



UNIVERSIDADE FEDERAL DO CEARÁ
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO, ATUÁRIA E CONTABILIDADE
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

ERNESTO DOS SANTOS VASCONCELOS

ENSAIOS SOBRE AVALIAÇÃO DO ENSINO FUNDAMENTAL NO BRASIL

FORTALEZA

2019

ERNESTO DOS SANTOS VASCONCELOS

ENSAIOS SOBRE AVALIAÇÃO DO ENSINO FUNDAMENTAL NO BRASIL

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Economia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de doutor em Economia. Área de concentração: Econometria.

Orientador: Prof. Ronaldo de Albuquerque e Arraes.

FORTALEZA

2019

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária

Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

- V45e Vasconcelos, Ernesto dos Santos.
Ensaio sobre avaliação do ensino fundamental no Brasil / Ernesto dos Santos Vasconcelos. – 2019.
114 f. : il. color.
- Tese (doutorado) – Universidade Federal do Ceará, Faculdade de Economia, Administração, Atuária e Contabilidade, Programa de Pós-Graduação em Economia, Fortaleza, 2019.
Orientação: Prof. Dr. Ronaldo de Albuquerque e Arraes.
1. Avaliação educacional. 2. Gerações sobrepostas na educação. 3. Desigualdades na educação. 4. Modelos multinomiais. I. Título.

CDD 330

ERNESTO DOS SANTOS VASCONCELOS

ENSAIOS SOBRE AVALIAÇÃO NO ENSINO FUNDAMENTAL DO BRASIL

Tese apresentada ao Programa de Pós-Graduação em Economia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de doutor em Economia. Área de concentração: Econometria.

Aprovada em: ___/___/_____.

BANCA EXAMINADORA

Prof. Dr. Ronaldo de Albuquerque e Arraes (Orientador)
Universidade Federal do Ceará(UFC)

Profa. Dra. Francisca Zilania Mariano
Universidade Federal do Ceará (UFC – Campus Sobral)

Profa. Dra. Vera Lúcia da Silva
Universidade de Fortaleza (UNIFOR)

Dra. Maitê Rimekká Shirasu
Universidade Federal do Ceará (UFC)

Profa. Dra. Carla Cíntia Andrade de Abreu
Universidade Federal do Ceará (UFC)

Ao Senhor DEUS, que criou o céu e a terra.

À minha linda e doce família, Aurilene
(Leninha), Gabriel e Natália.

Aos meus pais, Zé Maria e Francisca.

AGRADECIMENTOS

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico – Brasil (CNPQ) e à Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico – Ceará (FUNCAP), pelo apoio financeiro com a manutenção da bolsa de auxílio durante o período de Doutorado.

Ao Prof. Dr. Ronaldo de Albuquerque e Arraes, pela excelente orientação, tantas vezes oportuna e precisa.

Aos professores participantes da banca examinadora, Profa. Dra. Francisca Zilania Mariano, Profa. Dra. Vera Lúcia da Silva, Dra. Maitê Rimekká Shirasu e Dra. Carla Cíntia Andrade de Abreu, pelas valiosas colaborações e sugestões.

Aos professores do corpo de docentes do CAEN, pela excelência de seus conhecimentos e experiência repassados, durante a formação no Doutorado.

Ao corpo de funcionários do Programa de Pós-Graduação em Economia – CAEN, por disponibilidade, atenção, zelo e auxílio concedidos sempre que se fez necessário.

Aos colegas da turma de doutorado, pelas reflexões, críticas e sugestões recebidas.

RESUMO

A presente tese reúne capítulos, cujo tema central é a trajetória do desempenho escolar na educação básica no Brasil. Tem por objetivo identificar os fatores que explicam o desempenho escolar dos alunos do ensino fundamental analisando se existe um acúmulo de conhecimento entre os ciclos escolares, se os fatores que explicam o desempenho escolar persistem entre diferentes gerações e se as desigualdades de desempenho escolar podem ser devidas à diferenças socioeconômicas dos alunos. Com esse propósito, nos três capítulos, utilizou-se a regressão logística ordinal na modelagem econométrica. No primeiro, avaliou-se a relação e os efeitos dos fatores extraescolares – gênero, raça, classe social, região geográfica e rede (pública, privada) – no desempenho de proficiência em leitura de estudantes do Ensino Fundamental no Brasil, no interstício 1999 a 2017 abrangendo o período em que os resultados alcançados pelos estudantes brasileiros no Programa Internacional de Avaliação de Alunos (PISA) em proficiência em leitura apresentaram-se estagnados. Para tanto, utilizou-se de uma amostra composta por 3.851.542 alunos(as), divididos em duas etapas do Ensino Fundamental, e procedeu-se uma análise estatística descritiva e estimativas do modelo econométrico. Os resultados apontam para um menor desempenho em leitura acerca dos estudantes do segundo ciclo do Ensino Fundamental nos níveis superiores de proficiência. Constata-se quanto à desigualdade de desempenho escolar: ocorre entre estudantes de diferentes raças; quanto ao fator classe econômica: apresenta efeito positivo com chances maiores para os níveis mais altos de proficiências; quanto às Regiões Norte e Nordeste e à rede pública de ensino: apresentam efeito negativo e chances menores para níveis elevados de proficiência. No segundo capítulo, em uma abordagem de gerações sobrepostas e de modelagem econométrica com dados em corte transversal agrupados, avaliou-se a relação e o efeito dos fatores: gênero, raça, classe social e região geográfica, sobre desempenho escolar, para fins de comparação entre as gerações. Os resultados sugerem não haver mudanças significativamente positivas entre as gerações em relação a desempenho escolar (nível de proficiência em leitura), além de evidenciar uma desigualdade de ordem educacional, racial, regional e de classe. Nessa perspectiva, no último capítulo, conduziu-se uma análise sobre as condições socioeconômicas de crianças e adolescentes do 5º ano do Ensino Fundamental submetidas à avaliação do SAEB/Prova Brasil - edição 2015. O propósito foi esclarecer elementos que apontassem possíveis desigualdades de renda, de educação e de raça entre crianças e adolescentes. Os resultados evidenciaram não haver desigualdade educacional entre estudantes de diferentes raças. Além disso, os estudantes de famílias com menor poder aquisitivo apresentaram efeito positivo para desempenho escolar, com razão de chances menores de serem observados nos níveis inferiores de proficiência em leitura. Estudantes que moram com os pais e são incentivados por estes a irem à escola,

estudarem e não trabalharem fora de casa apresentam maiores chances de serem observados nos níveis superiores de proficiência em leitura.

Palavras-chave: Avaliação educacional. Gerações sobrepostas na educação. Desigualdades na educação. Modelos multinomiais.

ABSTRACT

This thesis brings together essays whose central theme is the trajectory of school performance in basic education in Brazil. In all three trials, ordinal logistic regression was used for econometric modeling. In the first essay, the relationship and the effects of extracurricular factors - gender, race, social class, geographic region and network (public, private) - on the reading proficiency performance of elementary school students in Brazil, from 1999 to 2017, covering the period when the results achieved by Brazilian students in the International Student Assessment Program (PISA) in reading proficiency were stagnant. For this, we used a sample composed of 3,851,542 students, divided into two stages of elementary school, and a descriptive statistical analysis and estimates of the econometric model were performed. The results point to a lower reading performance about the students of the second cycle of the Elementary School in the superior levels of proficiency. As for the inequality of school performance: it occurs among students of different races; As for the economic class factor: it has a positive effect with higher chances for higher proficiency levels; As for the North and Northeast Regions and the public school system: they have negative effect and lower chances for high levels of proficiency. In the second essay, using an overlapping generation and econometric modeling approach with clustered cross-sectional data, the relationship and effect of the factors: gender, race, social class and geographic region on school performance were evaluated for comparison purposes. between generations. The results suggest that there are no significant positive changes between generations regarding school performance (reading proficiency level), as well as evidence of educational, racial, regional and class inequality. In the third essay, an analysis was conducted on the socioeconomic conditions of children and adolescents of the 5th grade of elementary school submitted to the evaluation of SAEB / Prova Brasil - 2015 edition. The purpose was to clarify elements that pointed to possible inequalities of income, education and of race among children and adolescents. The results showed no educational inequality among students of different races, students with lower purchasing power had a positive effect on school performance with lower chances for lower levels of reading proficiency. Students who live with their parents and are encouraged to go to school, study, and not work outside the home are more likely to be observed at higher levels of reading proficiency.

Keywords: Educational evaluation. Overlapping generation and education. Inequalities in education. Multinomial models.

LISTA DE ILUSTRAÇÕES

Figura 1 – Desempenho das proficiências por gerações para o Ensino Fundamental	62
Quadro 1 – Quantidade de alunos (as) por ano e fases do Ensino Fundamental	25
Quadro 2 – Características; Variável; Codificação da categoria; <i>Dummy</i>	27
Quadro 3 – Descritores da escala de Proficiência em LP – 4ª Série/5º Ano	113
Quadro 4 – Descritores da escala de Proficiência em LP – 8ª Série/9º Ano	115
Quadro 5 – Escala de Proficiência em Língua Portuguesa – 4ª Série/5º Ano	29
Quadro 6 – Escala de Proficiência em Língua Portuguesa – 8ª Série/9º Ano	30
Quadro 7 – Tamanho amostral e censitário por ano e fases do Ensino Fundamental	54
Quadro 8 – Estrutura de gerações sobrepostas para o Ensino Fundamental	58
Quadro 9 – Demonstrativo de variáveis, Categorias/Codificação e <i>Dummy</i>	81
Quadro 10 – Estimativa para a renda média domiciliar para os estratos socioeconômicos	110

LISTA DE GRÁFICOS

Gráfico 1 – Proporção média de alunos por nível de proficiência na escala.....	30
Gráfico 2 – Proporção média por nível de proficiência (4ª Série/5º Ano)	37
Gráfico 3 – Proporção média por nível de proficiência (8ª Série/9º Ano)	37
Gráfico 4 – Proporção média de alunos por nível de proficiência de 1999 a 2017	38
Gráfico 5 – Desempenho por gênero (sexo) na 4ª série/5ºano – 1999/2017	105
Gráfico 6 – Desempenho por gênero (sexo) na 8ª série/9ºano – 1999/2017	105
Gráfico 7 – Desempenho por cor da pele (raça) na 4ª série/5ºano – 1999/2017	106
Gráfico 8 – Desempenho por cor da pele (raça) na 8ª série/9ºano – 1999/2017	106
Gráfico 9 – Desempenho por classe social na 4ª série/5ºano – 1999/2017	107
Gráfico 10 – Desempenho por classe social na 8ª série/9ºano – 1999/2017	107
Gráfico 11 – Desempenho por região na 4ª série/5ºano – 1999/2017	108
Gráfico 12 – Desempenho por região na 8ª série/9ºano – 1999/2017	108
Gráfico 13 – Desempenho por rede de ensino na 4ª série/5ºano – 1999/2017	109
Gráfico 14 – Desempenho por rede de ensino na 8ª série/9ºano – 1999/2017	109
Gráfico 15 – Demonstrativo de proficiências em leitura por gerações	63
Gráfico 16 – Demonstrativo de desempenho por sexo (gênero)	64
Gráfico 17 – Demonstrativo de desempenho por raça (cor da pele)	64
Gráfico 18 – Demonstrativo de desempenho por classe econômica (social)	65
Gráfico 19 – Demonstrativo de desempenho por região administrativa	66
Gráfico 20 – Proporção de estudantes no nível de proficiência em leitura 1 por classe econômica, raça	87
Gráfico 21 – Proporção de estudantes no nível de proficiência em leitura 2 por classe econômica, raça	88
Gráfico 22 – Proporção de estudantes no nível de proficiência em leitura 3 por classe econômica, raça	88
Gráfico 23 – Proporção de estudantes no nível de proficiência em leitura 4 por classe econômica, raça	89

LISTA DE TABELAS

Tabela 1 – Demonstrativo do tamanho da amostra por ano em etapas do Ensino Fundamental	26
Tabela 2 – Proporção nos níveis de proficiência (4ª Série e 5º Ano)	104
Tabela 3 – Proporção dos níveis de proficiência (8ª Série e 9º Ano)	104
Tabela 4 – Distribuição percentual por nível de proficiência – 1999/2017	36
Tabela 5 – Desempenho no Ensino Fundamental por fatores – 1999/2017	39
Tabela 6 – Estimativas dos efeitos esperados pelos fatores em análise do modelo econométrico para a 4ª série / 5º ano – 1999/2017	41
Tabela 7 – Efeito marginal dos fatores associados por nível de proficiência para a 4ª série /5º ano – 1999/2017	42
Tabela 8 – Estimativas dos efeitos esperados pelos fatores em análise do modelo econométrico para a 8ª série / 9º ano – 1999/2017	43
Tabela 9 – Efeito marginal dos fatores associados por nível de proficiência para a 8ª série /9º ano – 1999/2017	44
Tabela 10 – Demonstrativo de tamanho das subamostras por geração e ano	55
Tabela 11 – Resultados dos testes de <i>Hausman</i> para gerações	59
Tabela 12 – Proporção de alunos(as) nos níveis de proficiência (Série/Ano	61
Tabela 13 – Estimativas dos efeitos esperados pelos fatores associados no modelo econométrico para as gerações 1 a 6	67
Tabela 14 – Efeito marginal dos fatores associados por nível de proficiência	68
Tabela 15 – Probabilidades previstas por nível de proficiência	69
Tabela 16 – Proporção de alunos(as) na amostra por categorias de cada variável	85
Tabela 17 – Proporção de estudantes por classe econômica por nível de proficiência e raça	87
Tabela 18 – Estimativa dos efeitos esperados pelos fatores do modelo econométrico empírico	90
Tabela 19 – Efeito marginal dos fatores associados por nível de proficiência	92
Tabela 20 – Ponderações de amostras e subamostras 4ª série/5º ano	111
Tabela 21 – Ponderações de amostras e subamostras 8ª série/9º ano	112

SUMÁRIO

CAPÍTULO I - EFEITO DOS FATORES EXTRAESCOLARES NO DESEMPENHO DE PROFICIÊNCIA EM LEITURA NO ENSINO FUNDAMENTAL

1.1	INTRODUÇÃO	17
1.2	AVALIAÇÃO DA EDUCAÇÃO BÁSICA	20
1.2.1	Avaliação educacional	20
1.2.2	O Sistema de Avaliação da Educação Básica (SAEB)	21
1.2.3	Estudos empíricos sobre fatores associados ao desempenho escolar	22
1.3	METODOLOGIA	25
1.3.1	Descrição dos dados e das variáveis	25
1.3.2	Modelo teórico – Regressão Logística Ordinal	30
1.3.3	Estratégia empírica – Modelo econométrico	34
1.4	RESULTADOS E DISCUSSÃO	36
1.4.1	Análise descritiva dos dados	36
1.4.2	Análise e discussão dos resultados econométricos	40
1.5	CONSIDERAÇÕES FINAIS	45

CAPÍTULO II - EFEITO DE FATORES ASSOCIADOS AO DESEMPENHO ESCOLAR NO ENSINO FUNDAMENTAL ENTRE GERAÇÕES NO BRASIL

2.1	INTRODUÇÃO	48
2.2	DESIGN DE GERAÇÕES SOBREPOSTAS	51
2.2.1	O modelo de gerações sobrepostas	51
2.3	METODOLOGIA	54
2.3.1	Fontes dos dados e das variáveis	54
2.3.2	O modelo de regressão logística ordinal	56
2.3.3	Modelo empírico de gerações sobrepostas para o Ensino Fundamental	58
2.3.4	Modelo econométrico empírico – regressão logística ordinal	59
2.4	RESULTADOS E DISCUSSÃO	61
2.4.1	Análises descritivas dos dados amostrais	61
2.4.2	Resultados e análise do método econométrico	66
2.5	CONSIDERAÇÕES FINAIS	72

**CAPÍTULO III - INFÂNCIA E ADOLESCÊNCIA NO ENSINO
FUNDAMENTAL: EDUCAÇÃO E DESIGUALDADE
SOCIOECONÔMICA NO BRASIL**

3.1	INTRODUÇÃO	75
3.2	DESIGUALDADE SOCIOECONÔMICA E RACIAL NO BRASIL	77
3.2.1	Desigualdade de renda, raça e educação	77
3.2.2	Estudos sobre desigualdade racial e de classes	79
3.3	METODOLOGIA	81
3.3.1	Descrição dos dados e das variáveis	81
3.3.2	Modelo empírico econométrico - regressão logística ordinal	83
3.4	RESULTADOS E DISCUSSÃO	85
3.4.1	Resultados estatísticos descritivos	85
3.4.2	Resultados do modelo econométrico	90
3.5	CONSIDERAÇÕES FINAIS	95
	CONCLUSÃO GERAL	97
	REFERÊNCIAS	99
	APÊNDICE A – TABELAS DEMONSTRATIVAS DA PROPORÇÃO DE ALUNOS(AS) NOS NÍVEIS DE PROFICIÊNCIA CONFORME ESCALA SAEB PARA A 4ª SÉRIE/5º ANO E 8ª SÉRIE/9º ANO	104
	APÊNDICE B – GRÁFICOS DEMONSTRATIVOS DE DESEMPENHO POR GÊNERO (SEXO)	105
	APÊNDICE C – GRÁFICOS DEMONSTRATIVOS DE DESEMPENHO POR COR DA PELE OU RAÇA	106
	APÊNDICE D – GRÁFICOS DEMONSTRATIVOS DE DESEMPENHO POR CLASSE SOCIAL	107
	APÊNDICE E – GRÁFICOS DEMONSTRATIVOS DE DESEMPENHO POR REGIÃO GEOGRÁFICA	108
	APÊNDICE F – GRÁFICOS DEMONSTRATIVOS DE DESEMPENHO POR REDE DE ENSINO	109
	APÊNDICE G – CRITÉRIO DE CLASSIFICAÇÃO ECONÔMICA DO BRASIL	110
	APÊNDICE H – TABELAS DE PONDERAÇÕES DE AMOSTRAS E SUBAMOSTRAS	111

ANEXO A – DESCRITORES DO NÍVEL DE PROFICIÊNCIA DO SAEB PARA A 4ª SÉRIE/5º ANO	113
ANEXO B – DESCRITORES DO NÍVEL DE PROFICIÊNCIA DO SAEB PARA A 8ª SÉRIE/9º ANO	115

CAPÍTULO I

EFEITO DOS FATORES EXTRAESCOLARES NO DESEMPENHO DE PROFICIÊNCIA EM LEITURA NO ENSINO FUNDAMENTAL

RESUMO

No presente trabalho, avaliam-se a relação e os efeitos dos fatores extraescolares – gênero, raça, classe social, região geográfica e rede (pública e privada) – no desempenho de estudantes do Ensino Fundamental no Brasil no período de 1999 a 2017. A regressão logística ordinal é utilizada como método econométrico empírico por permitir a verificação do efeito de cada fator não somente no valor médio de desempenho, mas ao longo da distribuição de proficiências dos estudantes em duas fases do Ensino Fundamental: 4ª série/5º ano e 8ª série/9º ano. Para a verificação empírica, foram utilizadas 20 bases de microdados do SAEB/Prova Brasil, totalizando 3.851.542 alunos(as). Esses dados foram dispostos em corte transversal agrupados, em ambas as fases do Ensino Fundamental. Os resultados evidenciam um melhor desempenho do sexo feminino em relação ao masculino no Ensino Fundamental. Estudantes da raça branca apresentam tendências favoráveis a níveis superiores de proficiência nos dois ciclos do Ensino Fundamental. Alunos(as) das raças parda e amarela apresentam tendências desfavoráveis a nível superior de proficiência em leitura na 4ª série/5º ano e favoráveis na 8ª série/9º ano. Os estudantes da raça negra apresentam chances maiores para as menores proficiências em leitura nos dois ciclos do Ensino Fundamental. Em relação às classes econômicas, quanto maior o poder aquisitivo da família do estudante maiores são as chances de alcançar níveis mais altos de proficiência; as Regiões Norte e Nordeste apresentam para as séries iniciais e finais do Ensino Fundamental tendência aos níveis menores de proficiência, o mesmo ocorre em relação aos alunos(as) da rede pública.

