenciada por resultados confusos e conflitantes. Nesse trabalho, foi empregado um painel de dados de estudantes do Texas para identificar quais as fontes de diferenças no desempenho do aluno e a relevância de uma ampla classe de políticas relacionadas aos recursos da escola. A análise empírica iniciou-se com a obtenção de estimativas de diferenças nas contribuições do professor para o aprendizado do aluno, o que, segundo os autores, eliminaria as principais fontes da possível contaminação das práticas de seleção de estudantes ou de atribuição de professor¹¹. Os resultados revelaram grandes diferenças entre os professores em seus impactos sobre o rendimento e mostram que o ensino de alta qualidade nas escolas primárias poderia compensar substancialmente as desvantagens associadas ao baixo nível socioeconômico. No entanto, essas diferenças não são facilmente medidas. Assim, grandes diferenças na qualidade do ensino excluiriam a possibilidade de que as diferenças observadas sejam movidas por fatores familiares.

Um estudo para a China aplicou um modelo multinível para tentar identificar quais variáveis são significativas no desempenho em Matemática em escolas primárias. Os resultados encontrados por Zhao *et al.* (2011) mostraram que as características dos estudantes são

¹¹ Essa questão envolve o fato de que o pesquisador desconhece algumas informações, como, por exemplo, pais que escolhem bairros e escolas, diretores que designam quais alunos vão para determinada turma, professores que decidem onde trabalhar.

importantes preditoras da variância total no desempenho em Matemática, mas as variáveis relativas à escola e às turmas não foram satisfatórias para essa explicação.

Outra conclusão importante foi encontrada por Chen (2013), a partir de dados do TIMSS de 2007, para Cingapura. A pesquisa encontrou que o sexo do aluno, o número de livros em casa, a frequência do dever de Matemática, a autoconfiança do aluno na aprendizagem, a oportunidade de desenvolvimento profissional dos professores e a percepção do diretor sobre o clima escolar, entre outros fatores, influenciam no resultado daquele teste.

Já o trabalho de Shera (2014) aplicou um modelo hierárquico linear a dados de escolas albanesas do PISA de 2009, para os testes de leitura, encontrando a importância do nível socioeconômico e do gênero no resultado da prova e, ainda, encontrou diferenças entre o setor público e privado, o nível socioeconômico da escola e o ambiente da sala de aula. Assim, conclui-se (com cautela) que o sistema educacional albanês não é nem eficaz nem equitativo no que tange ao desempenho de seus alunos em leitura.

No caso brasileiro, o trabalho de Barros *et al.* (2001), ao investigar os determinantes do desempenho educacional para áreas urbanas das regiões Nordeste e Sudeste, com base na PNAD de 1996 e na Pesquisa sobre Padrões de Vida (PPV) de 1996/97, mostrou a importância das características familiares para explicar o desempenho educacional, em particular, a escolaridade dos pais. Em termos regionais, o desempenho educacional dos jovens do Nordeste seria significativamente mais afetado pela escolaridade dos pais do que o desempenho dos estudantes do Sudeste. Assim, o autor conclui que, uma vez que a escolaridade dos pais é um fator importante na explicação da própria escolaridade dos filhos, os resultados por ele encontrados sugerem um mecanismo de geração de desigualdade e de transmissão da pobreza entre as gerações.

Albernaz, Ferreira e Franco (2002) aferiram o efeito de características escolares no desempenho dos estudantes, por meio de modelos hierárquicos, a partir de dados do SAEB, para a 8^a série em 1999. Os resultados sugerem que cerca de 80% da variância de desempenho médio entre as escolas é explicada pelas diferenças na composição socioeconômica de seus alunos, ou seja, as características individuais e familiares dos alunos têm um forte poder explicativo sobre a variância do desempenho, tanto entre os alunos, bem como entre as escolas.

Soares (2003) utilizou modelos hierárquicos (aluno, turma) com o objetivo de avaliar o impacto que atributos do professor e do ambiente em sala de aula têm sobre a proficiência dos alunos das 4^a séries do ensino fundamental, em testes de Língua Portuguesa, para o ano de 2002, em Minas Gerais. Em um trabalho posterior, Soares (2005a) ampliou o modelo de regressão

hierárquico em três níveis (aluno, turma, escola). Segundo o autor, a dificuldade em utilizar três níveis está no fato de que grande parte das escolas apresenta, para uma mesma série, um número de turmas reduzido, o que dificulta a distinção entre o efeito relativo à turma e o efeito relativo à escola, além de poder afetar a estimativa dos parâmetros do modelo devido à falta de informação.

Gremaud, Felício e Biondi (2007) construíram um Indicador do Efeito da Escola (IEE) como um avanço dos meios de análise daquilo que a escola acrescenta aos seus estudantes. Segundo os autores, essa metodologia permite a identificação das escolas que tiveram um desempenho positivo em termos médios, levando em conta os diversos perfis socioeconômicos e as características da região à qual pertencem os alunos. Assim, a referida metodologia é útil para a pesquisa acerca do ambiente escolar de aprendizado, podendo explicar em quais aspectos uma escola se diferencia de outra, mesmo trabalhando em situações similares, além de poder auxiliar no julgamento ou construção de políticas educacionais. Destaca-se que os mais altos IEE envolvem as escolas que acrescentam mais ao conhecimento de seus estudantes, em comparação àquelas com um perfil socioeconômico similar.

Para De Andrade e Laros (2007), ainda que seja evidente que as condições socioeconômicas dos estudantes exerçam diretamente influência no desempenho escolar, verifica-se que a escola também tem importância, pois alunos que têm uma mesma situação socioeconômica podem ter desempenho escolar distinto por estudarem em escolas diferentes. Portanto, para os autores, o desempenho escolar está associado a vários fatores, estando isso sujeito ao que o aluno traz para o ambiente escolar, bem como às condições que a escola oferece para ele.

Franco *et al.* (2007), por meio de um modelo hierárquico linear, examinaram quais atributos escolares estão relacionadas ao aumento da performance média das escolas, o que foi avaliado por meio dos testes de Matemática, da 4^a série do Ensino Fundamental, com dados SAEB 2001. Encontrou-se que a escola faz diferença, principalmente, devido ao clima acadêmico, pela liderança do diretor e da escola e pela disponibilidade de recursos disponíveis.

Andrade e Soares (2008) também buscaram caracterizar a contribuição das escolas de educação básica sobre a proficiência dos alunos com a ajuda de modelos hierárquicos de regressão. A base de dados também é a do SAEB, mas, para o ciclo entre 1995 e 2003, referente a alunos da quarta e oitava séries do ensino fundamental e da terceira série do ensino médio. Concluiu-se que há escolas com efeitos diferenciados no país, principalmente, no que tange à

gestão pedagógica de seus recursos humanos e físicos, que poderiam ser aplicados a outras escolas, com vistas a melhorar a qualidade da educação básica brasileira.

Soares e Sátyro (2008), por sua vez, propõem a construção de um modelo explicativo da taxa de defasagem escolar das escolas brasileiras que oferecem ensino fundamental, com foco na infraestrutura. A metodologia principal advém do modelo de regressão linear para dados em painel, com base nos dados do Censo Escolar e do Censo Demográfico. Os resultados indicam que há um impacto significativo dos insumos escolares sobre a taxa de distorção idade-série escolar, entre 1998 e 2005, no Brasil. Observou-se ainda que o desempenho educacional é afetado pela maior precariedade das condições de infraestruturas, tendo havido queda nas taxas de repetência até o final dos anos 1990, mantendo-se constante nos anos seguintes. Além disso, não existem grandes diferenças entre as escolas estaduais e particulares no que se refere a condições materiais, entretanto, os resultados entre elas são bem distintos.

Machado *et al.* (2008) analisaram os determinantes do desempenho de alunos do ensino básico da rede pública estadual em Minas Gerais, na disciplina Matemática, para o ano de 2003. Os autores encontraram que as características do aluno e do *background* familiar têm maior efeito sobre a variabilidade observada no aproveitamento do aluno, quando comparado ao efeito da escola e do município de residência.

Considerando os dados do SAEB, Menezes-Filho (2011) observou uma significativa heterogeneidade nas notas obtidas por unidade da federação, mesmo considerando características das famílias dos estudantes, o que, para ele, revela a importante função que a gestão escolar possui. Viu-se que os estudantes da rede privada possuem desempenho melhor que os alunos das escolas públicas. A parcela explicada pelas diferenças entre escolas ficou entre 10% e 30%, enquanto o restante corresponde a diferenças entre os alunos e suas famílias. Em particular, o número de horas-aula mostrou-se como a única variável relacionada à escola que afeta consistentemente o desempenho do aluno.

Neto *et al.* (2013) construíram uma escala de infraestrutura escolar por meio da Teoria de Resposta ao Item, tendo como base os dados do Censo Escolar da Educação Básica 2011. Observou-se uma profunda desigualdade existente entre as escolas brasileiras, sobretudo, entre as áreas urbana e rural, e, também, entre as regiões Norte e Nordeste e as demais regiões do país.

Palermo, Silva e Novellino (2014) analisaram o caso do Rio de Janeiro, quanto a alunos do 5º ano do ensino fundamental, por meio de modelos hierárquicos com três níveis, tendo

concluído que os fatores que mais incidem no desempenho escolar são provenientes de características dos alunos, sejam elas próprias ou advindas da família.

Outra abordagem é a utilização de modelos de regressão quantílica. Moraes e Belluzzo (2014) utilizaram essa metodologia para analisar o diferencial existente entre escolas públicas e privadas no Brasil. Segundo os autores, o diferencial é favorável às escolas privadas em todos os *quantis* da distribuição, destacando a importância da condição socioeconômica da família no processo de escolha entre escola pública e privada. Em particular, os autores observaram que: estudantes do gênero feminino têm notas piores em Matemática do que os meninos; estudantes não brancos tendem a ter notas inferiores aos demais, principalmente, para escolas públicas; e a repetência tem um impacto negativo no desempenho.

Curi e Souza (2015) criaram um indicador de qualidade das escolas por meio de um modelo estrutural para cada uma das três séries da educação básica analisadas no SAEB, entre 1997 e 2005. A ideia dos autores foi ordenar as escolas conforme uma variável latente correlacionada às características da gestão escolar, com o objetivo de evitar problemas relacionados à subestimação da qualidade dos fatores escolares e de estimativas sensíveis a erros de mensuração. Conclui-se que há diferenças ao se classificarem as escolas pela boa gestão no que tange ao desempenho. Assim, o uso de tais indicadores tende a recompensar muito mais as escolas que elegem os melhores alunos do que aquelas que apresentam melhores desempenhos em termos de gestão.

Conforme foi apresentado, grande parte dos estudos que utilizaram o modelo multinível para explicar os fatores associados ao desempenho dos estudantes no país está baseada nos dados do SAEB, em sua maioria, analisando-se a prova de Matemática. O Quadro 1, abaixo, apresenta um resumo das evidências empíricas encontradas.

Quadro 1: Estudos Multinível de Fatores Associados ao Desempenho Escolar

Autores	Local	Base de dados	Disciplina	Principais variáveis explicativas
Carrasco (2005)	Peru	PISA 2000	Matemática	Capital econômico, cultural e rede de ensino.
Barbosa e Fernandes (2000) e (2001)	Região Sudeste do Brasil	SAEB 1997	Matemática (8ª série EF) e (4ª série EF)	Escolaridade do pai, gênero, rede de ensino. Infraestrutura, equipamentos escolares, características associadas aos professores.
Albernaz, Ferreira e Franco (2002)	Brasil	SAEB 1999	Matemática (8ª série EF)	NSE médio, proporção de repetentes, rede de ensino, raça.
Soares, (2005a)	Minas Gerais	SIMAVE	Português (4ª série EF)	Bolsa-escola, sexo, raça, NSE, professor participa de formação continuada, índice de posse de equipamentos.
Dias e Ferrão (2006)	Portugal	PISA 2000	Matemática	Nível de escolaridade, sexo, interesse acadêmico dos pais, proporção de meninas, comportamento do professor.
Andrade e Laros (2007)	Brasil	SAEB 2001	Português e Matemática (3ª série EM)	NSE agregado, atraso, gênero, relação da família com a escola, comparação com os colegas.
Biondi e Felício (2007)	Brasil	SAEB 1999 a 2003	Matemática (4ª série EF)	Ausência de rotatividade dos professores, experiência média dos professores superior a dois anos, a existência na escola de conexão com a internet.
Franco <i>et al.</i> (2007)	Brasil	SAEB 2001	Matemática (4ª série EF)	NSE médio, dever, existência e conservação de equipamentos da escola, biblioteca.
Laros e Marciano (2008a)	Estados Unidos	NELS 1988	Matemática	NSE, etnia, dever de casa, NSE agregado e etnia agregada.
Machado <i>et al.</i> (2008)	Minas Gerais	SIMAVE	Matemática	Reprovação, hábito de leitura, trabalho, dever, presença de coleta de lixo, localização da escola.
Marri e Racchumi (2012)	Minas Gerais	PROEB 2010	Português e Matemática (4ª e 8ª série EF, 3ª série EM)	Condições de infraestrutura das escolas, gênero, raça, escolaridade.
Moreira e Jacinto (2013)	Rio Grande do Sul	SAERS 2007	Matemática (5ª série EF e 1ª série EM)	Banheiro na residência, tempo de direção, dever, bolsa família, ter livros em casa.
Palermo, Silva e Novellino (2014)	Rio de Janeiro	Prova Brasil 2007	Matemática	Escolaridade dos pais, raça, trabalho, gênero, professores corrigem trabalhos, nível de capital incorporado na turma.

2.2 Críticas e Desafios aos Métodos de Avaliação Padronizada no Brasil

Mesmo diante dos avanços observados na educação brasileira e nas suas formas de avaliação, ainda são notáveis os dilemas por ela enfrentados. O Brasil ainda necessita acabar com o déficit histórico acumulado na educação. Eliminar o analfabetismo seria um primeiro passo na constituição de um sistema educacional integrador e capaz de conduzir o

desenvolvimento da sociedade. Nos diversos trabalhos estudados, ficou evidenciada a preocupação dos autores com os diversos desafios que o país enfrenta.

A universalização da educação básica com melhoria da qualidade oferecida é apenas um desses desafios. É por esse motivo que se tem focado nos indicadores de matrícula, nos níveis de educação infantil e ensino médio, bem como nos indicadores de qualidade da educação para o ensino fundamental.

Segundo o INEP, na década de 1990, o estudo dos desafios da educação de qualidade para todos mostrou que a parcela mais discriminada em termos de escolaridade estava relacionada à pobreza, sugerindo que a baixa renda familiar seria incompatível com a educação formal. Mas com os esforços empreendidos, pode-se assegurar o ensino fundamental obrigatório de oito anos (hoje a obrigatoriedade é de nove anos), possibilitando a progressiva universalização do ensino médio e tornando a educação infantil um direito assegurado pelo Estado (matrícula obrigatória a partir dos seis anos de idade). O impacto disso foi a elevação da oferta educacional pública nos diferentes níveis de ensino, além do declínio das taxas de analfabetismo. No entanto, reconhece-se que ainda é preciso avançar na melhoria dos serviços prestados pelas escolas (INEP, 2004).

Segundo Menezes-Filho (2011), uma questão importante seria como aumentar a qualidade da educação oferecida pelas escolas públicas brasileiras sem expulsar as crianças mais pobres do sistema. Para Fernandes (2007), o ponto fraco estava nas elevadas taxas de repetência e de abandono escolar na educação básica, além da baixa taxa de proficiência alcançada pelos estudantes nos testes padronizados. Nesse sentido, o autor propôs a construção do IDEB como uma combinação entre dois indicadores: indicadores de fluxo (promoção, repetência e evasão) e as notas obtidas nos exames padronizados por alunos ao final de um ciclo do sistema de ensino (4^a e 8^a séries do ensino fundamental e 3º ano do ensino médio). Pressupõe-se que há uma complementaridade entre essas duas dimensões, pois não é desejável que notas elevadas sejam combinadas com reprovações sistemáticas. Portanto, o que se esperava é que todas as crianças e adolescentes tivessem acesso à escola, não deixassem de frequentar as aulas, e, assim, aprendessem o conteúdo ministrado.

Soares e Nascimento (2012) concordam que, ultrapassado o problema da expansão das redes de ensino, a baixa qualidade de ensino e a repetência estão entre os maiores desafios da educação no Brasil. Os autores ressaltam ainda que um sistema educacional não deve ser avaliado como bom se os piores alunos aprendem pouco. Os autores concluem que, nos anos 2000, houve avanços significativos para os jovens entre 15 e 16 anos, principalmente, no

aperfeiçoamento das habilidades em leitura, Matemática e Ciências, devendo-se ressaltar que essa disciplinas que têm grande valor para o desenvolvimento de outras competências e habilidades necessárias para a vida em sociedade.

Franco *et al.* (2007) apontam que o desempenho dos alunos brasileiros, resultante de avaliações de larga escala, está bem próximo de grande parte dos países da América Latina. Contudo, ao se comparar com os demais países da OCDE, a performance dos estudantes brasileiros tem sido muito baixa. Ainda segundo os autores, em pesquisa feita pela OCDE, o Brasil se apresenta como uma das nações onde a correlação entre o nível socioeconômico e cultural dos educandos e as condições escolares relacionadas à eficácia escolar tem maior dimensão.

Para Castro (2009), considerando-se a extensão dos currículos escolares, a elaboração de provas nacionais exige que se estabeleça qual conteúdo é fundamental para a aprendizagem. Com ressalva, à exceção da Capes, que realiza a avaliação da pós-graduação brasileira, inexistiam avaliações sistemáticas para que as políticas educacionais tivessem em que se basearem. Desse modo, não era possível avaliar se as políticas vigentes geravam os efeitos esperados, pois, até os anos iniciais da década de 1990, não existiam meios de avaliação da aprendizagem com evidências empíricas concretas a respeito da qualidade da rede educacional brasileira. Além disso, mesmo diante de avanços no que se refere aos sistemas avaliativos, ainda existem dificuldades na utilização das informações, o que atrapalha o uso eficiente dos resultados das provas para melhorar tanto os recursos físicos, bem como os recursos humanos. Assim, a inexistência de um roteiro para se alcançar o resultado esperado em termos de qualidade educacional dificulta os avanços pretendidos.

De acordo com Brooke e Soares (2008), a consolidação e a regularidade do SAEB geraram uma mudança significativa no modo de análise do sistema de ensino brasileiro, pois permitiram aos pesquisadores brasileiros compreender, de modo sistemático, fatores relevantes para o desempenho em testes padronizados em larga escala, o que possibilitou que os sistemas educacionais fossem avaliados, tanto no que se refere a características intrínsecas à aprendizagem, bem como a aspectos socioeconômicos.

A mudança no modo de mensuração da qualidade da educação observada é bastante positiva. Essa mudança aponta que os alunos pertencentes a famílias em condições socioeconômicas inferiores entraram no sistema educacional com desvantagens no que se refere ao acesso a bens culturais e sociais. Por isso, tentar relacionar somente o número de anos de

estudo e o acesso ao ensino com o domínio das competências básicas não produz uma resposta que melhor descreva a realidade (CASTRO, 2009).

Ademais, ressalta-se que a interação professor-aluno não é uma relação que se possa padronizar. Mesmo que a proficiência seja vista como algo fundamental no papel da escola, não se deve esquecer que ela também busca a preparação da criança para a vida, em um meio social complexo (SOARES, 2009).

Diante disso, é preciso muito cuidado ao fazer comparações com resultados advindos de outros países. Países que possuem um sistema educacional não equitativo, como no caso brasileiro, em um contexto de grandes desigualdades regionais e com a enorme diferença existente entre as várias redes de ensino, podem apresentar especificidades quanto às possíveis influências no desempenho. Além disso, destaca-se que, no Brasil, a responsabilidade pela oferta da educação básica é separada entre estados e municípios (NETO *et al.*, 2013).

Por fim, Neri (2009, p. 17) destaca que “a educação constitui o verdadeiro custo de oportunidade da sociedade”. O desafio está em tentar, por meio das bases de dados existentes, conhecer a percepção de pais e alunos sobre o tempo passado na escola e as motivações para tanto, além de tentar convencê-los do alto custo de abandoná-la.

2.3 Fatores Associados à Trajetória Escolar

Ao longo da nossa história, marcada pela diferença entre classes sociais e entre países, pode-se notar que a educação, geralmente, é tratada como um meio para se alcançarem múltiplos objetivos. Nos dias atuais, ela se formalizou e se tornou cada vez mais acessível, sendo, geralmente, comandada pelo Estado. Na sociedade capitalista, fortemente marcada pela acelerada revolução tecnológica, os trabalhadores veem na educação formal um meio de ingresso e permanência no mercado de trabalho, e, portanto, uma condição para a sobrevivência (PIRES, 2005).

Para estudar os determinantes do atraso escolar, é necessário, primeiro, analisar os condicionantes da trajetória escolar. Marcada por um período de transições, é na juventude que se decide sobre a vida escolar, o trabalho, o processo de independência e a formação de uma nova família. Assim, a juventude é uma fase da vida caracterizada pela permanente necessidade de tomada de decisão, tanto no âmbito pessoal, bem como profissional.

Mesmo diante das melhorias em termos de acesso e frequência escolar observadas nas décadas recentes, a sociedade brasileira ainda tem alguns desafios pela frente. Ribeiro e

Cacciamali (2012) destacam, por exemplo, que a defasagem idade-série ainda é algo preocupante na nossa realidade. Assim, diante da necessidade de especialização do jovem para o ingresso no mercado de trabalho, a alocação do tempo entre trabalho e frequência à escola tem sido objeto de várias investigações.

As capacidades obtidas pelos jovens durante sua vida, assim como as oportunidades e dificuldades enfrentadas, podem influenciar a transição para a sua vida adulta, tendo consequências importantes sobre a posição que desempenharão social e economicamente no futuro. Alguns dos obstáculos fazem parte dessa fase da vida, sendo outras partes do momento socioeconômico do país. A necessidade de fazer escolhas nesse período da vida não está condicionada ao amadurecimento dos jovens, que, na maioria dos casos, não possuem experiências e informações necessárias para escolherem as melhores opções (CAMARANO; MELLO, 2006).

Muitos trabalhos têm pesquisado a temática acerca das trajetórias percorridas pelos jovens, uma vez que a desocupação deles está entre as preocupações fundamentais de países desenvolvidos e em desenvolvimento. Basicamente, os jovens têm como possibilidades de escolhas ocupacionais apenas estudar, estudar e trabalhar, e somente trabalhar, ou não exercer nenhuma dessas ocupações. A análise dessas preferências busca entender as particularidades do processo de transição entre a escola e o mercado de trabalho.

Os fatores que condicionam a decisão de frequentar a escola e trabalhar podem ser explicados em diferentes recortes teóricos. A literatura econômica oferece várias explicações, estando as mais frequentes relacionadas ao nível econômico das famílias (famílias mais pobres são levadas a tirar seus filhos da escola e enviá-los para o trabalho para aumentar a renda domiciliar), à baixa qualidade do ensino e ao elevado custo da escolaridade, que induz os pais a reduzirem o seu investimento na educação de seus filhos. Além disso, a cultura de valorização do trabalho como meio de amadurecimento e passagem para a vida adulta pode explicar o crescimento da probabilidade de trabalhar e estudar, mesmo para jovens que pertençam a contextos familiares com maiores recursos.

Ribeiro e Neder (2009) destacam que, mesmo que haja uma maior proporção de jovens que estudam e trabalham, ocorre por parte deles uma saída prematura da escola, confirmado a subtração da escola na juventude em troca de trabalho. Além disso, muitos jovens, ao constituírem um novo grupo familiar, assumem responsabilidades que os pressionam a desempenhar qualquer tipo de ocupação. Portanto, pode-se inferir que os jovens com menor nível escolar formem famílias vulneráveis financeiramente.

No geral, grande parte desses estudos apontam para a renda familiar e a escolaridade dos pais como as principais variáveis na decisão entre estudar e/ou trabalhar. Leme e Wajnman (2000) afirmam que pais mais educados e famílias mais ricas aumentam a probabilidade de o jovem frequentar a escola. Essas mesmas variáveis são as que mais afetam, mas em sentido oposto, a decisão de apenas trabalhar.

Nesse sentido, Becker (1994), ao estudar os determinantes das decisões familiares, destaca que os pais têm grande influência sobre a educação, a estabilidade conjugal e muitas outras dimensões da vida dos seus filhos. A enorme influência da família parece implicar uma relação muito estreita entre os rendimentos, educação e ocupações de pais e filhos. Segundo o autor, evidências empíricas mostraram que os ganhos entre pais e filhos parecem ser um pouco mais fortemente relacionados quando os pais são mais pobres, visto que famílias mais ricas podem pagar pela educação de seus filhos, assim, esses passam mais tempo em treinamento do que no trabalho.

Segundo Lima (2010), a antecipação do ingresso no mercado de trabalho e a combinação entre estudo e trabalho são algumas das principais particularidades que distinguem o padrão de transição adotado pelos jovens no Brasil. O caminho desejado e estabelecido socialmente é marcado pela condição da entrada no mercado de trabalho como o início da vida adulta.

Gomes-Neto e Hanushek (1994) apontaram a combinação de elevadas taxas de reprovação e abandono como umas das principais falhas do sistema educacional brasileiro. Ao estudar os determinantes da reprovação, utilizando um modelo de dados em painel para estudantes do Nordeste brasileiro, em específico, para as áreas rurais, os autores encontraram entre os fatores mais importantes o desempenho do aluno e políticas governamentais na escola (políticas de acesso e de melhoria na qualidade escolar).

Psacharopoulos (1997) mostra que, para o caso da Bolívia, um dos principais determinantes do sucesso escolar está relacionado ao fato de o jovem estar ou não trabalhando. Isso porque, segundo o autor, cada hora alocada para o trabalho diminui o tempo do aluno disponível para os estudos. Além disso, para o autor, o status de trabalho e o nível de escolaridade dos jovens são determinados, conjuntamente, por características dos pais, tal como, nível de escolaridade, bem como pela renda familiar. Contudo, para o caso boliviano, destacam-se as características que contribuem significativamente para a decisão do jovem de trabalhar: idade superior, sexo masculino, origem indígena e o fato de uma mulher ser chefe da família. Em comparação, na Venezuela, a decisão de trabalhar tem como determinantes: a idade, sexo masculino, uma mulher ser chefe da família e a renda familiar. Por outro lado, são

fatores que contribuem para anos de escolaridade mais elevados: sexo feminino, alta renda familiar, famílias menores e residentes em área urbana. É importante ter em mente que, nos países em desenvolvimento, alguns determinantes que levam os jovens a abandonarem a escola para trabalhar podem não ser mensuráveis, como, por exemplo, normas, tradições e crenças.

Em outro estudo, Patrinos e Psacharopoulos (1997) destacam que, muitas vezes, os pesquisadores consideram escola e trabalho como categorias mutuamente exclusivas, mas muitos jovens em países em desenvolvimento trabalham, independentemente de matrícula escolar. Além disso, aponta-se que o fato da pobreza estar correlacionada com baixos níveis de escolaridade, não se pode esperar, em se tratando de famílias pobres, que os pais proporcionem aos seus filhos todo o auxílio significativo nas atividades escolares e ainda proporcionem um bom ambiente para a aprendizagem.

Silva e Hasenbalg (2002), considerando um modelo logístico, analisaram os determinantes da progressão escolar. Segundo os autores, fatores socioeconômicos considerados em diferentes estágios do processo de escolarização têm efeito total no meio do ciclo escolar básico.

Nesse mesmo sentido, Rios-Neto, Cesar e Riani (2002) analisaram os determinantes do desempenho escolar para alunos do ensino fundamental, sob a ótica da progressão por série, por meio de um modelo logístico hierárquico. Com base nos dados da PNAD, o principal achado desses autores é que: i) a escolaridade do professor impacta positivamente sobre a progressão escolar na 1^a série; ii) a escolaridade da mãe afeta negativamente o progresso escolar. Assim, conclui-se que a escolaridade do professor seria um substituto para a escolaridade da mãe, indicando que professores mais qualificados elevariam o desempenho e diminuiriam a desigualdade educacional, ou seja, os professores conseguiriam diminuir a influência da origem socioeconômica.

Leon e Menezes-Filho (2002) desenvolveram um modelo que estudou os determinantes da reprovação, avanço e evasão escolar no Brasil, para o ensino fundamental e médio, com dados da PME, considerando o período entre 1984-1997. O foco da pesquisa se centrou na análise das características familiares e os seus efeitos sobre as probabilidades de transição, considerando um modelo de dois estágios sequenciais. Os principais resultados foram encontrados para às séries mais avançadas, observando-se uma perda de importância das variáveis familiares ao longo do tempo.

Já Ferrão, Beltrão e Dos Santos (2002) aplicaram modelos hierárquicos para os dados da SAEB-1999, referentes à 4^a série do ensino fundamental, para a região Sudeste, com o objetivo

de avaliar o efeito de políticas de não repetência. Os resultados sugerem que o desempenho acadêmico dos estudantes com idade superior à adequada para a série é inferior ao dos alunos em idade adequada e, em algumas escolas, a proficiência é mais penalizada. No entanto, não foram encontradas evidências de que haja diferenças no desempenho de alunos em atraso e daqueles em idade esperada devido ao regime de organização do ensino na escola (seriado ou progressão automática).

Vasconcellos (2005) estimou o impacto da renda familiar na frequência escolar no Brasil, no período entre 1981 e 1999. A autora destaca que, mesmo que exista correlação entre a renda familiar e frequência escolar, a renda pode captar o efeito de outras variáveis relacionadas ao contexto familiar que também impactam na educação.

Riani (2005) avaliou os determinantes do resultado educacional nos níveis de ensino fundamental e médio no Brasil, por meio da probabilidade de frequentar a escola na idade adequada e da probabilidade de progressão por série, utilizando modelos hierárquicos e espaciais. As análises da autora revelam que a educação da mãe é o fator mais relevante e que a trajetória escolar do estudante é fortemente relacionada à sua origem social, comprovando, assim, a estratificação educacional existente no Brasil.

Honda (2006) também estudou a influência das características familiares no atraso escolar, porém, por meio do método da quase verossimilhança, em dois estágios, com dados da PNAD de 1996. A autora encontrou que a educação dos pais é importante para reduzir o atraso escolar e que a renda também influencia nesse aspecto.

Por sua vez, Pontili e Kassouf (2007) fizeram uma análise da influência de características escolares e dos estudantes sobre a frequência e o atraso escolar no ensino fundamental, para os estados de São Paulo e Pernambuco. Os resultados sugerem que políticas voltadas para melhorar a escolaridade dos chefes de família e/ou a renda familiar podem reduzir o atraso e elevar a frequência escolar. Já as ações destinadas a melhorar a infraestrutura das escolas precisam considerar as especificidades locais e a meta final a ser atingida.

Hanushek, Lavy e Hitomi (2008) procuraram investigar a tomada de decisão individual sobre a frequência escolar, utilizando uma base de dados longitudinal com crianças em idade escolar no Egito. O foco da pesquisa foi centrado no modo pelo qual a qualidade da escola afeta as decisões dos alunos nos países em desenvolvimento de abandonar a escola, considerando a necessidade de esforços para aumentar os níveis de escolaridade nesses locais. O principal resultado encontrado é que os jovens parecem ser fortemente influenciados em suas decisões escolares pela qualidade de sua escola prospectiva. Observou-se que crianças com maior

desempenho tendem a ficar na escola, enquanto aquelas com habilidades individuais mais baixas tendem a abandonar a escola precocemente. Entretanto, mantidos constantes a habilidade e o desempenho individual, a qualidade da escola será o principal fator a afetar a decisão de continuar a estudar devido ao reconhecimento da qualidade escolar pelos alunos.

Neri (2009) procura ir além dos motivos tradicionais na tentativa de responder o porquê de um adolescente abandonar a escola, avaliando o quanto os distintos níveis educacionais afetam a renda e a ocupação no mercado de trabalho. O autor encontrou que, entre as principais motivações dos jovens para não permanecerem na escola estão: falta intrínseca de interesse; necessidade de trabalho e geração de renda; dificuldade de acesso à escola, entre outros. Além disso, entre as causas da evasão ligadas às dificuldades do lado da oferta estão a falta de vagas, custos mais elevados, distância da escola, falta de documentação, além da carência de escolas especiais para alunos com algum tipo de deficiência.

Observou-se que, além dos efeitos da educação na renda, os estudantes consideram os impactos da escolaridade sobre a saúde percebida pelas pessoas, isto é, maiores anos de estudo podem ter influência em outras variáveis, como a expectativa de vida, a fertilidade e a criminalidade. Além disso, segundo o autor, um dos motivos que levam os jovens a deixarem a escola diz respeito ao fato de que os ganhos auferidos pela educação virão no longo prazo, assim, poucos podem esperar para recebê-los no futuro. Esse fato é agravado pelas restrições no mercado de crédito, as quais frustram a decisão das famílias menos abastadas de continuar estudando.

Com base nos dados da Prova Brasil de 2009, Ortigão e Aguiar (2013) analisaram as principais características dos estudantes e da família que influenciam a probabilidade de repetência dos estudantes do 5º ano. Aplicando o modelo de regressão logística, notou-se que fazer dever de casa, ter frequentado a pré-escola e receber apoio da família são variáveis relevantes para reduzir o risco de reprovação.

Do mesmo modo, Shirasu e Arraes (2016), utilizando um modelo logístico multinível para os dados do Ceará, entre 2009 e 2011, analisaram os determinantes da evasão e da repetência escolar no ensino médio. Os resultados indicam que as maiores chances de abandono ocorrem em virtude da repetência contínua, do desinteresse pelos estudos e do elevado atraso escolar.

Por fim, percebe-se que ainda há controvérsias nesse debate, principalmente, pelo uso de diferentes bases de dados e pelo emprego de metodologias diversas, além da forma de

tratamento das variáveis. Assim, espera-se, com este estudo, complementar as discussões acerca de um importante desafio para a educação brasileira, que é a elevada taxa de atraso escolar.

CAPÍTULO 3 – Metodologia Econométrica

Neste trabalho, a metodologia econométrica adota modelos de regressão hierárquicos. O uso desses modelos tem se popularizado em diversas áreas de pesquisa, seja na psicologia, nas ciências sociais, na área da educação ou nas ciências médicas.

Uma definição útil é que esse é um modelo estatístico aplicado a dados coletados em mais de um nível, com vistas a esclarecer relações existentes entre os diversos níveis de hierarquia. Cabe destacar que esses modelos são frequentemente conhecidos, também, por modelos multinível, modelos de coeficientes randômicos, modelos de efeitos mistos, modelos de estrutura de covariância, entre outros (LUKE, 2004). Muitas vezes, essa diversidade de denominações pode causar certa confusão na literatura sobre esses modelos, entretanto, na maioria dos casos, esses diferentes termos são utilizados com o mesmo sentido.