Palavras-chave: Proficiência escolar. Ensino Fundamental. Regressão Logística Ordinal.

1.1 INTRODUÇÃO

É fato o baixo desempenho do Brasil em termos educacionais, particularmente no Ensino Fundamental, e, conforme Arretche et al. (2015), a educação brasileira ao longo do tempo tem sido motivo de preocupação e assunto recorrente nos mais diversos segmentos da sociedade. A razão para tamanha repercussão, seja na academia seja nos meios midiáticos, em relação à educação é a estagnação do desempenho educacional brasileiro nos principais indicadores nacionais e internacionais, como os apresentados pelo relatório de 2015 do Programa Internacional de Avaliação de Alunos (PISA), o qual revela que o Brasil não tem apresentado mudança na trajetória de desempenho nas avaliações desde 2000, mantendo-se nos mais baixos níveis de proficiência em leitura.

Segundo Gonçalves (2013), a leitura tem uma importância fundamental para a construção de novas aprendizagens, possibilitando o fortalecimento de ideias e ações que permitem ampliar e adquirir conhecimentos gerais e específicos, elevando o desempenho cognitivo e a análise crítica da realidade.

Três problemas graves que atingem a educação básica no Brasil são apontados por Helene (2013): 1) a diminuta taxa de atendimento na Educação Infantil em torno de 20% na faixa etária até os 4 anos de idade; 2) a altíssima evasão escolar no Ensino Fundamental, na ordem de 30%; 3) e a baixa taxa de conclusão do Ensino Médio, em que apenas 50% dos ingressantes concluem esse nível educacional. O autor ainda enfatiza que, no período de 1990 a 2000, as taxas de matrículas e de conclusão do Ensino Fundamental e Médio aumentaram significativamente, porém o desempenho médio neste interstício diminuiu. Já a partir de 2000, os indicadores de matrícula e conclusão praticamente estagnaram ainda que em patamares baixos. Mas, com a aplicação de mais recursos destinados à educação pública, os índices de desenvolvimento da educação básica melhoraram, apresentando entre 2005 e 2009 um crescimento médio de 3,6 para 4,1.

Nessa perspectiva o Índice de Desenvolvimento da Educação Básica (IDEB) de 2017, divulgado em 2018 pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), mostra que este indicador educacional se encontra estagnado, apresentando um índice médio de 3,5 entre 2015 e 2017, embora a meta brasileira seja igual a 6, equivalente à melhor posição no Programa Internacional de Avaliação de Alunos.

Segundo o diagnóstico de Almeida e Packard (2018), o Brasil enfrenta uma crise de aprendizagem, pois, apesar da expansão das matrículas nos últimos anos, o país levaria mais

de 260 anos para atingir as notas médias dos países da Organização para a Cooperação e Desenvolvimento Econômico (OCDE) de desempenho na leitura.

Para Dias (2010), a linguagem e o processo de comunicação estão diretamente relacionados ao progresso da inteligência sensório-motora e cognitiva nos anos iniciais de vida, em concordância com Coll et al. (2004), que ressalta a linguagem enquanto o instrumento mais importante que as pessoas empregam para se comunicar, configurando este um sistema de natureza simbólica que permite “representar” a realidade, não sendo possível dissociar desenvolvimento cognitivo e comunicativo.

A partir do contexto apresentado, surge este trabalho com o objetivo de investigar os fatores socioeconômicos: sexo, raça, classe social, região de domicílio e dependência administrativa ou rede de ensino; associados ao nível de desempenho (proficiência em leitura) de crianças e adolescentes do Ensino Fundamental e que foram submetidas às avaliações do Sistema de Avaliação da Educação Básica (SAEB) no âmbito do INEP, no período compreendido de 1999 a 2017, interstício em que o PISA realça desempenho de proficiência em leitura estagnado no Brasil.

As avaliações de larga escala têm sido base de informações para várias pesquisas científicas, cujo objetivo é identificar os fatores mais influentes no desempenho escolar, entre estes podemos destacar pesquisadores como: Alves e Soares (2008), que estudaram o efeito das escolas no aprendizado dos alunos no Ensino Fundamental, e Laros, Marciano e Andrade (2013), que pesquisaram sobre os fatores que afetam o desempenho na prova de Matemática do SAEB; entre outros. Uma característica comum na maioria destes estudos é que se utilizaram de dados em corte transversal, abrangendo um mesmo período de tempo e local geográfico.

A contribuição do presente trabalho, diferentemente de outros estudos de mesmo direcionamento, reside no fato de analisar determinantes extraescolares relacionados ao desempenho escolar no Ensino Fundamental no Brasil, em uma perspectiva temporal, abrangendo todo o período em que o PISA avalia o desempenho escolar brasileiro e ressaltando uma estagnação nos resultados de desempenho escolar. A importância deste trabalho se posiciona também no sentido de buscar entender por que alunos(as) com diferentes características e realidades apresentam diferentes proficiências.

Dessa forma, almeja-se apurar indícios de aumento de proficiência em leitura entre as fases inicial e final do Ensino Fundamental. Conseqüentemente, busca-se possibilitar o estabelecimento de subsídios e/ou adaptações de políticas públicas sociais e educacionais práticas e efetivas que levem à melhoria do desempenho dos alunos do Ensino Fundamental.

O texto está organizado de modo a abordar na segunda seção aspectos teóricos em relação às avaliações da educação básica. Na terceira seção, detalha-se a Metodologia adotada; na quarta seção os Resultados e as Discussões e na última seção as Considerações finais.

1.2 AVALIAÇÃO DA EDUCAÇÃO BÁSICA

Nesta seção, serão abordados aspectos relacionados à avaliação educacional, o sistema de avaliação da educação básica e em seguida apresentados alguns estudos empíricos sobre fatores associados ou determinantes, ou ainda, causais manifestos ao desempenho escolar no Ensino Fundamental.

1.2.1 Avaliação educacional

Conforme Carvalho (2017), a bibliografia relativa à avaliação educacional se apresenta nos mais diversos pressupostos políticos, filosóficos e metodológicos e, ao tratar dessa temática, sugere a perspectiva de gerações adotada por Guba e Lincoln (2011), segundo os quais numa mesma geração podem coexistir diferentes abordagens ou modelos. Acerca disso, destacam-se as seguintes gerações: 1) formulações, com ênfase nos métodos quantitativos; 2) modelos descritivos, com uma abordagem mais abrangente do que simplesmente a medição, tendo como foco a qualidade e a necessidade dos alunos; 3) juízo de valor, com foco nos objetivos educacionais, abrangendo todas as dimensões do processo de aprendizagem; 4) construtivista, na busca por crescimento cognitivo e desenvolvimento pessoal; 5) empoderamento com raízes na psicologia e na antropologia.

Na quinta e última geração, conforme Guba e Lincoln (2011), a avaliação para o empoderamento tem raízes na psicologia e na antropologia, posto que a psicologia trata dos estados e das disposições mentais relacionados ao comportamento e das suas interações com o ambiente físico e social, tendo na psicometria os meios para sua medição; e o estudo e a reflexão do ser humano, dando ênfase às diferenças e às variações com base nas características biológicas e socioculturais dos diversos grupos, são próprios da antropologia. Assim, é nessa configuração que se desenha atualmente a maioria dos trabalhos avaliativos existentes na literatura ao tratar dos fatores associados ao desempenho escolar na educação.

As discussões sobre a necessidade de melhoria da qualidade do ensino ofertado na educação básica na América Latina datam do final da década de 80 e início da década de 90, quando países como México, Argentina, Chile, Colômbia e Brasil desenvolveram avaliações de sistemas de ensino ou avaliações de larga escala (CARVALHO, 2017). No Brasil, em particular, a avaliação educacional veio com o objetivo de aferir a proficiência cognitiva dos alunos, com o propósito de prestar contas dos investimentos realizados neste âmbito e subsidiar as políticas públicas em educação. Assim, foram criados o Sistema Nacional de Avaliação da

Educação Básica (SAEB), a Avaliação Nacional da Alfabetização (ANA), o Exame Nacional de Desempenho dos Estudantes (ENADE), entre outros.

1.2.2 O Sistema de Avaliação da Educação Básica (SAEB)

A Portaria nº 931, de 21 de março de 2005 do Ministério da Educação, que instituiu o Sistema de Avaliação da Educação Básica, ressalta que este é composto por dois processos de avaliação: a Avaliação Nacional da Educação Básica (ANEAB) e a Avaliação Nacional de Rendimento Escolar (ANRESC).

A ANEB tem como objetivo principal avaliar a qualidade, a equidade e a eficiência da educação brasileira. Caracteriza-se por ser uma avaliação bianual, por amostragem, de larga escala, externa aos sistemas de ensino público e privado; a qual se utiliza de procedimentos metodológicos formais e científicos para coletar e sistematizar dados, além de produzir informações sobre as condições intra e extraescolares que incidem sobre o processo ensino-aprendizagem. Essas informações produzidas pela ANEB fornecem subsídios para a formulação de políticas públicas educacionais, com vista à melhoria da qualidade da educação.

A ANRESC é uma avaliação censitária bianual que tem como objetivos avaliar a qualidade do ensino ministrado nas escolas, de forma que cada unidade escolar receba o resultado global, e contribuir para o desenvolvimento, em todos os níveis educativos, de uma cultura avaliativa que estimule a melhoria dos padrões de qualidade e equidade da educação brasileira e adequados controles sociais de seus resultados.

Segundo o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, desde de 1990, o SAEB passou por algumas reestruturações. De 1990 a 1995, o público-alvo eram alunos(as) da 1ª, 3ª, 5ª e 7ª séries do Ensino Fundamental das escolas públicas e as áreas de conhecimento avaliadas eram: Língua Portuguesa, Redação, Matemática e Ciências Naturais. Em 1995, foi adotada a metodologia de construção de testes e análises de resultados conforme a Teoria de Resposta ao Item (TRI). Essa teoria estabelece que a probabilidade de resposta a um item é modelada como função da proficiência do aluno e de parâmetros dos itens, de forma que quanto maior é a proficiência, maior é a probabilidade de o aluno acertar os itens.

De 1997 até 2011, o público eram alunos(as) de 4ª e 8ª séries do Ensino Fundamental e 3ª série do Ensino Médio, abrangendo escolas públicas e privadas. Somente em 2001, as áreas de conhecimento avaliadas passaram a ser somente Língua Portuguesa e Matemática. Em 2005, o SAEB passou por uma reestruturação, conforme a Portaria Ministerial nº 931, de 21 de março, introduzindo um segmento censitário e passando a ser composto por

duas avaliações, a ANEB e a ANRESC, conhecido como Prova Brasil. Em 2013 até 2017, o público-alvo passou a ser alunos(as) do 5º e do 9º ano do Ensino Fundamental e da 3ª série do Ensino Médio.

A Portaria nº 1100, de 28 de dezembro de 2018, estabelece as diretrizes para o SAEB 2019, que deverá abranger: escolas públicas e escolas privadas localizadas em zonas urbanas e rurais; alunos(as) do 5º e 9º ano do Ensino Fundamental; e alunos(as) da 3ª e 4ª série do Ensino Médio tradicional e integrado.

1.2.3 Estudos empíricos sobre fatores associados ao desempenho escolar

Os fatores associados ao desempenho escolar são um dos temas mais pesquisados na área de educação desde o Relatório Coleman de 1966. Segundo Rocha (2014), esse foi o primeiro estudo que se propôs a relacionar desempenho escolar com insumos educacionais, cuja ideia inicial foi verificar a disparidade na qualidade das escolas frequentadas por alunos negros e brancos nos Estados Unidos, ficando comprovado que os principais determinantes eram as características familiares ou *background* familiar.

A literatura existente sobre os fatores que interferem no desempenho escolar reconhece o impacto do *background* familiar na performance educacional dos indivíduos e observa nessa relação a reprodução da desigualdade social. Na concepção de Luz (2006), considerar os fatores escolares inócuos em relação aos fatores familiares no desempenho escolar enfraquece as políticas governamentais na contenção do *déficit* educacional intergeracional e socioeconômico entre os estratos sociais. Acerca disso, tratando do papel exercido pelas características familiares e escolares na determinação do desempenho escolar, o autor obtém como fechamento a evidência de diferenças significativas quando consideradas características como renda, região geográfica e escolaridade dos pais.

Examinando a influência de diferentes categorias de *status* socioeconômico de estudantes com idades de 13 e 14 anos e seus desempenhos no Ensino Médio na Índia, Chandra e Azimuddin (2013) concluíram que existem correlação positiva e diferenças entre grupos com *status* socioeconômicos altos, médios e baixos.

Por meio do método *Propensity Score Matching* (PSM), Silva (2016) investigou o efeito da herança educacional dos pais sobre os desempenhos dos alunos, do 9º ano do Ensino Fundamental nas proficiências em Língua Portuguesa e Matemática, aferidos pelo SAEB de 2011 para as Regiões Sul/Sudeste e Nordeste do Brasil. Os resultados estimados, por meio do

método PSM, mostraram que a dotação (renda) educacional das famílias impacta de forma significativa no rendimento dos estudantes nas regiões analisadas.

De forma a contribuir com estudos sobre diferenciais de desempenho entre o ensino profissionalizante público e privado, Mariano (2016), utilizando regressão quantílica incondicional com dados do ENEM 2014, observou diferenciais positivos em todas as áreas para os estudantes de escolas privadas de ensino profissional, principalmente em Redação, concluindo que os alunos de escolas privadas são mais capazes de converter insumos educacionais em maiores pontuações em testes padronizados.

Compreender as relações entre o desempenho e as variáveis socioeconômicas constitui-se em importante fonte de informações com vistas a sua utilização na definição de políticas públicas. Com esta visão, Kleinke (2017) analisou grupos socioeducacionais que apresentaram melhor desempenho nas provas de Ciências da Natureza e Matemática nas edições do Enem de 2009 a 2012, comparando o decil superior com a totalidade dos candidatos, chegando à conclusão que as desigualdades de gênero, de raça, de renda e regionais, muito marcantes no Brasil, se manifestam no desempenho escolar.

Em um estudo transversal que explora o quociente de inteligência de crianças de 11 a 17 anos de idade e os fatores associados à desnutrição, à educação e ao *status* socioeconômico, Kumari, Shivgotra e Kumar (2017), utilizando a metodologia de regressão logística ordinal, concluíram que o nível de proficiências está associado à condição nutricional, socioeconômica e educacional da mãe. Além disso, crianças de famílias com menor renda apresentam baixo nível de proficiência.

Analisando os impactos das características de alunos sobre o desempenho escolar com base em dados desagregados por alunos no período de 1997 a 2005, Franco e Filho (2018) concluíram que o *background* familiar se mostrou muito importante para explicar desempenho escolar e as análises apontaram para efeitos diferenciados das variáveis explicativas sobre o fato de a escola ser pública ou privada.

O desempenho escolar demonstra a capacidade dos alunos de expressar sua aprendizagem e seu conhecimento adquirido no processo ensino-aprendizagem e, segundo Fernandes et al. (2018), deve ser compreendido como um fenômeno multideterminado e influenciado por aspectos dos estudantes no seu contexto familiar, escolar, sociocultural, institucional, político e econômico. Neste contexto de investigação, os autores realizaram um estudo correlacional preditivo com alunos de ambos os sexos de escolas públicas do Estado do Rio de Janeiro em 2018 e, por meio de regressão linear múltipla, observaram a existência de correlações positivas fracas entre habilidades sociais e percepção de apoio social da família,

com desempenho escolar dos estudantes. Os autores concluíram que os últimos anos do Ensino Fundamental são um momento crítico na trajetória acadêmica, aumentando a chance de abandono e evasão.

Nesse cenário, convergindo com estudos e pesquisas nacionais e internacionais nesta seção apresentados, é que este trabalho se propõe a investigar os efeitos dos fatores socioeconômicos associados ao desempenho escolar (níveis de proficiência em leitura) no Ensino Fundamental que possibilite encontrar evidências e nos leve a compreender por que o Brasil apresenta resultados tão baixos nos níveis de proficiência em leitura nos indicadores dos relatórios PISA. Assim, busca-se contribuir com uma abordagem mais abrangente a partir de dados em corte transversal, agrupados em uma perspectiva temporal diferenciada, e um método não tradicionalmente usual, como a regressão logística ordinal para o processo de estimação.

1.3 METODOLOGIA

O objetivo desta seção é descrever os dados e as variáveis elencadas para este estudo, bem como apresentar os aspectos teóricos relacionados à regressão logística ordinal e à estratégia empírica econométrica para este ensaio.

1.3.1 Descrição dos dados e variáveis

Os dados de natureza secundária foram obtidos na base de Microdados do INEP referente ao Sistema de Avaliação da Educação Básica (SAEB/Prova Brasil), compreendido no período de 1999 a 2017. Os dados referentes ao período de 1999 a 2003 e 2013 a 2017 são amostrais e os de 2005 a 2011 são censitários. O quadro 1 mostra a quantidade de alunos(as) avaliados pelo SAEB por ano e série/ano.

Quadro 1 – Quantidade de alunos (as) por ano e fases do Ensino Fundamental.

Ano	4ª Série/5º Ano	8ª Série/9º Ano
1999	20.862	17.052
2001	55.559	38.852
2003	44.711	27.721
2005	40.726	24.727
2007	2.217.984	1.257.759
2009	2.449.662	1.497.597
2011	2.226.381	1.546.163
2013	2.013.929	1.515.151
2015	1.986.425	1.463.531
2017	2.177.164	1.806.777

Fonte: Microdados SAEB 1999/2017.