As primeiras aplicações desse modelo foram utilizadas na pesquisa educacional nos anos 1980. A partir da análise do desempenho de estudantes, os pesquisadores concluíram que as observações de estudantes de uma mesma classe não eram independentes umas das outras. Uma vez que as ferramentas de pesquisa estatística padrão assumem a independência das observações, esses instrumentos não seriam apropriados para a pesquisa da performance estudantil. Nesse caso, tem-se que os estudantes de uma mesma classe podem ser descritos como uma pequena hierarquia na qual os estudantes estão agrupados em turmas. Nesse sentido, tem-se um exemplo de estrutura de dados de dois níveis, sendo o primeiro formado por estudantes, e o segundo, composto pelas turmas. Além disso, as turmas também estão agrupadas em escolas, o que configura um terceiro nível de análise. Portanto, a ideia geral da análise multinível é que a hierarquia dos dados é considerada, isto é, leva-se em conta a dependência das observações (TWISK, 2006).

Desse modo, ao levar em conta a estrutura hierárquica dos dados, a análise multinível apresenta-se como uma alternativa viável, pois considera os diversos níveis de agregação, corrigindo os erros-padrão, testes de hipóteses e intervalos de confiança. Ainda, acredita-se que, talvez, a vantagem mais evidente desses modelos venha do seu poder preditivo, pois o modelo hierárquico é capaz de gerar previsões mais consistentes quando comparado a regressões não agrupadas, em particular, quando prevê as médias de grupo (GELMAN, 2006).

Os dados do sistema educacional, geralmente, são organizados hierarquicamente ou em níveis, podendo-se encontrar vários níveis de hierarquia, tais como: alunos, turmas, escolas, municípios e estados. Um conjunto de alunos, por exemplo, forma uma turma, e um conjunto

de turmas forma uma escola, a qual está agrupada em uma dada localidade. Assim sendo, a estrutura hierárquica presente nos dados torna pouco plausível admitir a independência de observações individuais.

Conforme Goldstein (1995), uma hierarquia consiste em unidades agrupadas em diferentes níveis. Desse modo, filhos podem ser o nível 1 de unidades em uma estrutura de nível 2, em que as unidades de nível 2 são as famílias; já os alunos podem ser as unidades de nível 1 agrupados dentro das escolas, que estão no nível 2. Segundo o autor, não se pode ignorar essa relação, pois há riscos de se perder a importância dos efeitos de grupo, o que também pode tornar inválidas muitas das técnicas de análise estatísticas tradicionalmente utilizadas.

Barbosa e Fernandes (2000) destacam que, em virtude de motivos econômicos e geográficos, entre outros, a distribuição dos alunos não ocorre de modo aleatório. Diante disso, espera-se que os alunos de uma mesma escola apresentem características semelhantes entre si, se comparados aos estudantes de outras escolas.

Portanto, o chamado modelo multinível ou modelo hierárquico considera a estrutura de agrupamento dos dados. Assim, diferentemente do modelo de regressão clássico, em que o intercepto e o coeficiente de inclinação são parâmetros fixos, no modelo hierárquico, o coeficiente de inclinação e o intercepto são considerados parâmetros aleatórios.

Soares (2005a) destaca que esses modelos consideram a estrutura de agrupamento dos dados, aceitando que cada nível tenha um modelo de regressão próprio. Em compensação, esses modelos podem sofrer influência de variáveis do mesmo nível de agregação e de variáveis de níveis de agregação superiores, isto é, uma determinada característica do professor pode afetar o aluno de vários modos. Assim sendo, com a utilização de modelos hierárquicos, cada nível de agregação pode ser representado pelo seu modelo específico.

Portanto, os modelos multinível dizem respeito a modelos de coeficientes aleatórios, adequados para o exame de variáveis aleatórias cujas distribuições não sejam independentes e idênticas, mas, sim, agrupadas de modo não aleatório (ALBERNAZ; FERREIRA; FRANCO, 2002). Assim, a análise multinível permite controlar o grau de dependência das variáveis de diferentes níveis hierárquicos e suas interações.

Essas análises têm como objetivo considerar, ao mesmo tempo, o efeito de características dos alunos e das escolas no resultado obtido pelos alunos, considerando também que alunos que pertencem à mesma escola são dependentes uns dos outros (ALBERNAZ; FERREIRA; FRANCO, 2002). Assim, considerando a estrutura hierárquica presente nos dados

educacionais, esse modelo proporciona uma escolha metodológica adequada para a análise que se propõe neste trabalho.

3.1 O Modelo Hierárquico Linear

No presente trabalho, a estratégia empírica irá considerar a estrutura hierárquica presente nos dados educacionais, portanto, o modelo multinível se apresenta como uma alternativa metodologicamente apropriada para a análise proposta no presente estudo. O modelo econométrico irá considerar dois níveis hierárquicos: serão incluídas as variáveis relacionadas aos alunos (nível 1) e as relacionadas às escolas (nível 2). Dessa maneira, será possível considerar a variabilidade existente entre os alunos e entre as escolas, ou seja, leva-se em conta a possibilidade de variação de interceptos e inclinações entre as escolas.

De acordo com Laros e Marciano (2008a), ao tratar variáveis em diferentes níveis, o modelo de regressão clássico pode não ser o mais apropriado, por não levar em conta a presença de correlação entre indivíduos relacionados em um mesmo nível de agregação. No caso desta pesquisa, por exemplo, seria representado pela correlação entre alunos de uma mesma escola.

O argumento mais simples a favor do uso de técnicas de análise multinível expõe que, devido ao fato de que muito do que se é estudado tem uma natureza hierárquica, devem-se utilizar teorias e técnicas de análise que considerem essa estrutura. Ao ignorar o contexto, surge o problema de que o modelo assumirá que os coeficientes de regressão se aplicam igualmente a todos os contextos (LUKE, 2004).

Diante disso, infere-se que os modelos lineares hierárquicos são úteis para estudar estrutura de dados como os educacionais, uma vez que esses permitem que cada nível de hierarquia tenha uma especificação separada, que é reunida, posteriormente, em um modelo único. Além disso, esses modelos permitem incorporar efeitos aleatórios relacionados a cada um dos níveis, que representam as diferenças existentes entre as unidades de cada nível em relação à variável de interesse do estudo. No caso do sistema educacional, quando se deseja avaliar a proficiência em Língua Portuguesa dos estudantes de uma dada série, pode-se considerar que os alunos, em cada escola, sejam o primeiro nível do modelo e que as escolas sejam o segundo. O modelo irá permitir que variáveis como, por exemplo, nível socioeconômico, sexo, escolaridade dos pais, entre outras, possam ser incorporadas no primeiro nível do modelo, enquanto que variáveis relacionadas às escolas, como, por exemplo, escolaridade dos professores, infraestrutura e custo por aluno são incorporadas no segundo

nível, ou seja, cada nível do modelo irá incluir variáveis que busquem explicar as fontes de variabilidade da proficiência entre os estudantes (NATIS, 2001).

Em outras palavras, o modelo de regressão hierárquico explicado em dois níveis considera o aluno como a unidade do nível 1 (determinado pelo subscrito i) e a escola, como unidade do nível 2 (determinado pelo subscrito j). Assume-se a presença de J escolas, $j = 1, 2, \dots, J$, cada uma delas com n_j alunos, $i = 1, 2, \dots, n_j$. As duas fontes de variação (aluno e escola) podem ser incorporadas em um modelo em dois níveis da seguinte forma, conforme Barbosa e Fernandes (2000), Natis (2001), Luke (2004) e Machado *et al.* (2008):

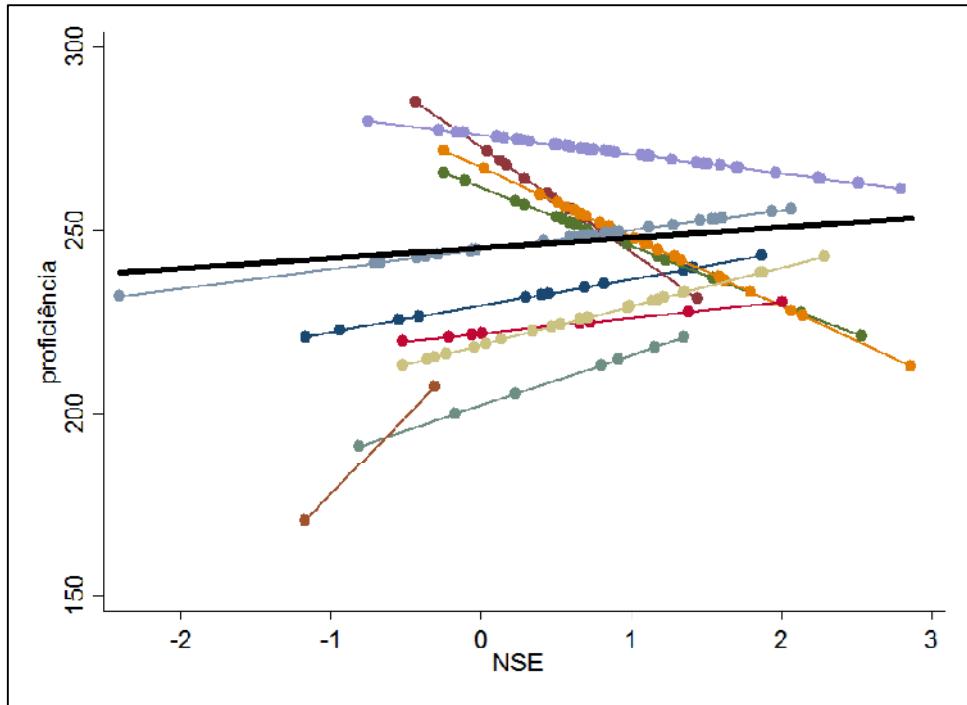
$$\begin{aligned} \text{Nível 1} \quad y_{ij} &= \beta_{0j} + \beta_{1j}(X_{ij}) + \varepsilon_{ij} \\ i &= 1, 2, \dots, n_j \text{ e } j = 1, 2, \dots, J \end{aligned} \quad (2)$$

Onde: y_{ij} representa a proficiência do i -ésimo aluno da escola j ; X_{ij} uma matriz de variáveis explicativas para a proficiência dos alunos e ε_{ij} um vetor de termos de erro que agrupa os fatores não observados que impactam a proficiência dos alunos da escola j . O intercepto β_{0j} corresponde à proficiência esperada de um aluno da escola j , quando X_{ij} é igual à sua média amostral global; já β_{1j} é a variação esperada na proficiência quando X_{1j} aumenta uma unidade em relação à média, mantidas as demais X_{kj} ($k > 1$) variáveis explicativas fixas.

A equação (2) se assemelha à estrutura típica de regressão de MQO. No entanto, o subscrito j mostra que um diferente modelo de nível 1 é, inicialmente, estimado para cada unidade j no segundo nível. Cada escola (nível 2) no estudo, por exemplo, pode ter uma pontuação média em uma determinada disciplina diferente (β_{0j}) e um diferente efeito do tempo de estudo na pontuação (β_{1j}). Dessa maneira, permite-se que o intercepto e a inclinação variem entre as unidades de nível 2. Dito de outro modo, essas equações mostram que as escolas possuem interceptos e inclinações diferentes, de modo que os efeitos aleatórios ajudam a explicar essa diferenciação entre as unidades de ensino. A inclusão de variáveis explicativas pode ser feita tanto no nível 1, bem como no nível 2.

Para ilustrar tal fato a Figura 5 exemplifica o caso de 10 escolas, com 10 retas de regressões ajustadas para cada escola tendo o Nível socioeconômico como preditor da proficiência. Na figura as linhas coloridas representam as escolas, enquanto que a linha preta representa a reta de regressão total calculada para todas as escolas juntas. Comparando as retas de regressão individuais com a reta de regressão global, nota-se que as escolas possuem inclinações e interceptos diferentes.

Figura 5: Retas de regressão para dez escolas



Fonte: Elaboração própria a partir de dados da Prova Brasil / INEP 2013

A inclusão de variáveis explicativas pode ser feita tanto no nível 1, bem como no nível 2. Assim, um modelo alternativo para o nível 2 pode ser apresentado da seguinte forma:

$$\mathbf{Nível\ 2} \quad \beta_{0j} = \gamma_{00} + \gamma_{01}Z_j + u_{0j} \quad (3)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}Z_j + u_{1j} \quad (4)$$

Onde: γ_{00} corresponde ao valor esperado dos interceptos da população de escolas; Z_j , uma matriz de variáveis explicativas para a proficiência da escola; γ_{10} é a estimativa das inclinações na população de escolas; u_{0j} e u_{1j} representam os efeitos aleatórios das escolas.

As equações (3) e (4) indicam: como cada um dos parâmetros do nível 1 são funções das variáveis explicativas e da variabilidade do nível 2; β_{0j} é o intercepto no nível 1 na unidade j do nível 2; γ_{00} é o valor médio da variável dependente do nível 1, mantendo constante os preditores do nível 2 (Z_j); γ_{01} é o efeito (inclinação) do preditor do nível 2 (Z_j); e u_{0j} é o erro, ou a variabilidade não modelada, para cada unidade j . A interpretação para a quarta equação é similar, mas nela estão sendo modelados os efeitos do nível 2 sobre a inclinação de X_{ij} ; β_{1j} é a inclinação no nível 1 na unidade j do nível 2; γ_{10} é o valor médio da inclinação do nível 1,

mantendo constante os preditores do nível 2 (Z_j); γ_{11} é o efeito das variáveis explicativas do nível 2; e u_{1j} é o erro para unidade j .

Substituindo as equações e rearranjando-as, temos uma equação que descreve o modelo geral:

$$Y_{ij} = [\gamma_{00} + \gamma_{10}X_{ij} + \gamma_{01}Z_j + \gamma_{11}Z_jX_{ij}] + [u_{0j} + u_{1j}X_{ij} + \varepsilon_{ij}] \quad (5)$$

Fixa
Aleatória

Repare-se que os β 's não são diretamente estimados, mas são estimados indiretamente através dos γ .

A seguir, é apresentada a estratégia de estimação do modelo hierárquico por meio de cinco passos, conforme sugerido por Hox (2002). Primeiramente, é realizada a análise de um modelo vazio ou nulo, isto é, a estimativa é feita apenas com o intercepto e sem nenhuma variável explicativa, conforme abaixo:

$$Y_{ij} = \gamma_{00} + u_{0j} + e_{ij} \quad (6)$$

Onde: Y_{ij} representa a proficiência do aluno i na escola j ; γ_{00} representa o intercepto da regressão; u_{0j} e e_{ij} são, respectivamente, os resíduos dos níveis da escola e do nível do aluno. Assume-se que os resíduos sigam distribuição normal, com média zero e variância σ_{u0}^2 e σ_e^2 .

Esse primeiro passo ajuda na tomada de decisão acerca da necessidade de se utilizar a modelagem multinível, isto é, ajuda a entender como características do nível superior (escola) influenciam no comportamento do nível individual (aluno). Dito de outra forma, o modelo nulo é útil porque permite calcular uma estimativa da correlação intraclasse (grau de correlação entre observações empilhadas) ρ , que nos dá a proporção da variância entre as escolas frente à variância total (o quanto da variação da proficiência escolar entre os alunos é explicado por diferenças presentes entre as escolas a que pertencem). O índice de correlação intraclasse (ICC) é dado pela seguinte equação:

$$\rho = \frac{\sigma_{u0}^2}{\sigma_{u0}^2 + \sigma_e^2} \quad (7)$$

Onde: σ_{u0}^2 é a variância dos resíduos u_{0j} do nível da escola e σ_e^2 é a variância dos resíduos e_{ij} do nível do aluno.

A partir do coeficiente de intracorrelação, pode-se estimar o grau de agrupamento existente em uma amostra (GOLDSTEIN, 1995). Esse coeficiente varia entre 0 e 1, e quanto mais próximo de zero, maior o indício de não haver estrutura de agrupamento, isto é, mais homogêneas serão as escolas, representando uma baixa influência no desempenho do estudante. Enquanto que, quanto mais o índice for próximo de 1, maior a estrutura de agrupamento, indicando que a variação no desempenho do estudante se deve a diferenças existentes entre as escolas, isto é, as características próprias do estudante não afetam o seu desempenho (BARBOSA; FERNANDES, 2000). O cálculo dessa estatística é, geralmente, utilizado para justificar a aplicação de modelos hierárquicos ao invés de modelos clássicos de regressão.

No segundo passo, analisam-se todas as variáveis explicativas no nível 1 (nível do aluno), considerando-as como efeitos fixos, conforme apresentado abaixo:

$$Y_{ij} = \gamma_{00} + \gamma_{p0}X_{pij} + u_{0j} + e_{ij} \quad (8)$$

em que X_{pij} corresponde a uma matriz das p variáveis explicativas para a proficiência dos alunos. A escolha por iniciar com as variáveis do primeiro nível ocorre em virtude do maior número de observações disponíveis nesse nível, gerando coeficientes melhores. Além disso, é estimada a contribuição dessas variáveis para a redução da variância (SOUZA; LAROS, 2010).

Feito isso, segue-se para o terceiro passo, quando se incorporam q variáveis explicativas para o nível 2 (nível da escola), representadas por Z_{qj} , conforme a equação (9).

$$Y_{ij} = \gamma_{00} + \gamma_{p0}X_{pij} + \gamma_{q0}Z_{qj} + u_{0j} + e_{ij} \quad (9)$$

Os modelos presentes na equação (8) e (9) são conhecidos como modelos de componente da variância, visto que eles decompõem a variância do intercepto em diferentes componentes de variância, para cada nível de hierarquia. Nesses modelos, considera-se que os coeficientes de regressão são fixos, mas que o intercepto varia entre as escolas (LAROS; MARCIANO, 2008a).

Na próxima etapa da estimação, é feita a inclusão da interação entre os resíduos do nível da escola (u_{pj}) e as variáveis explicativas do nível do aluno (X_{pij}) para que se possa verificar

se algum dos coeficientes de regressão das variáveis explicativas do menor nível possui variância significativa. Esse modelo é conhecido como modelo de coeficientes randômicos, sendo a equação representativa descrita da seguinte forma:

$$Y_{ij} = \gamma_{00} + \gamma_{p0}X_{pij} + \gamma_{q0}Z_{qj} + u_{pj}X_{pij} + u_{0j} + e_{ij} \quad (10)$$

No quinto passo, incluem-se no modelo as interações entre variáveis explicativas do nível da escola e as variáveis explicativas do nível do aluno, as quais tiveram significância estatística no passo anterior. Assim, tem-se o modelo completo descrito na equação (11):

$$Y_{ij} = \gamma_{00} + \gamma_{p0}X_{pij} + \gamma_{q0}Z_{qj} + \gamma_{pq}Z_{qj}X_{pij} + u_{pj}X_{pij} + u_{0j} + e_{ij} \quad (11)$$

Por fim, essas equações mostram que as escolas possuem interceptos e inclinações diferentes, de modo que os efeitos aleatórios ajudam a explicar essa diferenciação entre as unidades de ensino. Conforme apresentado, a inclusão de variáveis explicativas pode ser feita tanto no nível 1, bem como no nível 2.

3.2 O modelo Logístico Hierárquico

O segundo exercício econométrico a ser feito nesta tese considera como variável resposta (y) uma *dummy* para atraso escolar, a qual assume valores 0 ou 1. Nesse caso, não é indicado o uso de modelos hierárquicos lineares.

A abordagem linear não é adequada porque o termo de erro do nível 1 não segue os pressupostos de normalidade e homogeneidade, uma vez que o valor predito é considerado em termos de probabilidade de ocorrência de um evento. O erro do nível 1 assumirá um determinado valor quando a variável dependente for igual a um, e outro valor, quando ela for igual a zero (RIANI, 2005).

Assim, o modelo logístico multinível padrão considera que as observações em y são aleatórias e independentes, com distribuição Bernoulli, estando condicionadas a um vetor de efeitos aleatórios. A modelagem usual para o caso de respostas binárias utiliza uma função *logit* (MCMAHON; POUGET; TORTU, 2006; FIELDING; GOLDSTEIN, 2006), conforme equação 12:

$$\text{logit} (p_{ij}) = \log \left(\frac{p_{ij}}{1 - p_{ij}} \right) \quad (12)$$

Onde: p_{ij} é a probabilidade de ocorrência do evento, para o indivíduo i no grupo j ; η_{ij} é o logaritmo da probabilidade.

Neste trabalho, optou-se por um modelo de 2 níveis no contexto de atraso escolar. Têm-se, portanto, estudantes (i) que podem estar com dois anos ou mais além da idade adequada para a série ($y_{ij} = 1$) ou com idade regular ($y_{ij} = 0$), provenientes de escolas (j). O interesse está na probabilidade de o aluno se encontrar em atraso escolar. Destaca-se que essa probabilidade variará entre os alunos, dependendo dos valores de um conjunto de variáveis explicativas e do efeito da escola. Observa-se que o logaritmo das probabilidades de atraso é definido como a razão entre a probabilidade de atraso e a probabilidade de não atraso. O modelo logístico no nível 1 assume a seguinte forma:

$$\eta_{ij} = \beta_{0j} + \beta_{1j}x_{ij} \quad (13)$$

Onde: β_{0j} representa o valor de η_{ij} quando a variável independente assume o valor zero; β_{1j} é a inclinação da η_{ij} em x_{ij} para a escola j .

Note-se que o termo de erro é omitido pelo modelo, visto que sua variância decorre diretamente da probabilidade de sucesso (segue uma distribuição binomial), a qual é assumida como fixa na transformação das chances de sucesso. Assim, não há um parâmetro de variância separado (FIELDING; GOLDSTEIN, 2006). Com isso, a formulação padrão do índice de correlação intraclasse pode não ser válida nesse caso (MCMAHON; POUGET; TORTU, 2006). Shirasu e Arraes (2016) sugerem utilizar o cálculo do Coeficiente de Partição da Variância (CPV), que é igual ao índice de correlação intraclasse em um modelo multinível simples, considerando a variância residual do nível 1 como $\sigma_e^2 = \frac{\pi^2}{3} = 3,29$ em virtude da modelagem logística adotada.

Expandindo o modelo para o nível 2, tem-se que:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}z_j + u_{0j} \quad (14)$$

$$\beta_{1j} = \gamma_{10} \quad (15)$$

em que: γ_{00} é o valor médio do intercepto; γ_{01} é a inclinação das variáveis explicativas z_j ; u_{0j} representa os efeitos aleatórios do nível 2; e γ_{10} representa o efeito fixo entre os grupos.

Nesse caso, tem-se que, no primeiro nível, define-se um modelo *logit* usual para cada contexto (escola), permitindo-se que os coeficientes individuais variem ao longo das escolas. Substituindo as equações de nível 2 no modelo de nível 1, obtém-se a seguinte equação geral:

$$\eta_{ij} = \gamma_{00} + \gamma_{01}z_j + \gamma_{10}x_{ij} + u_{0j} \quad (16)$$

Desse modo, nesse modelo, u_{0j} representa o único efeito aleatório (considerado na literatura como o efeito-escola), sendo os demais termos da equação considerados fixos.

Por fim, no uso desses modelos, comumente, utilizam-se transformações para se chegar a um estimador linear. E, no caso do modelo hierárquico, é a partir desse estimador linear que as estruturas multiníveis se manifestam por meio de efeitos aleatórios (FIELDING; GOLDSTEIN, 2006). Na maioria das vezes, utiliza-se a função inversa para o *logit* por meio da transformação exponencial (HOX, 2010).

CAPÍTULO 4 – Determinantes da Proficiência em Minas Gerais

Segundo a revisão da literatura apresentada anteriormente, a análise do desempenho escolar deve levar em conta a condição socioeconômica dos alunos. No entanto, aspectos sociodemográficos (gênero, raça, idade) também são representativos e afetam o resultado dos estudantes. Além desses, o desempenho pode ser influenciado pelo impacto dos fatores sociais do coletivo vivenciado pelo estudante. Espera-se que escolas ou classes que possuam alunos com posição social elevada ou com maior nível cognitivo gozem de diversas vantagens, como, por exemplo, maior participação dos pais, menores conflitos disciplinares, atração e manutenção de melhores professores (ANDRADE; SOARES, 2008).

Schiefelbein e Simmons (1980) sugerem que a pesquisa empírica sobre as possíveis influências do desempenho escolar deveria considerar como unidade de observação a escola, a sala de aula e os alunos, com a finalidade de entender os efeitos que possam ter sido omitidos ou subestimados nos trabalhos antecedentes.

A metodologia do SAEB baseia-se na aplicação de testes padronizados de Língua Portuguesa e Matemática. Os testes padronizados buscam diagnosticar as habilidades cognitivas desenvolvidas pelos estudantes nas séries iniciais e finais do ensino fundamental e no primeiro ano do ensino médio. Entre os objetivos desse sistema de avaliação está a produção de informações sobre o desempenho dos estudantes e sobre as condições internas e externas à escola que recaem sobre o processo de ensino e aprendizagem.

Além da prova, os estudantes respondem a questionários sobre o ambiente e nível socioeconômico familiar, hábitos de estudo e de leitura, motivação, trajetória escolar, entre outros. Os professores e diretores também respondem a perguntas sobre seu nível socioeconômico, formação profissional, forma de gestão, práticas pedagógicas, dentre outros. Além disso, o aplicador da prova preenche um questionário com informações da escola relacionadas à infraestrutura, ambiente físico, recursos, entre outros.

Para alcançar o objetivo proposto neste trabalho, são utilizados os microdados do SAEB referentes ao ano de 2013, em específico, com informações da Prova Brasil para o estado de Minas Gerais relativas aos alunos da 4^a série/5^º ano do Ensino Fundamental. As escolas participantes são urbanas e rurais, da rede pública, com, no mínimo, vinte alunos matriculados nas séries avaliadas e que estejam declaradas no Censo Escolar de 2013.

A escolha da 4^a série/5^º ano tem relação com o fato de que, nessa série, os alunos têm um único professor, o que ajuda a analisar a sua influência no aprendizado e no desempenho dos

discentes. Além disso, como os pais evitam transferir os filhos de escola antes do término do primeiro ciclo do ensino fundamental, a proficiência alcançada ao final da 4^a série/ 5º ano reflete, em parte, a qualidade da escola atual. Além disso, optou-se pela utilização da prova de Matemática porque se espera que o estudo da Língua Portuguesa seja mais sensível às variáveis familiares, enquanto o estudo de Matemática sofre mais influência da qualidade das aulas e do ambiente escolar, sobretudo, nos anos iniciais de ensino (PALERMO, SILVA, NOVELLINO, 2014).

4.1 Variáveis Utilizadas

A variável dependente é a proficiência do estudante na prova de Matemática, sendo essa construída em uma escala de desempenho que vai de 0 a 500 pontos, formada por diferentes níveis, os quais compreendem intervalos de 25 pontos, em que cada nível envolve um grupo de habilidades que os estudantes que o compõem, provavelmente, possuem. Para cada disciplina e série, há uma escala de desempenho para a interpretação das notas¹². A ideia é que quanto mais o estudante avance no sistema de ensino, mais habilidades ele adquire. Os níveis finais da escala são compostos pelas habilidades mais elevadas previstas nas Matrizes de Referência, pois a cada nível são somadas as habilidades constantes nos grupos anteriores.

Os testes da Prova Brasil são elaborados para avaliar o sistema de ensino como um todo, e não os alunos, em particular. A intenção é levantar as notas dos estudantes brasileiros nas disciplinas avaliadas e, a partir disso, estabelecer metas para as escolas. Para efeito de comparação entre os resultados obtidos nos anos anteriores, a proficiência do aluno na disciplina é transformada na escala única do SAEB de 1997 (primeiro ano da avaliação), com média igual a 250 e desvio padrão igual a 50. As avaliações são comparáveis ao longo do tempo, porque os resultados são analisados com base no modelo unidimensional logístico de três parâmetros da Teoria da Resposta ao Item. Esses modelos fornecem os procedimentos para a obtenção de informações sobre indivíduos, itens e testes. Há dois fatores que afetam a resposta ao item: probabilidade de um indivíduo responder de uma maneira particular (traço psicológico) e o nível de dificuldade do item (INEP, 2014a). Assim, a modelagem via TRI é capaz de

¹² As escalas de proficiência para o ensino fundamental são divididas em níveis (nove para Língua Portuguesa e 10 para Matemática) que descrevem o que os alunos conseguem fazer em cada nível e exemplos de competências. Para maiores informações, consultar INEP (2016).

expressar, em uma única escala, a dificuldade do item, a proficiência (SOARES, 2002) e a discriminação do item.

No questionário, os alunos responderam a uma gama de itens relativos a características individuais (sexo, raça, idade), além de questões relacionadas à posse de determinados bens, à escolaridade dos pais ou responsáveis, à participação dos responsáveis nos estudos e suas experiências em casa e na escola, como, por exemplo, se trabalha ou não fora de casa, informação sobre abandono e reprovação, entre outros. De posse dessas informações e com base na revisão bibliográfica, procurou-se obter um maior entendimento da nota, considerando o contexto em que ela ocorreu, isto é, escolheram-se variáveis que representassem a origem social dos estudantes e as condições dos estabelecimentos de ensino.

A Tabela 5 apresenta as variáveis, referentes aos alunos, que foram escolhidas para compor as análises, conforme a revisão bibliográfica e a disponibilidade na base de dados. As características pessoais dos estudantes utilizadas nas regressões foram: gênero, cor, atraso escolar e sua condição em relação ao trabalho. Também foram utilizadas variáveis representativas do contexto socioeconômico dos estudantes: a escolaridade da mãe, se mora com os pais, nível socioeconômico da família e um indicador da atitude dos pais em relação aos estudos dos seus filhos (nível de capital social).

A literatura tem apontado a relevância de se considerar o gênero nos estudos educacionais, visto que, geralmente, ele aparece na forma de variável indicadora para o sexo masculino (ALBERNAZ, FERREIRA, FRANCO, 2002; SOARES, 2005a; FRANCO *et al.*, 2007; ANDRADE, LAROS, 2007; MARRI, RACCHUMI, 2012; VERNIER; BAGOLIN; JACINTO, 2015). A inclusão do fator racial também é comum na literatura (ALBERNAZ, FERREIRA, FRANCO, 2002; SOARES, 2005a; ANDRADE, LAROS, 2007; PALERMO, SILVA, NOVELLINO, 2014). A opção de separação entre brancos e não brancos segue a distribuição dos dados na amostra e também a bibliografia encontrada (MACHADO *et al.*, 2008; LAROS, MARCIANO, 2008b; MARRI, RACCHUMI, 2012; ALMEIDA, 2014).

A trajetória escolar do aluno também é importante na explicação dos resultados por ele alcançados. Nesse sentido, as análises, geralmente, empregam variáveis relacionadas à defasagem idade-série, atraso escolar, reprovação e abandono. A defasagem escolar é obtida, considerando-se a diferença entre a idade do aluno e a idade recomendada para a série que ele está cursando¹³. No presente trabalho, a escolha de utilizar uma variável indicadora para atraso

¹³ Considerando que o aluno entre na escola aos sete anos de idade (a partir de 2006, o aluno poderia entrar no primeiro ano com 6 anos de idade, e o ensino fundamental passa a ter 9 anos), a idade esperada para a 4^a série é de 10 anos.

tem como base, principalmente, os trabalhos de Barbosa e Fernandes (2001), Soares (2005a), Andrade e Laros (2007) e Almeida (2014). Considera-se em atraso escolar os alunos que tenham dois anos ou mais de diferença da idade adequada para a série.

Tabela 5: Variáveis Explicativas no Nível do Aluno e Descrição

Variável	Descrição
Homem	Assume o valor “1” se o aluno é do sexo masculino e “0” para o caso contrário (sexo feminino).
Branco	Assume o valor “1” se o aluno é branco e “0” para o caso contrário (pardo, preto, amarelo, indígena, não sabe).
Atraso	Assume o valor “1” se o estudante tem dois anos ou mais de diferença da idade adequada para a série e “0” para o caso contrário.
Mora com mãe e pai	Assume o valor “1” se o aluno mora com a mãe e o pai e “0” caso contrário (só com a mãe, só com o pai ou com outra pessoa responsável).
<i>Escolaridade da mãe (categoria base = mãe não estudou)</i>	
Até 4º série do fundamental	Assume o valor “1” se a mãe do aluno completou a 4ª série/5º ano, mas não completou a 8ª série/9º ano e “0” caso contrário.
Ensino fundamental completo	Assume o valor “1” se a mãe do aluno completou a 8ª série/9º ano, mas não completou o Ensino Médio e “0” caso contrário.
Ensino médio completo	Assume o valor “1” se a mãe do aluno completou o Ensino Médio, mas não completou a faculdade e “0” caso contrário.
Superior completo	Assume o valor “1” se a mãe do aluno completou a faculdade e “0” caso contrário.
Trabalho doméstico	Assume o valor “1” se o aluno exerce algum trabalho doméstico e “0” caso contrário.
Trabalha fora	Assume o valor “1” se o aluno exerce algum trabalho fora de casa (recebendo ou não um salário) e “0” caso contrário.
Tempo de lazer	Quantidade média de horas gastas, em dias de aula, assistindo televisão, navegando na internet ou com jogos eletrônicos.
Dever de Matemática	Assume o valor “1” se o aluno faz o dever de casa de Matemática sempre ou quase sempre e “0” caso contrário.
NSE	Nível socioeconômico do aluno obtido por TRI com respostas graduais, a partir dos seguintes itens do questionário do aluno: televisão; rádio; DVD; geladeira; geladeira duplex; freezer; máquina; carro; banheiro; empregada; escolaridade da mãe e do pai.
NCS	Nível de capital social do aluno obtido por TRI com respostas graduais, a partir dos seguintes itens do questionário do aluno: frequência com que os responsáveis vão às reuniões; pais incentivam a estudar; pais incentivam a fazer o dever de casa; pais incentivam a ler; pais incentivam a não faltar; se há diálogo sobre o que acontece na escola.

Fonte: Elaboração própria a partir de dados da Prova Brasil / INEP 2013

O fato de residir com os pais representa um indicador importante para um bom desempenho dos estudantes, já que pode dar indícios da influência do ambiente familiar no resultado obtido pelos estudantes. Os trabalhos de Soares (2005a), Menezes-Filho (2011) e Almeida (2014) são exemplos de estudos que consideraram esse aspecto.

Outra variável relevante e bastante utilizada é a escolaridade dos pais. Menezes-Filho (2011), por exemplo, incluiu diferentes *dummies* para a escolaridade da mãe na sua análise e

afirma que ela está entre as variáveis explicativas mais importantes no nível do aluno. Já Machado *et al.* (2008) utilizaram como variável indicadora se a mãe do aluno possui, no mínimo, o ensino médio completo. Barbosa e Fernandes (2000) utilizaram a escolaridade do pai (6 categorias possíveis) como medida de controle de nível socioeconômico do agregado familiar. Diante disso, consideraram-se seis categorias de *dummies* para a escolaridade da mãe: 1) nunca estudou e completou a 3º série do fundamental (categoria base); 2) completou a 4º série do fundamental; 3) completou o ensino fundamental; 4) completou o ensino médio; 5) possui superior completo; 6) aluno não sabe (acabou sendo excluída durante a junção dos bancos de dados).

A condição socioeconômica da família também pode ser representada pelo possível impacto negativo na nota do aluno que exerce algum trabalho fora (remunerado ou não) ou dentro de casa. Análises similares às realizadas neste trabalho podem ser encontradas em Machado *et al.* (2008); Menezes-Filho (2011); Almeida (2014); Palermo, Silva e Novellino (2014).

Ademais, considera-se o tempo gasto pelo estudante, em termos de horas por dia, assistindo à televisão, navegando na internet ou em jogos eletrônicos, como um fator com potencial impacto sobre o desempenho acadêmico. Conforme Palermo, Silva e Novellino (2014), esse tipo de tempo gasto com lazer representaria aspectos do capital cultural da família, visto que, em um primeiro momento, ele teria influência positiva sobre o desempenho, mas, posteriormente, essa tendência se inverteria, levando a efeitos negativos sobre a proficiência. A variável que indica se o aluno realiza dever de casa sempre ou quase sempre é utilizada para representar as atitudes do aluno e seu empenho com os estudos (MACHADO *et al.*, 2008; PALERMO; SILVA; NOVELLINO, 2014; VERNIER; BAGOLIN; JACINTO, 2015).

As duas últimas variáveis que medem o Nível Socioeconômico (NSE) e o Nível de Capital Social (NCS) da família são índices construídos a partir da Teoria da Resposta ao Item¹⁴. A Prova Brasil não contempla informações sobre a renda familiar, tornando-se complicada a aplicação de questões nesse sentido, visto que grande parte dos estudantes (alunos da 4ª série) desconhece essas informações.

No caso do NSE, em um primeiro momento, a escolha era trabalhar com sua construção baseada no Critério Padrão de Classificação Econômica Brasil. Esse critério, elaborado pela Associação Brasileira de Empresas de Pesquisa (ABEP) de 1º de janeiro de 2014, tem como objetivo desenvolver um sistema de pontuação em termos da posse de bens representativos da

¹⁴ Uma descrição mais detalhada do modelo de TRI é apresentada no apêndice.

capacidade de consumo das famílias (ABEP, 2014). Segundo Alves, Soares e Xavier (2014), essa forma de definição do índice socioeconômico tem apresentado resultados bastante confiáveis em relação ao poder de compra das famílias urbanas, com renda de até trinta salários mínimos, que engloba grande parte da nossa população.

Entretanto, nessa abordagem, é preciso que todos os itens sejam preenchidos, isto é, quando há valores faltantes, não são computados os escores. Se o estudante deixar de responder um único item, as suas respostas são excluídas da amostra para a construção do NSE. No presente trabalho, esse problema gerou uma perda de, aproximadamente, 60% das observações.

Diante disso, procurou-se um método alternativo para a construção dos indicadores de interesse que não perdesse tantas observações. Uma escolha alternativa foi a utilização da TRI, que fornece indicadores capazes de discriminar melhor os indivíduos ao extrair mais informações dos dados. A TRI propõe-se a modelar características não observáveis por meio de variáveis indicadoras que estejam diretamente relacionadas a essas características (PEREIRA, 2004). O escore obtido foi construído com base em um modelo de respostas graduadas, considerando uma série de questões sobre a posse de bens e escolaridade dos pais, ressaltando-se que, quanto maior for o seu valor, melhor será a condição socioeconômica da família do aluno.

A variável que mede o Nível de Capital Social (NCS) diz respeito ao interesse dos pais na educação de seus filhos. Barbosa e Fernandes (2001) empregaram em sua análise uma variável para representar a relação casa-escola, indicando a frequência com que há diálogo entre o aluno e seus pais sobre o que se passa na escola. As questões escolhidas para comporem o escore do NCS seguem o trabalho de Palermo, Silva e Novellino (2014), adotando como base as respostas dadas pelos alunos em relação à atitude dos seus pais quanto à frequência às reuniões, incentivo aos estudos, leitura, realização de dever, presença nas aulas e conversa sobre o cotidiano escolar. Espera-se que, quanto maior o valor do escore do NCS, maior seja a influência positiva dos pais sobre o desempenho dos estudantes.

As variáveis relacionadas ao ambiente escolar que buscam explicar as fontes de variabilidade da proficiência dos estudantes de diferentes escolas são variáveis relativas a características dos professores, diretores e ao ambiente físico da escola. As variáveis escolhidas para comporem a análise no nível da escola são descritas na Tabela 6.

O fato de a escola se localizar na área urbana pode representar boa qualidade de ensino, em virtude do melhor acesso a serviços públicos (MACHADO *et al.*, 2008). Também, procura-

se investigar se há alguma diferenciação entre as escolas da rede municipal e estadual de ensino¹⁵.

Tabela 6: Variáveis Explicativas no Nível da Escola e Descrição

Variável	Descrição
Escola	
Municipal	Assume o valor “1” se a escola pertence à rede municipal e “0” caso contrário.
Urbana	Assume o valor “1” se a escola está localizada na área urbana e “0” caso contrário.
Matrículas 5º ano	Número de alunos matriculados no 5º ano, conforme o censo 2013.
IPE	Índice de posse de Equipamentos da escola obtida por TRI com respostas graduais, a partir dos seguintes itens: computador, copiadora, impressora, projetor de slides e televisão.
ICE	Índice de conservação da escola obtida por TRI com respostas graduais, a partir dos seguintes itens: Telhado; Paredes; Piso; Entradas do prédio; Pátio; Corredores; Salas de aula; Portas; Janelas; Banheiros; Cozinha; Instalações hidráulicas; Instalações elétricas.
NSE da escola	Média do nível socioeconômico dos alunos da escola.
Diretor	
Assumi a direção via eleição	Assume o valor “1” se o diretor assumiu a direção por meio de eleição ou processo seletivo mais eleição e “0” caso contrário.
Tem mais de 5 anos de experiência	Assume o valor “1” se o diretor está no cargo de direção há mais de 5 anos e “0” caso contrário.
Está há mais de 5 anos na direção da escola	Assume o valor “1” se o diretor tem mais de 5 anos na direção da escola e “0” caso contrário.
Há programa de redução de abandono	Assume o valor “1” se a escola adota algum programa para reduzir taxas de abandono/evasão e “0” caso contrário.
Professor	
Branco	Assume o valor “1” se o professor é branco e “0” para o caso contrário (pardo, preto, amarelo, indígena, não declarou, não sabe).
Tem ensino superior	Assume o valor “1” se o professor tem ensino superior e “0” para o caso contrário.
Tem superior em Matemática	Assume o valor “1” se o professor tem ensino superior com Licenciatura em Matemática e “0” para o caso contrário.
Tem 5 anos ou mais de experiência	Assume o valor “1” se o professor tem mais de 5 anos como professor e “0” para o caso contrário.
Escola tem projeto pedagógico	Assume o valor “1” se o professor indicou que foi desenvolvido projeto pedagógico na escola e “0” para o caso contrário.

Fonte: Elaboração própria a partir de dados da Prova Brasil / INEP 2013

A variável “Matrículas 5º ano” é usada como uma *proxy* para o tamanho da escola, que é corretamente considerada nas análises empíricas como controle para heterogeneidade escolar. Segundo Soares (2005), alunos com baixos níveis de capital cultural aprendem mais em turmas menores (ou escolas menores), devido à maior proximidade com o professor, principalmente,

¹⁵ As escolas federais foram excluídas da análise para evitar problemas na estimação, visto que nessas escolas a forma de ingresso acontece por meio de exames que selecionam os melhores alunos. No banco de dados do SAEB, havia 27 escolas federais em Minas Gerais, representando somente 0,47% das escolas totais.

nas séries iniciais. Na estimação do modelo, foi incluída a variável “Matrículas 5º ano” ao quadrado para tentar capturar um possível efeito não linear dessa variável (MENEZES-FILHO, 2011).

Na consideração do ambiente escolar, a infraestrutura e a qualidade de equipamentos escolares, geralmente, são estudadas por meio da construção de índices, conforme foi discutido nos trabalhos de Barbosa e Fernandes (2001), Soares (2005), Soares (2005a), Franco *et al.* (2007), Andrade e Laros (2007), dentre outros. No presente estudo, o Índice de Posse de Equipamentos da escola (IPE) foi construído por TRI, indicando que, quanto maior o escore obtido, melhor é a qualidade dos equipamentos da escola e seu possível efeito positivo sobre o desempenho dos estudantes. O Índice de Conservação da Escola (ICE), que visa a caracterizar a conservação da infraestrutura dos estabelecimentos de ensino, foi inspirado no trabalho de Almeida (2014). O escore do ICE também foi obtido via TRI, e, nesse caso, quanto maior o valor encontrado, melhor é a qualidade da infraestrutura da escola.

O NSE médio da escola representa uma variável de controle a fim de assegurar que não haja sobrestimação dos efeitos de algumas das variáveis relativas à escola. Barbosa e Fernandes (2001), por exemplo, ressaltam que a falta de um controle adequado pode alterar o efeito e a magnitude das diversas esferas administrativas da organização escolar (estadual, municipal, privada).

No que se refere aos diretores, busca-se analisar suas características particulares e de medidas adotadas na sua gestão. A forma com que o diretor assumiu a direção e o tempo que ele exerce essa função na escola já foi estudado, por exemplo, em Menezes-Filho (2011) e Almeida (2014). Espera-se que formas mais democráticas de seleção do diretor levem a melhores resultados nas escolas, uma vez que isso reduziria o clientelismo político dentro das escolas e teria mais em vista a capacidade do indivíduo para exercer o cargo (SANTOS; SAMPAIO; SAMPAIO, 2016). A resposta do diretor, ao ser questionado com a finalidade de identificar se a escola adota algum programa para reduzir taxas de abandono/evasão, também é considerada. O exame da influência de programas para redução de abandono é encontrado nos trabalhos de Menezes-Filho (2011) e Vernier, Bagolin e Jacinto (2015).

Quanto às características dos professores que devem contribuir para o desempenho escolar, a escolaridade do docente tem sido usualmente empregada em estudos que visam analisar os determinantes do desempenho escolar e, geralmente, assume a forma de variável categórica para indicar se o professor possui ou não curso superior (BARBOSA, FERNANDES, 2001; ALBERNAZ, FERREIRA, FRANCO, 2002; SOARES, 2005; BIONDI,

FELÍCIO, 2007; MACHADO *et al.*, 2008; MENEZES-FILHO, 2011; ALMEIDA, 2014). Além disso, a experiência do professor e o tempo trabalhado na escola também têm sido tema de interesse nas investigações empíricas (SOARES, 2005; ANDRADE, LAROS, 2007; MENEZES-FILHO, 2011). Por fim, é importante analisar se a escola adota algum projeto pedagógico, visto a sua importância na definição do clima em que se dará o aprendizado (SOARES, 2005).

4.2 Estatísticas Descritivas

Nesta seção, são apresentadas as estatísticas descritivas das variáveis utilizadas no modelo econométrico. Ao filtrar os dados da Prova Brasil 2013 para os alunos de Minas Gerais da 4^a série/5^º ano do ensino fundamental que preencheram a prova, preencheram o questionário, estavam em concordância com o censo escolar e que pertenciam à rede municipal ou estadual de ensino, o banco de dados ficou com 227.296 mil observações. As informações para os professores dessa série totalizaram 12.560 mil e, para os diretores, foram 5.419 mil observações, de um total de 5.444 mil escolas. Após a junção dos três bancos de dados (alunos, professores e diretores), obtiveram-se 2.306 observações de professores, 2.334 observações de diretores, 2.353 observações de escolas e 41.610 observações de alunos. Quando estimado o modelo, a amostra final totalizou 32.987 mil observações de alunos e 2.174 mil observações de professores, diretores e de escolas.

Ademais, foram realizados testes de diferenças de médias e proporções das variáveis de alunos, professores, diretores e escolas analisadas neste trabalho, entre a amostra da Prova Brasil como um todo e a amostra do modelo final para verificar se as análises realizadas teriam validade externa¹⁶. Os resultados do teste sugerem que: os alunos da amostra da PB parecem ter características socioeconômicas de maior vulnerabilidade/desvantagem em relação àqueles que compõem a amostra final; há uma proporção maior de diretores eleitos e com maior tempo de experiência na amostra PB; há uma proporção maior de escolas municipais, na região urbana e, em média, com 2,4 alunos a menos matriculados na 4^a série/ 5^º ano na amostra PB. As diferenças quanto às características dos professores não se mostraram estatisticamente significantes. Contudo, as diferenças encontradas nos aspectos alunos, diretores e escolas se

¹⁶ As tabelas com os resultados dos testes foram suprimidas por questão de espaço e estão disponíveis sob requisição à autora.

mostraram muito pequenas em termos de magnitudes¹⁷. Além disso, a significância estatística nos testes de diferenças de médias e proporções muito se deve ao tamanho da amostra, que reduz artificialmente o erro padrão das estatísticas de teste (convergindo para zero). Assim, embora estatisticamente significantes, acredita-se que essas diferenças não comprometem a validade externa dos resultados aqui apresentados para o sistema educacional de Minas Gerais.

A Tabela 7, abaixo, traz as estatísticas descritivas das características dos alunos que compõem a amostra final¹⁸. Essas variáveis resumem os fatores que são relevantes para o desempenho escolar, sendo importante levar em conta não só o seu valor individual, mas, também, o seu valor no grupo da escola, o que é conhecido como efeito dos pares (SOARES; ALVES, 2013).

Tabela 7: Estatísticas Descritivas - Alunos, 5º Ano, Minas Gerais

Variável	Nº obs.	Média	Mínimo	Máximo	Desvio Padrão	CV
Proficiência em Matemática	32.987	239,51	80,64	341,25	49,69	0,21
Homem	32.987	0,50	0	1	-	-
Branco	32.987	0,30	0	1	-	-
Atraso	32.987	0,11	0	1	-	-
Mora com mãe e pai	32.987	0,71	0	1	-	-
<i>Escolaridade da mãe</i>						
Até 4º série do fundamental	32.987	0,22	0	1	-	-
Ensino fundamental completo	32.987	0,17	0	1	-	-
Ensino médio completo	32.987	0,24	0	1	-	-
Superior completo	32.987	0,20	0	1	-	-
Trabalho doméstico	32.987	0,80	0	1	-	-
Trabalho fora	32.987	0,12	0	1	-	-
Tempo de lazer	32.987	2,34	0	4	1,26	0,54
Dever de Matemática	32.987	0,85	0	1	-	-
Nível Socioeconômico	32.987	0,31	-3,18	4	0,87	2,83
Nível de Capital Social	32.987	-0,04	-2,76	0,32	0,57	-14,17

Fonte: Elaboração própria com base nos dados da Prova Brasil / INEP 2013

Conforme a Tabela 7, a proficiência média em Matemática foi de 239,51, com desvio-padrão de 49,69 e um coeficiente de variação da ordem de 21%. À guisa de referência, cabe aqui mencionar o que seria uma pontuação adequada para a série: há um movimento denominado Todos Pela Educação (TPE) que define cinco metas para a educação, dentre as

¹⁷ Os testes de diferenças de médias não foram realizados para as variáveis construídas por TRI (NSE, NCS, IPE e ICE) devido à grande perda de observações que a construção dessas variáveis acarreta.

¹⁸ Para as variáveis binárias entende-se a média como a proporção de sucesso, isto é, a proporção da variável *dummy* quando ela é igual a 1.

quais está a meta de desempenho que determina um nível mínimo esperado para as notas obtidas nas avaliações do SAEB, ao final de cada ciclo do ensino básico¹⁹. Segundo o TPE, para a proficiência em Matemática, a nota esperada para a 4^a série do ensino fundamental deve estar acima de 225 pontos para que se considere que o estudante esteja no nível adequado de desempenho (TPE, 2008). Levando isso em consideração, a nota média de 239,51 indica que os estudantes mineiros da 4^a série/5º ano, em média, conseguiram cumprir com o nível satisfatório da meta de desempenho do TPE, mas a alta heterogeneidade observada alerta para o fato de que muitos ainda estão aquém desse nível mínimo adequado. Assim, há espaço e necessidade de se promover, por meio de políticas educacionais, uma maior equalização do desempenho dos alunos dentro da rede pública no estado, no sentido de todos obterem resultados adequados para a série. Além disso, identificar quais fatores contribuem para diminuir a disparidade observada entre as escolas é um importante passo nessa direção.

No que se refere às características dos alunos, observa-se que, entre os estudantes, têm-se 50% de meninos. Como sugere Machado *et al.* (2008), na rede pública de ensino, é esperada uma menor proporção de brancos em relação a outras etnias, pois, nessa rede, esses grupos demográficos prevalecem. Isso é verificado na amostra pesquisada, visto que 30% dos alunos se declararam brancos.

Em termos médios, 11% dos alunos estão em atraso escolar. Grande parte da amostra é composta por alunos que pertencem a famílias tradicionais (71%). Na média, 80% dos estudantes realizam tarefas domésticas, enquanto que 12% trabalham fora de casa. No que se refere à escolaridade da mãe, em média, 22% têm até a 4^º série do fundamental, 17% têm ensino fundamental completo, 24% têm ensino médio completo e 20% têm superior completo. Ainda, os estudantes gastam, em média, 2,3 horas (em dias de aula) assistindo à televisão, navegando na internet ou jogando jogos eletrônicos, sendo o desvio-padrão 1,26 horas e dos estudantes da amostra, 85% declararam que realizam dever de Matemática na maioria das vezes.

O NSE médio do aluno foi de 0,31, com um coeficiente de variação de 2,83, indicando uma alta variabilidade dos dados em relação à média, além da heterogeneidade da condição econômica dos estudantes. A média do NCS foi de -0,04, com coeficiente de variação de 14,17, indicando que há bastante heterogeneidade quanto ao grau de participação e atenção dos pais no acompanhamento escolar dos filhos.

¹⁹ Esse movimento é uma organização sem fins lucrativos, formado por representantes de diferentes setores da sociedade civil, iniciativa privada, gestores públicos e organizações sociais, com objetivo de garantir o direito à educação básica de qualidade até 2022 (ano em que se comemora o bicentenário da independência do Brasil). Segundo Soares e Alves (2013), esse movimento tem sido aceito entre os especialistas em avaliação educacional como critério de referência para análise do desempenho escolar.

A Tabela 8 contém as estatísticas descritivas das variáveis relativas aos professores, aos diretores e à escola. Na amostra, tem-se que 63% das escolas pertencem à rede municipal e grande parte delas estão localizadas na área urbana (93%). As escolas têm, em média, 63,14 alunos matriculados no 5º ano, havendo uma heterogeneidade importante nesse aspecto (coeficiente de variação de 0,55).

Tabela 8: Estatísticas Descritivas – Professor, Diretor e Escola, 5º Ano, Minas Gerais

Nome da variável	Nº obs.	Média	Mínimo	Máximo	Desvio Padrão	CV
Escola						
Municipal	2.174	0,63	0	1	-	-
Urbana	2.174	0,93	0	1	-	-
Matrículas 5º ano	2.174	63,14	20	269	34,80	0,55
IPE	2.174	0,06	-3,10	0,78	0,73	12,10
ICE	2.174	-0,07	-2,80	0,97	0,83	-11,79
NSE da escola	2.174	0,22	-2,32	1,55	0,45	2,07
Diretor						
Assumi a direção via eleição	2.174	0,49	0	1	-	-
Tem mais de 5 anos de experiência	2.174	0,33	0	1	-	-
Está há mais de 5 anos na direção da escola	2.174	0,26	0	1	-	-
Há programa de redução de abandono	2.174	0,71	0	1	-	-
Professor						
Branco	2.174	0,43	0	1	-	-
Tem ensino superior	2.174	0,94	0	1	-	-
Tem superior em Matemática	2.174	0,03	0	1	-	-
Tem 5 anos ou mais de experiência	2.174	0,84	0	1	-	-
Escola tem projeto pedagógico	2.174	0,87	0	1	-	-

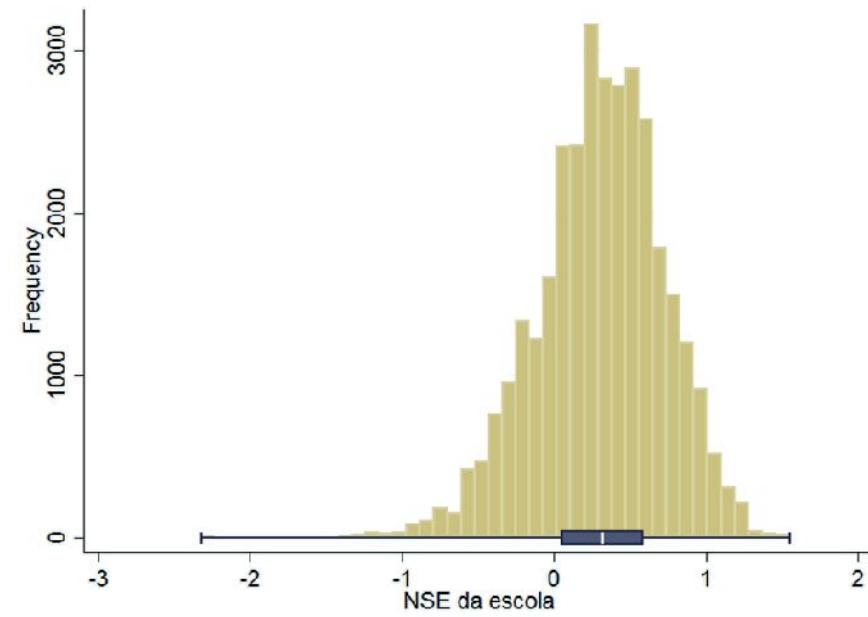
Fonte: Elaboração própria a partir de dados da Prova Brasil / INEP 2013

No que se refere ao IPE, observa-se que a sua média foi de 0,06, indicando que a maioria das escolas possuem os equipamentos analisados em boas condições, mas o elevado valor encontrado para o coeficiente de variação (12,10) indica que há bastante heterogeneidade quanto ao grau de conservação dos equipamentos entre as escolas. Já a média do ICE foi de -0,07, com desvio-padrão de 0,83, sugerindo que há uma alta dispersão na condição de conservação dos estabelecimentos de ensino entre as escolas.

O NSE médio da escola foi de 0,22, com desvio-padrão de 0,45, indicando que as escolas possuem médias de nível socioeconômico bem diferentes. O Gráfico 2 apresenta a distribuição dos valores do NSE em um histograma. Observa-se que os dados seguem uma distribuição normal, com maior concentração em torno da média, o que é esperado, dado que, na amostra, estão somente as escolas da rede pública, as quais são compostas, em grande parte, por alunos

provenientes de famílias com menor nível socioeconômico, e por seguir o Teorema do Limite Central.

Gráfico 2: Histograma do NSE da Escola

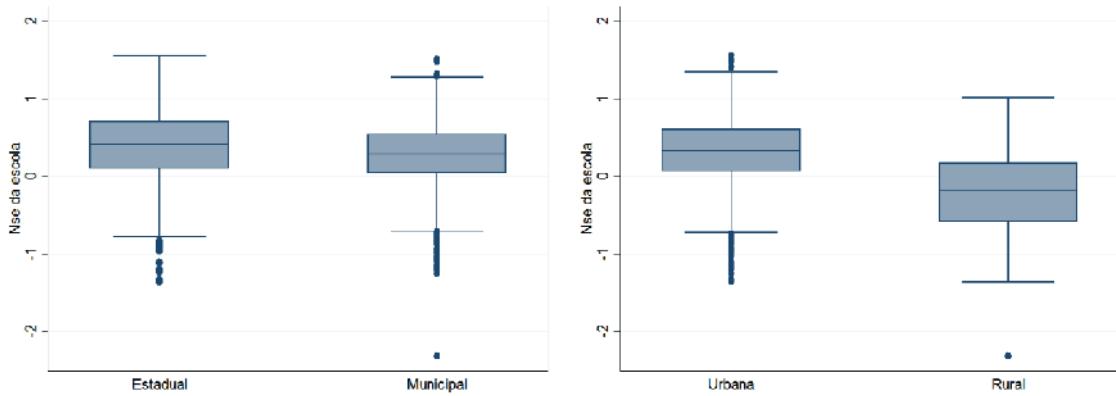


Fonte: Elaboração própria a partir de dados da Prova Brasil / INEP 2013

O Gráfico 3 diferencia o NSE da escola conforme a rede de ensino a que ela pertence. Nota-se que as redes municipais e estaduais possuem diferenças em termos do NSE dos seus alunos, e os alunos das escolas estaduais tendem a ter um NSE maior que os das escolas municipais. Isso é verificado pela linha no interior da caixa que corresponde à mediana dos dados, a qual é mais elevada nas escolas estaduais.

Já o tamanho da caixa do *boxplot* indica que o NSE das escolas estaduais é mais heterogêneo, sugerindo maior dispersão dos dados. Além disso, nota-se a presença de *outliers* nos valores mais baixos em ambas as redes de ensino, o que indica a presença de alunos com níveis socioeconômicos muito baixos. Destaca-se ainda que o NSE das escolas é diferenciado também conforme a localização, visto que os alunos das escolas da zona rural possuem um nível socioeconômico inferior aos da área urbana. Observa-se ainda que o NSE das escolas rurais tem uma dispersão maior, apontando que as condições econômicas dos alunos nessas escolas podem ser bem distintas.

Gráfico 3: Boxplot do NSE da Escola por Localização e Rede de Ensino



Fonte: Elaboração própria a partir de dados da Prova Brasil / INEP 2013

No que se refere aos diretores, 49% assumiram a direção por meio de eleição. Ainda, 33% desses estão há mais de cinco anos nessa função, e 26% estão no cargo de direção na escola avaliada há mais de cinco anos. Os dados mostram ainda que 71% das escolas apresentam algum programa de redução das taxas de abandono.

As informações quanto os professores indicam que 43% se declararam brancos. Já a proporção de professores com ensino superior é de 94%, mas somente 3% dos professores possuem ensino superior com licenciatura em Matemática. Dentre esses profissionais, 84% trabalham como professor há mais de cinco anos e 87% dos professores declararam que a escola em que trabalham desenvolve algum projeto pedagógico.

4.3 Resultado da Análise Econométrica

Na pesquisa educacional, é muito importante levar em conta a natureza dos dados. Estudantes de uma mesma escola tendem a ter características similares em virtude do método de seleção utilizado, do contexto em que se insere e da história que lhes é comum (LAROS; MARCIANO, 2008a).

O fato de os estudantes compartilharem do mesmo contexto leva a vários desafios estatísticos para o uso da regressão de mínimos quadrados ordinários, incluindo cálculos de erro padrão incorreto, heterogeneidade dos parâmetros de regressão e viés de agregação. O uso de regressão linear simples nessa estrutura de dados pode gerar erros padrão menores e problemas

de mensuração da relação entre as variáveis explicativas e independentes (modelada como tendo efeito constante entre os grupos), causando perda de informação (BICKEL, 2007).

Ao considerar a natureza hierárquica dos dados educacionais, os modelos multinível proveem ferramentas apropriadas para modelar as relações entre escolas e dentro delas (BRYK; RAUDENBUSH, 1986). Essa forma de modelagem calcula uma equação de regressão separada para cada nível dos dados, o que resulta na estimação correta dos erros-padrão. Ainda, essa modelagem permite que os efeitos de regressão sejam calculados em mais de um nível, de modo que os pesquisadores possam aferir o quanto as variáveis do nível do aluno variam entre as escolas. Além disso, a possibilidade de conjugar efeitos de níveis individuais e efeitos contextuais na mesma análise é uma razão convincente para que a modelagem multinível tenha se tornado tão visível no estudo do desempenho dos alunos (BICKEL, 2007).

O primeiro passo do exercício econométrico foi a estimação do modelo incondicional para verificar se a modelagem multinível é necessária, sendo também a base de comparação do ajuste para os modelos seguintes. Esse modelo é conhecido também como modelo nulo, sendo ele estimado apenas com a variável dependente e a variável de agrupamento, sem variáveis explicativas.

A partir desse modelo, obtém-se a média geral do desempenho do estudante na disciplina e também os componentes de variância nos níveis dos alunos (nível individual) e das escolas (nível coletivo). Considerando a hierarquia presente nos dados, decompõe-se a variância total da proficiência em duas partes: uma relativa aos estudantes (dentro das escolas) e outra relativa ao nível da escola (entre escolas). Na estimação do modelo utilizou-se o programa Stata 14.0. A construção do modelo nulo gerou os resultados expostos na Tabela 9.

Tabela 9: Modelo Nulo (sem variáveis explicativas)

VARIÁVEIS EXPLICATIVAS		MODELO NULO (M1)		
Efeito Fixo		Estimativa	Erro Padrão	P-valor
Intercepto (γ_{00})		236, 09	0,50	0.000
Efeito Aleatório		Variância	Erro Padrão	Razão-t
Variância do intercepto (σ_{u0}^2)		338,76	15,68	21,6
Variância residual (σ_e^2)		2.121,6	17,05	124,4
Correlação intraclasse		0,1377		
Deviance M1		348.713,05		
Número de parâmetros		3		
Número de observações		32.987		
Número de grupos		2.174		

Nesse modelo, a estimativa da média global da proficiência em Matemática é de 236,09 pontos. Essa média não é a mesma entre as escolas, pois o teste de Wald mostra que a variância do intercepto é estatisticamente significativa, visto que a razão-t obtida foi de 21,6 (maior que o valor crítico de 1,96). Existem, portanto, diferenças no desempenho médio em Matemática das escolas, em relação à média estadual, que justificam o uso da modelagem hierárquica para a análise dos fatores determinantes do desempenho dos alunos em Matemática, na 4^a série/ 5^º ano.

A decomposição da variância total da proficiência nos níveis indica que a variabilidade relativa aos estudantes foi de 2.121,6 e a variância relativa ao nível da escola foi de 338,76. O índice de correlação intraclasse aponta que 14% da variância total do desempenho em Matemática ocorre entre as escolas e 86% da variabilidade da nota ocorre devido a diferenças entre os estudantes dentro das escolas.

Franco *et al.* (2007) encontraram um ICC em torno de 39% para o Brasil com dados do Saeb 2001 (porém, sua amostra inclui escolas privadas, o que aumenta o grau de heterogeneidade nas características escolares). Segundo esse autor, em países europeus e nos Estados Unidos, essa porcentagem fica em torno de 20%, indicando a baixa estratificação de seus sistemas educacionais, visto a baixa desigualdade existente entre as escolas.

No presente caso, o valor encontrado de 14% pode ser justificado pelo fato de a amostra ser composta de apenas escolas públicas da rede municipal e estadual, as quais podem ser consideradas mais homogêneas em termos de gestão escolar. Esse valor encontrado para o índice de correlação intraclasse justifica a utilização de uma abordagem multinível de análise em detrimento do uso de uma regressão tradicional.

Entre as medidas existentes para se analisar o grau de ajuste dos modelos multiníveis têm-se a estatística *deviance* (-2 *Log Likelihood*) e os critérios de informação (IC), como Akaike e Schwartz, que medem a variabilidade total observada na variável dependente não explicada pelo modelo. Assim, quanto menor o valor dessas estatísticas, melhor o ajuste da modelagem multinível (BICKEL, 2007, p. 94-98).

A estatística *deviance* é utilizada para a comparação entre modelos, não podendo ser analisada isoladamente, isto é, essa estatística observa se as inclusões de variáveis explicativas contribuirão para diminuir o valor da estatística -2LL. Para testar se a redução no valor da estatística *deviance*, de um modelo ao outro, é estatisticamente significante, procede-se a um teste de hipóteses cuja estatística de teste é construída como:

$$\text{Estatística Deviance} = -2(LL \text{ de modelo nulo} - LL \text{ modelo com variáveis explicativas}) \quad (17)$$

Sob a hipótese nula de que os coeficientes das variáveis explicativas adicionais são iguais a zero (TWISK, 2006), a estatística de *deviance* segue uma distribuição Qui-quadrada com graus de liberdade iguais à diferença entre os números de parâmetros (k) usados na estimação dos dois modelos que se quer comparar (BICKEL, 2007, p. 94). Valores altos dessa estatística são tomados como evidência de que a hipótese nula é implausível e o modelo nulo é, portanto, uma descrição muito simples dos dados (RAUDENBUSH; BRYK, 2002).

A distinção entre usar IC ou -2LL está no fato de a primeira, em suas fórmulas, punir análises que usam parâmetros adicionais na estimativa de coeficientes, enquanto que a estatística -2LL não considera nenhuma perda de eficiência para o uso de parâmetros adicionais (BICKEL, 2007). Além disso, o IC teria vantagem ao comparar dois modelos não-aninhados (modelos que não constituem uma forma reduzida do outro) (LUKE, 2004).

Ainda no que se refere ao grau de ajuste, na modelagem multinível, há também uma estatística R^2 para cada nível do modelo²⁰. Uma abordagem alternativa para calcular e interpretar mensurações do R^2 para cada nível é feita pela estatística R_1^2 . No entanto, ao invés de relatar a quantidade de variabilidade na variável dependente de nível individual explicada por uma equação de regressão aleatória, o R_1^2 mostra a redução proporcional nos erros de predição, quando o modelo é comparado com o modelo incondicional (BICKEL, 2007). O cálculo da variância explicada no nível do aluno é feito da seguinte forma²¹:

$$R_1^2 = [1 - (\hat{\sigma}_{e/c}^2 + \hat{\sigma}_{u_0/c}^2 | \hat{\sigma}_{e/n}^2 + \hat{\sigma}_{u_0/n}^2)] * 100 \quad (18)$$

Onde: $\hat{\sigma}_{e/n}^2$ corresponde à variância residual do nível de aluno para o modelo nulo; $\hat{\sigma}_{e/c}^2$ é a variância residual do nível de aluno no modelo de comparação.

Por fim, outro aspecto importante e específico do procedimento de análise utilizado em modelos hierárquicos envolve a centralização das variáveis explicativas (RAUDENBUSH; BRYK, 2002). A centralização é, simplesmente, o processo de transformar linearmente uma variável X, subtraindo-se uma constante significativa, muitas vezes, algum tipo de média de X.

Na centralização pela grande média, as variáveis brutas são transformadas, subtraindo-se delas o valor da média amostral global. Esse procedimento, muitas vezes, melhora a interpretação dos coeficientes e permite o uso de termos de interação, sem gerar problemas de estimação. Além disso, outra razão para se utilizar a centralização pela grande média é o fato

²⁰ A estatística R^2 convencional não é usada com modelos que contêm coeficientes aleatórios (BICKEL, 2007).

²¹ Vale destacar que essa análise só pode ser feita quando se permite que o intercepto varie, não havendo outros efeitos aleatórios.

de essa minimizar os efeitos adversos que se surgem quando interceptos aleatórios e inclinações aleatórias estão fortemente correlacionadas. A centralização pela grande média é a mais recomendada porque é mais parcimoniosa e por fornecer estimativas mais precisas, especialmente, para amostras menores. Sugere-se que outras formas de centralização, como a centralização das variáveis pela média do grupo (escola), somente sejam usadas quando houver um bom referencial teórico ou uma forte razão para sua escolha, já que a interpretação dos coeficientes relativos às variáveis contextuais se torna mais complexa (BICKEL, 2007, p.134-144). Assim, no presente trabalho, todas as variáveis explicativas que entraram na regressão foram centradas na grande média.