A base de dados deste trabalho de investigação é composta por 20 conjuntos de dados divididos em duas etapas do Ensino Fundamental, sendo 10 para a 4ª série/5º ano e 10 para a 8ª série/9º ano. Cada conjunto consiste de informações de alunos(as) submetidos(as) à avaliação do SAEB, tomados em cada ano.

Para o cálculo das subamostras, foi tomada a quantidade de observações de cada ano e fase do Ensino Fundamental e adotado um nível de confiança de 95% ($z = 1,96$), com erro amostral igual a 5%, admitindo-se uma proporção amostral p igual a 0,50 (pelo fato de sob condições de completa incerteza, esse revelará a maior quantidade de erro possível). Quanto ao

tamanho da subamostra necessária para este estudo, ele foi calculado por meio da seguinte expressão da amostragem para população finita:

$$n = \frac{z^2 p(1-p)N}{e^2(N-1) + z^2 p(1-p)} \quad (1)$$

Em que,

N – tamanho de cada amostra (população)

n – tamanho da subamostra

z – estatística da distribuição normal

p – proporção amostral

e – erro amostral

Procedida uma análise de consistência e padronização nas variáveis categóricas dos dados, o tamanho da subamostra para cada ano e séries considerados está disposto na tabela 1. Já os pesos amostrais e subamostrais são apresentados nas tabelas 20 e 21 constantes no apêndice H.

Tabela 1 – Demonstrativo do tamanho da amostra por ano em etapas do Ensino Fundamental.

Ano	4ª Série/5º Ano	8ª Série/9º Ano
1999	3.213	2.895
2001	6.267	6.526
2003	5.066	4.787
2005	4.640	4.365
2007	390.369	215.451
2009	482.505	259.772
2011	394.638	265.102
2013	312.825	266.983
2015	310.041	259.230
2017	349.166	307.701
Total	2.258.730	1.592.812

Fonte: Elaborado pelo autor com base nos Microdados SAEB 1999/2017.

Conforme a estrutura dos dados amostrais, o design econométrico de dados de corte transversal agrupados foi adotado em razão de as análises se restringirem aos fatores: sexo, raça, classe social, região, rede (pública e privada) e desempenho escolar; estratificados em cada período, atenuando possíveis problemas de heterogeneidade, e também pelo fato de os dados possibilitarem a ordenação cronológica do conjunto de fatores.

Para muitas finalidades avaliativas, se o objetivo é concentrar-se na oportunidade real de o indivíduo alcançar seus objetivos, deve-se levar em conta não apenas os bens materiais, mas também as características pessoais relevantes. Para Sen (2010), a relação entre renda e capacidade seria acentuadamente afetada por gênero, classe social, localização geográfica [região] e por outras variações sobre as quais uma pessoa poderia ter controle total ou apenas parcial, como estudar na escola pública ou na privada. Ao contrastar grupos populacionais classificados segundo essas variações paramétricas, estes são particularmente importantes.

As variáveis independentes utilizadas neste estudo são as especificadas no quadro 2; e estas são componentes dos questionários aplicados pelo INEP por meio do SAEB/Prova Brasil, que coletam informações dos(as) alunos (as) sobre alguns aspectos da vida escolar, familiar, estudantil e também sobre o nível socioeconômico, o capital social e cultural. As *dummy* foram construídas de forma a permitir a inclusão das variáveis explanatórias nos aplicativos computacionais para obtenção das estimações.

Quadro 2 – Características; Variável; Codificação da categoria; *Dummy*.

Características	Variável	Codificação da categoria no SAEB	Dummy
Gênero	Sexo	1 – Masculino, 2 - Feminino	1 – Masculino, 0 - Feminino
Raça/Cor da pele	Raça	1 – Branco (a) 2 – Pardo (a) 3 – Preto (a) 4 – Amarelo (a) 5 - Indígena	1 – Branco (a), 0 - Demais 1 – Pardo (a), 0 - Demais 1 – Preto (a), 0 - Demais 1 – Amarelo (a), 0 - Demais
Classificação Econômica (Classe Social)	Classe	1 – Classe A 2 – Classe B 3 – Classe C 4 – Classe D 5 – Classe E	1 – Classe A, 0 - Demais 1 – Classe B, 0 - Demais 1 – Classe C, 0 - Demais 1 – Classe D, 0 - Demais
Regiões geográficas	Região	1 – Norte 2 – Nordeste 3 – Sudeste 4 – Sul 5 – Centro Oeste	1 – Norte, 0 – Demais 1 – Nordeste, 0 – Demais 1 – Sudeste, 0 – Demais 1 – Sul, 0 – Demais
Rede	Rede	1 – Pública 2 – Privada	1 – Pública, 0 – Privada

Fonte: Elaborado pelo autor com base nos Microdados SAEB 1999/2017.

A variável “classe” relacionada à característica classe econômica (classe social) foi construída como uma *proxy* para renda (poder aquisitivo) familiar dos alunos(as) submetidos

(as) à avaliação SAEB/Prova Brasil. A motivação para a construção desta variável foi a ausência de informações explícitas no questionário socioeconômico a respeito da renda familiar. Constatou-se, no entanto, a presença no questionário socioeconômico do SAEB/Prova Brasil de informações que possibilitaram uma classificação (segmentação por poder aquisitivo) segundo as classes econômicas (sociais) individuais dos(as) alunos(as), conforme o critério de classificação econômica da Associação Brasileira de Empresas de Pesquisa (ABEP), para o período de 1999 a 2017¹.

Esse critério de classificação econômica se constitui de um sistema de pontos para posses de bens domiciliares e grau de instrução do chefe de família, que, ao serem totalizados por domicílios, são estratificados em classes (cortes do critério Brasil) que atendem às necessidades de segmentação por poder aquisitivo. No apêndice G, é apresentado um histórico dos cortes do Critério de Classificação Econômica Brasil (CCEB) da ABEP desde de 1999.

O SAEB tem como medida de desempenho escolar dos(as) alunos(as), a escala ou o nível de proficiência que agrupa valores originalmente em uma escala contínua, que podem ser interpretados de maneira normativa para indicar as habilidades e as competências em cada nível. Essa escala contínua varia no intervalo de 0 a 500 e sua interpretação está disposta nos quadros 3 e 4 dos anexos A e B.

A Escala de proficiência em Língua Portuguesa é baseada em parâmetros estabelecidos pelos descritores que associam competências e habilidades desejáveis para cada série nos itens aplicados nos testes de desempenho e são calculados conforme a Teoria de Resposta ao Item que oferece uma explicação probabilística sobre as habilidades demonstradas em cada intervalo da escala.

Dessa forma, ela demonstra, de forma progressiva e acumulativa, habilidades como: localizar informações, estabelecer relações, identificar temáticas, reconhecer relações de causa e efeito, inferir, interpretar, distinguir, comparar, compreender, etc.; agrupados em quatro eixos cognitivos: reconhecer, analisar, avaliar e produzir.

Conforme UNESCO (2017), o Ministério da Educação (MEC) não tem uma recomendação oficial do desempenho desejável para os participantes do SAEB. E para entender o significado dos padrões de desempenho é divulgada uma descrição dos conteúdos de língua portuguesa (ou leitura) que os alunos provavelmente devem dominar segundo níveis da escala de proficiência ordenados em intervalos de 50 pontos. No 5º ano (4ª série) do ensino

¹ Foram consideradas as mudanças e as atualizações do quadro de classificação econômica a cada ano, no período em análise.

fundamental a escala é composta de nove níveis e no 9º ano (8ª série) a escala possui oito níveis.²

A escala do SAEB em Língua Portuguesa para as séries iniciais e finais no Ensino Fundamental é a mesma de 1999 a 2011, sendo que em 2013, a escala de proficiência passa por readequação pedagógica na métrica em relação ao eixo cognitivo para a 8ª série/9º ano, não havendo comprometimento (ruptura) do acompanhamento do desempenho escolar através do nível de proficiência ao longo do tempo em relação à escala do SAEB adotada no período de 1999 a 2011. Isso significa que são as mesmas, as habilidades cognitivas adquiridas a cada nível de proficiência, tanto para a 4ª série/5º ano, quanto para a 8ª série/9º ano.

Levando em consideração que os níveis de proficiência em Língua Portuguesa observados são de naturezas ordinais, optou-se por definir NIVPROC, conforme os quadros 5 e 6, como variável dependente no intuito de averiguar possíveis mudanças ou diferenças nas trajetórias de desempenho escolar em relação a 4ª série/5º ano e a 8ª série/9º ao longo do tempo e possibilitar estimar os efeitos, considerando tais mudanças³.

Quadro 5 – Escala de Proficiência em Língua Portuguesa – 4ª Série/5º Ano.

Escala	Nível (SAEB)	NIVPROC
Abaixo de 150	1	1
150 a 175	2	
175 a 200	3	2
200 a 225	4	
225 a 250	5	3
250 a 275	6	
275 a 300	7	4
300 a 325	8	
325 a 350	9	

Fonte: Microdados SAEB 1999/2011.

² No anexo A e B, os quadros 3 e 4 trazem o detalhamento dos descritores e respectivos níveis de proficiência da 4ª série/5º ano e 8ª série/9º ano, respectivamente.

³ As tabelas 2 e 3, que constam no apêndice A, apresentam a proporção média e os respectivos desvios padrões para todos os níveis de proficiência de alunos (as) submetidos (as) à avaliação do SAEB de 1999 a 2017.

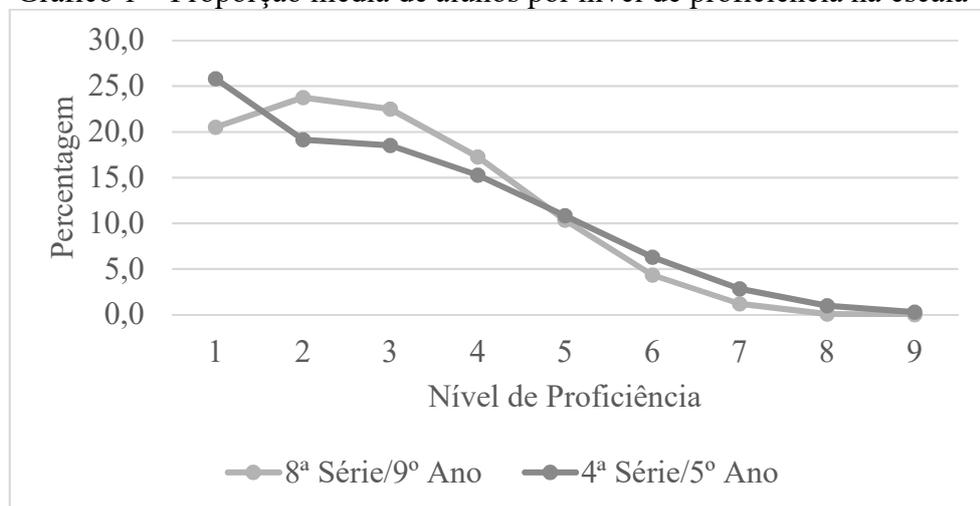
Quadro 6 – Escala de Proficiência em Língua Portuguesa – 8ª Série/9º Ano.

Escala	Nível (SAEB)	NIVPROC
200 a 225	1	1
225 a 250	2	
250 a 275	3	2
275 a 300	4	
300 a 325	5	3
325 a 350	6	
350 a 375	7	4
375 a 400	8	

Fonte: Microdados SAEB 2013/2017.

O gráfico 1 a seguir mostra o desempenho dos(as) alunos (as) por nível de proficiência conforme escala do SAEB para as duas etapas do Ensino Fundamental no período. Ele evidencia uma inversão no desempenho entre os níveis 1 - 2 e 5 - 6 nas etapas de ensino analisadas.

Gráfico 1 – Proporção média de alunos por nível de proficiência na escala



Fonte: Microdados SAEB/Prova Brasil – 1999 a 2017 (Elaborado pelo autor).

1.3.2 Modelo teórico - Regressão Logística Ordinal

O principal objetivo do design amostral adotado pelo SAEB/Prova Brasil na avaliação de Língua Portuguesa é produzir estimativas confiáveis para a média e para os respectivos erros de medidas da habilidade em leitura dos alunos, conjugando-se testes

cognitivos com questionários que levantam dados e informações de forma hierárquica sobre alunos(as) – nível hierárquico primário, gestores, escolas e docentes – nível hierárquico agregado. Neste trabalho em particular, as unidades experimentais se restringem ao nível hierárquico primário, ou seja, aos alunos(as).

Conforme Barbetta, Andrade e Tavares (2018), os modelos hierárquicos lineares se destacam na maioria dos estudos de fatores associados ao desempenho escolar quando este é medido por uma escala contínua como a do SAEB. A questão que se coloca é que os modelos relacionam o valor médio aos diversos fatores e estes podem afetar diferentemente estudantes de baixo e de alto desempenhos. Na visão de Fullerton e Xu (2016), a suposição de "linearidade nos parâmetros", ou efeitos constantes ao longo da distribuição dos resultados, pode levar a resultados enviesados e os modelos tradicionais de regressão lineares não são apropriados para resultados ordinais.

A motivação para adotar a regressão logística ordinal é avaliar os fatores associados ao longo da distribuição das escalas de proficiências. Precisamente, para este estudo, são consideradas as informações individuais dos(as) alunos(as) especificadas no quadro 2 como variáveis independentes; e como variável dependente, o nível de proficiência em Língua Portuguesa (nivproc) conforme apresentados nos quadros 5 e 6. É neste contexto que a regressão logística ordinal se apresenta como alternativa metodológica.

Até o início da década de 1960, o estágio de desenvolvimento do método estatístico de análise de dados categóricos era relativamente primitivo. Somente a partir da década de 1980, conforme Agresti (2010), tem havido uma crescente ênfase na análise de dados de variáveis categóricas ordenadas. Fullerton e Xu (2016) enfatizam que Snell, em 1964, estendeu o modelo *logit* ao estudo dos resultados ordinais, e Duncan, em 1967, foi o precursor da aplicação do modelo *logit* aos resultados ordinais, usando dados em nível individual. À Snell e McCullagh, em 1964 e 1980, respectivamente, é creditado o desenvolvimento do modelo.

Para muitas variáveis de estudo na educação e nas ciências sociais, O'Connell (2006) salienta que as escalas ordinais fornecem uma maneira simples de distinguir entre seus possíveis resultados os que podem ser considerados classificados. O modelo ordinal, segundo Fullerton e Xu (2016), restringe os efeitos de variáveis de forma a manter uma "ordenação estocástica". Dessa forma, a regressão logística ordinal atende às condições metodológicas para este trabalho, dado que as variáveis consideradas possuem os atributos necessários.

Agresti (2010), ao se referir a *Logits* Cumulativo, ressalta que para "c" resultados categóricos com probabilidades π_1, \dots, π_c associadas a cada categoria "j" em particular; os *logits* cumulativos são definidos como:

$$\begin{aligned}
 \text{logit} [P(Y \leq j)] &= \log \frac{P(Y \leq j)}{1 - P(Y \leq j)} \\
 &= \log \frac{\pi_1 + \dots + \pi_c}{\pi_{j+1} + \dots + \pi_c} \quad j = 1, \dots, c - 1
 \end{aligned} \tag{2}$$

O modelo de probabilidades, acumulativas incorporando variáveis explanatórias, tem a forma:

$$\text{logit} [P(Y_i \leq j)] = \alpha_j + \boldsymbol{\beta} \mathbf{x}_i = \alpha_j + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots \tag{3}$$

Para $j = 1, \dots, c - 1$ categorias o vetor coluna de parâmetros $\boldsymbol{\beta}$ descreve o efeito das variáveis explanatórias; \mathbf{x}_i denota o vetor coluna de valores das variáveis explanatórias ou independentes e α_j é o intercepto que representa o parâmetro de localização para as $j = 1, \dots, c - 1$ classes da variável dependente.

Uma expressão equivalente à anterior é dada por:

$$\text{logit} [P(Y \leq j)] = \frac{\exp(\alpha_j + \boldsymbol{\beta} \mathbf{x}_i)}{1 + \exp(\alpha_j + \boldsymbol{\beta} \mathbf{x}_i)}, \quad j = 1, \dots, c - 1 \tag{4}$$

O modelo *logit* acumulativo pode também ser expresso como:

$$\text{logit} [P(Y_i \leq j)] = \alpha_j - \boldsymbol{\beta} \mathbf{x}_i \tag{5}$$

Para esta parametrização com um sinal negativo que precede $\beta_k x_{ik}$ para o preditor k , o sinal de β_k tem o significado direcional usual. Especificamente, se $\beta_k > 0$, então à medida que x aumenta, cada *logit* cumulativo diminui. Portanto, as probabilidades cumulativas correspondentes diminuem. Então, relativamente menos probabilidade, a massa de probabilidade cai no limite inferior da escala de resposta, e Y é menos provável que caia na extremidade baixa e mais provável que caia na extremidade alta da escala. Em outras palavras, se $\beta_k > 0$ à medida que x aumenta, o *logit* $[P(Y_i \leq j)]$ diminui e consequentemente y aumenta, caso contrário se $\beta_k < 0$ à medida que x aumenta, o *logit* $[P(Y_i \leq j)]$ aumenta e a escala de y diminui. A parametrização com o sinal negativo precedendo $-\boldsymbol{\beta} \mathbf{x}_i$ é a forma como a regressão ordinal é modelada em grande parte dos aplicativos estatísticos, especificamente nos *softwares* SPSS e STATA.

O modelo de regressão logística ordinal pode ser expresso como um modelo de variável latente conforme: Agresti (2010); Wooldridge (2010); e Kumari et al. (2017).

Fullerton e Xu (2016) salientam que Mckelvey e Zavoina em 1975 foram os primeiros a utilizar estrutura de variável latente em modelos *probit*.

O modelo ordenado probit, conforme Wooldridge (2010), pode ser derivado da expressão para a variável latente y^* determinada por:

$$y^* = \mathbf{x}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (6)$$

Em que \mathbf{x} é um vetor de variáveis independentes, $\boldsymbol{\beta}$ é um vetor de coeficientes, e $\boldsymbol{\varepsilon}$ é o termo erro que segue uma distribuição normal; $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{1})$.