A segunda coluna da Tabela 10 traz os resultados do modelo 2, com a inclusão de variáveis explicativas para o nível do aluno. Como as variáveis são centradas na grande média, a estimativa do intercepto de 237,3 representa a média da proficiência estimada em Matemática na escola, quando todas as variáveis de nível 1 são iguais às médias amostrais.

Todas as variáveis do nível do aluno incluídas no modelo 2 foram estatisticamente significativas. Apenas o atraso escolar, trabalho doméstico e trabalho fora de casa exercem um efeito negativo sobre o desempenho na prova de Matemática. Chamam a atenção as estimativas encontradas para essas variáveis, mesmo após o controle pelo nível socioeconômico dos estudantes. O aluno em atraso, por exemplo, pontua, em média, 23,5 pontos a menos do que um aluno que não esteja em atraso.

Em termos de gênero, encontrou-se que os meninos têm um desempenho melhor, em média, de quase três pontos na prova de Matemática quando comparado às meninas. Esse é um fato recorrente na literatura, talvez porque existem diferenças no desenvolvimento cognitivo entre os gêneros (SOARES, 2005a).

No que se refere à cor, observou-se que os alunos brancos têm um desempenho superior, em média, de 4,4 pontos em comparação àqueles que se declararam pardos, pretos, amarelos, indígenas ou que não souberam informar. Isso pode estar relacionado à condição econômica, mas, também, a outras características da família do estudante. Esse resultado já é conhecido na literatura empírica e, conforme ressalva Couri (2010), os não brancos, além de possuírem menores níveis de proficiência, ainda enfrentam maior atraso escolar e maiores taxas de repetência. Segundo a autora, espera-se que a diferença de desempenho entre eles seja maior quanto maior for o nível socioeconômico, tanto do aluno como da escola.

Tabela 10: Modelos com Variáveis Explicativas Aluno e Escola

VARIÁVEIS EXPLICATIVAS	MODELO 2			MODELO 3			
	Efeito Fixo	Estimativa	E. P.	P-valor	Estimativa	E. P.	P-valor
Intercepto (γ_{00})		237,30	0,40	0.000	237,96	0,40	0.000
Aluno							
Homem (γ_{10})		2,85	0,50	0.000	2,87	0,50	0.000
Branco (γ_{20})		4,41	0,54	0.000	4,19	0,54	0.000
Atraso (γ_{30})		-23,52	0,81	0.000	-23,26	0,81	0.000
Mora com mãe e pai (γ_{40})		4,88	0,54	0.000	4,94	0,54	0.000
<i>Escolaridade da mãe: (categoria base = mãe não estudou)</i>							
- Até 4º série do fundamental (γ_{50})		7,64	0,78	0.000	7,56	0,78	0.000
- Ensino fundamental completo (γ_{60})		12,24	0,87	0.000	12,14	0,87	0.000
- Ensino médio completo (γ_{70})		22,62	0,86	0.000	22,31	0,86	0.000
- Superior completo (γ_{80})		15,53	0,97	0.000	15,32	0,97	0.000
Trabalho doméstico (γ_{90})		-3,51	0,62	0.000	-3,43	0,62	0.000
Trabalho fora de casa (γ_{100})		-20,41	0,77	0.000	-20,18	0,77	0.000
Tempo de lazer (γ_{110})		3,73	0,20	0.000	3,68	0,20	0.000
Dever de Matemática (γ_{120})		17,41	0,71	0.000	17,25	0,71	0.000
NSE (γ_{130})		1,19	0,36	0.001	0,55	0,38	0.142
NCS (γ_{140})		7,45	0,45	0.000	7,40	0,44	0.000
Escola							
Municipal (γ_{01})		-	-	-	-5,18	0,94	0.000
Urbana (γ_{02})		-	-	-	-1,46	1,93	0.449
Matrículas 5º ano (γ_{03})		-	-	-	0,02	0,04	0.652
Matrículas ² (γ_{04})		-	-	-	0,00	0,00	0.829
IPE (γ_{05})		-	-	-	0,46	0,58	0.431
ICE (γ_{06})		-	-	-	1,46	0,49	0.003
NSE da escola (γ_{07})		-	-	-	7,27	1,04	0.000
Diretor							
Assumiu a direção via eleição (γ_{08})		-	-	-	-4,76	0,90	0.000
Tem mais de 5 anos de experiência (γ_{09})		-	-	-	-3,67	1,47	0.012
Está há mais de 5 anos na direção da escola (γ_{010})		-	-	-	5,85	1,61	0.000
Há programa de redução de abandono (γ_{011})		-	-	-	-3,33	0,86	0.000
Professor							
Branco (γ_{012})		-	-	-	2,42	0,75	0.001
Tem ensino superior (γ_{013})		-	-	-	-0,44	1,61	0.783
Tem superior em Matemática (γ_{014})		-	-	-	5,36	2,01	0.008
Tem 5 anos ou mais de experiência (γ_{015})		-	-	-	1,36	1,03	0.188
Escola tem projeto pedagógico (γ_{016})		-	-	-	2,48	1,12	0.027
Efeito Aleatório	Variância	E. P.	Razão-t	Variância	E. P.	Razão-t	
Intercepto (σ_{u0}^2)	191,00	10,24	18,6	161,05	9,27	17,4	
Residual (σ_e^2)	1.821,23	14,63	124,5	1.822,27	14,64	124,5	
Correlação intraclasse	0,0949			0,0812			
<i>Deviance</i>	343.138,24			342.961,21			
Número de parâmetros	17			33			
Deviance M1 – Deviance M2	5.574,81			-			
Diferença de parâmetros	14			-			
Deviance M2 – Deviance M3	-			177,04			
Diferença de parâmetros	-			16			

O fato de pertencer a uma família tradicional (formada pelo pai e mãe) também gera resultados positivos (4,88 pontos a mais, em média, do que aqueles que não moram junto com ambos os pais), o que assinala a importância da presença dos pais nos anos iniciais de ensino. Destaca-se ainda a importância da escolaridade da mãe, visto que as quatro variáveis indicadoras que entraram na regressão foram positivas e estatisticamente significativas. Em comparação aos alunos com mães que não estudaram ou não completaram o primeiro ciclo do ensino fundamental (categoria base), a proficiência em Matemática é aumentada, em média, pela escolaridade da mãe, em 7,64 pontos para estudantes com mães que estudaram até a 4^a série, 12,24 pontos para alunos com mães que completaram o ensino fundamental, 22,62 pontos para estudantes com mães que estudaram todo o ensino médio e 15,53 pontos para alunos com mães que possuem curso superior completo. Segundo Menezes-Filho (2011), acredita-se que mães com maior escolaridade pressionem mais as escolas pela melhoria do ensino ou que isso facilite o aprendizado pelo melhor *background* familiar dos alunos.

O tempo gasto em televisão, internet ou jogos eletrônicos também reflete em resultados positivos sobre a proficiência. Nota-se que, para cada hora adicional com esse tipo de lazer por dia, a proficiência em Matemática tende a aumentar 3,73 pontos, em média. Pode ser que existam conexões entre aquilo que é ensinado e o que a criança absorve nesse tempo de lazer. Talvez, alguns jogos, por exemplo, desenvolvam habilidades que facilitem o aprendizado (concentração, estratégia, raciocínio lógico) de Matemática.

Além disso, estudantes que têm um maior compromisso com a escola, isto é, que fazem o dever de Matemática na maioria das vezes, têm notas maiores. Assim, estudantes que realizam o dever têm um incremento na nota de 17,4 pontos, em média, se comparados aos alunos que não fazem as tarefas de Matemática. Isso sugere que as escolas poderiam atuar no sentido de elaborarem projetos que visem a uma maior sistematização da realização de atividades extraescolares.

A estimativa do NSE de 1,19 aponta uma relação diretamente proporcional entre o nível socioeconômico e o desempenho do aluno, controlados os efeitos das demais variáveis. Do mesmo modo, aqueles estudantes que têm um NCS maior apresentam melhor desempenho na prova de Matemática quando controladas as demais variáveis, indo ao encontro do que aponta a literatura (PALERMO; SILVA; NOVELLINO, 2014). Isso significa que o estímulo dos pais ao estudo, à leitura, bem como o interesse desses pelo cotidiano escolar do filho, se mostram importantes para o alcance de bons resultados, principalmente, nos anos iniciais do ensino fundamental (ALMEIDA, 2014).

No modelo 2, os componentes de variância assumem um significado diferente daquele presente no modelo nulo. Como foram adicionadas variáveis explicativas de nível do aluno ao modelo, os componentes de variância são denominados componentes condicionais. Observou-se que o componente condicional para a variância dentro das escolas (variância residual) se modificou de 2.121,6 no modelo nulo para 1.821,23 no modelo 2. Já o componente de variação entre as escolas foi de 338,76 para 191,0. Conforme esperado, houve uma diminuição desses valores devido ao fato de que as variáveis incluídas explicaram uma significativa parcela da variação de escola para escola no desempenho médio em Matemática. Diante disso, a correlação intraclasse diminui de 0,1377 para 0,0949.

A inclusão de variáveis explicativas de nível 1 ao modelo 2 reduziu o *deviance*, sendo essa redução estatisticamente significativa, visto que a diferença entre os *deviance* de 5.574,81 com 14 graus de liberdade é superior ao valor crítico ao nível de 5% (23,68). Isso indica que o modelo 2 se ajusta melhor aos dados, ou seja, as variáveis de nível 1 aqui consideradas conseguiram explicar uma parte da heterogeneidade de notas observada dentro das escolas e entre as escolas. O R^2_1 calculado entre o modelo nulo e o modelo 2 foi de 18,21%. A inclusão das variáveis explicativas no nível 1 no modelo reduz os erros na previsão da proficiência em Matemática em 18,21%, quando comparado ao modelo nulo.

O componente de variância residual coloca um limite máximo efetivo da quantidade de variação nas médias escolares que jamais vai ser explicado por um fator de nível da escola (nível 2). Ao incluir o fator nível da escola no modelo seguinte, espera-se reduzir o tamanho desse componente de variância, indicando que já foi esclarecida parte da variação explicável.

A terceira coluna da Tabela 10 traz os resultados da estimação do modelo 3 com a inclusão de variáveis explicativas para o nível da escola. As estimativas dos parâmetros relativas aos efeitos fixos das variáveis de aluno foram muito semelhantes àquelas obtidas pelo modelo 2, com exceção do NSE do aluno, que deixou de ser significativo após feito o controle pelo NSE médio da escola. Essa não significância do NSE individual do estudante após o controle pelo NSE agregado também foi encontrada em outros estudos. Conforme Laros e Marciano (2008b), provavelmente, esse resultado ocorre em virtude da correlação existente entre as duas variáveis (0,5).

Das variáveis relativas à escola, aquelas que não apresentaram significância estatística nesse modelo 3 foram: a localização da escola; matrículas 5^a ano e matrículas ao quadrado; IPE; professor com ensino superior e professor com 5 anos ou mais de experiência. O aluno da rede

municipal de ensino tem desempenho menor se comparado aos da rede estadual, em média, de 5,18 pontos, quando as demais variáveis estão no valor médio da amostra.

Por outro lado, a nota do aluno é acrescida pela melhor infraestrutura das escolas: para cada ponto da escola acima da média amostral do ICE das escolas (o que corresponde a uma melhora no índice), a nota do aluno aumenta, em média, 1,46 ponto, mantendo as demais variáveis constantes. Observa-se a importância das condições físicas e de funcionamento dos estabelecimentos de ensino para o bom desempenho acadêmico, em concordância com o trabalho de Barbosa e Fernandes (2001). Do mesmo modo, o incremento de uma unidade acima da média amostral do NSE da escola eleva, em média, a proficiência em Matemática em 7,27 pontos dos alunos dessa escola, sendo as outras variáveis constantes.

Das variáveis relacionadas ao diretor, somente o fato de esse exercer a função de direção há cinco anos ou mais teve relação positiva com a proficiência dos alunos. Desse modo, alunos provenientes de escolas com diretores que estão nessa função há cinco anos ou mais têm a proficiência em Matemática aumentada, em média, em 5,85 pontos, quando comparados a alunos de escolas com diretores com menos de cinco anos no cargo. Isso sugere que a gestão de diretores com a maior experiência no cargo resulta em um ambiente escolar com efeitos positivos sobre o desempenho acadêmico dos estudantes. Talvez a alternância de profissionais no cargo de direção seja um fato ruim, já que isso pode levar a instabilidades no ambiente de trabalho.

Esperava-se que a forma de escolha do diretor via eleição contribuísse para um melhor aproveitamento dos estudantes. No entanto, para o presente caso, diretores eleitos parecem conduzir a resultados piores, dado que a proficiência em Matemática dos estudantes é reduzida, em média, em 4,76 pontos nas escolas em que os diretores assumiram a direção por meio de eleição. Almeida (2014), por meio de regressões quantílicas, encontrou relação negativa entre diretores escolhidos por meio de um processo de seleção e/ou eleição e a proficiência, nos casos dos percentis representativos dos estudantes com piores desempenhos na fase inicial do ensino fundamental. Em conformidade com esses resultados, Santos, Sampaio e Sampaio (2016) sugerem que isso possa estar relacionado à falta de incentivos para selecionar bons candidatos, já que, no geral, são poucos concorrentes à eleição. Além disso, há um maior peso dos votos dos professores, o que leva à formação de corporativismo e à necessidade de apoio durante a eleição; além das diferenças existentes entre esse processo nas escolas estaduais e municipais.

Do mesmo modo, a existência de programas de redução de taxas de abandono impacta negativamente na nota do aluno na prova de Matemática, o que também foi encontrado por

Menezes-Filho (2011). Observa-se que, em média, a nota do aluno é reduzida em 3,33 pontos nas escolas em que os diretores informaram desenvolver algum programa de redução de taxas de abandono, mantendo as demais variáveis constantes. Assim, cabe destacar que a simples existência de ações nesse sentido não leva aos resultados esperados no desempenho escolar, principalmente, para o público-alvo desses programas, que são estudantes que precisam de uma atenção maior por apresentarem maiores dificuldades em prosseguir com a vida escolar. Esses programas podem conseguir manter os estudantes nas escolas, mas tendem a reduzir a média escolar porque esses alunos possivelmente são os que estão em atraso e/ou são aqueles que têm um desempenho menor.

No que se refere às características dos professores, nota-se que, nas escolas que têm professores com licenciatura em Matemática, para cada ponto adicional, em média, a nota esperada em Matemática aumenta em 5,36 pontos, sendo as outras variáveis constantes. Isso representa um aumento significativo de 0,1 desvio-padrão na distribuição de notas, haja vista que, na amostra, em média, somente 3% dos professores possuem ensino superior com licenciatura em Matemática. Chama a atenção, portanto, a necessidade de políticas educacionais que considerem a importância da formação em Matemática dos docentes nas séries iniciais do ensino fundamental. Uma estratégia a ser seguida poderia, por exemplo, considerar a divisão do conteúdo nessas séries entre dois professores com diferentes formações. Outra explicação reside no próprio interesse do docente pelo gosto da Matemática que pode levar a dificuldades no ensino dessa disciplina, já que muitas vezes os professores com outras formações escolhem a licenciatura em Pedagogia ou Letras, por exemplo, pela aversão aos cálculos. Uma saída seria o oferecimento de cursos de capacitação para práticas didáticas do ensino de Matemática.

Além disso, observou-se que as escolas que possuem professores brancos, em média, para cada ponto adicional, a proficiência em Matemática aumenta em 2,42 pontos, *ceteris paribus*. Isso pode estar relacionado a características não observáveis nos dados da amostra, como, por exemplo, a possibilidade de que os professores brancos tenham maior acesso aos recursos culturais e econômicos, tenham maior acesso ao ensino de qualidade, melhores oportunidades de emprego e salários, façam parte de escolas com maior nível socioeconômico, entre outros fatores.

A declaração do professor sobre a existência de projeto pedagógico também aponta para resultados positivos sobre a proficiência. Espera-se que, nas escolas com projeto pedagógico, os alunos têm a proficiência aumentada, em média, em 2,48 pontos. A importância do projeto pedagógico está no auxílio ao processo de ensino-aprendizagem, já que nele são apresentados

os objetivos e prioridades para o ensino de qualidade. Assim, é um instrumento útil na condução da gestão escolar, servindo como parâmetro de avaliação e planejamento das atividades dentro da escola. Por fim, duas das variáveis referentes aos professores não tiveram significância estatística: o fato de o professor possuir curso superior (94% da amostra) e ter cinco ou mais anos de experiência (84% da amostra).

Alguns desses resultados sobre as características dos professores também são confirmados no trabalho de Menezes-Filho (2011). Segundo o autor, apesar de algumas características observáveis do professor não explicarem o desempenho dos alunos, não se pode descartar a importância dos docentes, pois, muitas características dos professores que podem impactar no desempenho dos estudantes não são observadas nos dados. Rios-Neto, César e Riani (2002) encontraram que melhores professores geram resultados positivos sobre o desempenho médio dos alunos, além de minimizar as desigualdades provenientes da origem socioeconômica dos estudantes.

A parte relativa aos efeitos aleatórios é apresentada pela estrutura de covariância estimada para os parâmetros. Como se percebe, houve redução na variância do nível da escola de 191,0 para 161,05, conforme esperado. A variância residual foi de 1.822,27, havendo uma redução da correlação intraclasse para 0,0812. O *deviance* observado para o modelo 3 foi menor que aquele observado do modelo 2 e a estatística *deviance* calculada foi de 177,04, com 16 graus de liberdade, sendo superior ao valor crítico de 23,30 a 5% de significância. Portanto, rejeita-se a hipótese nula de que as variáveis contextuais de escola, diretores e professores não acrescentam poder explicativo ao modelo.

O R^2_1 indica o quanto se conseguiu reduzir a variabilidade observada das notas quando o modelo é comparado com o modelo nulo. O valor observado do R^2_1 entre o modelo nulo e o modelo 3 indica que houve uma redução de 19,39% na variabilidade das notas ao incluir as variáveis do nível 1 e nível 2.

Destaca-se que, mesmo após o controle dos efeitos de seletividade e de composição escolar, a variância do intercepto continuou relativamente alta e estatisticamente significativa. Laros e Marciano (2008b) sugerem que a estimativa do intercepto, depois de feito o controle para as variáveis de composição e seletividade escolar, pode ser explicado como a eficácia escolar, pois, antes desse controle, o intercepto representa uma parte apreciável do NSE do aluno. Assim, pode-se interpretar a alta variância do intercepto como um indicador da existência de divergências significativas na eficácia do ensino escolar de Matemática.

Avançando na busca de se entender a heterogeneidade entre escolas, o passo seguinte da modelagem do desempenho em Matemática foi conferir se algumas das variáveis explicativas apresentavam inclinações aleatórias e não fixas, como suposto no modelo 3.

Observando-se a importância de fatores relacionados ao ambiente familiar, pode-se supor que os efeitos de algumas variáveis não são os mesmos entre as escolas. Para definir qual variável do nível 1 possuía algum efeito aleatório, observou-se a literatura empírica e os resultados encontrados no modelo 3, tendo sido testadas as seguintes variáveis: trabalha fora (ANDRADE; LAROS, 2007); atraso escolar (LAROS; MARCIANO, 2008b); NCS (COURI, 2010); escolaridade da mãe (RIANI, 2005) e cumprimento do dever (ANDRADE; LAROS, 2007). Somente apresentaram coeficientes aleatórios significativos as variáveis para o NCS e trabalha fora. Numa linguagem mais econômica, é como se o “preço” ou a “remuneração” pelo fato de o aluno trabalhar fora ou não, e ter um bom capital social em casa ou não, fosse diferenciado de escola para escola, em termos de desempenho alcançável pelo aluno.

Com o objetivo de identificar variáveis do nível 1 que trariam uma contribuição substancial para explicar a variabilidade na variável dependente que ocorre dentro das escolas (BICKEL, 2007), foram estimadas duas especificações de coeficientes aleatórios: o modelo 4 considera que apenas os coeficientes de inclinação associados ao NCS e à variável “trabalha fora” são aleatórios, e o modelo 5 que incorpora também as interações entre essas variáveis e variáveis do nível da escola (ou variáveis contextuais). Essas interações buscam captar quais variáveis da escola ajudam a explicar as diferentes inclinações observadas para as variáveis trabalha fora e NCS.

As colunas 2 e 3 da Tabela 11 trazem as estimativas referentes aos modelos 4 e 5, respectivamente. Observa-se que os efeitos fixos do nível 1 e 2 foram muito similares aos obtidos anteriormente nos modelos 2 e 3, tanto nas estimativas do modelo 4, bem como nas do modelo 5.

O componente aleatório aponta o quanto o componente fixo varia de escola para escola, tanto em se tratando de intercepto, bem como em se tratando de coeficientes de inclinação com um componente aleatório. Levando isso em consideração, a interpretação dos componentes aleatórios deve ser feita com muita cautela. Sugere-se, nesse caso, a análise da abrangência dos coeficientes aleatórios por meio de intervalos de confiança. Assim, os intervalos de confiança são construídos considerando que 95% da distribuição esteja incluída em um intervalo que comprehende a faixa de 1,96 desvio-padrão acima e abaixo do componente fixo do coeficiente aleatório de interesse (BICKEL, 2007), conforme uma distribuição normal.

Tabela 11: Modelos com Interações e Coeficientes Aleatórios (continua)

VARIÁVEIS EXPLICATIVAS		MODELO 4			MODELO 5		
Efeito Fixo		Estimativa	E. P.	P-valor	Estimativa	E. P.	P-valor
Intercepto (γ_{00})		237,95	0,40	0,000	237,95	0,40	0,000
Aluno							
Homem (γ_{10})		2,86	0,50	0,000	2,89	0,50	0,000
Branco (γ_{20})		4,21	0,54	0,000	4,12	0,54	0,000
Atraso (γ_{30})		-23,34	0,81	0,000	-23,62	0,81	0,000
Mora com mãe e pai (γ_{40})		4,92	0,54	0,000	4,90	0,54	0,000
<i>Escolaridade da mãe: (categoria base = mãe não estudou)</i>							
- Até 4º série do fundamental (γ_{50})		7,57	0,78	0,000	7,50	0,78	0,000
- Ensino fundamental completo (γ_{60})		12,16	0,87	0,000	12,09	0,87	0,000
- Ensino médio completo (γ_{70})		22,31	0,86	0,000	22,26	0,86	0,000
- Superior completo (γ_{80})		15,34	0,97	0,000	15,39	0,97	0,000
Trabalho doméstico (γ_{90})		-3,38	0,62	0,000	-3,34	0,62	0,000
Trabalho fora de casa (γ_{100})		-20,40	0,83	0,000	-20,35	0,83	0,000
Tempo de lazer (γ_{110})		3,67	0,20	0,000	3,67	0,20	0,000
Dever de Matemática (γ_{120})		17,27	0,71	0,000	17,25	0,71	0,000
NSE (γ_{130})		0,52	0,38	0,164	0,51	0,38	0,178
NCS (γ_{140})		7,45	0,46	0,000	7,43	0,46	0,000
Escola							
Municipal (γ_{01})		-5,17	0,94	0,000	-5,32	0,93	0,000
Urbana (γ_{02})		-1,48	1,93	0,442	-1,38	1,92	0,475
Matrículas 5º ano (γ_{03})		0,02	0,04	0,678	0,02	0,04	0,642
Matrículas ² (γ_{04})		0,00	0,00	0,855	0,00	0,00	0,830
IPE (γ_{05})		0,45	0,58	0,441	0,42	0,58	0,464
ICE (γ_{06})		1,45	0,49	0,004	1,44	0,49	0,004
NSE da escola (γ_{07})		7,29	1,04	0,000	7,27	1,05	0,000
Diretor							
Assumiu a direção via eleição (γ_{08})		-4,80	0,90	0,000	-4,78	0,90	0,000
Tem mais de 5 anos de experiência (γ_{09})		-3,66	1,46	0,012	-3,66	1,46	0,012
Está a mais de 5 anos na direção da escola (γ_{010})		5,84	1,61	0,000	5,85	1,61	0,000
Há programa de redução de abandono (γ_{011})		-3,33	0,85	0,000	-3,31	0,85	0,000
Professor							
Branco (γ_{012})		2,43	0,75	0,001	2,22	0,75	0,003
Tem ensino superior (γ_{013})		-0,37	1,61	0,817	-0,39	1,61	0,807
Tem superior em Matemática (γ_{014})		5,36	2,01	0,008	5,33	2,01	0,008
Tem 5 anos ou mais de experiência (γ_{015})		1,36	1,03	0,189	1,31	1,03	0,205
Escola tem projeto pedagógico (γ_{016})		2,45	1,12	0,029	2,40	1,12	0,032
Interações							
NSE escola * NCS aluno		-	-	-	-3,82	1,06	0,000
Atraso escolar * NCS aluno		-	-	-	-18,10	3,86	0,000
Proporção de brancos * trabalha fora		-	-	-	19,17	4,86	0,000
Proporção de mães com superior * trabalha fora		-	-	-	-17,90	6,16	0,004
Efeito Aleatório		Variância	E. P.	Razão-t	Variância	E. P.	Razão-t
Trabalha fora		105,93	32,95	3,2	97,82	32,60	3,0
NCS		20,40	10,28	2,0	16,27	9,95	1,6
Intercepto		160,37	9,26	17,3	158,65	9,20	32,6
Residual		1.805,28	15,05	119,9	1.805,05	15,04	120,0

Tabela 11: Modelos com Interações e Coeficientes Aleatórios (conclusão)

VARIÁVEIS EXPLICATIVAS	MODELO 4	MODELO 5
Correlação intraclasse	0,0816	0,0808
<i>Deviance</i>	342.942,64	342.892,86
Número de parâmetros	35	39
Deviance M3 – Deviance M4	18,57	-
Diferença de parâmetros	2	-
Deviance M4 – Deviance M5	-	49,78
Diferença de parâmetros	-	4

A variabilidade com relação a um coeficiente aleatório mostra os efeitos mínimos e máximos esperados de cada variável ou média escolar, no caso do intercepto. Construindo o intervalo de confiança para o intercepto no modelo 4, percebe-se que, no limite inferior, têm-se escolas com a proficiência média de 213,13 pontos, e, no limite superior, a proficiência média chega a 262,77 pontos.

A análise dos efeitos aleatórios no modelo 4 mostra que a relação entre o fato de o aluno trabalhar fora e a proficiência dentro das escolas, de fato, varia significativamente em toda a população de escolas. Da análise do intervalo de confiança, têm-se que: uma inclinação correspondente ao limite inferior do intervalo indicaria que, se o aluno trabalha fora, a proficiência se reduz em 40,57 pontos, quando comparado com um estudante que não trabalha fora; já uma inclinação correspondente ao limite superior do intervalo significaria que, se o aluno trabalha fora de casa, a proficiência se reduz em 0,23 ponto, quando comparado com estudantes que não trabalham fora. Nota-se uma significativa diferença na magnitude e no impacto das inclinações para alunos que trabalham fora.

Igualmente, a relação entre o NCS da família do estudante e a proficiência tem efeitos diferenciados entre as escolas. O intervalo de confiança para esse caso indica que, em média, no limite inferior, o NCS reduz a proficiência em 1,40 ponto, e, no limite superior, a proficiência se eleva em 16,30 pontos. Isso mostra quão diferentes podem ser as escolas em termos de retornos do nível de capital social da família para o desempenho dos estudantes.

Como se percebe, a significância desses coeficientes aleatórios indica que existe um alto grau de diferenciação entre escolas quanto ao retorno, em termos de desempenho em Matemática, do nível de capital social do aluno e do fato de o aluno trabalhar fora ou não. Os testes de interação entre essas variáveis e algumas características escolares, que são realizados em seguida, buscam explicar parte dessa remuneração diferenciada.

A melhora no ajuste do modelo 4 em comparação ao modelo anterior foi significativa, visto que o valor da estatística teste qui-quadrado calculada pela diferença dos *deviance* foi de

18,57, superior ao valor crítico a 5% com dois graus de liberdade (5,99). Rejeita-se, portanto, a hipótese nula de que os graus de ajuste dos modelos sejam iguais, considerando-se efeitos aleatórios ou não.

Como foi visto na seção anterior, o último passo da modelagem multinível proposto por Hox (2002) é verificar efeitos de interação entre os níveis. Isso é feito para se tentar explicar o porquê do retorno diferenciado dessas variáveis em termos de notas. Em termos sintéticos, pode-se dizer que o efeito de interação entre variáveis do nível micro e do nível macro representam os efeitos diferentes para grupos distintos (LAROS; MARCIANO, 2008a).

No quinto passo, as interações entre variáveis dos níveis do aluno e da escola que tiveram significância foram: 1) Nível de capital social do aluno com o Nível socioeconômico da escola; 2) Nível de capital social do aluno com a proporção de alunos em atraso escolar na escola; 3) Trabalha fora no nível do aluno com a proporção de alunos brancos na escola; 4) Trabalha fora no nível do aluno com a proporção de mães com ensino superior na escola.

Os resultados do modelo 5 são reportados na coluna 3 da Tabela 11. Quando as interações são adicionadas à análise, é interessante verificar os termos aleatórios do modelo. Comparando os resultados do modelo 5 com o modelo 4, nota-se que os componentes aleatórios das variáveis NCS e trabalha fora sofreram uma redução de 105,9 para 97,8, no primeiro caso, e de 20,4 para 16,7, no segundo. Assim, os intervalos de confiança também sofreram uma ligeira redução no modelo 5. Para a variável trabalha fora, o limite inferior do intervalo mostra que, se o aluno exerce algum trabalho fora de casa, a proficiência se reduz em 39,74 pontos quando comparado com um estudante que não trabalha fora. Já para o limite superior, o aluno que trabalha fora de casa tem a proficiência reduzida em 0,96 ponto quando comparado com estudantes que não trabalham fora. No caso do nível de capital social, tem-se que o intervalo indica que, no limite inferior, o NCS diminui a proficiência em 0,48 ponto, e, no limite superior, o NCS eleva a proficiência em 15,34 pontos. Por fim, para o intercepto no limite inferior, a proficiência média da escola é de 213,26 pontos, e, no limite superior, a proficiência média é de 262,64 pontos.

Esse resultado sugere que parte das diferenças no coeficiente de regressão para o NCS do aluno entre escolas esteja relacionada às diferenças existentes entre as escolas em termos do NSE agregado e atraso escolar. Da mesma forma, as diferenças existentes entre o retorno dos alunos que trabalham fora de casa e as escolas têm relação direta com a escolaridade média das mães e com a proporção de alunos brancos na escola, isto é, considera-se que o efeito agregado dessas variáveis também possa afetar os estudantes que interagem nas mesmas salas de aula ou escolas. Quando considerado no agregado, o efeito de uma variável do aluno sobre o

desempenho acadêmico dos seus colegas é conhecido por *peer effect* ou externalidade gerada pelos pares (OLIVEIRA, J. 2010).

O coeficiente para a primeira interação (*NSE escola*NCS aluno*) demonstra que, à medida que o nível socioeconômico agregado da escola aumenta, o retorno individual do nível de capital social do aluno, em média, experimenta uma vantagem menor no desempenho em Matemática. Assim, se todas as variáveis estivessem equivalentes às respectivas médias amostrais, inclusive, as variáveis contextuais de escola, para cada unidade a mais no NCS do aluno, o efeito esperado sobre a nota em Matemática seria de um aumento de 7,43. Contudo, para cada ponto percentual acima da média que a escola apresentar em termos de NSE agregado, esse efeito do NCS aluno se reduz para 3,61 pontos (diferença entre o efeito fixo estimado e o parâmetro estimado da interação) na proficiência. Esse resultado parece bastante intuitivo, mas traz poucas implicações em termos de política educacional.

O atraso escolar pode estar relacionado não só com a reprovação, mas, também, com as condições econômicas e sociais que vivenciam os estudantes de uma determinada escola. Assim, espera-se que o nível de capital social dos alunos exerça menor efeito em termos de ganhos de aprendizado em ambientes em que há uma elevada proporção de alunos em atraso escolar. O efeito encontrado para a segunda interação (*Atraso escolar*NCS aluno*) indica que, em escolas com maior proporção de alunos atrasados, em média, o efeito do NCS do aluno sobre a proficiência em Matemática é reduzido. Dito de outra forma, para cada ponto percentual na proporção de alunos em atraso escolar acima da média das escolas, cada unidade a mais de NCS do aluno leva, em média, a uma redução de 10,67 pontos na proficiência.

A conclusão é que, tanto em ambientes muito desfavoráveis (em termos da elevada taxa de crianças em atraso escolar) quanto em ambientes muito privilegiados (em termos de NSE agregado nas escolas), o retorno do NCS do aluno em termos de aprendizado é reduzido, chegando a se tornar negativo no primeiro caso. Assim, espera-se que as externalidades geradas no ambiente escolar se sobreponham às influências do ambiente familiar, sugerindo que a política educacional deva buscar mecanismos de atuação que minimizem os efeitos das desigualdades existentes entre os estudantes.

No caso da terceira interação (*Proporção de brancos*trabalha fora*), seu coeficiente aponta que, para um aumento de 1 ponto percentual acima da média na proporção de alunos brancos da escola, o fato de o aluno trabalhar reduz a proficiência, em média, em 1,18 ponto, isto é, o custo individual de se trabalhar fora, em termos de aprendizado, é bastante ou praticamente anulado, visto que o efeito negativo de se exercer algum trabalho fora de casa (-

20,35) é compensado nas escolas com maior proporção de brancos (19,17). Se considerarmos a proporção de brancos na escola como uma *proxy* para uma boa condição social dos alunos na média, podemos dizer que, nessas escolas, os efeitos negativos do trabalho seriam reduzidos pelos efeitos positivos da convivência com estudantes com uma melhor condição social. Do contrário, pode-se inferir que o tipo de trabalho com que esses alunos estariam envolvidos tenderia a ser menos prejudicial aos seus estudos ou apresentariam uma jornada de trabalho menor que não impactasse tanto no tempo de estudo.

A literatura empírica tem mostrado a importante relação existente entre o desempenho educacional e a escolaridade dos pais. Barros *et al.* (2001), por exemplo, concluem que a escolaridade da mãe é o fator de maior valor entre os determinantes do desempenho educacional. Em conformidade, Ermisch e Francesconi (2001) sugerem que isso ocorre porque as mães atribuem maior peso à renda futura da criança do que o pai.

O coeficiente estimado para a última interação (*Proporção de mães com superior*trabalha fora*) sugere que, para cada ponto percentual acima da média na proporção de mães com ensino superior na escola, o fato de o aluno trabalhar reduz a proficiência, em média, em 38,25 pontos, isto é, o coeficiente negativo da interação indica que a desvantagem sobre a proficiência em Matemática para alunos que trabalham fora é acentuada quando a percentagem de estudantes com mães com ensino superior na escola aumenta. Considerando a estreita relação existente entre a escolaridade e a renda familiar, pode-se supor que os filhos de mães mais escolarizadas não precisem trabalhar, além de viverem em um ambiente com maior capital cultural familiar que levam a menores custos e dificuldades de aprendizado. Assim, esses alunos têm uma vantagem maior frente aos estudantes que precisam trabalhar fora, os quais, possivelmente, pertencem a famílias mais fragilizadas economicamente, tendo menos tempo e incentivos para se dedicarem aos estudos.

Em resumo, observou-se que os efeitos negativos do trabalho infantil podem ser quase anulados nas escolas com maior proporção de brancos, mas o custo individual de o aluno trabalhar fora, em escolas com maior proporção de mães com educação superior, é bastante prejudicial para o aprendizado. A dificuldade em se conhecer melhor essa relação dos *peer effects* reside no fato de que o questionário da Prova Brasil não oferece informações sobre o tipo de atividade e em que condições ocorre o trabalho desses estudantes. Desse modo, sugere-se que a escola desenvolva alguma forma de acompanhamento dos alunos que trabalham, na tentativa de evitar maiores custos associados à sua aprendizagem.

Por fim, comparando os resultados do modelo 4 com o modelo 5, percebe-se que a variância total foi reduzida, assim como o índice de correlação intraclasse. Da análise do *deviance*, pode-se concluir que o modelo 5 está mais ajustado aos dados, já que o valor de 49,78 encontrado é maior do que o valor crítico a 5% com 4 graus de liberdade (9,49).

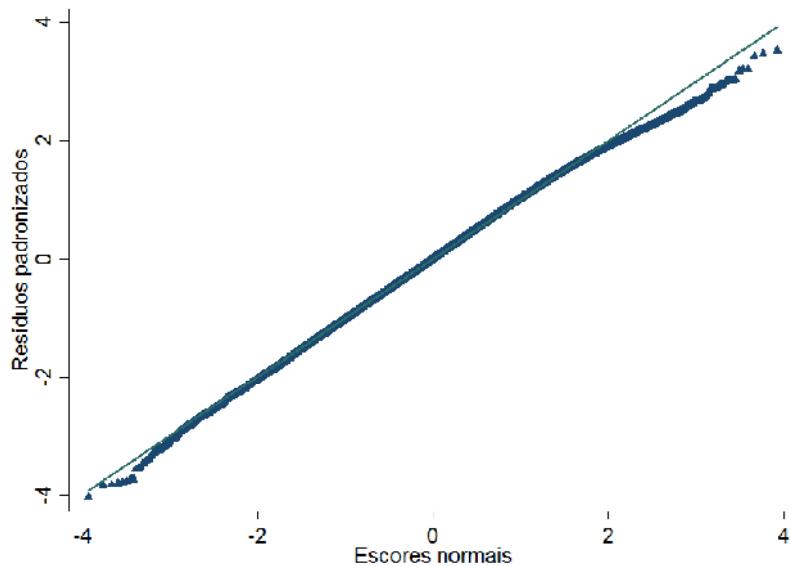
4.3.1 Avaliação de Diagnóstico do Modelo

É importante conferir se as pressuposições do modelo multinível são válidas para os dados do modelo final. Como em outros modelos de regressão, os resíduos desempenham um papel exploratório importante para a avaliação de diagnóstico em modelos multiníveis (SNIJDERS; BERKHOF, 2008). Empiricamente, pela análise gráfica dos resíduos, podem-se testar os seguintes pressupostos: (i) os erros são independentes e normalmente distribuídos com média zero no nível 1; (ii) os efeitos aleatórios são normalmente distribuídos com média zero e independentes entre grupos (LUKE, 2004).

Para o nível 1, a normalidade dos resíduos poderá ser verificada pela interpretação dos gráficos de dispersão dos resíduos contra a variável explanatória ou contra valores ajustados, de tal forma que se tenha uma variância constante para a distribuição utilizada. O padrão desses gráficos é uma distribuição dos resíduos em torno de zero, com amplitude constante, isto é, se o modelo estiver adequado, não deve haver nenhum padrão para os resíduos plotados em relação aos valores ajustados.

A normalidade dos resíduos para o modelo 5 é constatada no Gráfico 4, que mostra a relação entre os resíduos padronizados e os escores gerados aleatoriamente em uma distribuição normal. O referido gráfico mostra que os resíduos estão dispostos em uma linha reta ao longo dos valores, não exibindo grandes *outliers* (LAROS; MARCIANO, 2008a).

Gráfico 4: Verificação da Normalidade dos Resíduos com Escores Normais (Modelo 5)



Além disso, como verificado no Gráfico 5, os resíduos estão distribuídos uniformemente em torno da média. O gráfico de dispersão dos resíduos padronizados contra os valores ajustados é útil para avaliar, também, problemas de heterocedasticidade. Nesse sentido, os resíduos parecem estar centrados em zero e não apresentam ter problemas de heterocedasticidade. Além disso, a variabilidade dos resíduos parece ser, aproximadamente, a mesma para ambas as partes (LUKE, 2004).

Os mesmos gráficos, igualmente, podem ser utilizados para conferir os pressupostos para os efeitos aleatórios do nível 2 (LUKE, 2004). Como o modelo 5 tem três efeitos aleatórios (trabalha fora, NCS e intercepto), é preciso examinar cada um deles. O Gráfico 6 mostra como se comportam os resíduos para cada um desses efeitos aleatórios. No geral, grande parte dos valores está disposta linearmente, com alguns *outliers* nos extremos e nos valores mais baixos. Para verificar como os valores aleatórios estão centrados, utiliza-se uma matriz de dispersão, como disposto no Gráfico 7. Nesse gráfico, é possível determinar se os efeitos aleatórios são independentes uns dos outros entre os grupos (LUKE, 2004). Nota-se que os efeitos parecem estar centrados no zero, sugerindo que não existe uma forte dependência entre eles.

Gráfico 5: Dispersão dos Resíduos Padronizados e Valores Ajustados (Modelo 5)

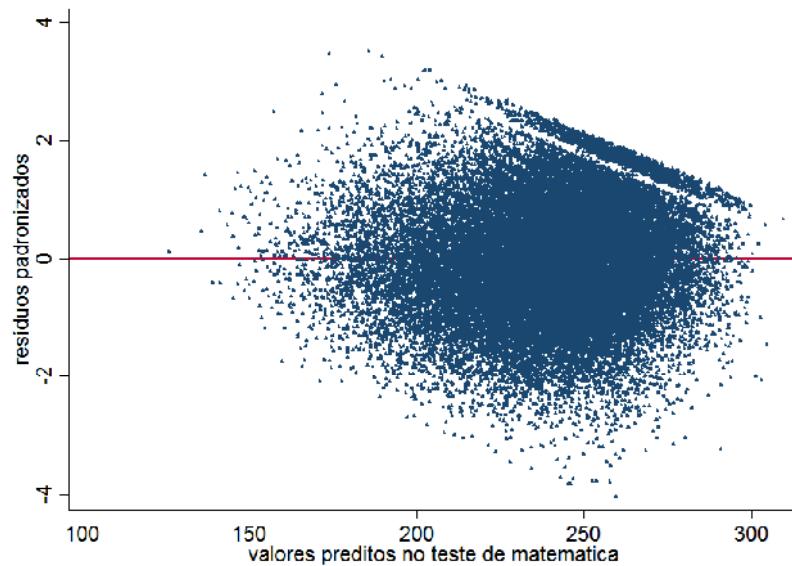


Gráfico 6: Normalidade dos Resíduos para Efeitos Aleatórios (Modelo 5)

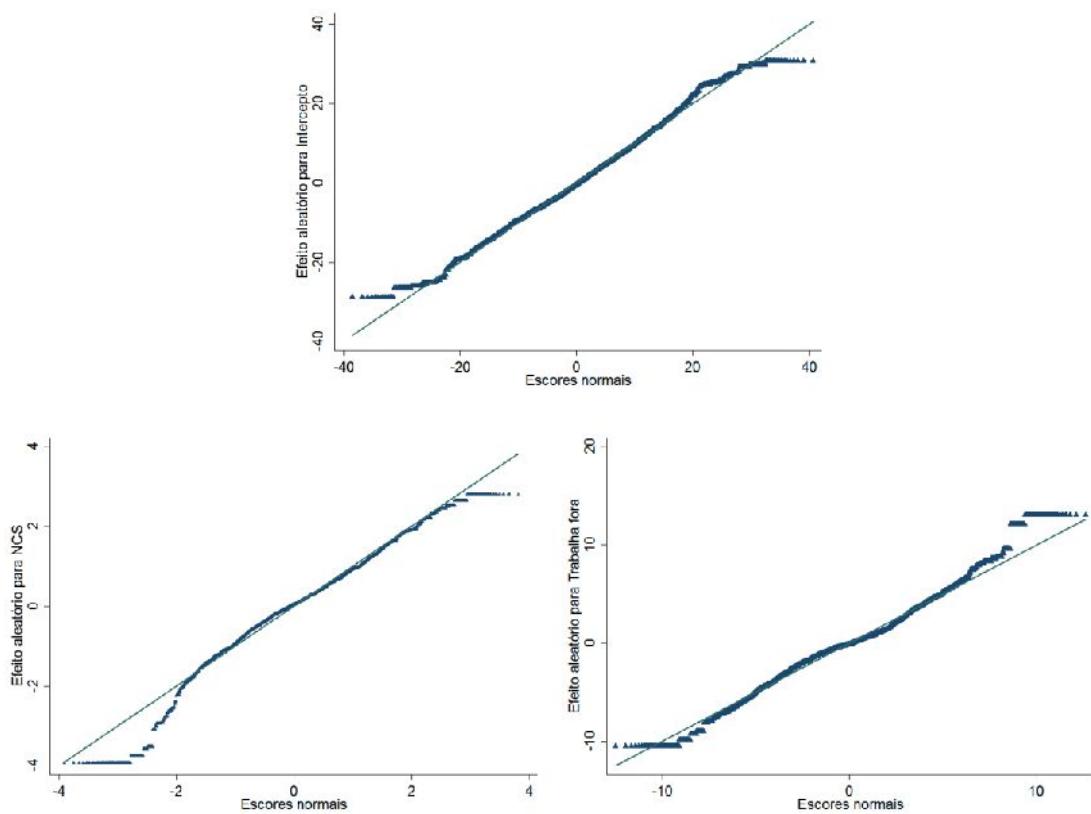
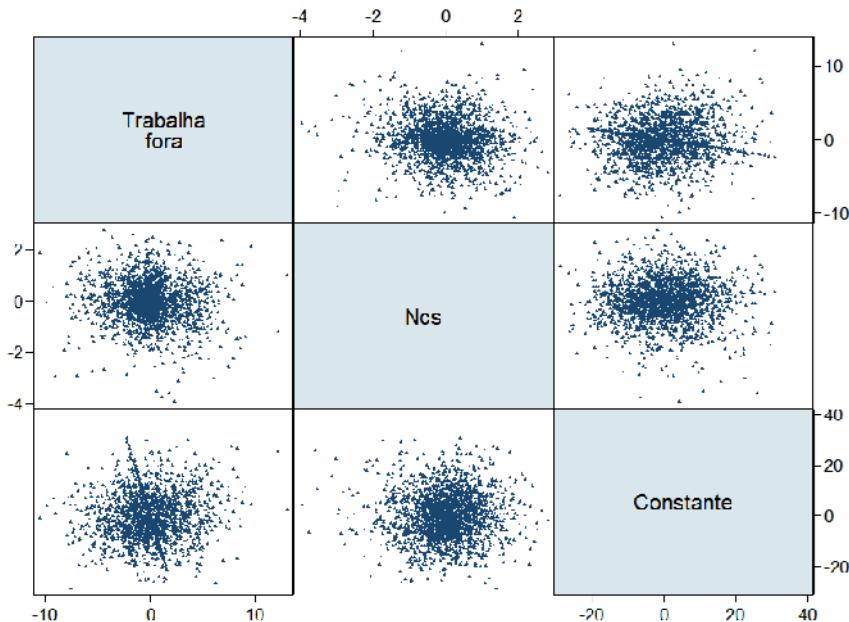


Gráfico 7: Matriz de Dispersão dos Efeitos Aleatórios E Estimados (Modelo 5)



A dependência linear das variáveis, a presença de valores atípicos e os efeitos potenciais de variáveis omitidas também podem ser verificados pela análise gráfica dos resíduos em relação às variáveis explicativas. Do mesmo modo, a hipótese de homocedasticidade pode ser verificada, analisando-se se os resíduos padronizados possuem uma distribuição normal com média 0 e variância σ^2 , mas isso somente deve ser feito depois de verificada a dependência linear nas variáveis explicativas dos efeitos fixos (SNIJDERS; BERKHOF, 2008).

4.3.2 Considerações Gerais sobre os Resultados do Modelo

É evidente que a escolha da escola não é algo aleatório. As condições socioeconômicas e culturais são determinantes na escolha da escola que o aluno irá frequentar (ANDRADE; LAROS, 2007). Os modelos hierárquicos têm se consolidado como um respeitável instrumento de análise dos resultados de pesquisas sobre avaliação educacional, uma vez que esses modelos são capazes de estimar a influência que variáveis contextuais exercem sobre o resultado obtido pelos estudantes em testes de proficiência.

O uso dessa modelagem para os dados da Prova Brasil 2013, de estudantes do 5º ano do ensino fundamental, no estado de Minas Gerais, foi justificado pelo modelo nulo que confirmou que há diferenças no desempenho médio das escolas na proficiência em Matemática, sendo 14% da variabilidade das notas explicada por diferenças entre as escolas e 86%, por diferenças entre os alunos dentro das escolas.

Como observado na estimação do modelo, todas as variáveis relativas aos alunos foram estatisticamente significativas. Desse modo, estudantes do sexo masculino, brancos, que moram com os pais, têm mães mais escolarizadas, desfrutam de algum tempo de lazer (televisão, internet ou jogos eletrônicos), fazem o dever de Matemática e possuem um NSE e um NCS acima da média apresentam melhor desempenho na prova de Matemática. De outro lado, há um efeito negativo sobre o desempenho daqueles alunos que estão em atraso escolar e realizam algum trabalho doméstico ou fora de casa.

Em relação às características das escolas, observou-se que estudantes que pertencem a escolas com melhores níveis de ICE (*proxy* para a qualidade da infraestrutura das escolas) e têm maior nível socioeconômico, matriculados em escolas com diretores que estão há mais de cinco anos nessa função, com maior proporção de professores brancos e professores com curso superior em Matemática (*proxy* para qualidade dos recursos humanos) e que desenvolvem projeto pedagógico, apresentam melhor desempenho na proficiência em Matemática. Do contrário, alunos de escolas municipais, com diretores eleitos (e não indicados ou nomeados por concurso) e que desenvolvem algum programa de redução de taxas de abandono, apresentam resultados piores na prova de Matemática. Além disso, a elevada variância da constante indica a existência de diferenças significativas na eficácia escolar em relação ao desempenho em Matemática, diferenças essas não explicadas pelas variáveis aqui consideradas.

Buscando entender melhor os fatores explicativos que levam à heterogeneidade observada entre as escolas, foram considerados também alguns coeficientes de inclinação aleatória no modelo, tendo sido identificados coeficientes aleatórios às variáveis capital social e trabalha fora, ambas no nível do aluno.

Da análise dos efeitos aleatórios para o aluno que trabalha fora, observou-se que há uma variabilidade entre essa variável e a proficiência em Matemática, havendo escolas em que o limite inferior para um aluno que trabalha fora corresponde a uma redução de 40,57 pontos na nota, e outras em que o limite superior do trabalho infantil equivale a uma redução da proficiência de 0,23 ponto. Já o efeito diferenciado observado entre o NCS do aluno e a proficiência foi de um efeito mínimo do NCS que reduz em 1,4 ponto a proficiência em Matemática, e um efeito máximo do NCS que eleva a nota em 16,3 pontos. Observa-se, portanto, a existência de um “*pay off*” diferenciado entre as escolas para alunos que trabalham fora ou possuem um NCS maior, em termos do desempenho que o estudante pode atingir.

Em busca de maior compreensão da heterogeneidade existente entre as escolas, procedeu-se ainda à inclusão no modelo de termos de interação formados por variáveis do nível do aluno

e variáveis contextuais da escola. Constatou-se que, tanto em escolas com maiores proporções de alunos em atraso escolar, bem como em escolas com elevado NSE, o efeito do retorno do NCS do aluno tende a ser menor sobre o desempenho estudantil, indicando a importância do ambiente escolar para o aprendizado. Além disso, há uma redução dos efeitos negativos sobre a proficiência dos alunos que trabalham, naquelas escolas com maior proporção de estudantes brancos, e um efeito negativo sobre a nota dos estudantes que trabalham fora, nas escolas com maior proporção de mães com ensino superior. Esse efeito distinto das externalidades do grupo pode estar relacionado às diferenças de condições do trabalho infantil.

Em termos de ajuste, os modelos, a cada passo, foram se tornando cada vez mais factíveis aos dados quando comparados às estimativas anteriores. Assim, conclui-se que o modelo 5 é o que melhor se ajusta aos dados.

Importante destacar que a literatura da economia da educação tem encontrado que os resultados educacionais são influenciados por uma série de fatores. A insistência da pesquisa acadêmica sobre o tema tem-se justificado na tentativa de suscitar a elaboração de políticas educacionais para melhorar a qualidade do ensino e os níveis de aprendizado.

A metodologia aqui proposta foi capaz de avaliar a importância relativa dos fatores determinantes do desempenho acadêmico relacionados às características do aluno, ao ambiente familiar e a características do meio escolar. Destaca-se a importância de variáveis ligadas à qualidade dos serviços educacionais (como a presença de professores com ensino superior em Matemática e a infraestrutura da escola), que são fatores passíveis de intervenção política. Assim, é preciso que ações sejam direcionadas no sentido de se conhecer melhor o contexto familiar e a trajetória escolar dos estudantes para que se possam reduzir as desigualdades existentes. A escolaridade dos pais e o NCS da família, por exemplo, são um indicador apropriado do valor que é atribuído à educação e como se dão as condições de aprendizado fora da escola. Por outro lado, uma elevada proporção na escola de alunos que trabalham fora parece diminuir o impacto positivo do capital social que o aluno traz de casa. Assim, as propostas de políticas para educação precisam considerar como compensar a carência social e educacional dos pais para melhorar o ambiente de aprendizado dentro das escolas, além de estimular o maior envolvimento dos pais com a vida escolar da criança e evitar a incidência de atraso escolar.

Enfim, os resultados apresentados indicam a importância de políticas públicas destinadas à melhoria da educação que: qualifiquem os professores para aperfeiçoar o ensino, atentando-se para os efeitos positivos do docente com licenciatura em Matemática; mitiguem os efeitos negativos do trabalho infantil para a trajetória escolar do aluno e de seus colegas; assegurem

uma boa infraestrutura da escola a fim de propiciar melhores condições de trabalho para os professores e aprendizado dos alunos. Além disso, tais políticas não podem ignorar a importância do papel da família para o desempenho acadêmico dos estudantes, devendo gerar uma maior integração dos pais e responsáveis com o ambiente escolar. Embora reconhecida a influência do contexto familiar, observou-se ainda que as características da escola e do meio escolar também impactam no resultado acadêmico dos alunos. Portanto, a formulação de políticas educacionais faz-se necessária, já que essas podem beneficiar os estudantes com melhoria da qualidade das escolas e do ensino ofertado, principalmente, para aqueles que estão em condições mais desiguais. Ademais, em alguns aspectos, o ambiente escolar pode reduzir a importância do ambiente familiar, no sentido de atenuar as diferenças educacionais advindas da trajetória escolar e da condição socioeconômica dos estudantes.

4.4 Uma Análise Comparativa ao Caso do Ceará

Nas últimas décadas, tem-se observado uma tendência de melhora no desempenho dos alunos cearenses nas avaliações padronizadas. Isso tem gerado intenso debate a respeito das medidas que vêm sendo adotadas por aquele estado para melhoria dos indicadores educacionais, principalmente, no que se refere à possibilidade de replicação dessas ações em outras localidades. Reconhece-se que, mesmo que haja uma ampla gama de estudos que discutem as condições do ensino nas escolas brasileiras, é preciso conhecer também como se encontra esse contexto no nível estadual, haja vista a heterogeneidade presente nas diversas regiões brasileiras.

Alguns fatos merecem destaque na trajetória da educação cearense. O Ceará está entre os primeiros estados que implantaram um sistema de avaliação da educação básica, na década de 1990, com o objetivo de definir políticas de avaliação e monitoramento do ensino público no estado, além de poder auxiliar os gestores na condução de políticas de incentivo e distribuição de recursos financeiros para escolas e municípios. Com base nessas avaliações, o estado realiza políticas de premiações para os profissionais envolvidos na educação (por meio de bônus salarial) e, também, para os melhores alunos do ensino médio, com a finalidade de elevar o padrão de ensino oferecido nas escolas (LIMA, 2012a).

O desenvolvimento de um sistema contínuo de avaliações em larga escala permitiu que o estado do Ceará tivesse à sua disposição um alto número de informações e, com isso, foi

possível o maior monitoramento das principais dificuldades e pontos de atuação, facilitando, assim, o desenvolvimento de políticas públicas focalizadas.

O estado também experimentou mudanças no processo de seleção de diretores e acelerou o processo de municipalização do ensino fundamental. Em relação à forma de escolha dos diretores, buscou-se reduzir a indicação política para cargos de direção nas escolas, na tentativa de escolher pessoas mais qualificadas para essa função. Ainda, a municipalização do ensino favoreceu a implantação do Fundo Nacional de Desenvolvimento do Ensino Fundamental e de Valorização do Magistério (FUNDEF), que depois se tornou um meio facilitador para o alcance das metas do estado para a educação, em virtude da vinculação de recursos às matrículas por rede de ensino (VIEIRA, 2007).

Além disso, na tentativa de melhorar a qualidade do ensino fundamental ofertado pelo estado, vinculou-se parte das receitas do ICMS aos resultados obtidos na educação, em moldes ao que já vinha sendo empregado nos estados de Minas Gerais (Lei Robin Hood) e Paraná (ICMS Ecológico). Uma parcela desse imposto (18%) foi direcionada para o ensino fundamental, em função dos resultados obtidos nos indicadores educacionais (LIMA, 2012b).

Nesse sentido, questiona-se quais fatores têm sido determinante para que as escolas garantam as metas e até que ponto tais ações são benéficas para o desenvolvimento de uma educação pública de qualidade. Assim, buscou-se comparar o caso do Ceará com o modelo apresentado anteriormente para Minas Gerais.

4.4.1 Estatísticas Descritivas

Nesta seção, são apresentadas as estatísticas descritivas das variáveis utilizadas no modelo hierárquico linear para o estado do Ceará. Após aplicados os filtros nos dados da Prova Brasil 2013 para os alunos do Ceará, do 5º ano do ensino fundamental, que preencheram a prova e o questionário, estavam em conformidade com as informações do censo escolar e que pertenciam à rede pública de ensino, o banco ficou com 99.273 observações. As informações para os professores dessa série totalizaram 5.672 e, para os diretores, foram 2.686, em um total de 2.684 escolas. Quanto às informações para esse estado, ocorreram problemas de ausência de dados no preenchimento dos questionários e, após a junção dos bancos de dados, obtiveram-se 580 observações de professores, 684 observações de diretores, 688 observações de escolas e 7.104 observações de alunos. Quando estimado o modelo, a amostra final ficou com 4.135 observações de alunos e 482 observações de professores, diretores e de escolas.

Em comparação ao caso mineiro, nota-se que tanto a amostra da Prova Brasil, como a amostra final do modelo de Minas Gerais, são maiores do que a amostra do Ceará. Conforme a Tabela 12, o percentual de alunos que restou na amostra final foi de 14,5% em Minas Gerais, enquanto que, para o Ceará, o valor encontrado foi de 4,2%. Já em relação às escolas, na amostra final do modelo para Minas, sobraram 39,9% das escolas da amostra da PB, e, no Ceará, somente 18%, indicando que o modelo para Minas Gerais seja mais representativo em comparação ao caso do Ceará.

Tabela 12: Comparação da Dimensão das Amostras entre Minas Gerais e Ceará

	Minas Gerais			Ceará		
	<i>Amostra PB</i>	<i>Amostra final</i>	<i>%</i>	<i>Amostra PB</i>	<i>Amostra final</i>	<i>%</i>
Alunos	227.296	32.987	14,5	99.273	4.135	4,2
Escolas	5.444	2.174	39,9	2.684	482	18,0

Fonte: Elaboração própria a partir de dados da Prova Brasil / INEP 2013

Assim como feito no modelo de Minas Gerais, realizou-se o teste de diferenças de médias e proporções entre a amostra da Prova Brasil e a amostra do modelo final para o estado do Ceará. Os resultados do teste sugerem que: os alunos da amostra da PB parecem ter médias piores em relação àqueles que compõem a amostra final; em comparação com a amostra final, na amostra da PB, há uma proporção maior de meninos, alunos em atraso, com mães menos escolarizadas, que trabalham fora de casa e uma proporção menor de estudantes que moram com os pais e realizam o dever de Matemática; há uma proporção maior de diretores eleitos e que realizam programas de combate ao abandono na amostra PB; há uma proporção maior de escolas municipais, na área urbana, e com média de 6,3 alunos a mais matriculados no 5º ano na amostra final. As diferenças quanto às características dos professores não se mostraram estatisticamente significantes. No entanto, as diferenças em termos de magnitudes entre as médias das variáveis relativas às características dos alunos, diretores e escolas foram muito baixas, ou seja, não apresentam diferenças economicamente relevantes. Assim, espera-se que essas diferenças não comprometam a validade externa dos resultados aqui expostos para o estado do Ceará.

A Tabela 13 apresenta as estatísticas descritivas das variáveis do nível do aluno para o Ceará. A média da proficiência em Matemática na Prova Brasil para o estado foi igual a 215,58 pontos, com desvio-padrão de 49,23 e coeficiente de variação de 23%. Esse valor está abaixo

da nota esperada para a 4^a série do ensino fundamental (a meta do TPE estabelece o desempenho mínimo de 225 pontos), indicando a variabilidade encontrada que as diferenças em termos de nota ainda são preocupantes.

Tabela 13: Estatísticas Descritivas – Alunos, 5º Ano, Ceará

Variável	Nº obs.	Média	Mínimo	Máximo	Desvio Padrão	CV
Proficiência em Matemática	4.135	215,58	80,56	341,25	49,23	0,23
Homem	4.135	0,49	0,00	1,00	-	-
Branco	4.135	0,20	0,00	1,00	-	-
Atraso	4.135	0,21	0,00	1,00	-	-
Mora com mãe e pai	4.135	0,67	0,00	1,00	-	-
<i>Escolaridade da mãe (categoria base = mãe não estudou)</i>						
Até 4º série do fundamental	4.135	0,23	0,00	1,00	-	-
Ensino fundamental completo	4.135	0,18	0,00	1,00	-	-
Ensino médio completo	4.135	0,17	0,00	1,00	-	-
Superior completo	4.135	0,12	0,00	1,00	-	-
Trabalho doméstico	4.135	0,76	0,00	1,00	-	-
Trabalho fora	4.135	0,13	0,00	1,00	-	-
Tempo de lazer	4.135	2,03	0,00	4,00	1,27	0,62
Dever de Matemática	4.135	0,78	0,00	1,00	-	-
Nível Socioeconômico	4.135	0,02	-3,09	3,62	0,81	51,42
Nível de Capital Social	4.135	0,09	-0,22	2,47	0,55	6,10

Fonte: Elaboração própria a partir de dados da Prova Brasil / INEP 2013

Entre os estudantes, 49% são meninos, 20% se declaram brancos, em média, 21% dos estudantes estão em atraso escolar e 67% moram com os pais. Quanto à escolaridade da mãe, na média, têm-se que: 23% têm até 4^a série do fundamental; 18% têm ensino fundamental completo; 17% têm ensino médio completo; e 12% têm superior completo.

Grande parte dos estudantes realizam tarefas domésticas (76%), enquanto que 13% trabalham fora de casa. Em média, os estudantes gastam 2,03 horas em lazer na internet, jogos eletrônicos ou televisão, com desvio-padrão de 1,27. Uma parcela significativa dos estudantes declarou fazer o dever de casa de Matemática (78%).

A média para o NSE do aluno foi de 0,02, com um elevado coeficiente de variação de 51,42, indicando uma alta variabilidade dos dados em relação à média e uma disparidade na condição econômica dos estudantes. A média do NCS foi de 0,09, com coeficiente de variação de 6,10, indicando que também há diferenças entre o nível de participação e atenção dos pais na vida escolar dos filhos.

Ao confrontar essas informações com aquelas referentes a Minas Gerais, constata-se que, na amostra mineira, há uma maior proporção de: i) estudantes brancos (30%); ii) alunos que moram com os pais (71%); iii) mães com ensino médio (24%) e superior (20%) completo; iv) alunos que fazem tarefas domésticas (80%) e dever de casa (85%). Ainda, constata-se uma proporção menor de estudantes em atraso escolar (11%). Os estudantes cearenses têm uma média de proficiência em Matemática e um NSE menor do que os estudantes mineiros. No entanto, a média do NCS do Ceará (0,09) foi um pouco maior que a de Minas (-0,04), sugerindo uma maior atenção dos pais no primeiro caso.

Na Tabela 14, estão as estatísticas descritivas para as variáveis de professores, diretores e escolas. Observa-se que, aproximadamente, 99% das escolas pertencem à esfera municipal. Já na amostra final, observaram-se somente 6 escolas estaduais. Quanto à localização das escolas, observou-se que 74% estão situadas na zona urbana. Com relação ao número de estudantes matriculados na série, tem-se que, em média, as escolas possuem 58 alunos matriculados no 5º ano do ensino fundamental, com desvio padrão de 37%.

Tabela 14: Estatísticas Descritivas – Professor, Diretor e Escola, 5º Ano, Ceará

Nome da variável	Nº obs.	Média	Mínimo	Máximo	Desvio Padrão	CV
Escola						
Municipal	482	0,99	0	1	-	-
Urbana	482	0,74	0	1	-	-
Matrículas 5º ano	482	57,99	20,00	240,00	36,95	0,64
IPE	482	-0,05	-0,81	2,41	0,74	-13,98
ICE	482	0,09	-1,16	2,03	0,85	9,48
NSE da escola	482	-0,01	-1,59	2,25	0,45	-66,36
Diretor						
Assumi a direção via eleição	482	0,02	0	1	-	-
Tem mais de 5 anos de experiência	482	0,35	0	1	-	-
Está há mais de 5 anos na direção da escola	482	0,18	0	1	-	-
Há programa de redução de abandono	482	0,85	0	1	-	-
Professor						
Branco	482	0,26	0	1	-	-
Tem ensino superior	482	0,91	0	1	-	-
Tem superior em Matemática	482	0,04	0	1	-	-
Tem 5 anos ou mais de experiência	482	0,84	0	1	-	-
Escola tem projeto pedagógico	482	0,81	0	1	-	-

Fonte: Elaboração própria a partir de dados da Prova Brasil / INEP 2013

Em relação à existência de equipamentos nas escolas cearenses, observa-se que a sua média foi de -0,05 e o coeficiente de variação de 13,99, indicando que existe uma alta heterogeneidade entre as escolas em termos de posse de equipamentos. Já a média para o ICE

foi de 0,09, com desvio-padrão de 0,85, apontando para uma alta dispersão dos dados. A média do NSE da escola foi de -0,01, com coeficiente de variação de 66,36, evidenciando que as escolas diferem enormemente em termos do nível socioeconômico de seus alunos.

No que se refere às características dos diretores, observou-se que: apenas 2% assumiram a direção via eleição; 35% têm mais de 5 anos de experiência; 18% estão há mais de 5 anos no cargo de diretor na escola avaliada; e 85% declararam desenvolver algum programa de redução de taxas de abandono.

Dos professores da amostra, tem-se: 26% se declararam brancos; grande parte possui curso superior (91%), sendo apenas 4% em Matemática; 84% têm cinco anos ou mais de experiência; e 81% afirmaram que, na escola em que trabalham, são desenvolvidos projetos pedagógicos.

Comparando as informações do Ceará com as estatísticas descritivas de Minas Gerais apresentadas anteriormente, destaca-se que, em Minas Gerais, há: i) menor municipalização das escolas (63%); ii) maior proporção de escolas urbanas (93%); iii) uma média maior de alunos matriculados no 5º ano (63,14); iv) uma média melhor do IPE (0,06) e do NSE (0,22) da escola, mas um ICE (-0,07) pior; v) elevada proporção de diretores eleitos (49%) e que estão há mais tempo no cargo de direção (26%); vi) menor proporção de escolas que realizam programas de combate ao abandono (73%); vii) maior proporção de professores brancos (43%) e com ensino superior (94%).

4.4.2 Resultados do Modelo

Os resultados da estimação do modelo nulo para o caso do Ceará são apresentados na Tabela 15. A estimativa da média global da proficiência em Matemática é de 216,62 pontos (19,47 pontos abaixo da média mineira). Já a variância do intercepto foi significativa, mostrando que: i) essa média difere entre as escolas que compõem a amostra; ii) a presença de diferenças na proficiência média em torno da média do estado justifica o uso do modelo hierárquico.

A variância total da proficiência é dividida em duas partes, sendo uma relativa aos estudantes (1.906,53) e outra referente ao nível da escola (556,89). A partir desses valores, encontrou-se o índice de correlação intraclasse (0,2263), o qual indica que 22,63% da variabilidade da nota ocorre entre as escolas e 77,37% da variância total da proficiência ocorre devido a diferenças entre os estudantes dentro das escolas.

Tabela 15: Modelo Nulo (Ceará)

VARIÁVEIS EXPLICATIVAS		MODELO NULO (M1)		
Efeito Fixo		Estimativa	Erro Padrão	P-valor
Intercepto (γ_{00})		216,62	1,37	0.000
Efeito Aleatório		Variância	Erro Padrão	Razão-t
Variância do intercepto (σ_{u0}^2)		556,89	57,2	9,74
Variância residual (σ_e^2)		1.906,53	44,47	42,87
Correlação intraclasse		0,2261		
Deviance M1		43.495,8		
Número de parâmetros		3		
Número de observações		4.135		
Número de grupos		482		

Conforme a revisão da literatura empírica feita por Ferrão e Fernandes (2003), a parcela de contribuição da escola para o aprendizado do estudante, mensurada pelo ICC, no caso brasileiro, situa-se em cerca de 30% para o modelo vazio. Mas, quando o modelo controla a condição econômica do estudante, esse valor se reduz consideravelmente. Segundo os autores, evidências empíricas mostraram que, quando controlado o NSE dos alunos, a região mais carente do país, o Nordeste (19%), teria um efeito-escola maior quando comparado à região mais desenvolvida, o Sul (7%).

No presente caso, em comparação a Minas Gerais (ICC = 0,1377), o valor encontrado para ICC do Ceará foi maior, sugerindo que as escolas cearenses são mais heterogêneas do que as escolas mineiras que compõem a amostra. Em outras palavras, os fatores escolares contribuem mais para o desempenho dos alunos cearenses do que para os estudantes mineiros.

Na segunda coluna da Tabela 16, são apresentados os resultados do modelo 2, que inclui somente variáveis do nível 1 (aluno), lembrando que as variáveis foram centralizadas, portanto, o valor estimado para o intercepto de 216,08 representa a média do desempenho em Matemática na escola, quando todas as variáveis explicativas são iguais às médias amostrais.