Considerando $\alpha_1 < \alpha_2 < \dots < \alpha_j$ pontos de corte ou limites, define-se como:

$$y = 0 \text{ if } y^* \leq \alpha_1 \quad (7)$$

$$y = 1 \text{ if } \alpha_1 < y^* \leq \alpha_2 \quad (8)$$

⋮

$$y = J \text{ if } y^* > \alpha_j \quad (9)$$

Assim, dada a suposição normal padrão para $\boldsymbol{\varepsilon}$, a distribuição condicional de y dado \mathbf{x} é expressa pela distribuição de probabilidade acumulativa.

$$P(y \leq j/\mathbf{x}) = F(\mathbf{x}\boldsymbol{\beta} + \boldsymbol{\varepsilon}), \quad j = 1, \dots, J - 1 \quad (10)$$

Em particular, se escrevermos a função F da expressão (9), com $(\boldsymbol{\varepsilon} / \mathbf{x})$ normal padrão, como uma função *logit* do modelo logístico ordenado, então a função F assume a função link *logit*:

$$\text{Logit}[P(y \leq j)] = \alpha_j - \boldsymbol{\beta} \mathbf{x}, \quad j = 1, \dots, J - 1 \quad (11)$$

Fullerton e Xu (2016) destacam vários métodos de interpretação dos resultados, incluindo: *odds ratios*/razões de chances, efeitos marginais médios (AME) e gráficos baseados em probabilidades previstas.

$$\text{odds ratio} = e^{-\boldsymbol{\beta}} \quad (12)$$

$$\text{AME} = \frac{1}{N} \sum_{i=1}^N \frac{\partial P(y_i = j/\mathbf{x} = \mathbf{x}_i)}{\partial x_k} \quad (13)$$

O modelo de regressão ordinal, em conformidade com Fullerton e Xu (2016), usa a estimação por máxima verossimilhança (EMV), amplamente utilizada para estimativas nos modelos de regressão com dados categóricos. E segundo Gujarati (2006), o motivo de maximizar a função de verossimilhança logarítmica é obter a máxima probabilidade dos dados observados. Para Greene (2012), a atratividade do EMV reside nas propriedades assintóticas: Consistência, Normalidade, Eficiência e Invariância.

1.3.3 Estratégia empírica – Modelo econométrico

O objetivo desta modelagem é relacionar a variável resposta ordinal *nivproc* com os fatores ou variáveis explicativas: sexo, raça, classe (classe social), região e rede (dependência administrativa). A variável dependente *nivproc* é definida pelas escalas de proficiência para a 4ª série/5º ano e para a 8ª série/9º ano, conforme os quadros 5 e 6; já os fatores são definidos como variáveis *dummy*, correspondentes ao quadro 2.

O modelo adotado neste trabalho após ajustado, demonstrando sua significância, é dado pela função genérica da função acumulada. A função neste caso em particular é a *Logit*, dada pela expressão (11) e as previsões das probabilidades acumuladas são expressas como:

$$P(Y \leq j) = \frac{1}{1 + e^{-(\alpha_j - X\beta)}} \quad (14)$$

Em que, $j = 1, \dots, 3$;

$$X\beta = \sum_k \beta_k \text{Sexo}_{ik} + \sum_k \beta_k \text{Raça}_{ik} + \sum_k \beta_k \text{Classe}_{ik} + \sum_k \beta_k \text{Região}_{ik} + \sum_k \beta_k \text{Rede}_{ik} \quad (15)$$

em que:

α_j é a constante;

Sexo_{ik} é o vetor de característica do aluno *i* no gênero *k*;

Raça_{ik} é o vetor de característica do aluno *i* na raça *k*;

Classe_{ik} é o vetor de característica do aluno *i* na classe social *k*;

Região_{ik} é o vetor de característica do aluno *i* na região *k*;

Rede_{ik} é o vetor de característica do aluno *i* na rede de ensino *k*;

β_k é o vetor de parâmetros de interesse que reporta as relações condicionais dos fatores acerca do desempenho escolar dos alunos(as).

De acordo com a função (6), a probabilidade do nível de proficiência em Língua Portuguesa dos(as) alunos(as) que frequentam a escola pode ser calculada como:

$$P(y \leq 1) = P(y^* \leq \alpha_1) = P(\mathbf{x}\boldsymbol{\beta} + \mathbf{e} \leq \alpha_1) = F[\alpha_1 - \mathbf{x}\boldsymbol{\beta}] \quad (16)$$

$$P(y \leq 2) = P(\alpha_1 < y^* \leq \alpha_2) = F[\alpha_2 - \mathbf{x}\boldsymbol{\beta}] - F[\alpha_1 - \mathbf{x}\boldsymbol{\beta}] \quad (17)$$

$$P(y \leq 3) = P(\alpha_2 < y^* \leq \alpha_3) = F[\alpha_3 - \mathbf{x}\boldsymbol{\beta}] - F[\alpha_2 - \mathbf{x}\boldsymbol{\beta}] \quad (18)$$

$$P(y \leq 4) = P(\alpha_3 < y^* \leq \infty) = 1 - F[\alpha_3 - \mathbf{x}\boldsymbol{\beta}] \quad (19)$$

1.4 RESULTADOS E DISCUSSÃO

Os resultados, as análises e as discussões das estatísticas descritivas dos dados e do modelo empírico econométrico são explicitados nesta seção.

1.4.1 Análise descritiva dos dados

A análise descritiva dos dados consistiu na consolidação dos dados amostrados. Nesse contexto, o passo inicial, a partir da tabulação dos dados, foi descrever a proporção de alunos(as) das séries iniciais 4ª série e/ou 5º ano e séries finais 8ª série e/ou 9º ano, nos níveis de proficiência em Língua Portuguesa no período de 1999 a 2017.

A tabela 4 mostra o percentual médio de alunos (as) por nível de proficiência a cada ano, juntamente com a média proporcional do período e respectivos desvios padrão; já os gráficos 2 e 3 apresentam uma linha contínua, que liga os pontos médios em cada nível de proficiência das séries iniciais 4ª série e/ou 5º ano e finais 8ª série e/ou 9º ano e suas respectivas variações.

Tabela 4 - Distribuição percentual por nível de proficiência – 1999/2017

Série/Ano	4ª Série / 5º Ano				8ª Série / 9º Ano				
	Nível	1	2	3	4	1	2	3	4
1999		54,8	30,9	11,9	2,4	49,2	37,4	12,1	1,3
2001		54,5	30,2	12,9	2,3	40,9	39,2	17,5	2,4
2003		49,1	33,7	14,8	2,5	43,3	38,5	16,6	1,7
2005		49,7	33,1	15,5	1,8	44,6	37,5	16,1	1,9
2007		53,5	35,0	10,4	1,1	55,6	36,3	7,7	0,3
2009		48,4	34,5	14,3	2,8	47,5	40,2	11,7	0,6
2011		42,4	36,3	17,4	3,9	46,5	40,0	12,7	0,8
2013		41,4	33,1	19,9	5,6	43,9	40,6	14,6	1,0
2015		29,1	36,8	26,4	7,7	39,8	42,4	16,6	1,3
2017		26,2	35,5	28,3	10,0	36,3	43,3	18,4	2,0
Média		44,9	33,9	17,2	4,0	44,8	39,5	14,4	1,3
Desvio Padrão		10,2	2,2	6,0	2,9	5,4	2,2	3,3	0,6

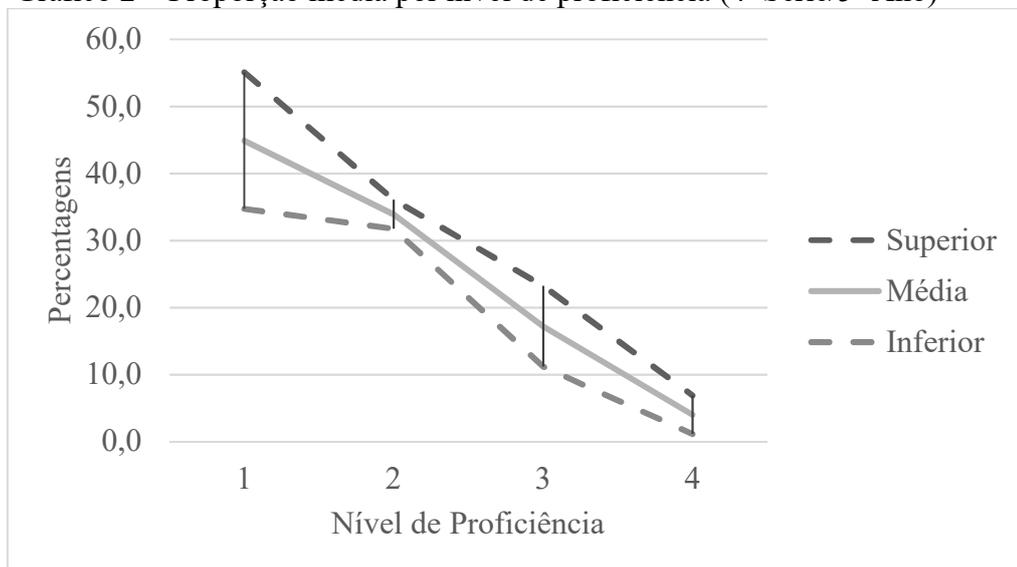
Fonte: Microdados SAEB/Prova Brasil (Elaborado pelo autor).

Conforme a tabela 4, evidencia-se um decréscimo na proporção média de alunos (as) da 4ª série / 5º ano no nível 1, ao longo do tempo, e um acréscimo proporcionalmente moderado nos níveis 2, 3 e 4.

Ainda de acordo com a tabela 4, verifica-se um decréscimo na proporção de alunos(as) da 8ª série / 9º ano no nível 1 a partir de 2009; acompanhado de acréscimo na proporção, mesmo que sutil nos níveis 2 e 3 com início em 2009 até 2017. Em 2007, percebe-se um aumento considerável na proporção média de alunos (as) no nível 1, acompanhado de declínio na proporção nos níveis 2 e 3.

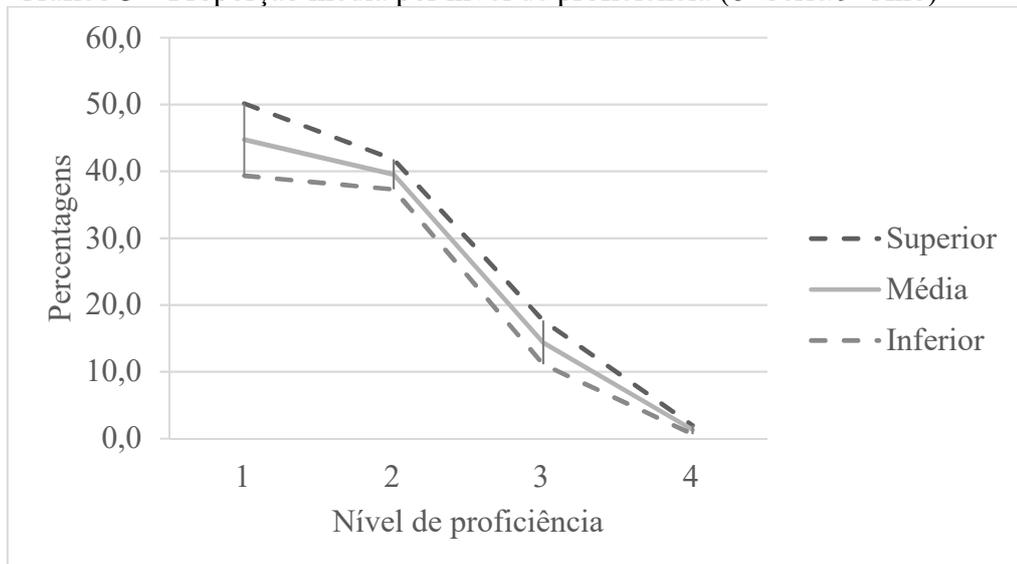
Uma análise da tabela 4 demonstra que, ao longo do tempo, ocorre uma aparente transição para os níveis maiores de proficiência em leitura para ambas a séries iniciais e finais do Ensino Fundamental.

Gráfico 2 – Proporção média por nível de proficiência (4ª Série/5º Ano)



Fonte: Microdados SAEB/Prova Brasil (Elaborado pelo autor).

Gráfico 3 – Proporção média por nível de proficiência (8ª Série/9º Ano)



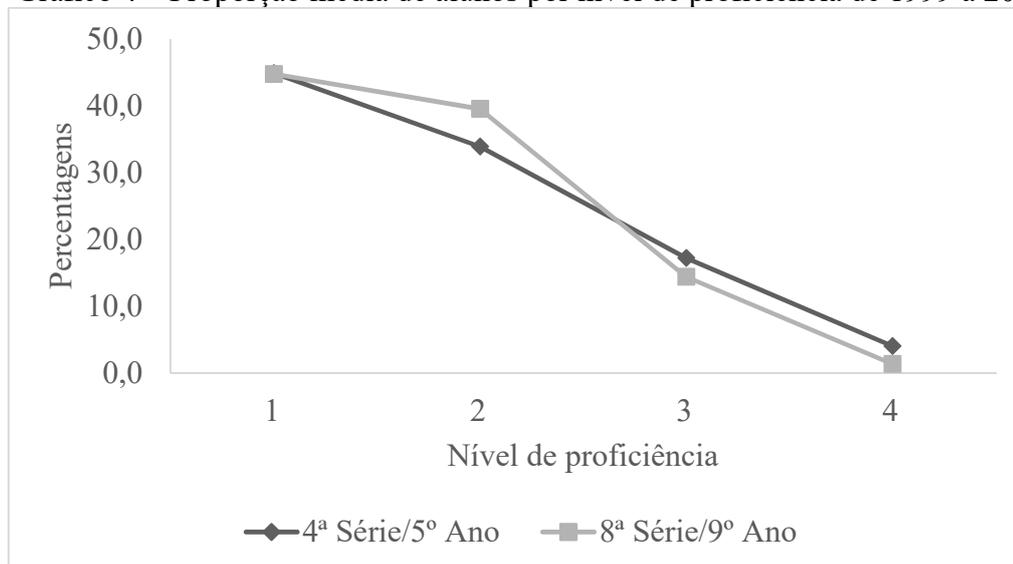
Fonte: Microdados SAEB/Prova Brasil (Elaborado pelo autor).

Os gráficos 2 e 3 realçam a proporção média e o desvio padrão por nível de proficiência, demonstrando de maneira geral, para ambas fases do Ensino Fundamental no período de 1999 a 2017, que o desempenho se apresenta decrescente nos níveis de proficiência.

Para verificar se há diferenças nas similaridades entre as proporções médias no desempenho nas duas etapas do Ensino Fundamental, tanto em relação aos níveis quanto ao fator temporal, foi realizado um teste de *Hausman*, que apresentou um $\chi^2(1) = 1,23$ com $Prob > \chi^2 = 0,2667$, a partir do qual pôde-se concluir que não se rejeita a hipótese de similaridade entre as distribuições de proficiências.

O gráfico 4 apresenta a sobreposição da proporção média para os níveis de proficiência em ambas as fases do Ensino Fundamental. Assim, percebe-se que houve um aumento na proporção de alunos (as) da 8ª série/9º ano, no nível 2 (dois), e um decréscimo nos níveis 3 (três) e 4 (quatro), em relação à 4ª série/5º ano. Além disso, verifica-se uma “inversão” na aquisição de habilidades cognitivas em Língua Portuguesa entre as duas etapas do Ensino Fundamental em relação aos níveis 3 e 4.

Gráfico 4 – Proporção média de alunos por nível de proficiência de 1999 a 2017



Fonte: Microdados SAEB/Prova Brasil – 1999 a 2017 (Elaborado pelo autor).

A tabela 5, por sua vez, expõe a proporção de alunos(as) nas duas etapas do Ensino Fundamental em cada nível de proficiência em Língua Portuguesa no período de 1999 até 2017, por fatores ou variáveis socioeconômicas.

Tabela 5 – Desempenho no Ensino Fundamental por fatores – 1999/2017

Variáveis ou Fatores		Nível de Proficiência (%)							
		4ª Série / 5º Ano				8ª Série / 9º Ano			
		1	2	3	4	1	2	3	4
Sexo	Masculino	44,7	34,3	16,8	4,1	47,9	38,6	12,5	1,0
	Feminino	34,3	37,4	22,2	6,1	40,7	42,6	15,5	1,2
Raça	Branco	35,8	34,9	22,4	6,8	37,3	42,7	18,3	1,6
	Pardo	39,1	36,7	19,4	4,8	46,4	40,2	12,5	0,9
	Preto	53,0	32,9	11,9	2,2	50,4	38,4	10,5	0,7
	Amarelo	46,1	32,3	17,1	4,5	42,8	42,1	14,2	1,0
	Indígena	40,3	36,0	19,3	4,4	47,3	41,1	11,0	0,6
Classe Econômica	Classe A	29,9	32,5	27,5	10,1	31,7	40,8	23,8	3,7
	Classe B	31,5	35,6	25,1	7,8	36,2	42,5	19,5	1,8
	Classe C	35,8	37,4	21,2	5,6	42,4	41,6	14,9	1,2
	Classe D	43,2	36,0	16,9	3,9	48,3	39,5	11,5	0,8
	Classe E	69,0	25,8	4,7	,5	68,8	27,5	3,6	0,1
Região	Norte	51,5	33,4	12,6	2,5	50,7	38,4	10,2	0,7
	Nordeste	56,8	30,0	11,0	2,3	53,3	36,5	9,5	0,7
	Sudeste	33,0	36,5	23,6	6,9	39,7	42,2	16,8	1,4
	Sul	30,8	40,1	23,2	5,9	39,3	43,9	15,6	1,2
	Centro Oeste	34,9	39,3	21,0	4,8	41,5	42,4	14,9	1,2
Rede	Pública	35,6	35,5	22,4	6,5	41,9	41,6	15,2	1,2
	Privada	14,0	32,3	37,9	15,8	18,3	41,5	35,0	5,2

Fonte: Microdados SAEB/Prova Brasil (Elaborado pelo autor).

Os gráficos 5 e 6, que constam no apêndice B, realçam o desempenho dos(as) alunos(as) avaliados para a variável sexo masculino e feminino. Verifica-se para ambos os gêneros uma distribuição assimétrica à direita em relação aos níveis de proficiência, porém o sexo feminino apresenta proporção média maior do que o sexo masculino a partir do nível 2, evidenciando que as alunas apresentam melhores resultados que os alunos.

O gráfico 7 do apêndice C, referente à tabela 5, mostra o desempenho da 4ª série/5ºano por raça segundo autodeclaração de cor de pele dos(as) alunos(as). Percebe-se desempenho decrescente para todas as raças. Os autodeclarados pretos(as) apresentam proporções médias maiores nos níveis inferiores de proficiência até nível 2 (dois) e menores proporções médias nos níveis mais altos de proficiência, acima de 3 (três). Praticamente não há diferenças no desempenho nas demais raças, somente a raça branca apresenta uma diferença discreta nos níveis 3 (três) e 4 (quatro).

O desempenho por raças da 8ª série/9ºano, demonstrado no gráfico 8 do apêndice C, atesta que a raça negra apresenta baixo desempenho comparado as demais raças a partir do nível

3. Observa-se que pardos, amarelos e indígenas têm praticamente o mesmo desempenho, e a raça branca tem melhor situação com maior proporção média de indivíduos nos níveis mais altos de proficiência.

Em relação à variável classe social, a proporção média de alunos(as) nos níveis de proficiência ao longo de 1999 até 2017 é demonstrada no gráfico 9 e 10, exibidos no apêndice D. Conforme os gráficos 9 e 10, para a 4ª série/5ºano e 8ª série/9ºano, respectivamente, é observada uma inversão na ordem de desempenho das classes sociais a partir do nível 2 (dois). A classe social E apresenta uma proporção média de alunos(as) que supera as demais classes no nível 1 (um) de proficiência em leitura, já a partir do nível 2 (dois) as classes sociais sequencialmente em ordem decrescente de E para A vão se apresentando com proporções maiores de alunos(as) nos níveis superiores de proficiência.

Em se tratando de desempenho nos níveis de proficiência das regiões brasileiras, os gráficos 11 e 12 do apêndice E exibem para as Regiões Norte e Nordeste maior proporção de alunos(as) no nível 1(um) e menores proporções acima de 1; as Regiões Sul, Sudeste e Centro-Oeste mostram maiores proporções nos níveis acima de 2, apesar disso percebe-se uma disparidade entre estas regiões e as Regiões Norte e Nordeste para a e 8ª série/9ºano do Ensino Fundamental.

Nos gráficos 13 e 14, presentes no apêndice F, é possível observar uma diferença bastante acentuada entre os desempenhos da rede pública e a privada. Acima do nível 2 (dois), a rede privada possui as maiores proporções, apresentando distribuição quase simétrica em relação aos níveis de proficiência, e a rede pública distribuição assimétrica à direita, tanto para a 4ª série/5ºano quanto para a 8ª série/9ºano.

1.4.2 Análise e discussão dos resultados econométricos

Como o design amostral para a 4ª série/ 5º ano é um conjunto de dados de corte transversal agrupados, foi realizado o teste de *Hausman*, que apresentou um $\chi^2(5) = 1179,83$ com $Prob > \chi^2 = 0,000$. Portanto, significativo concluindo que o modelo de efeito fixo é mais adequado do que o de efeitos aleatórios, assim se presume que as relações são as mesmas no tempo.

Nas tabelas 6 e 8, são apresentadas as estimativas e seus respectivos erros padrões, estatística robusta z, os p-value e a razão de chances.

Tabela 6 - Estimativas dos efeitos esperados pelos fatores em análise do modelo econométrico para a 4ª série /5º ano – 1999/2017

Variáveis/Fatores	Estimativa	Erro Padrão	Estatística Z	p-value	Razão de Chances
Sexo (Masc.)	- 0,468	0,004	- 113,47	0,000	1,60
Raça (Branca)	0,039	0,012	3,21	0,001	0,96
Raça (Parda)	- 0,053	0,012	- 4,40	0,000	1,05
Raça (Negra)	- 0,597	0,013	- 44,31	0,000	1,82
Raça (Amarela)	- 0,167	0,017	- 9,498	0,000	1,18
Classe Social (A)	1,146	0,037	30,49	0,000	0,32
Classe Social (B)	1,217	0,021	56,34	0,000	0,30
Classe Social (C)	1,144	0,020	54,74	0,000	0,32
Classe Social (D)	1,036	0,020	49,61	0,000	0,35
Região (Norte)	- 0,607	0,009	- 63,76	0,000	1,83
Região (Nordeste)	- 0,763	0,008	- 92,90	0,000	2,14
Região (Sul)	0,224	0,007	29,50	0,000	0,80
Região (Sudeste)	0,200	0,008	22,94	0,000	0,82
Rede (Pública)	- 1,024	0,014	- 72,47	0,000	2,78

Fonte: Elaborada pelo autor.

A tabela 6 mostra as estimativas dos fatores para a 4ª série/5º ano que se apresentaram estatisticamente significantes ao nível de significância de 0,05. Verifica-se que os fatores associados à raça (branca), às classes sociais e às Regiões (Sul e Sudeste) apresentam efeito positivo no desempenho escolar. Já os fatores associados a sexo (masculino), raças (parda, negra e amarela), Regiões (Norte e Nordeste) e rede (pública) apontam efeito negativo.

Conforme essa tabela, ocorre um aumento na razão de chance de estar nos níveis inferiores de proficiência os seguintes fatores: sexo masculino (1,60), raça parda (0,96), raça negra (1,82), raça amarela (1,18), Região Norte (1,83), Região Nordeste (2,14) e rede pública (2,78). Já para os fatores: raça branca (0,96), classe social A (0,32), classe social B (0,30), classe social C (0,32), classe social D (0,35), Região Sul (0,80) e Região Sudeste (0,82), ocorre uma diminuição nas chances de estar nos níveis menores de proficiência.

A tabela 7 apresenta as probabilidades dos fatores se encontrarem no ponto médio em cada nível de proficiência e os efeitos marginais dos fatores por nível de proficiência para a 4ª série / 5º ano. Seus resultados corroboram com os apresentados na tabela 6. Para o nível 1,

a probabilidade é de 29,5%; no nível 2 é de 39,1%; a probabilidade, no nível, é de 24,6% e no nível 4 é de 6,8%.

Tabela 7 - Efeito marginal dos fatores associados por nível de proficiência para a 4ª série/5º ano – 1999/2017.

Variáveis/Fatores	dy/dx (1)	dy/dx (2)	dy/dx (3)	dy/dx (4)
	0,295	0,391	0,246	0,068
Sexo (Masc.)	0,097	0,003	- 0,071	- 0,029
Raça (Branca)	- 0,008	- 0,000	0,006	0,003
Raça (Parda)	0,011	0,000	- 0,008	- 0,003
Raça (Negra)	0,135	- 0,018	- 0,085	- 0,031
Raça (Amarela)	0,036	- 0,001	- 0,025	- 0,009
Classe Social (A)	- 0,178	- 0,098	0,159	0,118
Classe Social (B)	- 0,202	- 0,087	0,173	0,115
Classe Social (C)	- 0,237	- 0,004	0,168	0,073
Classe Social (D)	- 0,199	- 0,033	0,155	0,077
Região (Norte)	0,137	- 0,019	- 0,086	- 0,031
Região (Nordeste)	0,169	- 0,018	- 0,109	- 0,041
Região (Sul)	- 0,046	- 0,002	0,034	0,014
Região (Sudeste)	- 0,040	- 0,004	0,031	0,013
Rede (Pública)	0,167	0,079	- 0,147	- 0,099

Fonte: Elaborado pelo autor.

O efeito negativo dado pelo sinal dos coeficientes marginais aponta a redução na probabilidade do fator no nível de proficiência conexo, e o sinal positivo o aumento na probabilidade do fator.

Destá forma, verifica-se redução nas probabilidades nos níveis de proficiência superiores para os fatores: sexo masculino, a partir do nível 3, raça (parda, negra e amarela), Regiões Norte e Nordeste, e rede pública; percebe-se aumento nas probabilidades para os fatores: raça branca, classes econômicas (A, B, C e D) a partir do nível 3, regiões (Sul e Sudeste).

O teste de *Hausman*, para os dados em corte transversal agrupados para a 8ª série/9º ano, apresentou um $\chi^2(5) = 948,65$ com $Prob > \chi^2 = 0,000$, portanto significativo, concluindo que o modelo de efeito fixo é a melhor opção a ser considerada.

A tabela 8 mostra as estimativas dos efeitos esperados dos fatores associados ao desempenho escolar dos(as) alunos(as) da 8ª série/9º ano do Ensino Fundamental. Ressalta-se que todas as estimativas dos fatores são estatisticamente significantes a 0,05.

Tabela 8 - Estimativas dos efeitos esperados dos fatores em análise do modelo econométrico para a 8ª série/9º ano – 1999/2017

Variáveis/Fatores	Estimativa	Erro Padrão	Estatística Z	p-value	Razão de Chances
Sexo (Masc.)	- 0,338	0,004	- 72,52	0,000	1,40
Raça (Branca)	0,368	0,015	24,47	0,000	0,69
Raça (Parda)	0,070	0,014	4,75	0,000	0,93
Raça (Negra)	- 0,105	0,016	- 6,57	0,000	1,11
Raça (Amarela)	0,208	0,018	11,34	0,000	0,81
Classe Social (A)	1,087	0,045	23,85	0,000	0,34
Classe Social (B)	1,108	0,032	34,17	0,000	0,33
Classe Social (C)	0,993	0,031	31,12	0,000	0,37
Classe Social (D)	0,878	0,031	27,52	0,000	0,42
Região (Norte)	- 0,374	0,011	- 33,59	0,000	1,45
Região (Nordeste)	- 0,423	0,009	- 45,53	0,000	1,53
Região (Sul)	0,088	0,008	10,21	0,000	0,92
Região (Sudeste)	- 0,021	0,009	- 2,13	0,033	1,02
Rede (Pública)	- 1,132	0,014	- 76,45	0,000	3,10

Fonte: Elaborado pelo autor.

As estimativas dos fatores para a 8ª série/9º ano apresentaram efeito positivo em relação ao desempenho escolar para: raça (branca, parda e amarela), classe social (todas) e Região (Sul). Já os fatores associados a sexo (masculino), raça (negra), Regiões (Norte, Nordeste e Sudeste), bem como a rede (pública) exibiram efeito negativo.

Na tabela 8, ocorre um aumento na razão de chance de estarem nos níveis inferiores de proficiência os seguintes fatores: sexo masculino (1,40), raça negra (1,11), Região Norte (1,45), Região Nordeste (1,53), Região Sudeste (1,02) e rede pública (3,10). Já para os fatores: raça branca (0,69), raça parda (0,93), raça amarela (1,11), classe social A (0,34), classe social B (0,33), classe social C (0,37), classe social D (0,42) e Região Sul (0,92) ocorre uma diminuição nas chances de estar nos níveis menores de proficiência.

A tabela 9 mostra os efeitos marginais dos fatores associados por nível de proficiência para a 8ª série/9º ano, corroborando os resultados da tabela 8.

Tabela 9 - Efeito marginal dos fatores associados por nível de proficiência para a 8ª série / 9º ano – 1999/2017.

Variáveis/Fatores	dy/dx (1)	dy/dx (2)	dy/dx (3)	dy/dx (4)
	0,392	0,434	0,161	0,013
Sexo (Masc.)	0,081	- 0,033	- 0,044	- 0,004
Raça (Branca)	- 0,086	0,032	0,049	0,005
Raça (Parda)	- 0,017	0,007	0,009	0,001
Raça (Negra)	0,025	- 0,011	- 0,013	- 0,001
Raça (Amarela)	- 0,048	0,017	0,029	0,003
Classe Social (A)	- 0,214	0,004	0,185	0,025
Classe Social (B)	- 0,232	0,032	0,177	0,022
Classe Social (C)	- 0,234	0,092	0,129	0,013
Classe Social (D)	- 0,199	0,059	0,126	0,014
Região (Norte)	0,092	- 0,043	- 0,044	- 0,004
Região (Nordeste)	0,103	- 0,046	- 0,052	- 0,005
Região (Sul)	- 0,021	0,008	0,012	0,001
Região (Sudeste)	0,005	- 0,002	- 0,003	- 0,000
Rede (Pública)	0,222	- 0,005	- 0,192	- 0,026

Fonte: Elaborado pelo autor

A tabela 9 apresenta as probabilidades dos fatores se encontrarem no ponto médio de cada nível de proficiência. Assim, a probabilidade, no nível, 1 é de 39,2%; no nível 2 é de 43,4%, a probabilidade, no nível 3, é de 16,1% e no nível 4 é de apenas 1,3%.

Verifica-se que o efeito negativo dado pelo sinal dos coeficientes marginais realça a redução na probabilidade do fator no nível de proficiência relacionado, e o sinal positivo o aumento na probabilidade do fator.

Neste contexto, percebe-se redução nas probabilidades nos níveis superiores para os fatores: sexo masculino, raça negra, Regiões Norte, Nordeste e Sudeste, e rede pública; e aumento nas probabilidades para os fatores: raças (branca, parda e amarela), classes econômicas (A, B, C e D) e Região Sul.

1.5 CONSIDERAÇÕES FINAIS

Analisando-se as estimativas do modelo econométrico com os resultados estatísticos descritivos, vê-se que o modelo descreve adequadamente os fenômenos evidenciados nos resultados estatísticos descritivos para as duas etapas do Ensino Fundamental, ou seja:

O gênero feminino apresenta melhores resultados no desempenho escolar no Ensino Fundamental que o masculino.

Todas as classes econômicas apresentam efeito positivo em relação a desempenho escolar, contudo, nas classes de menor poder aquisitivo, ocorre uma diminuição na magnitude deste efeito positivo sobre o desempenho escolar.

Estudantes das Regiões Norte e Nordeste, segundo suas estimativas, realçam efeito negativo em relação a desempenho escolar para as duas fases do Ensino Fundamental. O destaque fica em relação à Região Sudeste, que evidencia efeito negativo em relação a nível de proficiência em leitura para estudantes da 8ª série/9º ano.

A rede pública tem efeito negativo em relação a desempenho escolar e a razão de chances é de 2,78 para a 4ª série/5º ano e 3,10 para a 8ª série/9º ano, constatando que os alunos(as) da fase final do Ensino Fundamental apresentam chance maiores de estarem nos níveis menores de proficiência que alunos(as) da fase inicial.

Ao comparar as probabilidades no ponto médio de cada nível de proficiência entre as duas fases do ensino fundamental, percebe-se que a 8ª série/9º ano supera a 4ª série/5º ano no nível 1 em 10% e, no nível 2, em 4%, já nos níveis 3 e 4 a primeira fase supera a segunda fase em 8% e 5%, respectivamente.

Os resultados aqui apresentados atestam os estudos de Kleinke (2017), que, utilizando-se do método de regressão logística, analisou o efeito das dimensões: sexo, raça e renda familiar, no desempenho escolar de estudantes do Ensino Médio e Fundamental. Concluiu-se que ocorre desigualdades de gênero, de raça, de renda e regionais, muito marcantes no Brasil e que esses se manifestam no desempenho escolar.

Os resultados do presente trabalho confirmam também os trabalhos de Luz (2006), Chandra e Azimuddin (2013), Silva (2016), Kumari, Shivgotra e Kumar (2017), Franco e Filho (2018), Fernandes et al. (2018), entre outros que constataram efeitos semelhantes de fatores de mesma natureza no desempenho escolar.

Em todo o período analisado, as dimensões: sexo, raça, classe social, região geográfica apresentam em sua distribuição ao longo do tempo desempenho assimétrico à

direita; com exceção de rede (dependência administrativa), que mostra desempenho assimétrico à direita para rede pública e aproximadamente desempenho simétrico para rede privada, resultado este favorecido em termos de comprovação pelo trabalho de Mariano (2016), o qual conclui que os alunos de escolas privadas são mais capazes de converter insumos educacionais em melhores resultados em testes padronizados.

Diante dos resultados do modelo adotado, conclui-se que o desempenho, nos níveis de proficiência em Língua Portuguesa dos alunos(as) do Ensino Fundamental inicial 4ª série/5º ano e final 8ª série/9º ano, compreendido no período de 1999 a 2017 no Brasil, apresenta-se massificado nos níveis inferiores de proficiência e possivelmente por esta razão a avaliação do PISA, nas edições, ao longo deste período, não apresenta aceleração ou desaceleração no desempenho escolar brasileiro em leitura.

No geral, o que se pode inferir em relação aos resultados é que, considerando-se a linguagem como o principal processo de desenvolvimento do capital humano, as crianças e os adolescentes brasileiros não apresentaram no decorrer de todo o período provável ascensão aos níveis superiores de proficiência, o que é alarmante, pois o capital humano é um dos principais elementos de desenvolvimento do país.

A partir destas conclusões, sugere-se, como medida de ação ou efeito de induzir ou ativar o desempenho escolar, a implementação de políticas públicas (programa nacional) com ênfase nos fatores associados que apresentaram efeito negativo em relação ao desempenho escolar, com o firme objetivo de alavancar e melhorar os indicadores de desempenho escolar.

CAPÍTULO II

EFEITO DE FATORES ASSOCIADOS AO DESEMPENHO ESCOLAR NO ENSINO FUNDAMENTAL ENTRE GERAÇÕES NO BRASIL

RESUMO

O estudo visa analisar os fatores associados ao desempenho escolar, com base em uma abordagem de gerações, utilizando uma modelagem econométrica regressiva logística ordinal com dados em corte transversal agrupados. O exercício empírico é fundamentado em dados de avaliação em larga escala do Sistema de Avaliação do Ensino Básico (SAEB) do período de 2003 a 2017, os quais foram dispostos em seis gerações – 2003/2007, 2005/2009, 2007/2011, 2009/2013, 2011/2015, 2013/2017 – de alunos, estratificados em cinco fatores – gênero, raça, classe social, região geográfica e geração, para fins de comparação entre as gerações. Em uma perspectiva temporal, os resultados permitem concluir que os fatores: gênero, raça (com exceção da branca), regiões (Norte e Nordeste) apresentaram possibilidades maiores de serem observados nos níveis inferiores de proficiência em Língua Portuguesa; a primeira fase do ensino fundamental (4ª série/5º ano) apresenta um melhor desempenho em relação à segunda fase (8ª série/9º ano), no decorrer das gerações; o fator classe econômica (social) é o que melhor realça convergência de melhoria em relação ao desempenho escolar. De modo geral, os resultados não apresentam evidências de uma possível ascendência de proficiência em linguagem na transição entre a primeira e segunda fase do ensino fundamental entre as gerações.

Palavras-chaves: Evolução intergeracional na educação. Regressão logística ordinal em dados de corte transversal agrupados. Estimação de fatores no ensino fundamental.

2.1 INTRODUÇÃO

O Relatório Competências e Empregos do Word Bank (2018) enfatiza que a universalização do Ensino Fundamental, no Brasil, teve um significativo avanço nos últimos anos, ainda que o indicador taxa líquida de matrícula no Ensino Fundamental tenha evoluído positivamente apenas um ponto percentual no período entre 2012 e 2017, passando de 96,7 para 97,7, segundo o Anuário Brasileiro da Educação 2018. Não obstante, ainda há uma preocupação constante em relação à qualidade da educação e à relevância de competências adquiridas pelos estudantes, que se refletem nos indicadores avaliativos nacionais e internacionais.

Desde a primeira participação do Brasil, em 2000, no Programa Internacional de Avaliação de Alunos (PISA), o desempenho brasileiro não apresenta mudança estatisticamente significativa, mantendo-se praticamente estável ao longo do tempo. Segundo o relatório do PISA/2015, o desempenho escolar dos alunos brasileiros fica abaixo da média dos alunos de países signatários da Organização para a Cooperação e Desenvolvimento Econômico (OCDE).