Diferentemente do caso mineiro, em que todas as variáveis do nível do aluno haviam sido significativas no modelo 2, para o caso do Ceará, três variáveis não apresentaram significância estatística (morar com os pais; mãe com superior completo; e NSE). Um detalhe observado nas estatísticas descritivas foi o elevado valor encontrado para o coeficiente de variação do NSE (52,42), revelando que há uma alta heterogeneidade entre os estudantes em termos da sua condição socioeconômica.

Tabela 16: Modelos com Variáveis Explicativas Aluno e Escola (Ceará)

VARIÁVEIS EXPLICATIVAS	MODELO 2			MODELO 3			
	Efeito Fixo	Estimativa	E. P.	P-valor	Estimativa	E. P.	P-valor
Intercepto (γ_{00})		216,08	1,26	0,000	214,40	1,28	0,000
Aluno							
Homem (γ_{10})		7,16	1,38	0,006	7,10	1,38	0,000
Branco (γ_{20})		-4,61	1,68	0,000	-4,78	1,67	0,004
Atraso (γ_{30})		-14,64	1,74	0,000	-14,44	1,74	0,000
Mora com mãe e pai (γ_{40})		0,55	1,44	0,704	0,51	1,44	0,723
<i>Escolaridade da mãe: (categoria base = mãe não estudou)</i>							
- Até 4º série do fundamental (γ_{50})		6,54	1,88	0,001	6,60	1,88	0,000
- Ensino fundamental completo (γ_{60})		7,29	2,12	0,001	7,46	2,11	0,000
- Ensino médio completo (γ_{70})		12,89	2,23	0,000	12,97	2,23	0,000
- Superior completo (γ_{80})		-0,03	2,65	0,991	0,20	2,65	0,938
Trabalho doméstico (γ_{90})		-4,33	1,60	0,007	-4,43	1,60	0,006
Trabalho fora de casa (γ_{100})		-18,17	2,05	0,000	-18,29	2,05	0,000
Tempo de lazer (γ_{110})		2,49	0,54	0,000	2,57	0,54	0,000
Dever de Matemática (γ_{120})		13,31	1,69	0,000	13,24	1,69	0,000
NSE (γ_{130})		1,68	0,99	0,090	2,01	1,02	0,048
NCS (γ_{140})		-8,94	1,25	0,000	-8,91	1,25	0,000
Escola							
Municipal (γ_{01})		-	-	-	5,52	10,99	0,615
Urbana (γ_{02})		-	-	-	3,45	3,45	0,317
Matrículas 5º ano (γ_{03})		-	-	-	-0,11	0,12	0,346
Matrículas ² (γ_{04})		-	-	-	0,00	0,00	0,971
IPE (γ_{05})		-	-	-	-1,14	1,72	0,507
ICE (γ_{06})		-	-	-	0,06	1,46	0,965
NSE da escola (γ_{07})		-	-	-	-2,18	3,31	0,510
Diretor							
Assumiu a direção via eleição (γ_{08})		-	-	-	-8,76	8,56	0,306
Tem mais de 5 anos de experiência (γ_{09})		-	-	-	-3,58	3,26	0,272
Está há mais de 5 anos na direção da escola (γ_{010})		-	-	-	11,50	4,03	0,004
Há programa de redução de abandono (γ_{011})		-	-	-	2,55	3,45	0,460
Professor							
Branco (γ_{012})		-	-	-	0,69	2,64	0,793
Tem ensino superior (γ_{013})		-	-	-	-1,97	4,33	0,649
Tem superior em Matemática (γ_{014})		-	-	-	6,00	5,91	0,310
Tem 5 anos ou mais de experiência (γ_{015})		-	-	-	5,10	3,28	0,120
Escola tem projeto pedagógico (γ_{016})		-	-	-	5,94	2,99	0,047
Efeito Aleatório	Variância	E. P.	Razão-t	Variância	E. P.	Razão-t	
Intercepto (σ_{u0}^2)	452,25	47,85	9,45	397,16	44,13	9,00	
Residual (σ_e^2)	1.712,65	39,93	42,89	1.713,55	39,96	42,88	
Correlação intraclass	0,2089			0,1881			
Deviance	43.022,82			42.988,31			
Número de parâmetros	17			33			
Deviance M1 – Deviance M2	472,98			-			
Diferença de parâmetros	14			-			
Deviance M2 – Deviance M3	-			34,51			
Diferença de parâmetros	-			16			

Ainda, os resultados mostram que os meninos têm um desempenho melhor na prova de Matemática do que as meninas. Em média, a proficiência deles é aumentada em 7,16 pontos. Esse resultado está em conformidade com a literatura e também foi encontrado para o modelo de Minas.

Um resultado contrário ao esperado foi encontrado para a variável que indica a cor do estudante. De acordo com o modelo, alunos que se declararam brancos têm seu desempenho reduzido em 4,61 pontos em comparação aos alunos não brancos. Isso pode estar relacionado a características dos estudantes ou a efeitos não controlados na amostra.

Assim como acontece em Minas, os alunos que estão em atraso escolar pontuam, em média, 14,64 pontos a menos do que os alunos que estão em idade adequada para a série. Diferentemente, a maior escolaridade da mãe influencia positivamente o desempenho em Matemática dos estudantes do 5º ano do ensino fundamental no Ceará. Quando comparado aos alunos cujas mães não estudaram ou não completaram o primeiro ciclo do ensino fundamental, a nota em Matemática sobe, em média: i) 6,54 pontos quando as mães completaram a 4ª série; ii) 7,29 pontos quando mães têm ensino fundamental completo; iii) e 12,89 pontos se as mães concluíram o ensino médio.

Conforme esperado, alunos que exercem algum trabalho, seja ele doméstico ou fora de casa, têm seu desempenho penalizado. Nesse sentido, o trabalho doméstico reduz a proficiência em Matemática, em média, 4,33 pontos, em comparação aos estudantes que não trabalham. Já o trabalho fora de casa é mais preocupante, visto que o aluno que exerce alguma atividade fora de casa (remunerada ou não) tem sua nota reduzida, em média, em 18,17 pontos, indicando que é preciso pensar em ações que mitiguem os efeitos negativos do trabalho infantil na vida escolar.

O tempo que o estudante dispende com televisão, videogame ou computador impacta positivamente no desempenho acadêmico. Os resultados indicam que, para cada hora adicional, a proficiência aumenta, em média, 2,49 pontos. A realização de dever de casa também é importante para um melhor desempenho, visto a estreita relação existente entre a realização de dever e o comprometimento do estudante com a escola. Para o modelo do Ceará, encontrou-se que estudantes que afirmaram realizar o dever de Matemática sempre ou quase sempre têm um aumento na nota de 13,31 pontos, em média, quando comparado aos alunos que não fazem as tarefas escolares.

Outro resultado contrário ao esperado foi para a variável que representa o nível de capital social da família. Os resultados do modelo sugerem que alunos com NCS acima da média têm a sua nota reduzida em 8,94 pontos, em termos médios. Acredita-se que essa variável esteja

capturando algum efeito que não foi controlado no modelo. Além disso, ao analisar as estatísticas descritivas dessa variável, vê-se que o NCS tem uma média de -0,04 e que varia entre o mínimo de -2,76 ao valor máximo de 0,32, isto é, os alunos cearenses contam com um baixo nível de participação e atenção dos pais em relação aos seus estudos.

No modelo 2, os componentes condicionais de variância reduziram-se em comparação aos componentes de variância encontrados no modelo nulo. Isso era esperado, uma vez que a inclusão de variáveis do nível 1 explicaram parte da variação do desempenho médio na prova de Matemática, entre as escolas. Com isso, o ICC diminuiu, aproximadamente, de 0,23 para 0,21.

Do mesmo modo, a estatística *deviance* reduziu e é estatisticamente significante, dado que a diferença entre o *deviance* do modelo nulo e do modelo condicional foi igual a 472,98 com 14 graus de liberdade. Isso é maior que o valor crítico ao nível de 5% (23,68), sugerindo que o modelo 2 é o que melhor se ajusta aos dados, uma vez que as variáveis incluídas na regressão conseguiram explicar parte da desigualdade nas notas. O R^2_1 avaliado entre o modelo vazio e o modelo 2 foi de 12,12, ou seja, a proporção de redução nos erros preditos da proficiência em Matemática, após a inclusão de variáveis no nível 1, é de 12,12% quando comparado ao modelo nulo.

A terceira coluna da Tabela 16 apresenta os resultados do modelo 3 que inclui as variáveis relativas ao nível 2 (escola). A parte relativa aos efeitos fixos das variáveis do nível 1 foi muito semelhante ao modelo 2, exceto para o NSE, que passou a ser significativo para explicar a proficiência do aluno em Matemática. Assim, espera-se que estudantes que possuam um NSE acima da média tenham a nota aumentada em dois pontos, em média, se comparados aos alunos que tenham condições socioeconômicas inferiores à média (índicio de que o modelo 2 sofre de viés da variável omitida).

A maioria das variáveis do nível da escola não tiveram significância estatística, resultado diferente daquele encontrado no modelo para Minas Gerais. Somente explicam o desempenho dos alunos dois fatores: diretores que estão há mais de cinco anos na direção da escola e o desenvolvimento de projeto pedagógico. Talvez as características escolares que importam para o desempenho dos estudantes cearenses não sejam observáveis nos dados.

Os resultados do modelo 3 indicam que alunos pertencentes a escolas com diretores que estão à frente da direção da escola há cinco anos ou mais têm sua nota elevada, em termos médios, em 11,5 pontos, quando comparados a estudantes de escolas cujos diretores estão nos

anos iniciais do primeiro mandato. Isso sugere que o maior tempo do diretor na gestão da escola leva a resultados positivos no desempenho.

A existência de projeto pedagógico também influencia positivamente o resultado obtido pelos estudantes no teste de Matemática. Nas escolas em que há desenvolvimento de projeto pedagógico, espera-se que a proficiência do aluno seja, em média, maior em 5,94 pontos do que nas escolas em que não há plano pedagógico.

Analizando-se os efeitos aleatórios do modelo 3, observa-se que a inclusão das variáveis do nível 2 reduziu a variância do intercepto de 452,25 para 397,16, como era esperado. A variância residual foi de 1.713,55. Assim, a correlação intraclasse caiu para cerca de 0,19. Houve redução também do *deviance*, sendo a estatística *deviance* calculada de 34,51 (16 graus de liberdade), o que é superior ao valor crítico de 23,30 a 5% de significância. Desse modo, rejeita-se a hipótese nula de que o grau de ajuste dos modelos seja igual. O R^2_1 calculado entre o modelo vazio e o modelo 3 indica que houve uma redução de 14,32% na variância das notas, ao se incluírem variáveis relativas à escola e aos alunos.

O próximo passo da análise foi a inclusão de efeitos aleatórios, na tentativa de entender melhor as diferenças existentes entre as escolas no desempenho na prova de Matemática. Foram testadas diversas especificações, com diferentes variáveis no nível do aluno como efeito aleatório (trabalha fora, dever, NCS, escolaridade da mãe, NSE, atraso). Diferentemente do modelo para Minas Gerais, que teve os coeficientes de inclinação associados ao NCS e à variável trabalha fora como aleatórios, no caso do Ceará, somente a variável atraso teve coeficiente aleatório significativo.

Os resultados para o modelo 4 e 5 são apresentados na Tabela 17. As estimativas dos efeitos fixos do nível da escola e do aluno nesses modelos foram muito similares àqueles encontrados nos modelos 2 e 3.

A estimativa do coeficiente aleatório para a variável atraso foi estatisticamente significativa, segundo o teste de Wald. Como a variância é sempre positiva, esse valor do teste tem que ser comparado com o valor crítico de uma distribuição normal em um teste unilateral à direita (a 5% de significância, o valor crítico, nesse caso, é 1,645). Como o valor do teste para o atraso foi 1,82, que é maior que o valor crítico, rejeita-se a hipótese nula de que o componente aleatório seja nulo.

Tabela 17: Modelos com Coeficientes Aleatórios (Ceará)

VARIÁVEIS EXPLICATIVAS		MODELO 4			MODELO 5		
Efeito Fixo		Estimativa	E. P.	P-valor	Estimativa	E. P.	P-valor
Intercepto (γ_{00})		214,38	1,28	0.000	214,33	1,29	0.000
Aluno							
Homem (γ_{10})		7,19	1,38	0.000	6,98	1,38	0.000
Branco (γ_{20})		-4,71	1,67	0.005	-4,65	2	0.005
Atraso (γ_{30})		-14,63	1,91	0.000	-14,79	1,88	0.000
Mora com mãe e pai (γ_{40})		0,56	1,44	0.696	0,53	1,44	0.714
<i>Escolaridade da mãe: (categoria base = mãe não estudou)</i>							
- Até 4º série do fundamental (γ_{50})		6,80	1,88	0.000	6,84	1,88	0.000
- Ensino fundamental completo (γ_{60})		7,63	2,11	0.000	7,53	2,11	0.000
- Ensino médio completo (γ_{70})		13,07	2,22	0.000	13,04	2,22	0.000
- Superior completo (γ_{80})		0,33	2,64	0.900	0,31	2,64	0.905
Trabalho doméstico (γ_{90})		-4,30	1,59	0.007	-4,39	1,59	0.006
Trabalho fora de casa (γ_{100})		-18,32	2,05	0.000	-18,14	2,05	0.000
Tempo de lazer (γ_{110})		2,56	0,54	0.000	2,57	0,54	0.000
Dever de Matemática (γ_{120})		13,33	1,69	0.000	13,33	1,69	0.000
NSE (γ_{130})		2,04	1,01	0.044	2,07	1,01	0.041
NCS (γ_{140})		-8,89	1,25	0.000	-8,91	1,25	0.000
Escola							
Municipal (γ_{01})		5,42	10,98	0.622	4,27	11,02	0.699
Urbana (γ_{02})		3,39	3,45	0.325	3,30	3,46	0.339
Matrículas 5º ano (γ_{03})		-0,11	0,12	0.368	-0,10	0,12	0.392
Matrículas ² (γ_{04})		0,00	0,00	0.933	0,00	0,00	0.903
IPE (γ_{05})		-1,08	1,72	0.531	-1,09	1,72	0.527
ICE (γ_{06})		0,11	1,45	0.937	0,18	1,46	0.901
NSE da escola (γ_{07})		-2,25	3,31	0.497	-2,32	3,31	0.483
Diretor							
Assumiu a direção via eleição (γ_{08})		-8,58	8,55	0.316	-9,24	8,58	0.281
Tem mais de 5 anos de experiência (γ_{09})		-3,52	3,26	0.280	-3,36	3,27	0.305
Está a mais de 5 anos na direção da escola (γ_{010})		11,27	4,03	0.005	11,17	4,04	0.006
Há programa de redução de abandono (γ_{011})		2,65	3,45	0.442	2,58	3,46	0.456
Professor							
Branco (γ_{012})		0,64	2,64	0.808	0,65	2,64	0.807
Tem ensino superior (γ_{013})		-2,12	4,33	0.624	-1,88	4,34	0.664
Tem superior em Matemática (γ_{014})		6,15	5,91	0.298	6,07	5,92	0.305
Tem 5 anos ou mais de experiência (γ_{015})		5,15	3,28	0.117	4,95	3,29	0.133
Escola tem projeto pedagógico (γ_{016})		5,90	2,99	0.048	6,06	3,00	0.043
Interação							
Proporção de meninos na escola * atraso		-	-	-	24,11	11,30	0.033
Efeito Aleatório		Variância	E. P.	Razão-t	Variância	E. P.	Razão-t
Atraso		144,31	79,33	1,82	122,08	77,12	1,58
Intercepto		395,96	44,13	8,97	399,48	44,40	9,00
Residual		1.691,45	40,69	41,57	1.691,29	40,69	41,57
Correlação intraclasse		0,1896			0,1910		
<i>Deviance</i>		42.983,72			42.979,26		
Número de parâmetros		34			35		
Deviance M3 – Deviance M4		4,6			4,46		
Diferença de parâmetros		1			1		

Nesse sentido, pode-se inferir que há um “custo” ou um “preço” diferente entre as escolas, em termos de desempenho em Matemática, para o fato de o aluno estar ou não em atraso escolar. Analisando-se o intervalo de confiança, tem-se que, em média, no limite inferior, o atraso reduz a proficiência em 38,2 pontos, se comparado a um aluno que está com a idade esperada para a série, e, no limite superior, a proficiência é elevada em 8,91 pontos. Observa-se que há uma diferença expressiva em termos de magnitude e impacto das inclinações para os alunos que estão em atraso escolar, revelando quão diferentes podem ser as escolas nesse aspecto.

As escolas diferem também em termos de sua média de desempenho. No caso do intercepto, o seu intervalo de confiança para o modelo 4 mostra que, no limite inferior, têm-se escolas com a proficiência média de 175,38 pontos, e, no limite superior, a proficiência média é de 253,38 pontos.

Na análise dos efeitos aleatórios, observa-se que houve uma redução na variância residual e na variância do intercepto no modelo 4, além de uma queda no *deviance*. A diferença entre os *deviances* do modelo 4 e 3 foi de 4,6, superior ao valor crítico a 5%, com um grau de liberdade (3,84). Rejeita-se, portanto, a hipótese nula de que os graus de ajuste dos modelos sejam iguais.

Reconhecida a presença de um alto grau de diferenciação entre escolas quanto ao retorno (em termos das notas em Matemática) do atraso escolar, o último passo foi a tentativa de incluir interações no modelo para tentar explicar essas desigualdades. Novamente, vários testes foram realizados na busca de possíveis interações entre variáveis dos níveis do aluno e da escola, tendo sido estatisticamente significativa apenas a interação entre atraso no nível do aluno com a proporção de meninos na escola.

Os resultados do modelo 5 são reportados na terceira coluna da Tabela 17. Com a inclusão da interação, é preciso verificar se os componentes aleatórios do modelo sofreram alguma alteração. Comparando-se os resultados do modelo 5 com o modelo 4, nota-se que o componente aleatório da variável atraso reduziu de 144,31 para 122,08. Com isso, o intervalo de confiança se alterou da seguinte forma: no limite inferior, o aluno atrasado tem a proficiência reduzida em 36,44 pontos, em comparação ao estudante em idade regular; já no limite superior, a nota em Matemática se eleva em 6,87 pontos. Para o intercepto, a alteração do intervalo de confiança não foi expressiva em termos de magnitude (limite inferior = 175,16; limite superior = 253,51).

A inclusão da interação aponta que parte das diferenças no coeficiente estimado para o atraso do aluno entre escolas está relacionado com a proporção de meninos na escola. O coeficiente estimado para a interação (*Proporção de meninos*atraso*) sinaliza que, para um

aumento de 1 ponto percentual acima da média na proporção de meninos na escola, a proficiência do aluno em atraso escolar aumenta, em média, em 9,32 pontos. Isso significa dizer que o custo individual de estar em atraso, em termos de aprendizado, é reduzido, visto que o efeito negativo do atraso (-14,79) é compensado nas escolas com maior proporção de meninos (24,11), ressaltando-se que os meninos tendem a ter um desempenho superior ao das meninas na prova de Matemática. Assim, nas escolas com maior proporção de meninos, os efeitos negativos do atraso seriam mitigados pelos efeitos positivos da convivência com alunos que têm melhor desempenho nos testes de Matemática.

Enfim, comparando-se os resultados do modelo 5 com o modelo 4, nota-se que a estatística *deviance* foi igual a 4,46, superior ao valor crítico a 5%, com um grau de liberdade (3,84). Pode-se concluir, portanto, que o modelo 5 está mais ajustado aos dados.

4.4.3 Considerações Gerais Sobre os Resultados do Modelo

Vários fatores foram encontrados como fortes influências sobre a aprendizagem dos alunos cearenses do 5º ano do ensino fundamental. Esses incluem as influências de características pessoais e familiares (gênero, cor, escolaridade da mãe, condição em relação ao trabalho, cumprimento do dever, dentre outros), além de condições escolares (tempo que o diretor está no cargo e a existência de projeto pedagógico). Entretanto, grande parte da heterogeneidade escolar que explica o diferencial de nota (ICC =19%) não pode ser explicada pelo modelo. Dentre esses fatores, poucos são passíveis de mudanças no curto prazo.

Assim como no caso mineiro, no Ceará, os resultados da proficiência em Matemática nos anos iniciais do ensino fundamental são fortemente influenciados pelas variáveis do aluno. No entanto, a comparação entre os modelos mostra que, em Minas Gerais, as diferenças escolares em termos de notas são menores (ICC modelo nulo = 0,14 para ICC modelo 5 = 0,08) e, para o Ceará, o percentual residual da heterogeneidade escolar que o modelo não conseguiu explicar é maior (ICC modelo nulo = 0,23 para ICC modelo 5 = 0,19).

Das diversas variáveis escolares incluídas no modelo, somente duas foram significativas, quais sejam: tempo que o diretor está no cargo e a existência de projeto pedagógico. O modelo não conseguiu captar outros fatores escolares que expliquem o fato de algumas escolas terem um desempenho melhor do que outras. Observando-se os dados, tem-se que grande parte da amostra final do modelo contém escolas municipais, o que dificulta a diferenciação da gestão escolar nesse caso. Outros fatores, que não entraram no modelo e que talvez sejam específicos ao caso cearense, podem explicar a heterogeneidade escolar, como, por exemplo, a presença de

políticas de bonificação aos professores, a maior destinação de recursos que está atrelada a metas em avaliações educacionais feitas pelo próprio estado e as práticas pedagógicas adotadas (VIEIRA, 2007; LIMA, 2012a; LIMA, 2012b)

No caso cearense, não se verificou a existência de efeitos diferenciados em termos do nível de capital social e da condição de trabalho dos estudantes, como encontrado no modelo para Minas Gerais. Mas, nas escolas do Ceará, há diferenças em relação ao atraso escolar dos alunos, mostrando que o impacto desse atraso no desempenho escolar é diferenciado entre as escolas. Espera-se, por exemplo, que os efeitos negativos do atraso possam ser amenizados nas escolas com maior proporção de homens, sugerindo que o desempenho do aluno é influenciado pelo resultado geral da escola.

Por fim, os valores encontrados para a parcela da variância do desempenho que ocorre entre as escolas mostram que é preciso realizar avaliações em profundidade para entender quais são os fatores escolares que contribuem para o aprendizado, no caso cearense. As análises mostram que são necessárias políticas que possam minimizar a importância de características individuais do estudante e seu *background* familiar, considerando a enorme heterogeneidade de resultados observada e não explicada, entre as escolas.

Reconhecida a importância dos diversos fatores estudados nesse capítulo e diante da elevada magnitude encontrada para o coeficiente estimado do atraso escolar, tanto no modelo de Minas Gerais como no do Ceará, o próximo capítulo irá focar na análise dos fatores que afetam a probabilidade do aluno estar em idade diferente da recomendada para a série.

CAPÍTULO 5 – Os Fatores que Determinam o Atraso Escolar

Considerando-se os benefícios econômicos e sociais da educação, identificar as principais causas dos problemas escolares tem sido agenda de pesquisa de gestores e estudiosos. Na tentativa de entender o que é preciso para se obterem melhorias no aprendizado e na qualidade da educação, a pesquisa empírica tem se preocupado em entender quais fatores mais afetam o desempenho (BIONDI; FELÍCIO, 2007) e a trajetória escolar dos jovens.

A maioria dos países em desenvolvimento está preocupada com o problema das baixas taxas de conclusão escolar, tanto em virtude da perda de oportunidades de investimento para a sociedade, bem como pela ineficiência geral na oferta de educação pública (HANUSHEK; LAVY; HITOMI, 2008). Reconhecendo, portanto, a crucial importância da educação para o desenvolvimento econômico de um país, sugere-se que o ponto de partida para elevar a escolaridade média de uma nação seja o aumento da frequência escolar e a manutenção das crianças na escola, potencializando, assim, a melhoria dos níveis educacionais (PONTILI; KASSOUF, 2007).

A literatura da economia da educação tem se ocupado quanto ao entendimento da influência do contexto social sobre o comportamento individual, como, por exemplo, o impacto de características das escolas e salas de aulas ou a influência de aspectos econômicos do município sobre a performance acadêmica de estudantes.

Na identificação dos determinantes do desempenho mencionados no capítulo anterior, observaram-se diferenças relevantes no resultado dos alunos nos anos iniciais do ensino fundamental. Essa divergência no desempenho dos estudantes pode estar associada a diversos fatores, entre os quais, encontra-se o atraso escolar. Não ter a idade esperada para a série pode ser determinado por características individuais, da família, do meio social e do ambiente escolar. Tendo isso em mente, questiona-se o que afeta a probabilidade de um estudante estar em atraso escolar e em que sentido as escolas são importantes para evitar que esse fato ocorra.

O governo brasileiro, considerando a relevância dos níveis de escolaridade para a renda nacional e para o bem-estar social, implementou várias reformas, principalmente, nos anos 1990, tendo em vista a melhoria dos índices educacionais, entre os quais, estão os indicadores de atraso escolar (PONTILI; KASSOUF, 2007). Essas políticas representaram uma alteração de enfoque das políticas públicas, com a vinculação de parte das receitas públicas aos gastos em educação, em particular, no ensino fundamental, e com um nível mínimo de gasto por estudante, além de modificações na duração do ano letivo e na qualificação mínima exigida para os professores (VASCONCELLOS, 2005).

Nesse sentido, a educação brasileira passou por mudanças significativas, principalmente, em termos de melhorias de acesso ao ensino fundamental²². Em 2006, por exemplo, as crianças passaram a ingressar no ensino fundamental, cuja duração é de nove anos (o limite final para implementação da lei foi em 2010), aos 6 anos de idade, e, em 2009, a escolaridade obrigatória, que se limitava a esse nível de ensino, foi ampliada²³.

Mesmo diante dos avanços observados, a discussão acerca das desigualdades sociais existentes nas escolas ainda é tema recorrente nos debates educacionais (ORTIGÃO; AGUIAR, 2013). Ademais, as altas taxas de repetência e evasão levam a desigualdades sociais, visto sua relação com a renda futura. Sendo assim, são necessários esforços para reinserir os jovens que não concluíram a educação básica (IPEA, 2014), de modo a reduzir o elevado número de estudantes que estão com a idade acima da considerada adequada para a série, o que precisará ser acompanhado por melhorias e expansão da oferta no ensino médio.

Honda (2006) aponta que os dados do SAEB revelam que estudantes com defasagem idade-série obtinham resultados menores na proficiência, evidenciando a forte correlação entre as duas variáveis. No capítulo anterior desta tese, também foram apresentadas evidências de que os alunos que estão em atraso escolar (tanto em Minas Gerais como no Ceará) têm seu desempenho penalizado na Prova Brasil. No caso cearense, nota-se ainda que o impacto da diferença de dois anos ou mais na idade considerada adequada para a série é diferente entre as escolas. Diante disso, sugere-se que o atraso escolar também pode ser utilizado como indicador do desempenho escolar.

O atraso escolar difere conforme a realidade econômica e social dos estados e regiões (PONTILI; KASSOUF, 2007). De acordo com os dados do Censo Escolar, a taxa de distorção idade-série no Brasil, que era de 28,6%, em 2006, reduziu-se para 19,2%, em 2015²⁴. Entre as regiões, para o ano de 2015, a maior taxa encontrada foi a do Norte (28,3%), e a menor foi a da região Sudeste (12,7%). A maior proporção de alunos em atraso escolar é encontrada nas áreas rurais, principalmente, na rede pública de ensino e nos anos finais do ensino fundamental. Para o ano de 2015, em Minas Gerais, a taxa de distorção idade-série total, por exemplo, foi de 12,6%, sendo a taxa nos anos iniciais do ensino fundamental de 5,3%, enquanto que, nos anos finais, o percentual de alunos em atraso escolar encontrado foi de 20,6% (INEP, 2017).

²² Além da questão normativa, outro fator que é apontado para a melhoria do acesso à escola foi o da “transição demográfica”, haja vista a redução do crescimento populacional (HONDA, 2006), principalmente, de crianças e jovens (RIANI, 2005).

²³ Destaca-se que Minas Gerais antecipou essa política, sendo o primeiro estado a aceitar crianças de 6 anos de idade no ensino fundamental.

²⁴ A taxa de distorção idade-série indica a proporção de alunos com dois anos ou mais de atraso escolar, conforme a idade adequada para a série.

Segundo dados do IBGE (2016a), o atraso escolar afetou 40,7% dos alunos mais carentes no Brasil em 2015. Ainda, a taxa de distorção idade-série entre os 20% mais pobres foi 5,0 vezes maior do que entre os alunos 20% mais ricos conforme a distribuição do rendimento mensal domiciliar per capita. Além disso, observou-se que 1,6 milhão de jovens entre 15 a 17 anos de idade não estudavam, dentre os quais, 61,4% tinham o ensino fundamental incompleto e 16,4% abandonaram a escola sem concluir ensino médio, ou seja, grande parte da evasão escolar ocorre no ensino fundamental, principalmente, atingindo os jovens mais fragilizados economicamente.

Diante dessas informações, questiona-se quais fatores podem estar influenciando as principais mudanças na trajetória escolar dos estudantes. Assim, o objetivo aqui é identificar as principais variáveis relacionadas às características individuais, familiares e escolares que afetam a probabilidade de atraso escolar nas escolas públicas do estado de Minas Gerais. Acredita-se que muitos dos resultados gerais podem ser transferíveis para outras partes do Brasil. A variável dependente do modelo refere-se à condição de atraso escolar do aluno, sendo essa uma variável binária, que assume o valor 1 se o estudante tem dois anos ou mais de diferença da idade adequada para a série, e zero, caso contrário. Considerando a estrutura hierárquica presente nos dados educacionais, o modelo escolhido foi o *logit* hierárquico (ou multinível).

A preocupação com o atraso escolar deriva de questões orçamentárias e sociais. A importância de se analisarem os motivos que causam o elevado atraso escolar se justifica pela possibilidade de, por meio dessa análise, se vislumbrarem pontos de ação para políticas públicas que sejam capazes de reverter esse oneroso problema educacional. Mesmo que esse tema já venha sendo debatido, a possibilidade de estudar novas bases de dados e explorar novas técnicas estatísticas justificam seu estudo. Assim, podem-se acrescentar elementos importantes para o debate sobre como investir na melhoria da qualidade do ensino público.

5.1 Variáveis Utilizadas

A Tabela 18 apresenta a descrição das variáveis que entraram na análise do modelo *logit*, conforme a revisão da literatura e a disponibilidade na base de dados, e que não foram descritas no capítulo anterior. Em resumo, as variáveis escolhidas para compor o modelo de atraso escolar no nível do aluno foram: gênero, cor, condição em relação ao trabalho, mora com os pais, escolaridade da mãe, tempo de lazer, realização de dever, nível socioeconômico da família,

e nível de capital social. No nível da escola, as variáveis relacionadas ao contexto escolar que buscam explicar o atraso foram: esfera administrativa, localização, matrículas, ICE, IPE, proficiência média, adoção de ciclo de progressão, forma de composição das turmas, programa de redução de abandono, proporção de professores com superior, macrorregião.

Tabela 18: Variáveis explicativas no nível da escola

Variável	Descrição
Escola	
Proficiência média	Proficiência média da escola obtida na prova de Português e Matemática
Adota ciclo ²⁵	Assume o valor “1” se a escola organiza o ensino fundamental em ciclos e “0”, caso contrário.
Formação de turmas heterogêneas	Assume o valor “1” se a escola, no processo de formação de turmas, opta pela heterogeneidade quanto à idade e rendimento e “0”, caso contrário.
Há programa de redução de abandono	Assume o valor “1” se a escola adota algum programa para reduzir taxas de abandono/evasão e “0”, caso contrário.
Proporção professor com curso superior	Proporção de professores com curso superior na escola.
Alto Paranaíba; Central; Centro-Oeste; Jequitinhonha e Mucuri; Zona da Mata; Noroeste; Rio Doce; Sul; Triângulo.	Assume o valor “1” se a escola está localizada na região e “0”, caso contrário (categoria base = norte).

Há diferenças na organização do ensino entre as redes municipais e estaduais. No ensino fundamental, algumas escolas estaduais adotam dois ciclos de alfabetização: um inicial, de três anos, e um ciclo complementar, com duração de dois anos. Cabe aos municípios a decisão sobre a forma de organização do sistema de ensino em relação à quantidade de ciclos e sua duração (RIANI; DA SILVA; SOARES, 2012). A inclusão da variável relativa aos ciclos decorre da adoção desse sistema que impede a retenção de alunos em determinadas séries (OLIVEIRA; SOARES, 2012), o que, provavelmente, irá afetar a probabilidade de o estudante estar em idade adequada para a série. No presente estudo, essa variável indica se a escola tem ou não o sistema de ciclos.

Inclui-se no modelo uma variável que descreve qual o critério utilizado para a formação das turmas na escola, qual seja, se segue a homogeneidade ou heterogeneidade quanto à idade e ao rendimento. Um meio encontrado pelos gestores educacionais para tentar controlar variáveis do contexto escolar foi atuar na forma de composição das turmas, já que tal composição pode levar a diferenças em termos de rendimento. Espera-se, por exemplo, que a formação de turmas com estudantes de diferentes situações econômicas melhore o rendimento dos alunos menos privilegiados, mas há incertezas quanto ao verdadeiro efeito no desempenho

²⁵ As informações para essa variável foram extraídas dos microdados do Censo Escolar de 2013.

médio do grupo (SOARES, 2002). Foram incluídas, também, as macrorregiões do estado como variáveis de controle das diferenças culturais e regionais (PONTILI; KASSOUF, 2007).

Assim como feito anteriormente, as variáveis foram centradas na grande média. Para o modelo logístico multinível, essa transformação tem como vantagem a estimativa da probabilidade de atraso prevista no valor médio das variáveis independentes somente pelo intercepto, visto que, na média, os valores das variáveis explicativas são iguais a zero (RIOS-NETO; CEZAR; RIANI 2002).

5.2 Estatísticas Descritivas

As Tabelas 19 a 22 trazem as estatísticas descritivas: média, desvio-padrão, valor máximo e mínimo e coeficiente de variação (CV) das características dos alunos e das escolas que compõem a amostra final no 5º e 9º ano. Para não tornar exaustiva a leitura, optou-se pela discussão apenas das informações para o 9º ano, uma vez que os resultados do 5º ano foram muito similares aos apresentados no capítulo anterior.