O desempenho escolar, conforme Perrenoud (2003, *apud* FERNANDES, 2018), pode ser entendido como a capacidade que os alunos têm de expressar sua aprendizagem e seu conhecimento adquirido no processo ensino-aprendizagem. E, para averiguar mudanças no desempenho em termos de ganhos de aprendizagem (habilidades, competências), segundo o Sistema de Avaliação do Ensino Básico (SAEB, 2017), deve haver uma evolução da proficiência média em relação às últimas edições de avaliação.

Existem três tipos de competências: cognitivas, socioemocionais e técnicas; as quais interagem na formação do indivíduo e que permitem sua plena inclusão na sociedade e no mercado de trabalho. Dentre as três competências, a cognitiva é essencial para formar o início da aprendizagem.

A linguagem e o processo de comunicação, conforme Dias (2010), estão diretamente relacionados ao progresso da inteligência sensório-motora e cognitiva nos anos iniciais de vida. Em complementação, Coll et al. (2004) afirmam que a linguagem é o instrumento mais importante que as pessoas empregam para se comunicar, sendo este um sistema de natureza simbólica que permite “representar” a realidade, não sendo possível dissociar desenvolvimento cognitivo e comunicativo.

Há, segundo Heckman (2007), um desenvolvimento emergente na literatura econômica sobre a importância das condições ambientais iniciais sobre a evolução de habilidades cognitivas em adolescentes e adultos. Essas habilidades são formadas em uma

sequência previsível de períodos sensíveis às influências ambientais e são determinantes de sucesso escolar, econômico e social. As aptidões e as habilidades pessoais que permitem auferir renda são designadas por capital humano, o qual deriva, segundo Sandroni (1999), de aptidões naturais ou adquiridas no processo de aprendizagem.

No contexto dos conceitos econômicos, Blanchard (2011) conceitua capital humano como o conjunto de habilidades dos trabalhadores na economia. Clemente e Higachi (2000, *apud* GARCIA, 2004) enfatizam que o modelo econômico de Lucas considera capital humano como fonte primária de crescimento econômico de um país ou de uma região, ensejam ainda que quanto maior a taxa de acumulação de capital intelectual, maior o crescimento da economia a longo prazo. Schmidt et al. (2010) concordam com o mesmo pensamento sobre o modelo de Lucas quando conclui que a taxa de crescimento do capital humano gera um aumento permanente no crescimento do PIB *per capita*.

Trabalhos desenvolvidos por Schaie (1993, *apud* COLL, 2004), utilizando modelos sequenciais, mostraram que as diferenças nas habilidades cognitivas entre jovens se devem mais às diferenças de geração do que de idade. Para Coll (2004), existem fatores pessoais, ambientais e culturais que estão relacionados a maior ou menor intensidade do declínio intelectual.

Conforme Coll (2004), modelos transversais em pesquisas nas quais são comparadas pessoas que correspondem a diversas gerações tendem a subvalorizar a capacidade intelectual das pessoas que tiveram menos oportunidades educativas. Já pesquisas do tipo longitudinal tendem a ser mais otimistas em relação às capacidades dessas pessoas quando comparadas com elas mesmas ao longo do tempo. O modelo de geração sobreposta é adequado para captar as relações entre indivíduos de diferentes gerações convivendo entre si.

Habilidades cognitivas são medidas em testes avaliativos em larga escala como as do SAEB, e, segundo Viana (2012), são alvo de políticas públicas implementadas para a melhoria dos indicadores resultantes. No entanto, na análise dos determinantes da educação, a extensão longitudinal causa uma série de problemas de análises, que poderiam ser resolvidos por pesquisas do tipo painel.

A pergunta diretriz que desejamos esclarecer é: como os fatores extraescolares associados se relacionam ao desempenho escolar por meio das gerações de estudantes do Ensino Fundamental?

Dentro do contexto apresentado, este trabalho tem por objetivo investigar a relação e o efeito dos fatores socioeconômicos extraescolares – gênero, raça, classe social e região geográfica – no desempenho escolar de crianças e adolescentes do Ensino Fundamental, submetidas às avaliações do SAEB no período compreendido de 2003 a 2017. O arcabouço

metodológico seguirá procedimentos não usuais baseados em um design de gerações sobrepostas, com modelagem econométrica de regressão logística ordinal.

O presente trabalho é composto por cinco seções: introdução; design de gerações sobrepostas e regressão logística ordinal; metodologia; resultado e discussão; e considerações finais.

2.2 DESIGN DE GERAÇÕES SOBREPOSTAS

Nesta seção, é apresentado conceitualmente o modelo de gerações sobrepostas em sua configuração inicial e atual. Relacionam-se, também, estudos em educação que utilizaram empiricamente em sua modelagem gerações sobrepostas.

2.2.1 O modelo de gerações sobrepostas

Analisando um modelo em que a função de produção dependia do capital humano, de jovens e velhos, Kremer e Thomson (1998, *apud* BARRO e MARTIN, 2003) argumentaram que esses dois fatores humanos são complementares e, por o capital humano das primeiras gerações ser pequeno e perfeitamente móvel, as pessoas não tenderão a aumentar seu estoque de capital humano para um nível estacionário, pois a produtividade do jovem não será muito alta se os velhos tiverem pouco capital humano. Dessa forma, consideraram o acúmulo do capital humano gradual – este mecanismo equivale a introduzir o custo de ajustamento na acumulação de capital humano.

Ainda conforme Barro e Martin (2003), os primeiros modelos designados por geração de sobreposição são creditados a Samuelson (1958) e Diamond (1965) e os indivíduos nesses modelos assumem horizontes finitos – em razão de considerar que as pessoas vivem um número fixo de períodos distintos, com o período da idade adulta se sobrepondo ao período da infância – este arcabouço captura os efeitos de horizontes finitos.

O modelo de gerações sobrepostas ou modelo de Diamond, segundo La Fuente (2002), caracteriza-se por uma economia composta por sucessivas gerações de indivíduos de vida finita, de forma que os indivíduos vivem dois períodos. Nesse sentido, em qualquer instante do tempo, coexistem membros de duas ou mais gerações, que interagem uns com os outros, bem como com a realidade dos mercados competitivos de trabalho e capital. O comportamento das famílias no modelo conforme a utilidade dos indivíduos é uma função crescente do consumo no primeiro e no segundo períodos de sua vida.

Um modelo de gerações sobrepostas com agentes heterogêneos é apresentado por Glomm e Ravikumar (1992), que consideram o capital humano obtido com a educação formal como o motor do crescimento. Neste modelo, são utilizadas tecnologias, distribuição de renda e educação pública e privada para distinguir economias, os resultados evidenciam que a desigualdade de renda diminui mais rapidamente sob a educação pública. Em contrapartida, a

educação privada gera maiores rendas *per capita*; além disso, a sociedade escolhe educação pública se os agentes apresentam renda abaixo da média.

Um modelo de gerações sobrepostas em tempo contínuo de acumulação de capital físico e humano é proposto por Pessôa (2001), em que o capital humano é incorporado como um fator limitado, já internalizado pelo trabalhador e acumulado nos primeiros anos de vida. A partir desta proposição, mostra-se que a oferta de ensino público básico e gratuito financiado por impostos que não gerassem distorções na economia (distorcivos) não altera a escolha ótima individual de acumulação do capital humano.

Em um estudo a respeito de educação, igualdade social e crescimento econômico, Gylfason e Zoega (2003) descrevem como o crescimento endógeno pode surgir em um modelo de educação e crescimento de gerações sobrepostas, chegando à conclusão de que a educação parece estimular o crescimento econômico, não apenas aumentando e melhorando o capital humano, mas também o capital físico e o capital social, isto é, reduzindo a desigualdade.

O mecanismo da expectativa de vida explica em grande parte a relação entre desigualdade e acumulação de capital humano, isto é, o que concluem Castelló-Climent e Doménech (2008) em um artigo no qual apresentam um modelo de gerações sobrepostas em que a desigualdade afeta a renda *per capita* dos indivíduos que decidem investir em educação, considerando sua expectativa de vida, dependentes em alguma instância do capital humano de seus pais. Os resultados evidenciam ainda que, no estado estacionário baixo, as crianças nascidas em famílias pobres têm baixa expectativa de vida e trabalham como não habilitadas.

Examinando o nível de renda estadual e o crescimento da renda em um modelo de gerações sobrepostas usando capital físico, capital humano e terra como componentes da função de produção, Artige (2010), conclui que a acumulação do capital humano depende do financiamento parental da educação e do estoque de capital humano agregado.

Assim, de acordo com o contexto apresentado, o modelo de geração sobreposta manifesta-se adequado para captar as relações entre indivíduos de diferentes gerações, convivendo entre si.

Conforme a estrutura teórica do modelo de gerações sobrepostas, é possível idealizar um modelo de geração para o Ensino Fundamental no Brasil, já que, no sistema educacional brasileiro, ocorre um conjunto de transições educacionais para representar as carreiras educacionais.

O ensino básico fundamental, conforme Arretche et al. (2015), conta com duas transições ou ciclos. A primeira transição é composta por indivíduos que entraram no sistema

educacional e completaram os 4 (1ª, 2ª, 3ª e 4ª série; ou 1º, 2º, 3º e 4º ano) primeiros anos do ensino básico e a segunda é constituída de alunos(as) que completaram a 1ª transição e concluíram o ensino até a 8ª série/9º ano. Assim, este arcabouço do modelo de transição educacional assemelha-se ao modelo de Diamond (gerações sobrepostas) de horizonte finito.

Supondo que o modelo de transição educacional no Ensino Fundamental assuma uma estrutura teórica de gerações sobrepostas os indivíduos (jovens) das primeiras gerações (1º ciclo do ensino fundamental) não tenderão a aumentar seu nível de proficiência (capital humano), se o nível de proficiência dos velhos (alunos do 2º ciclo do Ensino Fundamental) forem baixo. Desta forma, assume-se como no modelo de Kremer e Thomson (1998, *apud* BARRO e MARTIN, 2003) que o acúmulo de capital humano (nível de proficiência em leitura) seja gradual.

Diante do contexto apresentado sobre os trabalhos de pesquisas considerando modelagem de gerações na educação, percebe-se uma escassez de trabalhos semelhantes no Brasil, em que, em geral, as pesquisas investigam a evolução do desempenho escolar entre gerações, avaliando mudanças na qualidade das oportunidades educacionais e comparando diferentes anos ou séries do Ensino Fundamental.

Assim, buscou-se estruturar os microdados do SAEB de 2003 a 2017 em dados de corte transversal agrupados em uma estrutura de gerações sobrepostas para o Ensino Fundamental de forma a investigar o efeito dos fatores extraescolares: sexo, raça, classe econômica (social) e região geográfica; no desempenho escolar (nível de proficiência em leitura). Neste contexto de investigação, a análise se faz sobre a evolução dos ciclos escolares a partir da comparação entre os fatores (características) na perspectiva da trajetória das gerações educacionais.

2.3 METODOLOGIA

A presente seção expõe: a fonte de dados e a caracterização das variáveis investigadas, o modelo de regressão logística ordinal, o design amostral que se apresenta numa configuração de gerações sobrepostas estruturadas em dados de cortes transversais agrupados do Ensino Fundamental, e o modelo econométrico empírico estabelecido pela regressão logística ordinal.

2.3.1 Fontes dos dados e das variáveis

A avaliação da educação básica tem como objetivo avaliar a qualidade, a equidade e a eficiência da educação brasileira, caracterizando-se por ser uma avaliação bianual, por amostragem; de larga escala. Utiliza-se de procedimentos e métodos formais, como a Teoria de Resposta ao Item (TRI), e científicos a fim de coletar e sistematizar dados para a produção de informações sobre condições intra e extraescolares que impactam o processo de aprendizagem.

Os dados secundários objeto desta investigação constam da base de microdados do SAEB, referente ao período de 2003 a 2017, disponibilizados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). Os dados de 2003 e 2013 a 2017 são amostrais, os de 2005 a 2011 são censitários. O período de 2003 a 2017 foi selecionado para esta investigação por apresentar diferenças nas trajetórias de desempenho para a 4ª série/5º ano e 8ª série/9º ano nos níveis de proficiência em leitura entre as gerações ao longo do tempo. A quantidade de alunos(as) na amostra e nos censos avaliados pelos SAEB por ano e série/ano são os constantes no quadro 7.

Quadro 7 – Tamanho amostral e censitário por ano e fases do Ensino Fundamental.

Ano	4ª Série/5º Ano	8ª Série/9º Ano
2003	44.711	-
2005	40.726	-
2007	2.217.984	1.257.759
2009	2.449.662	1.497.597
2011	2.226.381	1.546.163
2013	2.013.929	1.515.151
2015	-	1.463.531
2017	-	1.806.777

Fonte: Microdados SAEB 2003/2017.

A tabela 10 mostra o tamanho das subamostras calculadas a partir dos quantitativos do quadro 7. Totalizando 3.164.282 estudantes do Ensino Fundamental, divididos por gerações/ano e fases do ensino básico. As fases do ensino básico estão assim discriminadas: 4ª série/5º ano do período de 2003 até 2013 e 8ª série/9º ano no período de 2007 a 2017.

No cálculo do tamanho de cada subamostra, foi utilizada a expressão (1) de amostragem para população finita, adotando-se um nível de confiança de 95% ($z = 1,96$), erro amostral de 5% e uma proporção amostral p igual a 0,50. Nas tabelas 20 e 21 são apresentados os pesos subamostrais obtidos a partir das amostras e censos do SAEB, as tabelas constam no apêndice H.

Tabela 10 – Demonstrativo de tamanho das subamostras por geração e ano.

GERAÇÃO	ANO							
	2003	2005	2007	2009	2011	2013	2015	2017
1	5.066	-	215.451					
2		4.640	-	259.772				
3			390.369	-	265.102			
4				482.505	-	266.983		
5					394.638	-	259.230	
6						312.825		307.701
Total								3.164.282

Fonte: SAEB – 2003/2017 (Elaborada pelo autor).

As variáveis utilizadas neste trabalho de investigação foram selecionadas a partir dos questionários aplicados aos estudantes nas edições de avaliação do SAEB no período de 2003 a 2017. As características de interesse foram tratadas, passando por um processo de uniformização (codificação) em todo o período analisado, sendo em seguida dispostos em variáveis *dummy* para atender às especificações de entrada dos dados nos aplicativos computacionais STATA, versão 15 e SPSS, versão 20.

Os fatores ou as variáveis independentes, associados ao desempenho escolar, são os especificados conforme quadro 2 incluindo-se as variáveis *dummy* “gera” referente às gerações que permitirá verificar o efeito de possíveis mudanças ao longo do tempo; e excetuando-se da variável rede administrativa (pública, privada).

O desempenho escolar dos estudantes submetidos à avaliação do SAEB é medido por uma escala ou por um nível de proficiência baseado em parâmetros estabelecidos por descritores que associam competências e habilidades; oferecendo uma explicação probabilística

sobre as habilidades demonstradas em cada intervalo da escala. Quanto maior o nível de proficiência, melhor o desempenho do estudante na avaliação do SAEB. A variável dependente “nivproc” neste trabalho expressa o nível de proficiência em leitura ou o desempenho escolar dos estudantes e foram padronizadas e especificados conforme os quadros 5 e 6.

A tabela 10 da forma que se apresenta em seu delineamento amostral, e as variáveis agrupadas e categorizadas conforme os quadros 2; 5 e 6 possibilitam o tratamento dos dados embasados em modelos de regressão com dados em corte transversal agrupados.

Cabe ressaltar que os dados da forma que estão estruturados podem também ser tipificados como pseudo-painel, que trata da análise de dados ao longo do tempo utilizando-se de grupos de características dos indivíduos que se mostram estáveis no tempo (efeito fixo no tempo).

2.3.2 Modelo de regressão logística ordinal

O modelo de regressão logística ordinal, conforme Fullerton (2016), Kumari, Shivgotra e Kumar (2017); e Wooldridge (2010), pode ser derivado do modelo *probit* ordenado a partir de uma estrutura de variável latente com y contínuo.

$$y^* = \beta' x_i + \varepsilon \quad (20)$$

Em que, x_i é um vetor de variáveis explanatórias (independentes), β' é o vetor dos coeficientes de regressão e ε é o erro aleatório que segue distribuição normal padrão; $\varepsilon \sim N(0, 1)$.

Concentrando-se na distribuição cumulativa, os resultados contínuos latentes estão ligados ao resultado ordinal observado y^* através da seguinte equação de medição:

$$y \leq m \quad \text{se} \quad y^* < \tau_m \quad (1 \leq m \leq M) \quad (21)$$

Na qual, m é uma categoria de resultado de M categorias e τ_m são pontos de corte ou limitantes na escala contínua latente. As categorias de y correspondem a intervalos específicos de y^* limitados por pontos de corte, de forma que:

$$y = m \quad \text{se} \quad \tau_{m-1} \leq y^* < \tau_m \quad (1 \leq m \leq M) \quad (22)$$

Em que, τ_{m-1} é o limite inferior e τ_m é o limite superior.

Baseando-se em y^* , a equação para a probabilidade acumulativa pode se expressa como:

$$P(y \leq m/x) = P(y^* < \tau_m/x) \quad (23)$$

Substituindo y^* por $\beta' x_i + \varepsilon$:

$$P(y \leq m/x) = P(\beta' x_i + \varepsilon < \tau_m/x) \quad (24)$$

Subtraindo $\beta' x_i$ de ambos os lados da desigualdade, tem-se:

$$P(y \leq m/x) = P(\varepsilon < \tau_m - \beta' x_i/x) \quad (25)$$

A expressão da probabilidade acumulativa como uma função de τ_m e $\beta' x_i$ pode ser expressa como:

$$P(y \leq m/x) = F(\tau_m - \beta' x_i) \quad (26)$$

Em que, F é a função acumulativa de probabilidade.

A probabilidade para uma particular categoria m é a diferença entre as probabilidades cumulativas para as categorias m e $m - 1$:

$$P(y = m/x) = F(\tau_m - \beta' x_i) - F(\tau_{m-1} - \beta' x_i) \quad (27)$$

Conforme Wooldridge (2010), outras funções de distribuição podem ser usadas em substituição a F ; como a função *logit* que fornece o modelo *logit* ordenado. Assim quando F é a função acumulativa de probabilidade da distribuição logística padrão, com:

$$F(\varepsilon) = e^\varepsilon / 1 + e^\varepsilon \quad (28)$$

Então, F^{-1} é a função de link *logit* que é a inversa da expressão (26):

$$F^{-1}[P(y \leq m/x)] = \text{logit}[P(y \leq m/x)] = \tau_m - \beta' x_i \quad (29)$$

Substituído m por j e τ_m por α_j , conforme notação de Agresti (2010), tem-se função genérica da função acumulada, *Logit*, expressa por:

$$\text{logit}[P(y \leq j)] = \alpha_j - \beta' x_i \quad (30)$$

Neste contexto teórico da regressão logística ordinal, a expressão para o cálculo das previsões das probabilidades acumuladas do modelo logístico ordinal é dada pela expressão:

$$P(Y \leq j) = \frac{1}{1 + e^{-(\alpha_j - X\beta)}} \quad (31)$$

Em que, $j = 1, \dots, c - 1$;

Os programas Stata versão 15 e o SPSS versão 20 possibilitam a estimação em dados em corte transversal agrupados do modelo logístico ordenado.