Tabela 19: Estatísticas descritivas no nível do aluno, 5º ano, Minas Gerais

Variável	Nº obs.	Média	Mínimo	Máximo	Desvio Padrão	CV
Atraso	35.787	0,11	0,00	1,00	-	-
Homem	35.787	0,50	0,00	1,00	-	-
Branco	35.787	0,30	0,00	1,00	-	-
Mora com mãe e pai	35.787	0,71	0,00	1,00	-	-
<i>Escolaridade da mãe (categoria base = mãe não estudou)</i>						
Até 4º série do fundamental	35.787	0,22	0,00	1,00	-	-
Ensino fundamental completo	35.787	0,17	0,00	1,00	-	-
Ensino médio completo	35.787	0,24	0,00	1,00	-	-
Superior completo	35.787	0,20	0,00	1,00	-	-
Trabalho doméstico	35.787	0,80	0,00	1,00	-	-
Trabalho fora	35.787	0,12	0,00	1,00	-	-
Tempo de lazer	35.787	2,34	0,00	4,00	1,26	0,54
Dever de Matemática	35.787	0,85	0,00	1,00	-	-
Dever de Português	35.787	0,83	0,00	1,00	-	-
Nível Socioeconômico	35.787	0,31	-3,18	4,43	0,87	2,81
Nível de Capital Social	35.787	-0,04	-2,76	0,32	0,57	-14,31

Fonte: Elaboração própria a partir de dados da Prova Brasil / INEP 2013

De acordo com a Tabela 20, nota-se que, entre os estudantes do 9º ano, em média, 16% estão com idade diferente da recomendada para a série. Observa-se que, cerca de 47% dos estudantes são meninos e em relação à cor, encontrou-se que, aproximadamente, 31% dos estudantes se declararam brancos.

Tabela 20: Estatísticas Descritivas no Nível do Aluno, 9º Ano, Minas Gerais

Variável	Nº obs.	Média	Mínimo	Máximo	Desvio Padrão	CV
Atraso	54.624	0,16	0,00	1,00	-	-
Homem	54.624	0,47	0,00	1,00	-	-
Branco	54.624	0,31	0,00	1,00	-	-
Mora com mãe e pai	54.624	0,69	0,00	1,00	-	-
<i>Escolaridade da mãe (categoria base = mãe não estudou)</i>						
Até 4º série do fundamental	54.624	0,26	0,00	1,00	-	-
Ensino fundamental completo	54.624	0,17	0,00	1,00	-	-
Ensino médio completo	54.624	0,26	0,00	1,00	-	-
Superior completo	54.624	0,12	0,00	1,00	-	-
Trabalho doméstico	54.624	0,82	0,00	1,00	-	-
Trabalho fora	54.624	0,19	0,00	1,00	-	-
Tempo de lazer	54.624	2,73	0,00	4,00	1,24	0,45
Dever de Matemática	54.624	0,61	0,00	1,00	-	-
Dever de Português	54.624	0,61	0,00	1,00	-	-
Nível Socioeconômico	54.624	0,04	-3,73	5,05	0,88	23,37
Nível de Capital Social	54.624	-0,08	-2,80	0,43	0,66	-7,81

Fonte: Elaboração própria a partir de dados da Prova Brasil / INEP 2013

No que diz respeito à escolaridade materna, em média, 26% têm até 4º série do fundamental, 17% têm ensino fundamental completo, 26% têm ensino médio completo e 12% têm superior completo. Nota-se que 69% dos alunos moram com os pais, grande parte dos estudantes (82%), em média, realizam tarefas domésticas; já 19% dos alunos trabalham fora de casa. O tempo que os alunos gastam, em média, com TV, internet ou jogos eletrônicos, para a amostra, foi de 2,73 horas, com desvio-padrão de 1,24 horas. Na amostra, 61% dos alunos afirmaram que fazem as tarefas, tanto de Português como de Matemática, na maior parte das vezes.

Em média, os estudantes têm um NSE de 0,04, com um coeficiente de variação de 23,37, indicando uma elevada variabilidade dos dados em relação à média e uma alta heterogeneidade da condição socioeconômica dos discentes. A média do NCS foi de -0,08, com coeficiente de variação de 7,81, mostrando que há bastante heterogeneidade quanto ao nível de participação e atenção dos pais na vida escolar dos seus filhos.

Tabela 21: Estatísticas Descritivas no Nível da Escola, 5º Ano, Minas Gerais

Nome da variável	Nº obs.	Média	Mínimo	Máximo	Desvio Padrão	CV
Municipal	2.278	0,63	0,00	1,00	-	-
Urbana	2.278	0,93	0,00	1,00	-	-
Matrículas	2.278	62,88	20,00	269,00	34,73	0,55
ICE	2.278	-0,07	-2,80	0,97	0,83	-11,96
IPE	2.278	0,06	-3,10	0,78	0,73	11,63
Proficiência média	2.278	223,75	155,21	314,97	21,84	0,10
Adota ciclo	2.278	0,81	0,00	1,00	-	-
Formação de turmas heterogêneas	2.278	0,46	0,00	1,00	-	-
Há programa de redução de abandono	2.278	0,71	0,00	1,00	-	-
Prop. Professor com curso superior	2.278	0,94	0,00	1,00	-	-
<i>Macrorregiões</i>						
Alto Paranaíba	2.278	0,04	0,00	1,00	-	-
Central	2.278	0,31	0,00	1,00	-	-
Centro-Oeste	2.278	0,06	0,00	1,00	-	-
Jequitinhonha e Mucuri	2.278	0,06	0,00	1,00	-	-
Zona da Mata	2.278	0,12	0,00	1,00	-	-
Noroeste	2.278	0,02	0,00	1,00	-	-
Norte	2.278	0,08	0,00	1,00	-	-
Rio Doce	2.278	0,10	0,00	1,00	-	-
Sul	2.278	0,13	0,00	1,00	-	-
Triângulo	2.278	0,07	0,00	1,00	-	-

Fonte: Elaboração própria a partir de dados da Prova Brasil / INEP 2013

A Tabela 22 contém as estatísticas descritivas das variáveis relativas à escola para o 9º ano. Na amostra, tem-se que 33% das escolas são da rede municipal e a maior parte está localizada na área urbana (93%). A média de estudantes matriculados no 9º ano foi de 70,55 alunos, mas as escolas diferem quanto a esse aspecto (coeficiente de variação de 0,61).

No que diz respeito ao ICE, observa-se que a sua média foi de -0,19 e o desvio-padrão foi de 0,83, sugerindo que há uma volatilidade das condições de conservação da infraestrutura entre as escolas. O IPE teve uma média de -0,003 e um elevado coeficiente de variação de 235,25, sinalizando a presença de uma significativa heterogeneidade entre as escolas quanto à condição de conservação dos equipamentos. Em termos médios, a proficiência média em Português e Matemática foi de 259,56 pontos, com desvio-padrão de 18,50 pontos. O baixo valor encontrado para o coeficiente de variação (0,07) sugere que as escolas são homogêneas em termos da proficiência média obtida pelos seus estudantes nas duas disciplinas.

Tabela 22: Estatísticas Descritivas no Nível da Escola, 9º Ano, Minas Gerais

Nome da variável	Nº obs.	Média	Mínimo	Máximo	Desvio Padrão	CV
Municipal	1.793	0,33	0,00	1,00	-	-
Urbana	1.793	0,93	0,00	1,00	-	-
Matrículas	1.793	70,55	20,00	356,00	43,07	0,61
ICE	1.793	-0,19	-2,80	0,97	0,83	-4,46
IPE	1.793	-0,003	-3,10	0,78	0,74	-235,25
Proficiência média	1.793	259,56	192,50	317,67	18,50	0,07
Adota ciclo	1.793	0,88	0,00	1,00	-	-
Formação de turmas heterogêneas	1.793	0,43	0,00	1,00	-	-
Há programa de redução de abandono	1.793	0,89	0,00	1,00	-	-
Prop. Professor com curso superior	1.793	0,97	0,00	1,00	-	-
<i>Macrorregiões</i>						
Alto Paranaíba	1.793	0,04	0,00	1,00	-	-
Central	1.793	0,31	0,00	1,00	-	-
Centro-Oeste	1.793	0,05	0,00	1,00	-	-
Jequitinhonha e Mucuri	1.793	0,06	0,00	1,00	-	-
Zona da Mata	1.793	0,12	0,00	1,00	-	-
Noroeste	1.793	0,02	0,00	1,00	-	-
Norte	1.793	0,10	0,00	1,00	-	-
Rio Doce	1.793	0,10	0,00	1,00	-	-
Sul	1.793	0,13	0,00	1,00	-	-
Triângulo	1.793	0,07	0,00	1,00	-	-

Fonte: Elaboração própria a partir de dados da Prova Brasil / INEP 2013

Na amostra, em média, 88% das escolas organizam o ensino fundamental em ciclos. Nota-se que, em média, 43% das escolas adotam a opção pela heterogeneidade quanto à idade e rendimento na composição das turmas. De outro lado, grande parte das escolas (89%) tem algum programa de redução de taxas de abandono e observa-se ainda que a maioria dos professores possuem ensino superior (97%). Por fim, em termos regionais, percebe-se que a maior parte das escolas faz parte da região central do estado, estando incluída a Região Metropolitana de Belo Horizonte, que reúne o maior número de municípios.

5.3 Modelo de Regressão Hierárquica Linear para o Atraso Escolar

A Tabela 23 mostra os resultados do modelo nulo para a probabilidade de atraso escolar em Minas Gerais. Foram estimados modelos para os dados da Prova Brasil, do 5º ano e 9º ano do ensino fundamental, na tentativa de identificar se existem diferenças entre os fatores que afetam o atraso escolar nos anos iniciais e finais do ensino fundamental.

Ao exponenciar a estimativa do coeficiente de regressão, tem-se a razão de chances (*Odds Ratio – OR*), que representa como as chances de atraso se modificam quando se passa entre as diferentes categorias de um fator. Assim, para cada coeficiente das variáveis explicativas na regressão logística, a OR indica a diferença percentual nas chances de estar em atraso quando a variável independente aumenta em uma unidade, mantendo-se as demais variáveis constantes (ORTIGÃO; AGUIAR, 2013).

Quando a razão de chance for menor do que 1, a variável explicativa reduz a probabilidade de o evento ocorrer, e, quando for maior do que 1, aumenta-se a possibilidade de ocorrência do evento, ou seja, essa razão quantifica como cada variável explicativa afeta as probabilidades de cada valor de resposta da variável dependente.

A estimativa do intercepto indica que as chances de o aluno ter dois anos ou mais da idade adequada em relação à média das escolas ($u_{0j} = 0$) são de 0,10, no 5º ano, e de 0,18, no 9º ano. A variância do intercepto em log de chances do atraso é 0,787, com erro padrão de 0,06. Segundo o teste t, essa variância é estatisticamente significativa, mostrando que há diferenças entre as escolas. Já o Coeficiente de Partição da Variância (CPV) revela que cerca de 19% e 11% da variabilidade na probabilidade de atraso escolar ocorre devido às características escolares, respectivamente, no 5º ano e no 9º ano, recomendando-se a necessidade da modelagem hierárquica.

Tabela 23: Modelo Nulo – Probabilidade de Atraso Escolar, Minas Gerais

VARIÁVEIS EXPLICATIVAS	5º EF			9º EF				
	Efeito Fixo	Estimativa	EP	P-valor	Efeito Fixo	EP	P-valor	
Intercepto (γ_{00})		-2,33	0,003	0.000		-1,71	0,02	0.000
Efeito Aleatório		Variância	EP	Razão-t		Variância	EP	Razão-t
Variância do intercepto (σ_{u0}^2)		0,787	0,06	13,1		0,42	0,02	21,0
CPV		0,1929				0,1132		
Deviance M1		23.878,49				47.074,31		
Número de parâmetros		2				2		
Número de observações		35.787				54.624		
Número de grupos		2.278				1.793		

A Tabela 24 apresenta os resultados do modelo logístico hierárquico para a probabilidade de atraso escolar. Observa-se que, para a variável que indica o sexo do aluno, encontrou-se a razão de chance para o estudante que se declarou homem igual a 1,74. Esse valor indica que a chance de os meninos estarem em atraso é 74% maior do que para meninas, no caso do 5º ano do ensino fundamental, mantidas constantes as demais variáveis. No caso do 9º ano, as chances

de estar em atraso é 60% maior para os meninos. Esse resultado está em consonância com a literatura, que tem encontrado diferenças em termos de gênero na trajetória escolar.

Tabela 24: Modelo para a Probabilidade de Atraso no Ensino Fundamental, MG

VARIÁVEIS EXPLICATIVAS		5º EF			9º EF		
Efeito Fixo		Estimativa	P-valor	O.R	Estimativa	P-valor	O.R
Intercepto		-2,630	0,000	0,07	-1,947	0,000	0,14
Aluno							
Homem		0,056	0,000	1,74	0,468	0,000	1,60
Branco		-0,026	0,000	0,77	-0,178	0,000	0,84
Mora com mãe e pai		-0,059	0,000	0,56	-0,442	0,000	0,64
<i>Escolaridade da mãe: (categoria base= mãe não estudou)</i>							
- Até 4º série do fundamental		-0,025	0,000	0,78	-0,516	0,000	0,60
- Ensino fundamental completo		-0,050	0,000	0,61	-0,719	0,000	0,49
- Ensino médio completo		-1,125	0,000	0,32	-0,976	0,000	0,38
- Superior completo		-1,153	0,000	0,32	-1,140	0,000	0,32
Trabalho doméstico		0,002	0,619	1,02	-0,070	0,036	0,93
Trabalho fora de casa		0,071	0,000	2,03	0,638	0,000	1,89
Tempo de lazer		-0,005	0,001	0,95	-0,173	0,000	0,84
Dever de Matemática		-0,028	0,000	0,75	-0,322	0,000	0,72
Dever de Português		-0,024	0,000	0,79	-0,165	0,000	0,85
NSE		-0,006	0,048	0,94	-0,022	0,240	0,98
NCS		-0,020	0,000	0,82	-0,155	0,000	0,86
Escola							
Municipal		0,064	0,000	1,89	-0,130	0,009	0,88
Urbana		-0,007	0,596	0,94	-0,181	0,039	0,83
Matrículas		0,000	0,973	1,00	-0,001	0,182	1,00
ICE		0,001	0,742	1,01	0,026	0,302	1,03
IPE		-0,001	0,892	0,99	-0,029	0,312	0,97
Proficiência média		-0,001	0,000	0,99	-0,009	0,000	0,99
Adota ciclo		-0,040	0,000	0,67	-0,322	0,000	0,72
Formação de turmas heterogêneas		0,011	0,027	1,12	0,147	0,000	1,16
Há programa de redução de abandono		0,013	0,027	1,13	-0,018	0,778	0,98
Prop. Professor com curso superior		0,002	0,832	1,02	-0,018	0,885	0,98
<i>Macrorregiões (categoria base= Norte)</i>							
Alto Paranaíba		0,067	0,000	1,95	0,282	0,014	1,33
Central		0,041	0,000	1,51	0,550	0,000	1,73
Centro-Oeste		0,035	0,024	1,42	0,170	0,160	1,18
Jequitinhonha e Mucuri		0,058	0,000	1,79	0,489	0,000	1,63
Zona da Mata		0,078	0,000	2,18	0,611	0,000	1,84
Noroeste		0,034	0,093	1,40	0,028	0,863	1,03
Rio Doce		0,037	0,004	1,45	0,482	0,000	1,62
Sul		0,065	0,000	1,91	0,383	0,000	1,47
Triângulo		0,063	0,000	1,89	0,282	0,006	1,33
Efeito Aleatório		Variância	E. P.	Razão-t	Variância	E. P.	Razão-t
Intercepto (σ_{u0}^2)		0,383	0,039	9,94	0,308	0,023	13,19
NSE		0,082	0,028	2,93	0,046	0,013	3,59
CPV		0,1064			0,0857		

A probabilidade de atraso entre os alunos que se declararam brancos é menor do que os estudantes que se declararam pretos, pardos, amarelos e índios, em ambas as séries. Nota-se que a razão de chances de atraso para brancos é 16% menor do que para não brancos, no 9º ano, e de 23%, no 5º ano, refletindo a presença de desigualdades associadas à cor nas escolas.

Os resultados do modelo sugerem que alunos que moram com uma família tradicional, formada por pai e mãe, têm probabilidade menor de estarem em atraso escolar. Nos anos iniciais do ensino fundamental, por exemplo, a probabilidade de os alunos que moram com pai e mãe apresentarem atraso é 44% menor do que a dos estudantes que não moram com o pai e a mãe. Espera-se, portanto, que famílias formadas por pai, mãe e filhos exibam condições de socialização mais favoráveis aos jovens (SILVA; HASENBALG, 2002).

A educação da mãe, tida como uma referência do capital cultural familiar, é um fator importante para que os estudantes frequentem a escola na idade correta. Tomando-se como referência a categoria-base “mãe não estudou”, observa-se que, quanto mais escolarizada é a mãe, menor se torna a probabilidade de atraso escolar do filho, sendo o impacto maior para o 5º do que para o 9º ano, o que pode ser indicativo da maior influência materna nas decisões escolares das crianças nos anos iniciais de ensino.

Do resultado do modelo para o caso do 5º ano, em comparação aos alunos com mães que não estudaram ou não completaram o primeiro ciclo do ensino fundamental (categoria-base), as chances de atraso são, em média: 22% menor para alunos com mães que estudaram até a 4ª série, 39% menor para alunos com mães que completaram o ensino fundamental, 68% menores para estudantes com mães que estudaram todo o ensino médio e para alunos com mães que possuem curso superior completo. Rios-Neto, Cezar e Riani (2002), em suas conclusões, sugeriram que, nas séries iniciais e nos locais com maior proporção de mães com baixa escolaridade, deveriam ser alocados os melhores professores com a finalidade de melhorar a equidade da educação nessas regiões.

A variável relativa ao trabalho doméstico não teve significância estatística para alunos do 5º ano, mas foi significativa para os do 9º ano. Nesse último caso, observa-se que as chances de atraso são 7% vezes menores para alunos que realizam tarefas domésticas. Para o 5º ano, o trabalho fora de casa, remunerado ou não, aumenta as chances de atraso escolar em duas vezes mais do que para os alunos que não trabalham fora. Para o 9º ano, esse valor cai para 1,89. Desse modo, estudantes que trabalham têm menor probabilidade de avançar nos estudos se comparados aos que não trabalham (LEON; MENEZES-FILHO, 2002).

O tempo dispendido em lazer, seja assistindo à televisão, navegando na internet ou com jogos eletrônicos, reduz as chances de atraso em 5%, no 5º ano, e em 16%, no 9º ano. A realização de dever de casa pelo aluno é vista como um indicador do interesse ou motivação pelos estudos. Por isso, inclui-se no modelo a variável que capta se o aluno realiza ou não as lições de casa, nas duas disciplinas avaliadas na Prova Brasil. Como esperado, os estudantes que afirmaram fazer dever de casa têm chances menores de estarem em atraso escolar. Em relação à Matemática, as chances de o aluno que afirmou fazer o dever de casa são 0,75 (0,72) vezes menores no 5º ano (9º ano). No caso de Língua Portuguesa, as chances são 0,79 (0,85) vezes menores no 5º ano (9º ano), em comparação aos estudantes que não fazem as tarefas. Portanto, o empenho em realizar o dever indica a responsabilidade do discente com o aprendizado, o que pode reduzir a probabilidade de ele reprovar ou abandonar a escola, entre outros fatores que levam ao atraso nos estudos.

O NSE do aluno não foi estatisticamente significativo para explicar a probabilidade de atraso escolar em ambas as séries. Essa perda de significância está relacionada à inclusão do controle do contexto escolar no modelo. Por outro lado, o NCS que representa a importância do interesse dos pais pela vida escolar de seus filhos foi significativo, indicando que jovens que possuem um NCS acima da média têm menores chances de ter dois anos a mais de diferença da idade adequada para a série que cursa. No 5º ano do ensino fundamental, a chance de estar em atraso é 18% menor para os estudantes que têm um NCS acima da média escolar se comparado aos alunos que têm um capital social familiar abaixo da média. Para o 9º ano, nesse caso, a chance de atraso foi 14% menor. Assim, os pais que vão às reuniões, incentivam os estudos, a leitura, a realização de dever, marcam presença nas aulas e conversam sobre a escola influenciam positivamente a trajetória escolar de seus filhos. Esses resultados confirmam outros estudos que analisaram a importância de um ambiente familiar favorável à educação como determinante no desempenho escolar (ORTIGÃO; AGUIAR, 2013; PALERMO; SILVA; NOVELLINO, 2014).

Passando à análise das variáveis relacionadas ao ambiente escolar, observa-se que há diferenças entre a chance de estar em atraso escolar, conforme a esfera administrativa da escola e as séries. Para o 5º ano, a chance de um aluno de escola municipal estar atrasado é 89% maior, quando comparado a um discente de escola estadual. Já para o 9º ano, a chance de um aluno do ensino municipal estar em idade superior à adequada para a série é 12% menor, quando comparado aos estudantes da rede estadual.

Como esperado, a adoção de ciclos de progressão pela escola dificulta a probabilidade de atraso escolar, visto que esse sistema impede a reaprovação em determinadas séries. Para alunos do 5º ano que estudam em escolas que adotam a política de ciclos, a chance de estar em atraso escolar é reduzida em 33%. Já para os estudantes do 9º ano, a chance de atraso escolar é 28% menor nas escolas que adotam ciclos, quando comparadas às escolas que não adotam.

A variável urbana buscou captar os diferenciais entre escolas provenientes da zona urbana e rural quanto ao atraso escolar, não sendo essa variável significativa para o 5º ano. Nota-se que, para o último ano do ensino fundamental, a localização da escola na área urbana reduz em 17% a chance de atraso escolar, em comparação com as escolas rurais. Isso pode estar relacionado a dificuldades de transporte, à precariedade da infraestrutura da escola, à carência de professores nessas áreas, ao trabalho dos jovens (custo de oportunidade de continuar estudando) na propriedade e, também, à falta de políticas educacionais específicas para a área rural.

A variável “Matrículas” foi incluída para tentar captar o tamanho da escola e, com isso, controlar a heterogeneidade escolar. Para cada ano, utilizou-se o número de alunos matriculados na série. No entanto, no presente modelo, essa variável não apresentou significância estatística. Do mesmo modo, os indicadores de conservação (ICE) e de posse de equipamentos (IPE) também não foram estatisticamente significativos para explicar o atraso escolar.

Tentando verificar se há relação entre o desempenho individual do aluno e a probabilidade de ele estar com idade diferente da recomendada para a série, inclui-se no modelo a nota média do aluno obtida na prova de Matemática e de Língua Portuguesa. Em ambas as séries, a razão de chances encontrada foi de 0,99, indicando que cada ponto adicional na proficiência média do estudante reduz a chance de atraso, em média, em 1%. Assim, políticas que levem a melhorias no desempenho, principalmente, dos estudantes com menores notas, podem ter impactos significativos na redução das taxas de atraso escolar.

Outra variável incluída para controle do contexto escolar foi a forma de composição das turmas. Observou-se que escolas que adotam o critério da heterogeneidade (quanto à idade e rendimento) na formação de turmas têm as chances de atraso escolar maiores do que aquelas que adotam a homogeneidade (ou outro critério). A chance de atraso em escolas que adotam a heterogeneidade é 12% maior, no 5º ano, e 16% maior, no 9º ano. Nesse caso, a influência dos pares na trajetória escolar parece ter consequências negativas, sinalizando que as escolas devem estar atentas para que a forma de composição das turmas não cause elevadas diferenças no rendimento.

Outra questão analisada foi em relação às ações da escola para reduzir a taxa de abandono. Os resultados mostraram que, nas escolas que adotam programas de redução das taxas de abandono, as chances de atraso escolar são 13% maiores do que nas escolas que não adotam ou naquelas em que não há esse problema para o 5º ano (no 9º ano, não houve significância estatística). O que pode explicar essa associação positiva é que, possivelmente, essas ações levam à retenção de alunos com baixo desempenho que estão em atraso. Ou porque nas escolas onde se implementam essas ações já persista uma taxa muito elevada de evasão escolar.

Também foi incluída no modelo a proporção de professores com curso superior, com a finalidade de avaliar se a maior escolaridade do docente influencia no atraso escolar. No entanto, essa variável não apresentou significância estatística, talvez porque grande parte da amostra seja composta por professores com ensino superior (97%).

Por fim, as *dummies* para região buscaram captar as diferenças regionais na probabilidade de atraso. A região Norte de Minas foi escolhida como a categoria de base, visto que ela apresentou a menor proporção de alunos com atraso escolar no banco de dados. No geral, observa-se que há diferenças regionais em termos da probabilidade de atraso escolar, como, por exemplo, para o 5º ano, em comparação aos alunos residentes no Norte de Minas, as chances de estar em atraso é: 1,95 vezes maior para alunos do Alto Paranaíba; 1,51 vezes maior para residentes da região central do estado; 1,42 vezes maior para estudantes do Centro-oeste; 1,79 vezes maior para alunos de escolas do Vale do Jequitinhonha e Mucuri; 2,18 vezes maior para o aluno que reside na Zona da Mata; 1,45 vezes maior para discentes do Rio Doce; 1,91 vezes maior para alunos do Sul de Minas; e 1,89 vezes maior para estudantes do Triângulo Mineiro. Não tiveram significância estatística a *dummy* para a região Noroeste do estado, em ambas as séries, e para o Centro-oeste, no 9º ano.

Passando-se à análise dos efeitos aleatórios do modelo, ao considerar que o intercepto varia entre as escolas, assume-se que há um efeito diferenciado do atraso escolar, isto é, em alguns casos, o desempenho dos estudantes com atraso escolar é mais prejudicial do que em outros. Dos resultados do modelo, pode-se considerar que o intercepto e o nível socioeconômico do estudante tenham efeitos aleatórios. Comparando as duas séries, nota-se que o efeito aleatório é maior para o 5º ano do ensino fundamental, tanto no intercepto, bem como para o NSE. Portanto, a probabilidade média de um aluno estar em atraso escolar é diferenciada entre as escolas, assim como o NSE do aluno influencia diferentemente na probabilidade de ele frequentar a série na idade correta. Em outras palavras, os componentes de variância resumem, para cada estabelecimento de ensino, a influência com que a probabilidade de o aluno estar em

atraso difere sistematicamente dos estudantes (com características similares) matriculados em outras escolas.

Da construção de intervalos de confiança, por exemplo, para o 5º ano, para o intercepto, tem-se que, no limite inferior, as chances de o aluno estar em atraso escolar em relação à média escolar são de 0,02, e, no limite superior, as chances de atraso em relação à média escolar são de 0,24. Para o NSE, o intervalo de confiança mostra que, no limite inferior, cada ponto adicional no NSE acima da média reduz as chances de atraso em 43%, e, no limite superior, cada ponto a mais no NSE acima da média aumenta as chances de atraso em 74%.

Após realizado o controle do modelo por todas essas características apresentadas, o cálculo do CPV para o modelo final aponta que, aproximadamente, 11% da variabilidade na probabilidade de atraso escolar é explicada por fatores escolares no 5º ano e 9%, no 9º ano. Esse resultado sugere que políticas de intervenção nas escolas podem ser importantes para reduzir as taxas de atraso escolar.

5.4 Considerações Gerais Sobre os Resultados do Modelo

Neste capítulo, o objetivo é aprofundar o conhecimento a respeito dos motivadores do atraso escolar no ensino fundamental. Em resumo, a partir da análise da influência de características dos estudantes e de fatores escolares sobre o atraso, notou-se que têm maiores chances de atraso alunos do sexo masculino, não brancos, que trabalham fora de casa, estudam em turmas heterogêneas e em escolas que adotam programas de redução de taxas de abandono.

Em contraste, as chances de estar em atraso escolar são menores para os estudantes que moram com os pais, têm mães mais escolarizadas, têm tempo de lazer, fazem o dever de casa, têm pais que apoiam os estudos, estudam em escolas que adotam ciclos no ensino fundamental e em escolas que possuem níveis maiores na proficiência média das duas disciplinas. Além disso, há diferenças entre as regiões na probabilidade de atraso escolar.

A utilização dos modelos multinível permitiu a análise da influência mútua do nível familiar com o nível das políticas educacionais (RIOS-NETO; CEZAR; RIANI 2002). Assim, espera-se que a presente análise contribua para a discussão do efeito do atraso escolar sobre o aprendizado do aluno, destacando-se a relevância desses resultados para a melhoria da qualidade da educação pública em Minas Gerais que ainda precisa superar as desigualdades existentes. Portanto, sugere-se que políticas destinadas à redução das taxas de atraso escolar, as

quais levem em conta a importância dos fatores aqui apresentados, poderão levar a melhorias na escolaridade no estado de Minas Gerais.

CONSIDERAÇÕES FINAIS

Com o objetivo de analisar os principais determinantes do desempenho e da trajetória escolar dos estudantes em Minas Gerais, inicialmente, realizaram-se a apresentação dos instrumentos de avaliação da educação e a descrição de dados educacionais no estado. Assim foi possível compreender a importância que o desenvolvimento dos sistemas de avaliação teve para a melhora na formulação e condução das políticas na área educacional. Além disso, foi possível observar o contexto atual dos principais indicadores educacionais em Minas Gerais.

No segundo capítulo, apresentaram-se as principais teorias sobre economia da educação e evidências empíricas dos determinantes do aprendizado e da trajetória escolar. Mesmo sendo vastas as contribuições, a literatura ainda não encontrou consenso acerca de quais são os principais determinantes do desempenho dos estudantes em testes de proficiência, em particular, qual parcela é devida a fatores relacionados ao *background* familiar e a fatores escolares.

Na tentativa de compreender a importância de cada um desses elementos, optou-se por modelos hierárquicos que considerem o grau de dependência entre as observações e sejam capazes de suscitar previsões mais sólidas. Os resultados dos modelos econôméticos desenvolvidos no Capítulo 4 evidenciam que os determinantes do desempenho escolar no ensino fundamental estão relacionados às características individuais dos estudantes e da sua família. Mas os fatores escolares também são importantes para o aprendizado. A escolaridade da mãe, por exemplo, é uma variável importante para a proficiência, sendo o trabalho infantil prejudicial ao desempenho acadêmico dos estudantes. Ainda, ficou evidenciado que as escolas diferem umas das outras na sua forma de gestão escolar e, também, devido a características individuais de seus estudantes.

Mesmo após o controle por variáveis do nível do aluno e do nível da escola, o índice de correlação intraclasse no modelo final mostra que ainda há uma heterogeneidade explicada pelas diferenças entre as escolas. Esses valores podem ser explicados, em parte, pela falta de dados mais completos sobre o ambiente em que ocorre o aprendizado.

Uma vez que as variáveis escolares são as mais passíveis de intervenção no médio prazo, espera-se que a atuação de políticas públicas possa ser importante para reduzir a tendência de reprodução das desigualdades sociais nas escolas. Dos resultados encontrados, destaca-se a relevância de fatores ligados à qualidade dos serviços escolares e de seus recursos humanos, como, por exemplo, a infraestrutura escolar, a presença de professores com licenciatura em

Matemática, diretores que estão há mais tempo nessa função e a existência de projeto pedagógico na escola. Além disso, é preciso aprimorar as ferramentas de pesquisa educacionais em alguns aspectos, como os questionários socioeconômicos aplicados nas avaliações e, também, que os aplicadores reforcem a importância de os alunos preencherem corretamente essas informações.

Considerando a forte evidência encontrada no Capítulo 4 sobre os impactos negativos do atraso escolar no desempenho escolar, tanto no estudo para Minas Gerais, bem como no do Ceará, no Capítulo 5, procurou-se estudar o que afeta a probabilidade de o estudante estar em idade acima da recomendada para a série. Nesse caso, é indicado o uso do modelo logístico hierárquico, que tem sido aplicado em várias áreas da pesquisa acadêmica na tentativa de conhecer melhor os processos econômicos, demográficos e sociais por meio de respostas binárias.

Em particular, observou-se que o desempenho individual do aluno está relacionado com probabilidades de atraso escolar. Em termos de política, é importante oferecer uma visão sobre como analisar diferentes propostas para lidar com elevadas taxas de atraso escolar. Além disso, nota-se que o trabalho infantil duplica as chances de atraso escolar em comparação aos estudantes que não trabalham, mostrando quão prejudicial isso pode ser para a vida escolar desses jovens.

Em relação à opção pela adoção de ciclos ou sistemas de progressão continuada pelas escolas, observou-se que ela reduz a probabilidade de atraso escolar, como esperado, já que esse sistema evita a reprovação em algumas séries. No entanto, é preciso estar atento a que custos essas ações são positivas para a produção de um aprendizado de qualidade. Independentemente da proposta adotada para se tentar corrigir a defasagem idade-série, deve-se garantir a qualidade do ensino, de modo que os alunos que concluam o ensino fundamental tenham a idade esperada, mas que, também, dominem o conteúdo necessário para exercerem seus direitos e deveres de cidadão (FERRÃO; BELTRÃO; DOS SANTOS, 2002).

Os determinantes aqui analisados fazem parte de uma ampla gama de fatores que ainda precisam ser melhor estudados, mas os resultados encontrados neste trabalho podem servir de suporte para o debate acerca da qualidade da educação no Brasil e para a agenda de políticas educacionais. Em particular, espera-se que este trabalho possa contribuir na formação de políticas que levem à melhoria da qualidade da educação e à redução das desigualdades. Além disso, espera-se que esta pesquisa possa suscitar ideias para trabalhos futuros que adicionem novas variáveis que descrevam melhor o contexto escolar, ou que utilizem outros estados ou

períodos ou dados longitudinais. Ademais, sugere-se a estimação de um modelo de terceiro nível que busque captar fatores específicos dos municípios que possam ser importantes para explicar as diferenças entre as escolas.

REFERÊNCIAS

- ALBERNAZ, A.; FERREIRA, F.; FRANCO, C. Qualidade e equidade no ensino fundamental brasileiro. *Pesquisa e Planejamento Econômico* (Rio de Janeiro), v. 32, n.3, p. 453-476, 2002.
- ALMEIDA, A. T. C. A. Determinantes dos piores e melhores resultados educacionais dos alunos da rede pública de ensino fundamental no Brasil. *Planejamento e Políticas públicas*, n.42, jan./jun., p.147-182, 2014.
- ALVES, M. T. G.; SOARES, J. F.; XAVIER, F. P. Índice socioeconômico das escolas de educação básica brasileiras. *Ensaio: avaliação de políticas públicas da educação* (Rio de Janeiro), v. 22, n. 84, p. 671-704, jul/set, 2014.
- AMORIM, L. J. O capital humano no contexto da sociedade e do conhecimento. *Revista SEBRAE* (Brasília), v. 1, p. 1-16, 2009.
- ANDRADE, D. F.; TAVARES, H. R.; VALLE, R.C. Teoria da Resposta ao Item: Conceitos e Aplicações. São Paulo: *ABE – Associação Brasileira de Estatística*, 2000.
- ANDRADE, J. M.; LAROS, J. A. Fatores associados ao desempenho escolar: um estudo multinível com dados do SAEB/2001. *Psicologia: Teoria e Pesquisa*, Brasília, v.23, n.1, jan-mar, p. 33-42, 2007.
- ANDRADE, R. J.; SOARES, J. F. O efeito da escola básica brasileira. *Estudos em Avaliação Educacional*, v. 19, p. 379-406, 2008.
- AQUINO, J. A. As teorias da ação social de Coleman e de Bourdieu. *Humanidades e Ciências Sociais*, v. 2, n. 2, p. 17-29, 2000.
- ASSOCIAÇÃO BRASILEIRA DE EMPRESAS DE PESQUISA (ABEP). *Critério de Classificação Econômica Brasil*. 2014. Disponível em: <<http://www.abep.org/criterio-brasil>>. Acesso em: 22 ago. 2016.
- BARBOSA, M. E. F.; FERNANDES, C. Modelo multinível: uma aplicação a dados de avaliação educacional. *Estudos em Avaliação Educacional*, n. 22, p. 135-154, 2000.
- _____. A escola brasileira faz diferença? Uma investigação dos efeitos da escola na proficiência em Matemática dos alunos da 4^a série. *Promoção, ciclos e avaliação educacional*. Curitiba: ArtMed, p. 155-172, 2001.
- BARBOSA-FILHO, F. H; PESSOA, S. A. Educação e crescimento: O que a evidência empírica e teórica mostra? *Revista Economia*, Brasília (DF), v.11, n. 2, p.265–303, 2010.
- BARROS, R. P.; HENRIQUES, R.; MENDONÇA, R. S. Pelo fim das décadas perdidas: educação e desenvolvimento sustentado no Brasil. In: HENRIQUES, R. (Org.). *Desigualdade e Pobreza no Brasil*. Rio de Janeiro: IPEA, v.1, p. 405-423, 2000.
- BARROS, R. P.; MENDONÇA, R. S.; SANTOS, D.; QUINTAES, G. Determinantes do desempenho educacional no Brasil. *Pesquisa e Planejamento Econômico* (Rio de Janeiro), v. 31, n.1, p. 1-42, 2001.