2.3.3 Modelo empírico de gerações para o Ensino Fundamental

O design amostral, como se apresenta na tabela 10, designa um modelo empírico estruturado de gerações para o Ensino Fundamental no Brasil, no período de 2003 a 2017. Considerando que os(as) alunos(as) estudam por um período “finito” no Ensino Fundamental e as instituições educacionais existem por um período “infinito, a cada período é admitida uma nova geração e egressa uma outra, de forma que o número de gerações existentes em cada período é sempre igual.

Outra motivação para adotar o design de gerações sobreposta parte da suposição que a utilidade (capital humano medido pelo nível de proficiência em leitura) do indivíduo que ingressa no primeiro ciclo (fase inicial) do Ensino Fundamental, no ano “i” (que pertence a geração “t”) é uma função crescente do consumo (aprendizagem) no primeiro (4ª série/5º ano) e segundo (8ª série/9º ano) ciclos do Ensino Fundamental.

Assim, nesse contexto, evidencia-se o esboço da estrutura de gerações, conforme o quadro 8.

Quadro 8 – Estrutura de gerações sobrepostas para o Ensino Fundamental

GERAÇÃO	ANO (série/ano)							
	2003	2005	2007	2009	2011	2013	2015	2017
1	4ª	-	8ª					
2		4ª	-	8ª				
3			4ª	-	8ª			
4				4ª	-	9º		
5					4ª	-	9º	
6						5º		9º

Fonte: SAEB – 2003/2017 (Elaborado pelo autor).

Na construção do conjunto de dados utilizados neste estudo, cada geração foi constituída de um agrupamento independente de corte transversal, sendo 4ª série/5ºano no

primeiro grupo e 8ª série/9ºano no segundo grupo. Em seguida, as gerações foram agrupadas formando um único conjunto de dados. Assim, cada geração é composta de dois períodos (ciclos ou fases do Ensino Fundamental).

Agrupamentos independentes de cortes transversais possibilitam, conforme *Wooldridge* (2013), agrupar amostras aleatórias extraídas de uma mesma população em períodos diferentes de tempo possibilitando a obtenção de estimadores mais precisos e estatísticas de testes mais confiáveis. Isso, claro, se a relação entre a variável dependente e a independente permanecer constante ao longo do tempo (efeito fixo).

O teste de *Hausman* foi realizado para atestar efeito fixo para as características (fatores) em cada geração que abrange dois períodos, o mesmo sendo feito para o agrupamento único, os resultados indicam considerar efeito fixo para os fatores ao longo do tempo. O resultado dos testes são os relacionados na tabela 11.

Tabela 11 – Resultados dos testes de Hausman para gerações

Gerações		Teste de Hausman	
		χ^2	Prob > χ^2
1	2003 / 2007	155,35	0,000
2	2005 /2009	70,61	0,000
3	2007 /2011	32,46	0,000
4	2009 / 2013	584,45	0,000
5	2011 / 2015	31,00	0,000
6	2013 / 2017	70,14	0,000
Completa	2003 / 2017	504,91	0,000

Fonte: Elaborada pelo autor.

2.3.4 Modelo econométrico empírico – regressão logística ordinal

O modelo adotado neste trabalho é dado pela função genérica da função acumulada.

A função neste caso em particular é a *Logit*, dada por:

$$F[\alpha_j - \mathbf{X}^* \boldsymbol{\beta}] = P(Y \leq j) = \frac{1}{1 + e^{-(\alpha_j - \mathbf{X}\boldsymbol{\beta})}} \quad (32)$$

onde, $j = 1, \dots, 3$;

$$\mathbf{X}\boldsymbol{\beta} = \sum_k \beta_k \text{Sexo}_{ik} + \sum_k \beta_k \text{Raça}_{ik} + \sum_k \beta_k \text{Classe}_{ik} + \sum_k \beta_k \text{Região}_{ik} + \sum_k \beta_k \text{Gera}_{ik} \quad (33)$$

em que:

α_j é a constante;

$Sexo_{ik}$ é o vetor de característica do aluno i no gênero k ;

$Raça_{ik}$ é o vetor de característica do aluno i na raça k ;

$Classe_{ik}$ é o vetor de característica do aluno i na classe social k ;

$Região_{ik}$ é o vetor de característica do aluno i na região k ;

$Gera_{ik}$ é o vetor de característica da geração i no tempo k ;

β_k é o vetor de parâmetros de interesse que reporta as relações condicionais dos fatores sobre o desempenho escolar dos (as) alunos(as);

De acordo com a função (32), a probabilidade do nível de proficiência em Língua Portuguesa dos(as) alunos(as) que frequentam a escola pode ser calculada como:

$$P(y \leq 1) = P(y^* \leq \alpha_1) = P(\mathbf{x}\boldsymbol{\beta} + \mathbf{e} \leq \alpha_1) = F[\alpha_1 - \mathbf{x}\boldsymbol{\beta}] \quad (34)$$

$$P(y \leq 2) = P(\alpha_1 < y^* \leq \alpha_2) = F[\alpha_2 - \mathbf{x}\boldsymbol{\beta}] - F[\alpha_1 - \mathbf{x}\boldsymbol{\beta}] \quad (35)$$

$$P(y \leq 3) = P(\alpha_2 < y^* \leq \alpha_3) = F[\alpha_3 - \mathbf{x}\boldsymbol{\beta}] - F[\alpha_2 - \mathbf{x}\boldsymbol{\beta}] \quad (36)$$

$$P(y \leq 4) = P(\alpha_3 < y^* \leq \infty) = 1 - F[\alpha_3 - \mathbf{x}\boldsymbol{\beta}] \quad (37)$$

A regressão logística ordinal se mostra uma alternativa adequada para este trabalho de investigação por ser um modelo apropriado para resultados ordinais e também pelo fato de, conforme Wooldridge (2010), ser adaptável para dados em corte transversal agrupados, as estimativas calculadas a partir do modelo proposto serão geradas com a utilização do STATA versão 15.1 e SPSS versão 20.

2.4 RESULTADOS E DISCUSSÃO

A presente seção manifesta os resultados descritivos e econométricos.

2.4.1 Análises descritivas dos dados amostrais

A tabela 12 apresenta o demonstrativo da proporção de alunos(as) nos níveis de proficiência por geração, série/ano.

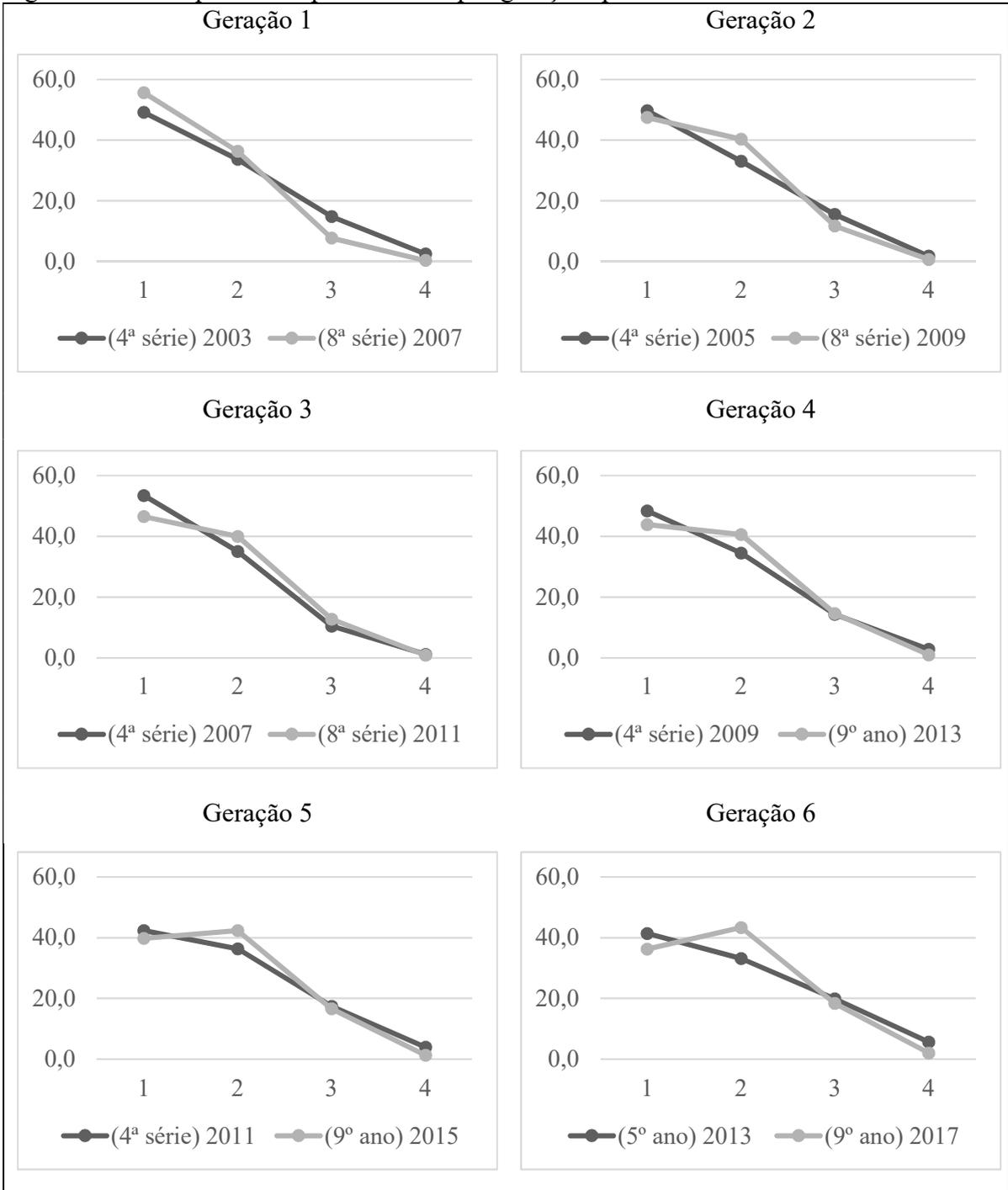
Tabela 12 – Proporção de alunos(as) nos níveis de proficiência (Série/Ano)

Geração	Nível	Série / Ano (%)	
		(4ª série) 2003	(8ª série) 2007
1	1	49,1	55,6
	2	33,7	36,3
	3	14,8	7,7
	4	2,5	0,3
		(4ª série) 2005	(8ª série) 2009
2	1	49,7	47,5
	2	33,1	40,2
	3	15,5	11,7
	4	1,8	0,6
		(4ª série) 2007	(8ª série) 2011
3	1	53,5	46,5
	2	35,0	40,0
	3	10,4	12,7
	4	1,1	0,8
		(4ª série) 2009	(9º ano) 2013
4	1	48,4	43,9
	2	34,5	40,6
	3	14,3	14,6
	4	2,8	1,0
		(4ª série) 2011	(9º ano) 2015
5	1	42,4	39,8
	2	36,3	42,4
	3	17,4	16,6
	4	3,9	1,3
		(5º ano) 2013	(9º ano) 2017
6	1	41,4	36,3
	2	33,1	43,3
	3	19,9	18,4
	4	5,6	2,0

Fonte: Microdados SAEB/Prova Brasil (Elaborada pelo autor).

A figura 1 ilustra o desempenho dos(as) alunos(as) por nível de proficiência para cada geração.

Figura 1 – Desempenho das proficiências por gerações para o Ensino Fundamental

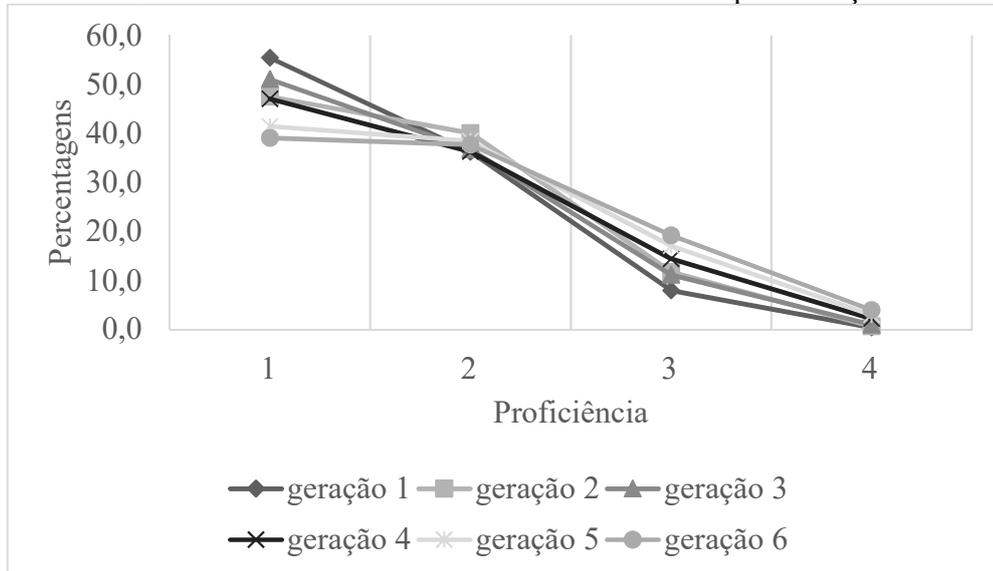


Fonte: SAEB 2003/2017 (Elaborada pelo autor).

A tabela 12 juntamente com as ilustrações da figura 1 evidenciam que, a cada geração, o desempenho dos(as) alunos(as) da 4ª série/5ºano evolui melhor nos níveis 3 e 4,

superando o desempenho dos alunos da 8ª série/9º ano. Já este último tem certa evolução do nível 1 para o nível 2 e 3, mas praticamente se mantém constante ao longo das gerações neste nível.

Gráfico 15 – Demonstrativo de Proficiências em leitura por Gerações



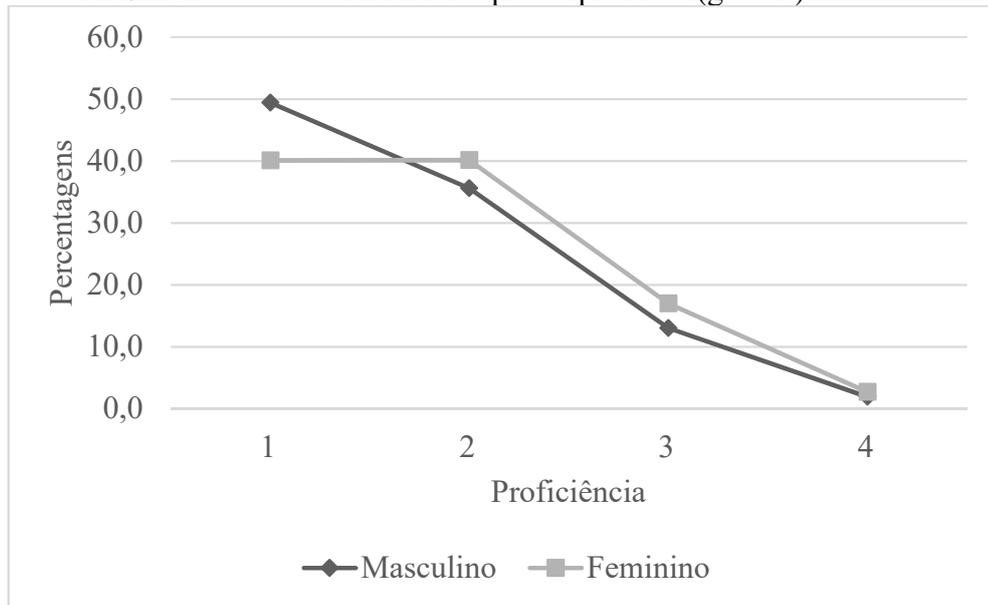
Fonte: Microdados SAEB/Prova Brasil (Elaborado pelo autor).

O que se percebe, conforme o gráfico 15, é que, com o transpassar das gerações, os (as) alunos(as) do Ensino Fundamental em sua proporção média apresentam uma evolução em desempenho nos níveis de proficiência, ainda que as diferenças sejam pequenas para os níveis maiores. Uma preocupação emerge com relação aos alunos(as) da 8ª série/9º ano, que são proporcionalmente superados pelos(as) alunos(as) da 4ª série/5º ano nos níveis 3 e 4, evidenciando não haver uma acumulação de capital humano ou cognitivo em relação à Língua Portuguesa, que se supõe ainda poder permanecer por um longo período com os indicadores do relatório PISA estacionários.

Os gráficos 16 a 19 mostram o desempenho médio proporcional calculado a partir da proporção média de cada geração de alunos(as) por nível de proficiência em leitura dos fatores (variáveis, características): gênero, raça, classe econômica e região geográfica.

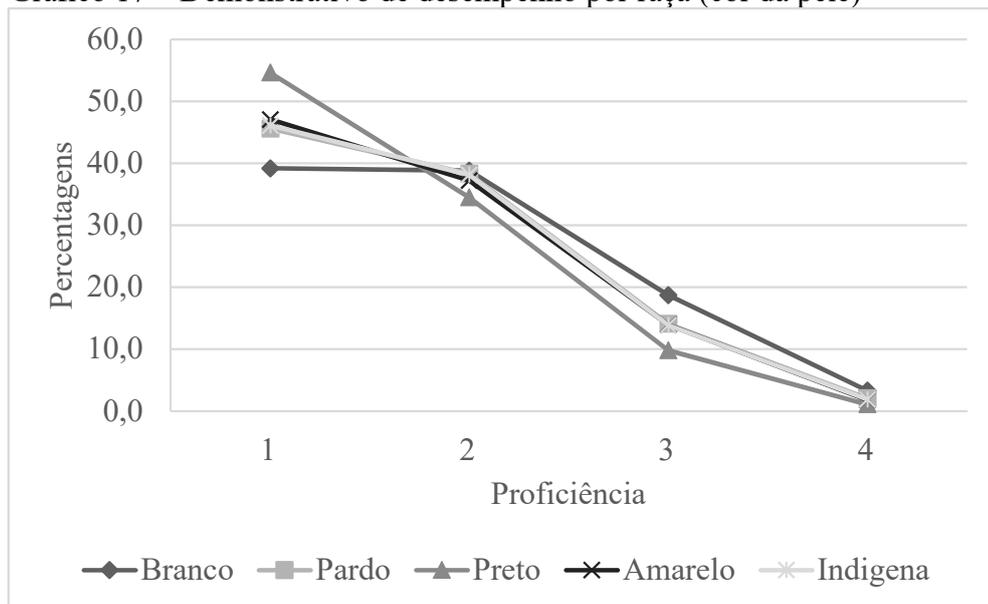
O gráfico 16 mostra que o desempenho do sexo feminino em relação ao masculino, no decorrer das gerações, se mantém proporcionalmente maior nos níveis 2, 3 e 4 em todas as gerações.