BECKER, G. S. Human Capital Revisited. In: BECKER, G. S. *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education* (3rd Edition). The University of Chicago Press, p. 15-28, 1994. Disponível em: <<http://www.nber.org/chapters/c11229>>. Acesso em: 14 jan. 2015.

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

BOURDIEU, P. *Escritos de educação*. Petrópolis: Vozes, 1998.

BRASIL. *Constituição* (1988). Constituição da República Federativa do Brasil. Brasília, DF: Senado, 1988. Disponível em: <<https://www.planalto.gov.br/>>.

BROOKE, N.; SOARES, J. F. *Pesquisa em eficácia escolar: origem e trajetórias*. Belo Horizonte: EdUFMG, 2008.

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

BRYK, S., RAUDENBUSH, W. A hierarquical model for studying school effects. *Sociology of Education*, v. 59, p. 1-17, 1986.

CAMARANO, A. A.; MELLO, J. L.; Introdução. In: CAMARANO, A. A. (Org.). *Transição para a vida adulta ou vida adulta em transição?* Rio de Janeiro: IPEA, p. 13-28, 2006.

CARRASCO, G. G. *Estimação das escalas dos construtos capital social, capital cultural e capital econômico e análise do efeito escola nos dados do Peru-PISA 2000*. Dissertação (Mestrado em Engenharia Elétrica), Departamento de Engenharia Elétrica. Rio de Janeiro: PUC-Rio, 2005.

CASTRO, C. M. Avaliar não é para amadores. In: Souza, A. M. (org.) *Dimensões da avaliação educacional*. Petrópolis, RJ: Vozes, 2005.

CASTRO, M. H. G. Sistemas de avaliação da educação no Brasil: avanços e novos desafios. *São Paulo em Perspectiva*, São Paulo, Fundação Seade, v. 23, n. 1, p. 5-18, jan./jun. 2009. Disponível em: <<http://www.seade.gov.br>>. Acesso em: 25 nov. 2014.

CHEN, Q. A multilevel analysis of mathematically low-achieving students in Singapore. In: *5th IEA International Research Conference*, p. 26-28, 2013.

COLEMAN, J. S.; CAMPBELL, E. Q.; HOBSON, C. J.; McPARTLAND, J.; MOOD, A. M.; WEINFELD, F. D.; YORK, R. L. *Equality of Educational Opportunity*. Washington: Office of education and Welfare, 1966.

COLEMAN, J. S. Social capital in the creation of human capital. *American Journal of Sociology*, p. S95-S120, 1988.

COURI, C. Nível socioeconômico e cor/raça em pesquisas sobre efeito-escola. *Estudos em Avaliação Educacional*. v.1, n.47, p.449-472, set./dez., 2010.

CURI, A. Z.; SOUZA, A. P. Medindo a qualidade das escolas: Evidências para o Brasil. *Economia Aplicada*, v. 19, n. 3, p. 541-574, 2015.

DA COSTA, F. M. *O modelo de Samejima no estudo da relação NSE × desempenho via teoria da resposta ao item*. Dissertação (Mestrado em Estatística), Programa de Pós-graduação em Matemática e Estatística. Belém: Universidade Federal do Pará, 2010.

DATAVIVA. *Cursos Básicos*. Disponível em: <http://www.dataviva.info/pt/>. Acesso em: 23 Set. 2016.

DEBEAUV AIS, M. The concept of human capital. In: *Economics of education. UNESCO International Social Science Journal*, vol. XIV, n. 4, 1962. Disponível em: <http://unesdoc.unesco.org/images/0001/000173/017399eo.pdf>. Acesso em: 14 jan. 2015.

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

DIAS, V. M.; FERRÃO, M. E. Modelo Multinível do Desempenho Escolar de Alunos Socialmente Desfavorecidos em Escolas Públicas/Privadas—Aplicação aos Dados Portugueses do PISA 2000. *Psicologia e Educação*, v. 5, n.2, 2006.

ERMISCH, J.; FRANCESCONI, M. Family matters: Impacts of family background on educational attainments. *Economica*, v. 68, n. 270, p. 137-156, 2001.

FERRÃO, M. E.; BELTRÃO, K. I.; DOS SANTOS, D. P. O impacto de políticas de não-repetência sobre o aprendizado dos alunos da 4ª série. *Pesquisa e Planejamento Econômico*, v. 32, n. 3, 2002.

FERRÃO, M. E.; FERNANDES, C. O efeito-escola e a mudança-dá para mudar? Evidências da investigação Brasileira. *REICE: Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, v. 1, n. 1, p. 1-13, 2003.

FERNANDES, R. Índice de Desenvolvimento da Educação Básica (IDEB) 2006. Brasília: INEP, 2007. (*Texto para Discussão*, n. 26). Acesso em: 03 dez. 2014.

FIELDING, A.; GOLDSTEIN, H. Cross-classified and multiple membership structures in multilevel models: An introduction and review. *Research Report n. 791*, University of Birmingham, 2006.

FRANCO, A. M. P. *Os determinantes da qualidade da educação no Brasil*. 2008. 149f. Tese (Doutorado em Economia) – Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, 2008.

FRANCO, C.; ALVES, F.; BONAMINO, A. Qualidade do ensino fundamental: políticas, suas possibilidades, seus limites. *Educ. Soc.*, Campinas, vol. 28, n. 100 - Especial, p. 989-1014, out. 2007.

FRANCO, C.; ORTIGÃO, I.; ALBERNAZ, A.; BONAMINO, A.; AGUIAR, G.; ALVES, F.; SÁTYRO, N. Qualidade e equidade em educação: reconsiderando o significado de “fatores intra-escolares”. *Ensaio: Avaliação e Políticas Públicas em Educação*, v. 15, n. 55, p. 277-298, 2007.

FREEMAN, C. The ‘National System of Innovation’ in historical perspective. *Cambridge Journal of Economics*, v. 19, p. 5-24, 1995.

FREITAS, L. C.; SORDI, M. R. L.; MALAVASI, M. M. S.; FREITAS, H. C. L. *Avaliação educacional: caminhando pela contramão*. Petrópolis, RJ: Vozes, 2009.

FURR, R. M.; BACHARACH, V. R. *Psychometrics: an introduction*. Sage, 2008.

GELMAN, A. Multilevel (hierarchical) modeling: what it can and cannot do. *Technometrics*, v. 48, n. 3, p. 432-435, 2006.

GOLDSTEIN, H. *Multilevel Statistical Models*. New York: John Wiley & Sons, 1995.

GOMES-NETO, J. B.; HANUSHEK, E. A. Causes and consequences of grade repetition: Evidence from Brazil. *Economic Development and Cultural Change*, v. 43, n. 1, p. 117-148, 1994.

GREMAUD, A. P.; FELÍCIO, F.; BIONDI, R. L. Indicador de efeito escola: uma metodologia para a identificação dos sucessos escolares a partir dos dados da Prova Brasil. Brasília: INEP, 2007. (*Texto para Discussão n. 27*).

HANUSHEK, E. A. Publicly provided education. *Handbook of public economics*, v. 4, p. 2045-2141, 2002.

HANUSHEK, E. A.; LAVY, V.; HITOMI, K. Do students care about school quality? Determinants of dropout behavior in developing countries. *Journal of Human Capital*, v. 2, n. 1, p. 69-105, 2008.

HEYNEMAN, S. P. Avaliação da qualidade da educação: lições para o Brasil. In: SOUZA, A. M. (org.) *Dimensões da avaliação educacional*. Petrópolis, RJ: Vozes, 2005.

HILL, P. W.; ROWE, K. J. Multilevel modelling in school effectiveness research. *School effectiveness and school improvement*, v. 7, n. 1, p. 1-34, 1996.

HONDA, K. M. Um estudo sobre os determinantes do atraso escolar. Dissertação (Mestrado em Economia Aplicada), Faculdade de Economia, Administração e Contabilidade. Ribeirão Preto: Universidade de São Paulo, 2006.

HOX, J. *Multilevel analysis: techniques and applications*. Mahwah: Lawrence Erlbaum Associates, 2002.

_____. *Multilevel analysis: techniques and applications*, 2^a ed. New York: Routledge, 2010.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). Síntese de Indicadores Sociais: uma análise das condições de vida da população brasileira. *IBGE, Rio de Janeiro*, n.36, 146p., 2016a.

_____. *Sistema IBGE de Recuperação Automática*: SIDRA. Disponível em: <<http://www.sidra.ibge.gov.br/>>. Acesso em: jul. 2016b.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (INEP). SAEB 2001: novas perspectivas. Brasília: INEP, 2001.

_____. O desafio de uma educação de qualidade para todos: educação no Brasil – 1990-2000. Brasília: INEP, 2004.

_____. *PISA*. Brasília. Disponível em: <<http://portal.inep.gov.br/pisa-programa-internacional-de-avaliacao-de-alunos>>. Acesso em: 23 dez. 2014.

_____. Nota explicativa resultados Prova Brasil 2013. Brasília: INEP, 2014a. Disponível em: <http://download.inep.gov.br/mailng/2014/nota_explcativa_prova_brasil_2013.pdf>. Acesso em: 08 jan. 2015.

_____. Nota Técnica Indicador de Nível Socioeconômico das Escolas de Educação Básica (Inse). Brasília: INEP, 2014b. Disponível em: <http://download.inep.gov.br/mailng/2014/nota_tecnica_inse.pdf>. Acesso em: 08 jan. 2015.

_____. Escalas da Prova Brasil e Saeb. Disponível em: <<http://provabrasil.inep.gov.br/escalas-da-prova-brasil-e-saeb>>. Acesso em: 27 set. 2016.

_____. *Taxas de distorção idade-série*. Disponível em: <http://download.inep.gov.br/informacoes_estatisticas/indicadores_educacionais/2015/distorc ao_idade_serie/tdi_brasil_regioes_UFs_2015.zip>. Acesso em: 16 jan. 2017.

INSTITUTO DE PESQUISA ECONÔMICA APLICADA (IPEA). *Ipeadata*: dados regionais. 2012. Disponível em: <<http://www.ipeadata.gov.br>>. Acesso em: 12 jul. 2016.

_____. Educação (Capítulo 4). In: *Boletim de Políticas Sociais: acompanhamento e análise*. – Brasília: Ipea, n. 22, 2014.

LAU, L. J.; JAMISON, D. T.; LOUAT, F. F. Education and productivity in developing countries: An aggregate production function approach. *World Bank Publications*, 1991.

LAROS, J. A.; MARCIANO, J. L. P. Análise multinível aplicada aos dados do NELS: 88. *Estudos em Avaliação Educacional*, v. 19, n. 40, p. 263-278, 2008a.

_____. Índices educacionais associados à proficiência em Língua Portuguesa: um estudo multinível. *Avaliação Psicológica*, v.7, n.3, p.371-359, 2008b.

LEE, V. E. Using hierarchical linear modeling to study social contexts: the case of school effects. *Educational Psychologist*, v.35, n.2, p. 125-141, 2000.

LEITE, D. B. Avaliação da educação superior. In: MOROSINI, M. (Org.). *Enciclopédia de pedagogia universitária - Glossário*. Brasília: INEP, v.2, p.459-506, 2006.

LEON, F. L. L.; MENEZES-FILHO, N. A. Reprovação, avanço e evasão escolar no Brasil. *Pesquisa e Planejamento Econômico*, v. 32, n. 3, 2002.

LEWIS, A. Education and economic development. In: *Economics of education. UNESCO International Social Science Journal*, v. XIV, n. 4, 1962. Disponível em: <<http://unesdoc.unesco.org/images/0001/000173/017399eo.pdf>>. Acesso em: 14 jan. 2015.

LEME, M. C.; WAJNMAM, S. A alocação de tempo dos adolescentes brasileiros entre o trabalho e a escola. In: *Anais do Encontro Nacional da Associação Brasileira de Estudos Populacionais*, 2000.

LIMA, A. C. Ciclo de avaliação da educação básica do Ceará: principais resultados. *Estudos em Avaliação Educacional*, v. 23, n. 53, p. 38-58, 2012a.

LIMA, A. E. M. *Efeito do desempenho educacional sobre o repasse da cota parte do ICMS para os municípios cearenses*. Dissertação (mestrado profissional em economia do setor público - Cariri), Programa de Pós Graduação em Economia, CAEN, Fortaleza: Universidade Federal do Ceará, 2012b.

LIMA, R. G. O que fazer? Entre a escola e o mercado de trabalho: as trajetórias percorridas pelos jovens brasileiros nas últimas décadas. In: *34º Encontro Anual da ANPOCS*, 2010.

LUKE, D. A. *Multilevel modeling. Quantitative applications in the social sciences*. Sage, v. 143, 2004.

MACHADO, A. F., MORO, S., MARTINS, L.; RIOS, J. Qualidade do ensino em Matemática: determinantes do desempenho de alunos em escolas públicas estaduais mineiras. *Revista da Anpec*, v. 9, n. 1, 2008.

MARRI, I.; RACCHUMI, J. Infraestrutura escolar e desempenho educacional em Minas Gerais: possíveis associações. In: *XVIII Encontro Nacional de Estudos Populacionais*, ABEP, Águas de Lindóia, p. 19-23, 2012.

MCMAHON, J. M.; POUGET, E. R.; TORTU, S. A guide for multilevel modeling of dyadic data with binary outcomes using SAS PROC NLMIXED. *Computational statistics & data analysis*, v. 50, n. 12, p. 3663-3680, 2006.

MENEZES-FILHO, N. A. Os determinantes do desempenho escolar do Brasil. In: DELFIN NETO, A. (coord.); DUARTE, P. G.; SILBER, S.; GUILHOTO, J. (Org.). *O Brasil do século XXI*. São Paulo: Saraiva, 2011, v. 1, p. 231-256.

MINCER, J. Investment in human capital and personal income distribution. *The Journal of Political Economy*, v. 66, n. 4, p. 281-302, 1958.

MORAES, A. G. E.; BELLUZZO, W. O diferencial de desempenho escolar entre escolas públicas e privadas no Brasil. *Nova Economia* (UFMG. Impresso), v. 24, p. 409-430, 2014.

MOREIRA, K. S. G.; JACINTO, P. A. Determinantes da proficiência em Matemática no Rio Grande do Sul: uma análise a partir de modelos hierárquicos. In: *XVI Encontro de Economia da Região Sul*, Curitiba, 2013. Disponível em: <http://www.anpec.org.br/sul/2013>. Acesso em: 06 jan. 2015.

NATIS, L. Modelos Hierárquicos Lineares. *Estudos em Avaliação Educacional*. n. 23, jan-jun, p. 3-29, 2001.

NERI, M. *Motivos da evasão escolar*. Brasília: Fundação Getúlio Vargas, 2009.

NETO, J. J. S.; JESUS, G. R.; KARINO, C. A.; ANDRADE, D. F. Uma escala para medir a infraestrutura escolar. *Estudos em Avaliação Educacional*, v. 24, n. 54, p. 78-99, 2013.

NOGUEIRA, M. A.; CATANI, A. Uma sociologia da produção do mundo cultural e escolar. In: NOGUEIRA, M. A.; CATANI, A. (Org.). *Escritos de Educação*. Petrópolis: Vozes, p. 7-15, 1998.

OLIVEIRA, J. M. Custo-efetividade de políticas de redução do tamanho da classe e ampliação da jornada escolar: uma aplicação de estimadores de matching. Rio de Janeiro: *Revista do BNDES*, n.33, p. 281-306, 2010.

OLIVEIRA, K. M. M. *Um estudo da evolução da GQT por meio do modelo de resposta gradual da teoria da resposta ao item*. Dissertação (Mestrado em Gestão Logística e Pesquisa Operacional) Pró-reitoria de Pesquisa e Pós-Graduação, Universidade Federal do Ceará, Fortaleza, 2010.

OLIVEIRA, L. F. B.; SOARES, S. Determinantes da Repetência Escolar no Brasil: Uma Análise do Painel dos Censos Escolares Entre 2007 e 2010. (*Texto para discussão 1706*). Instituto de Pesquisa Econômica Aplicada-IPEA, 2012.

ORTIGÃO, M. I. R.; AGUIAR, G. S. Repetência escolar nos anos iniciais do ensino fundamental: evidências a partir dos dados da Prova Brasil 2009. *Revista Brasileira de Estudos Pedagógicos*, v. 94, n. 237, 2013.

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 (Impresso)*, v. 31, p. 367-394, 2014.

PATRINOS, H. A, PSACHAROPOULOS, G. Family size, schooling and child labor in Peru: an empirical analysis. *Journal of Population Economics*, v. 10, n. 4, p.387-405, 1997.

PEREIRA, V. R. *Métodos Alternativos no Critério Brasil para Construção de Indicadores Sócio-Econômico: Teoria da Resposta ao Item*. Dissertação (Mestrado em Engenharia Elétrica), Departamento de Engenharia Elétrica. Rio de Janeiro: Pontifícia Universidade Católica do Rio de Janeiro, 2004.

PERRENOUD, P. *Avaliação: da excelência à regularização das aprendizagens: entre duas lógicas*. Porto Alegre: Artmed, 1999.

- PIRES, V. *Economia da Educação: para além do capital humano*. São Paulo: Cortez, 2005.
- PLÁ, N. B. El Modelo de Rasch aplicado a la medición multidimensional de la pobreza en España: bienaventurados los que tienen tele, aunque no coman carne. In: *XVI Encuentro de Economía Pública*, Granada, 2009.
- PONTILLI, R. M.; KASSOUF, A. L. Fatores que afetam a frequência e o atraso escolar, nos meios urbano e rural, de São Paulo e Pernambuco. *Revista de Economia e Sociologia Rural*, v. 45, n. 1, p. 27-47, 2007.
- PROGRAMA DAS NAÇÕES UNIDAS PARA O DESENVOLVIMENTO (PNUD). *Ranking IDHM Unidades da Federação 2010*. Disponível em: <<http://www.pnud.org.br/>>. Acesso: 12 jul. 2016.
- PSACHAROPOULOS, G. Child labor versus educational attainment Some evidence from Latin America. *Journal of Population Economics*, v. 10, n. 4, p. 377-386, 1997.
- RAUDENBUSH, S. W.; BRYK, A. S. *Hierarchical linear models: Applications and data analysis methods*. London: Sage, 2002.
- RIANI, J. L. R. *Determinantes do resultado educacional no Brasil: família, perfil escolar dos municípios e dividendo demográfico numa abordagem hierárquica e espacial*. Tese (Doutorado em Demografia), Centro de Desenvolvimento e Planejamento Regional. Belo Horizonte: Universidade Federal de Minas Gerais, 2005.
- RIANI, J. L. R.; DA SILVA, V. C.; SOARES, T. M. Repetir ou progredir? Uma análise da repetência nas escolas públicas de Minas Gerais. *Educação e Pesquisa*, São Paulo, v.38, n.03, p. 623-636, 2012.
- RIBEIRO, R. A.; CACCIAMALI; M. C. Defasagem Idade-Série a partir de distintas perspectivas teóricas. *Revista de Economia Política*, v. 32, n. 3 (128), p. 497-512, 2012.
- RIBEIRO, R. A.; NEDER, H. D. Juventude (s): desocupaçāo, pobreza e escolaridade. *Nova Economia (UFMG. Impresso)*, v. 19, p. 475-508, 2009.
- RIOS-NETO, E. L. G; CÉSAR, C. C.; RIANI, J. L. R. Estratificação educacional e progressão escolar por série no Brasil. *Pesquisa e Planejamento Econômico* (Rio de Janeiro), v. 32, n. 3, p. 395-415, 2002.
- RIVKIN, S. G.; HANUSHEK, E. A.; KAIN, J. F. Teachers, Schools, and Academic Achievement. *Econometrica*, v. 73, n. 2, p. 417-458, 2005.
- SAMEJIMA, F. A. Estimation of latent ability using a response pattern of graded scores. *Psychometric Monograph*, 17, 1969.
- SANTOS, F. J. S., SAMPAIO, R. M. B.; SAMPAIO, L. M. B. Eleição nas escolas: uma análise do impacto do diretor eleito sobre o desempenho educacional no estado da Bahia. *Pesquisa e Planejamento Econômico* (Rio de Janeiro), v. 46, n. 1, p. 177-209, 2016.

SCHIEFELBEIN, E.; SIMMONS, J. Os determinantes do desempenho escolar: uma revisão de pesquisas nos países em desenvolvimento. *Cadernos de Pesquisa*, n. 35, p. 53-71, 1980. Disponível em: <http://publicacoes.fcc.org.br/ojs/index.php/cp/article/view/1649>. Acesso: 18/03/2016.

SCHULTZ, T. W. Investment in human capital. *The American Economic Review*, v. 51, n. 1, p. 1-17, 1961.

SCHWARTZMAN, S. As avaliações de nova geração. In: SOUZA, A. M. (org.) *Dimensões da avaliação educacional*. Petrópolis, RJ: Vozes, 2005.

SHERA, P. School effects, gender and socioeconomic differences in reading performance: a multilevel analysis. *International Education Studies*, v. 7, n. 11, p. 28, 2014.

SHIRASU, M. R.; ARRAES, R. A. Determinantes da evasão e repetência escolar. In: Anais do *XLIII Encontro Nacional de Economia*. ANPEC. Florianópolis, 2016.

SILVA, N. V.; HASENBALG, C. Recursos familiares e transições educacionais. *Cad. saúde pública*, v. 18, n. suppl., p. 67-76, 2002.

SNIJDERS, T. A. B.; BERKHOF, J. Diagnostic checks for multilevel models. In: *Handbook of multilevel analysis*. Springer New York, p. 141-175, 2008.

SOARES, J. F. O efeito da escola no desempenho cognitivo de seus alunos. In: Souza, A. M. (org.) *Dimensões da avaliação educacional*. Petrópolis, RJ: Vozes, 2005.

_____. Avaliação da qualidade da educação escolar brasileira. In: SWARTZMAN, F. F.; SWARTZMAN, I. F.; SWARTZMAN, L. F; SWARTZMAN, M. L. (Orgs.). *O Sociólogo e as Políticas Públicas*. 1ed. Rio de Janeiro: FGV Editora, p. 215-241, 2009.

SOARES, J. F.; ALVES, M. T. G. Efeitos de escolas e municípios na qualidade do ensino fundamental. *Cadernos de Pesquisa*, São Paulo, v.43, n.149, p.492-517, 2013.

SOARES, S. Os fatores que determinam o sucesso educacional. *Pesquisa e Planejamento Econômico* (Rio de Janeiro), v. 32, n.3, p. 385-394, 2002.

SOARES, S.; SÁTYRO, N. O impacto da infra-estrutura escolar na taxa de distorção idade-série das escolas brasileiras de ensino fundamental – 1998-2005. Rio de Janeiro: IPEA, 2008. (*Texto para Discussão* n. 1338). Disponível em: <http://www.ipea.gov.br/portal/images/stories/PDFs/TDs/td_1338.pdf>. Acesso em: 25 nov. 2014.

SOARES, S.; NASCIMENTO, P. A. M. M. Evolução do desempenho cognitivo dos jovens brasileiros no PISA. *Cadernos de Pesquisa*, São Paulo, v. 42, n. 145, p. 68-87, 2012. Disponível em: <<http://www.scielo.br/pdf/cp/v42n145/06.pdf>>. Acesso em: 03 dez. 2014

SOARES, T. M. Influência do professor e do ambiente em sala de aula sobre a proficiência alcançada pelos alunos avaliados no SIMAVE-2002. *Estudos em Avaliação Educacional*. São Paulo, n. 28, p. 103-124, 2003.

_____. Modelo de três níveis hierárquicos para a proficiência dos alunos de 4^a série avaliados no teste de Língua Portuguesa do SIMAVE/ PROEB-2002. *Revista Brasileira de Educação*, São Paulo, v. 29, p. 73-87, 2005a.

_____. Utilização da Teoria da Resposta ao Item na produção de indicadores socioeconômicos. *Pesquisa Operacional*, Rio de Janeiro, v.25, n.1, 2005b.

SOUZA, M. M.; LAROS, J. A. Satisfação no trabalho e responsabilidade social empresarial: uma análise multinível. *Revista Psicologia: Organizações e Trabalho*, v. 10, n. 2, jul-dez, p. 21-37, 2010.

SURYADARMA, D.; SURYAHADI, A., SUMARTO, S., ROGERS, H. The determinants of student performance in Indonesian public primary schools: the role of teachers and schools. Washington, DC and Jakarta, Indonesia: *World Bank and SMERU*, 2004.

TEDDLIE, C.; REYNOLDS, D. *The International handbook of school effectiveness research*. Nova York: Falmer Press, 2000.

TODOS PELA EDUCAÇÃO (TPE). *De olho nas metas: Primeiro relatório de acompanhamento das 5 metas do Movimento Todos Pela Educação*. São Paulo: Movimento Todos Pela Educação, 2008.

TWISK, J. W. R. *Applied Longitudinal Data Analysis for Epidemiology: A Practical Guide*. New York: Cambridge University Press, 2006.

UNITED NATIONS EDUCATIONAL SCIENTIFIC AND CULTURAL ORGANIZATION (UNESCO). *Readings in the economics of education*. Paris: UNESCO, 1968.

_____. *Declaração mundial sobre educação para todos: satisfação das necessidades básicas de aprendizagem*. Jomtien, 1990.

VASCONCELLOS, L. A relação entre freqüência escolar e renda familiar no Brasil – 1981-1999. *Pesquisa e Planejamento Econômico*, v. 35, n. 2, p. 267-296, 2005.

VAIZEY, J. Introduction. In: *Economics of education*. UNESCO *International Social Science Journal*, vol. XIV, n. 4, 1962. Disponível em: <http://unesdoc.unesco.org/images/0001/000173/017399eo.pdf>. Acesso em: 14 jan. 2015.

VERNIER, L. D. S.; BAGOLIN, I. P.; JACINTO, P. A. Fatores que influenciam o desempenho escolar no estado do Rio Grande do Sul: uma análise com regressões quantílicas. *Análise Econômica*, v. 33, n. 64, p.143-170, 2015.

VIEIRA, S. L. Gestão, avaliação e sucesso escolar: recortes da trajetória cearense. *Estudos avançados*, v. 21, n. 60, p. 45-60, 2007.

WENGLINSKY, H. Teacher classroom practices and student performance: How schools can make a difference. *ETS Research Report Series*, v. 2001, n. 2, p. i-37, 2001.

WOLFF, L. Avaliações educacionais na América Latina: estágio atual e desafios futuros. Rio de Janeiro: Preal-Brasil; Fundação Getúlio Vargas, 1998. (*Documentos n.11*).

ZHAO, N.; VALCKE, M.; DESOETE, A.; VERHAEGHE, J.; XU, K. A multilevel analysis on predicting mathematics performance in Chinese primary schools: Implications for practice. *Asia-Pacific Education Researcher*, v. 20, n. 3, p. 503-520, 2011.

Apêndice: O modelo de Teoria de Resposta ao Item

Na década de 1960, Georg Rasch desenvolveu um modelo objetivando ter uma medida prática para os resultados de testes educacionais, por meio da avaliação do desempenho escolar de soldados dinamarqueses. Atualmente esse modelo tem sido muito utilizado, sendo aplicado a diversas áreas para mensurar fenômenos complexos do mundo social com base em vários determinantes. Entre as aplicações estão às ciências sociais em geral, a avaliação da qualidade, psicometria e educação. Acredita-se que este modelo seja um bom instrumento para medir variáveis latentes. Logo, tem sido usado extensivamente para avaliar um fenômeno latente (identificação de dimensões que estão presentes na realidade a ser estudada), não observável diretamente, a partir de respostas a diferentes perguntas sobre um teste (itens) em diferentes indivíduos (FURR, 2008).

Destaca-se que o que distingue a TRI de outros indicadores é o fato de considerar dois fatores para medir a variável latente: a capacidade individual e a dificuldade do item. A avaliação de cada indivíduo é feita baseada nas respostas positivas de cada item e também considerando o grau de dificuldade de cada questão. O modelo representa estes dois fatores por meio de duas famílias de parâmetros: parâmetros de dificuldade para os itens e parâmetros de capacidade para os indivíduos. O parâmetro de dificuldade cria um ranking dos fatores cuja falta pode contribuir principalmente para gerar um resultado negativo e descreve as desigualdades dos itens entre os indivíduos. Tem-se que, quanto maior for esse parâmetro, mais difundida é a propriedade do bem, e maior é o grau de desigualdade correspondente à sua falta. Já o parâmetro de capacidade é considerado como um parâmetro de posição, pois ele permite classificar os elementos de análise. Assim, quanto maior o parâmetro de posição, pior o desempenho (PLÁ, 2009).

O modelo mais comum seria o próprio modelo de Rasch ou modelo logístico de um parâmetro (1PL), em que a resposta de um indivíduo a um item binário é determinado pelo nível característica do indivíduo e da dificuldade do item. Esse modelo pode ser expresso em relação à probabilidade que um indivíduo com um nível de traço particular irá responder corretamente a um item que tem uma determinada dificuldade (FURR, 2008). Usualmente essa relação é expressa por:

$$P(X_{ij} = 1 | \theta_j, \beta_i) = \frac{e^{(\theta_j - \beta_i)}}{1 + e^{(\theta_j - \beta_i)}} \quad (19)$$

onde X_{ij} refere-se à resposta (X) dada pelo sujeito j ao item i ; θ_j refere-se ao nível de traço do sujeito j ; β_i equivale à dificuldade do item i ; $X_{ij} = 1$ refere-se à resposta “correta” ou aprovação do item; e é à base do logaritmo natural. Portanto, $P(X_{ij} = 1 | \theta_j, \beta_i)$ nos dá a probabilidade (P) que o sujeito j irá responder corretamente ao item i , condicionada ao nível de traço do sujeito e a dificuldade do item. Destaca-se que a avaliação de qual é o modelo mais ajustado dependerá do conjunto particular de dados.

Para a construção dos indicadores a partir das respostas aos questionários contextuais foi empregado o Modelo de Respostas Graduais de Samejima (1969). O objetivo do autor era obter maiores informações, além de se considerar se as respostas aos itens eram corretas ou incorretas (ANDRADE; TAVARES; VALLE, 2000). Assim, seu modelo assume que as categorias de respostas podem ser ordenadas entre si.

O modelo de respostas graduadas é uma generalização do modelo logístico de dois parâmetros (2PL) para o caso de mais de duas categorias ordenadas de respostas. Suponha que os escores das categorias de um item i são dispostos em ordem do menor para o maior e denotados por $k = 0, 1, \dots, m_i$ em que $(m_i + 1)$ corresponde ao número de categorias do i -ésimo item. Conforme descrito em Andrade, Tavares e Valle (2000), a probabilidade que um indivíduo j escolha uma determinada categoria do item i é expressa como segue:

$$P_{i,k}^+(\theta_j) = \frac{1}{1 + e^{-D a_i (\theta_j - b_{i,k})}} \quad (20)$$

onde $i = 1, 2, \dots, I$, $j = 1, 2, \dots, n$, e $k = 0, 1, \dots, m_i$; $P_{i,k}^+(\theta_j)$ representa a probabilidade de um indivíduo j escolher a categoria k ou outra mais alta do item i ; $b_{i,k}$ nos dá o parâmetro de dificuldade da k -ésima categoria do item i ; θ_j representa o traço latente do j -ésimo indivíduo; a_i é o parâmetro de discriminação do item i , com valor proporcional à curva de inclinação da curva e característica do item; e D é um fator de escala constante e igual a 1.

Espera-se que haja uma ordenação entre o nível de dificuldade das categorias de um dado item, conforme a classificação de seus escores. Assim, temos que a probabilidade de um indivíduo j receber um escore k é:

$$P_{i,k}(\theta_j) = P_{i,k}^+(\theta_j) - P_{i,k+1}^+(\theta_j) \quad (21)$$

Segundo Samejima (1969), $P_{i,0}^+(\theta_j) = 1$ e $P_{i,m_i+1}^+(\theta_j) = 0$, disso chegamos que:

$$P_{i,k}(\theta_j) = \frac{1}{1 + e^{-Da_i(\theta_j - b_{i,k})}} - \frac{1}{1 + e^{-Da_i(\theta_j - b_{i,k+1})}} \quad (22)$$

Assim, para cada item o número de parâmetros que serão estimados é representado pelo número de categorias de resposta. O modelo irá estimar o valor de θ conforme os valores que forem atribuídos as respostas dos itens (DA COSTA, 2010).

O *software* usado para a elaboração dos escores foi o R, sendo o processo de estimação o método de máxima verossimilhança. Foram construídas quatro variáveis por meio de TRI: nível socioeconômico e o nível de capital social do aluno, índice de posse de equipamentos e índice de conservação da escola. Os resultados foram validados quantitativamente, calculando-se a correlação polisserial. A correlação polisserial, utilizada no caso de itens politônicos²⁶ (obtidos a partir de variáveis categóricas), representa uma correlação de Pearson entre o escore do item e escore bruto do teste (SOARES, 2005b). Ou seja, mensura a capacidade de discriminação do item em relação ao resultado do teste, que examina se há correlação significativa entre o escore bruto produzido pelo conjunto de itens, mensurando segundo uma escala ordenada que as diferenças entre dois valores consecutivos seja sempre iguais. Desse processo obtêm-se os itens que possuem consistência interna e que tem boa associação com o escore produzido (OLIVEIRA, 2010).

Neste trabalho, os resultados obtidos apresentaram todas correlações superiores a 0,3 indicando que os itens apresentam uma boa correlação com o escore bruto produzido pelo conjunto de itens²⁷. Ressalva-se que o método exclui aqueles itens para os quais os indivíduos não fornecem resposta. Além disso, optou-se por não fazer nenhuma transformação na escala original.

²⁶ São itens que se caracterizam por todo um conjunto ordenado de respostas e não apenas pela presença do fator certo ou errado.

²⁷ Conforme sugere Soares (2005b) é usual aceitar valores superiores a 0,3 para a correlação no processo de escolha dos itens que apresentam consistência interna e se relacionam bem ao escore que será gerado.