Gráfico 16 – Demonstrativo de desempenho por sexo (gênero)



Fonte: Microdados SAEB/Prova Brasil (Elaborado pelo autor).

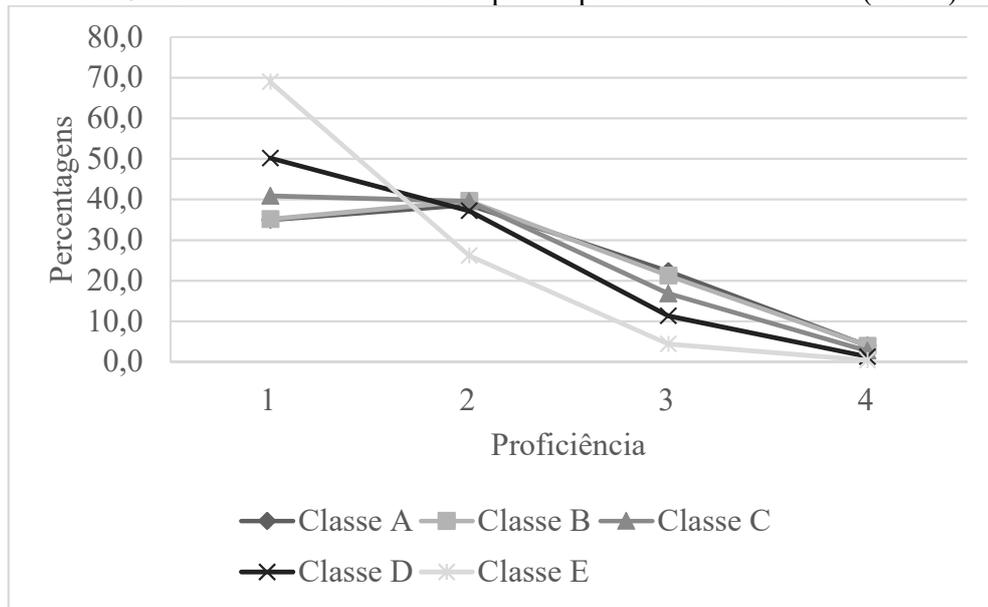
Gráfico 17 – Demonstrativo de desempenho por raça (cor da pele)



Fonte: Microdados SAEB/Prova Brasil (Elaborado pelo autor).

O gráfico 17 evidencia uma situação das mais críticas para os alunos(as) que se autoafirmaram serem de cor de pele preto(a), pois apresentam maior proporção no nível 1, sendo que nos níveis 2 a 4 apresenta-se sempre em menor proporção quando comparados às demais raças. Há uma segmentação evidente entre o desempenho relacionado à raça que se agrupa em: preto; pardo, amarelo e indígena; e branco. No contexto apresentado, conclui-se que ao longo das gerações persistem as desigualdades raciais e que estas possivelmente se refletem no desempenho escolar ao longo das gerações.

Gráfico 18 – Demonstrativo de desempenho por classe econômica (social)

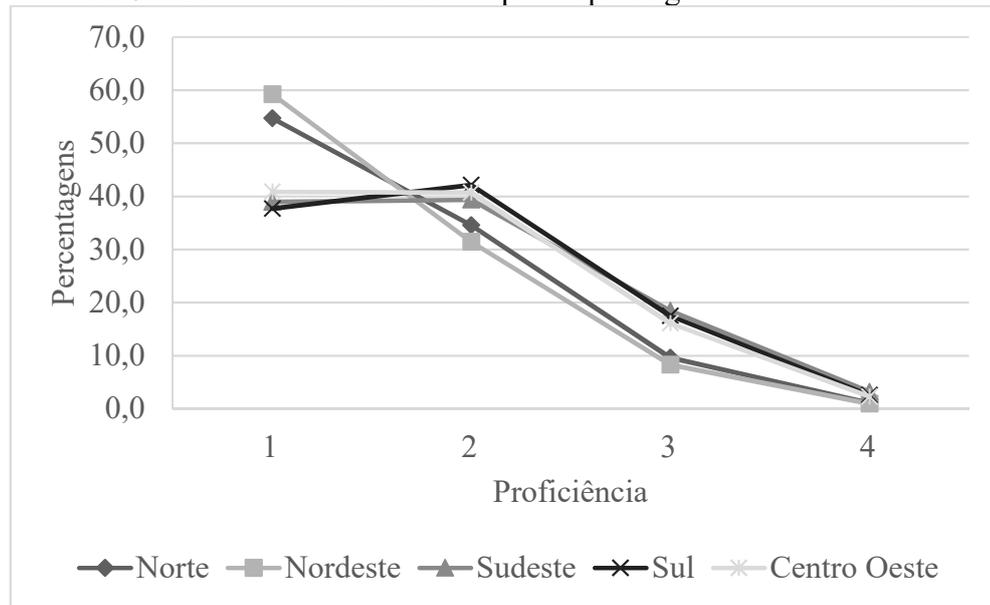


Fonte: Microdados SAEB/Prova Brasil (Elaborado pelo autor).

O desempenho escolar em relação às classes econômicas (sociais), demonstrado no gráfico 18, realça a diferença que há entre as classes econômicas. Percebe-se claramente que as classes econômicas com menor poder aquisitivo se apresentam em maiores proporções no nível 1 e a partir do nível 2 há uma inversão, com as classes mais altas apresentando proporções maiores nos níveis mais altos de proficiência. Ainda que a distribuição de alunos(as) se apresente assimétrica à direita para todas as classes sociais, a situação das classes de menor renda é mais crítica. Evidencia-se aqui a desigualdade persistente de renda ao longo das gerações, como no caso analisando anteriormente.

O gráfico 19 ilustra o desempenho dos(as) alunos (as) em relação às regiões administrativas brasileiras. Tem-se uma segmentação clara ao longo dos níveis de proficiência em que as Regiões Norte e Nordeste apresentam proporções maiores de alunos(as) no nível 1 e proporções menores nos níveis 2 a 3. As Regiões Sul, Sudeste e Centro Oeste se posicionam inversamente ao fenômeno observado para o Norte e Nordeste, sugerindo que as desigualdades regionais são persistentes em relação ao desempenho escolar ao longo das gerações.

Gráfico 19 – Demonstrativo de desempenho por região administrativa



Fonte: Microdados SAEB/Prova Brasil (Elaborado pelo autor).

Os resultados apontados através da análise descritiva são concordantes com os estudos de Luz (2006), que evidencia diferenças significativas nas características: renda e região geográficas quando relacionadas ao desempenho escolar e observa nesta relação reprodutibilidade da desigualdade social. Também concordam com o trabalho de Silva (2016), que chega à conclusão de que a renda impacta de forma positiva o rendimento dos alunos.

Os resultados aqui apresentados de forma descritiva também entram em concordância com pesquisas internacionais como os de Chandra e Azimuddin (2013) e de Kumari, Shivgotra e Kumar (2017), que apontam para o fato de que o nível de proficiência está associado à diferença entre grupos de diferentes condições socioeconômicas.

2.4.2 Resultados e análise do método econométrico

Na tabela 13, constam as estimativas dos efeitos esperados pelos fatores associados no modelo econométrico e a variável “geração” que foi incluída como variável *dummy* para todas as gerações com o propósito de avaliar se a população possui ou não distribuição diferente ao longo do tempo, já que se trata de agrupamentos de corte transversal.

Tabela 13: Estimativas dos efeitos esperados pelos fatores associados no modelo econométrico para as gerações 1 a 6.

Variáveis/Fatores	Estimativa	Erro Padrão	Estatística Z	p-value	Razão de Chances
Sexo (Masc.)	- 0,409	0,003	- 147,35	0,000	1,51
Raça (Branca)	0,177	0,008	21,71	0,000	0,84
Raça (Parda)	- 0,021	0,008	- 2,59	0,010	1,02
Raça (Negra)	- 0,369	0,009	- 41,54	0,000	1,45
Raça (Amarela)	- 0,040	0,010	- 3,68	0,000	1,04
Classe Social (A)	0,862	0,034	25,18	0,000	0,42
Classe Social (B)	0,974	0,011	87,87	0,000	0,38
Classe Social (C)	0,755	0,011	71,63	0,000	0,47
Classe Social (D)	0,459	0,011	43,13	0,000	0,63
Região (Norte)	- 0,461	0,007	- 68,99	0,000	1,59
Região (Nordeste)	- 0,592	0,006	- 104,50	0,000	1,81
Região (Sul)	0,149	0,005	28,73	0,000	0,86
Região (Sudeste)	0,032	0,006	5,48	0,000	0,97
Gera 1	- 1,150	0,006	- 186,20	0,000	3,16
Gera 2	- 0,779	0,006	- 138,28	0,000	2,18
Gera 3	- 0,743	0,004	- 166,30	0,000	2,10
Gera 4	- 0,526	0,004	- 118,96	0,000	1,69
Gera 5	- 0,195	0,004	- 43,84	0,000	1,22

Fonte: Elaborada pelo autor.

Todos os coeficientes das variáveis (fatores) estimadas através do modelo econométrico empírico mostraram-se estatisticamente significantes ao nível de 5%.

Conforme a tabela 13, têm-se os seguintes resultados:

- O fato de o indivíduo ser do sexo masculino aumenta em 1,51 a razão de chances de ser observado nas classes de menor ordem de proficiência em leitura ao longo das gerações;
- Os alunos da raça branca apresentam uma relação positiva se comparados aos níveis de proficiência, ou seja, possuir a pele de cor branca diminui em 0,84 a chance de serem observados nos níveis baixos de proficiência ao longo das gerações;
- As raças parda, negra e amarela apresentam uma relação negativa relacionada ao desempenho escolar ao longo das gerações com chances maiores de serem observados nos níveis mais baixos de proficiência possuindo respectivamente: 1,02, 1,45 e 1,04; razão de

chances de serem observados nos níveis menores. É evidenciada situação de persistência de desigualdade racial relacionada ao desempenho escolar ao longo das gerações;

- Todas as classes econômicas, A, B, C e D, segundo os coeficientes estimados, apresentam relação positiva, demonstrando razões de chances menores de serem observados nos níveis mais baixos de proficiência;
- As Regiões Norte e Nordeste possuem relação negativa acerca da proficiência em Língua Portuguesa, apresentando razão de chances maiores de serem observadas nos níveis menores de proficiência. A Região Nordeste com uma razão de chance de 1,81 e a Região Norte com 1,59. As Regiões Sul e Sudeste apresentam relação positiva em se tratando de desempenho nos níveis de proficiência, tendo diminuído conforme suas razões de chance serem observadas nos níveis menores de proficiência em Língua Portuguesa no decorrer das gerações.
- As gerações apresentam efeito negativo em relação a nível de proficiência em leitura. A geração 1 (3,16) tem três vezes mais chances de ser observada nos níveis menores de proficiência; a geração 2 (2,18) e a geração 3 (2,10) têm 2 vezes mais chances de se encontrarem nos menores níveis e as gerações 4 e 5, uma vez apenas na mesma situação das anteriores. Evidencia-se uma evolução em relação às gerações, no entanto a distribuição nas proficiências ao longo do tempo (gerações) reflete a mesma distribuição de aglomeração nos níveis menores de proficiência.

Os resultados dos efeitos marginais dos fatores em relação aos níveis de proficiências são os especificados na tabela 14.

Tabela 14 - Efeito marginal dos fatores associados por nível de proficiência. (continua)

Variáveis/Fatores	dy/dx (1)	dy/dx (2)	dy/dx (3)	dy/dx (4)
	0,413	0,413	0,152	0,022
Sexo (Masc.)	0,099	- 0,040	- 0,049	- 0,009
Raça (Branca)	- 0,042	0,017	0,022	0,004
Raça (Parda)	0,005	- 0,002	- 0,003	- 0,000
Raça (Negra)	0,091	- 0,042	- 0,041	- 0,007
Raça (Amarela)	0,009	- 0,004	- 0,005	- 0,001
Classe Social (A)	- 0,183	0,025	0,129	0,028
Classe Social (B)	- 0,212	0,042	0,141	0,029

Tabela 14 - Efeito marginal dos fatores associados por nível de proficiência. (continuação)

Variáveis/Fatores	dy/dx (1)	dy/dx (2)	dy/dx (3)	dy/dx (4)
	0,413	0,413	0,152	0,022
Classe Social (C)	- 0,181	0,077	0,089	0,016
Classe Social (D)	- 0,108	0,037	0,059	0,011
Região (Norte)	0,113	- 0,055	- 0,049	- 0,008
Região (Nordeste)	0,145	- 0,068	- 0,065	- 0,011
Região (Sul)	- 0,036	0,014	0,018	0,003
Região (Sudeste)	- 0,008	0,003	0,004	0,001
Gera 1	0,277	- 0,159	- 0,102	- 0,016
Gera 2	0,191	- 0,101	- 0,077	- 0,013
Gera 3	0,182	- 0,089	- 0,079	- 0,013
Gera 4	0,129	- 0,059	- 0,059	- 0,010
Gera 5	0,047	- 0,020	- 0,023	- 0,003

Fonte: Elaborada pelo autor.

A tabela 14 reflete a probabilidade de os fatores se encontrarem em cada nível de proficiência, assim: nos níveis 1 (um) e 2 (dois), têm-se 41,3%; no nível 3 (três) 15,2% e no nível 4 (quatro) 2,2%.

Percebe-se na tabela 14 que diminuem as probabilidades dos fatores: raça (branca), classes econômicas (A, B, C e D), regiões (Sul e Sudeste) de se encontrarem no nível 1; e aumenta para os fatores: raças (parda, negra e amarela), regiões (Norte e Nordeste) e gerações (1, 2, 3, 4 e 5). Nos níveis 2, 3 e 4 esta situação se inverte. Ainda nessa tabela, são evidenciados os resultados analíticos.

Uma forma mais atraente de leitura dos resultados em alternativa aos coeficientes estimados, à razão de chances e aos efeitos marginais se dá através das probabilidades previstas de associação a cada categoria dos fatores. As probabilidades previstas são as especificadas na tabela 15.

Tabela 15 - Probabilidades previstas por nível de proficiência (continua)

Variáveis	Categorias	Nível de Proficiência			
		1	2	3	4
Sexo	Masculino	46,6	37,0	14,2	2,2
	Feminino	38,0	41,0	18,1	2,9
Raça	Branco	36,6	39,9	19,9	3,6
	Pardo	43,5	39,3	15,0	2,2

Tabela 15 - Probabilidades previstas por nível de proficiência (continuação)

Variáveis	Categorias	Nível de Proficiência			
		1	2	3	4
	Preto	52,0	36,1	10,7	1,2
	Amarelo	44,1	38,9	15,1	2,0
	Indígena	43,8	39,3	14,8	2,1
Classe Social	Classe A	33,4	39,4	23,2	4,1
	Classe B	34,1	40,1	21,7	4,1
	Classe C	39,9	40,0	17,3	2,8
	Classe D	49,2	37,8	11,7	1,3
	Classe E	68,0	27,0	4,6	0,4
Região	Norte	51,1	36,9	10,8	1,3
	Nordeste	55,6	33,9	9,4	1,1
	Sudeste	36,0	40,6	19,9	3,5
	Sul	35,5	43,0	18,7	2,8
	Centro Oeste	38,5	41,8	17,3	2,5
Geração	1	54,7	36,7	8,2	0,4
	2	46,4	40,6	12,2	0,7
	3	46,7	39,4	12,7	1,2
	4	43,0	38,3	16,1	2,5
	5	37,4	40,1	19,1	3,4
	6	33,7	39,1	16,3	2,6

Fonte: Elaborada pelo autor.

Conforme a tabela 15, ficam evidenciados os seguintes resultados:

- Em relação ao gênero, o sexo masculino apresenta probabilidade prevista maior somente para o nível 1 quando comparado com o sexo feminino. A situação se inverte nos níveis de 2 a 4, em que o sexo feminino detém maioria;
- A raça branca apresenta probabilidades maiores que todas as raças nos níveis 2 a 4 e menor probabilidade prevista no nível 1. Entre as raças, a negra é a que apresenta menores probabilidades previstas nos níveis 2 a 4 e maior no nível 1;
- Em relação às classes econômicas no nível 1, quanto menor a classe, maior a probabilidade prevista. Nos demais níveis de proficiência 2 a 4, as probabilidades crescem à medida que aumenta o poder aquisitivo, ou seja, aumenta a classe econômica;
- As Regiões Norte e Nordeste mostram probabilidades previstas maiores no nível 1 e menores nos níveis 2 a 4, quando comparado com as outras regiões;
- As gerações de 1 a 6 apresentam probabilidades previstas decrescentes no nível 1, crescentes nos níveis 3 e 4, e estáveis no nível 2, oscilando entre 36,7 e 40,6.

No contexto apresentado em relação à leitura e à análise das tabelas consolidadas a partir do método econométrico e em comparação com os resultados estatístico descritivos, este trabalho se coloca em concordância com os trabalhos de outros autores como: Coll (2004), quando este enfatiza que existem fatores pessoais e ambientais que estão relacionados com a intensidade do nível intelectual; Castelló-Climent e Doménech (2008), quando ressalta que indivíduos de baixa classe econômica apresentam menores habilidades (proficiências); Silva (2016), que apresenta a renda como dotação educacional das famílias de forma significativa no rendimento escolar dos estudantes; e tantos outros autores que evidenciaram desigualdades persistentes de ordem racial, regional, de classe social e intergeracional e que se utilizaram de outros métodos quantitativos e qualitativos.

2.5 CONSIDERAÇÕES FINAIS

Conforme os resultados estatísticos, econométricos e analíticos da presente investigação, o que se percebe é que a cada geração de alunos(as) do Ensino Fundamental não ocorrem evidências de diferenças significativas nos níveis de proficiência em Língua Portuguesa. Ainda que sejam consideradas as pequenas proporções de acréscimos nos níveis 3 e 4 nos desempenhos escolares, a maioria dos(as) alunos(as) permanece de forma probabilística nos níveis inferiores de proficiência.

O gênero feminino, a cada geração, mantém-se com melhores resultados em relação ao gênero masculino, este segundo não apresenta expectativas a cada geração.

Os (as) alunos(as) que se autoafirmaram possuir cor de pele preto(a), ao longo das gerações, não aparentam mudanças positivas em direção aos níveis maiores de proficiência, sendo estes acompanhados das raças: parda, amarelo e indígena, o que se evidencia uma situação de desigualdade racial em relação à educação.

Em relação às classes sociais, há evidências de uma clara desigualdade de renda que afeta, sobremaneira, o desempenho escolar dos(as) alunos(as). No entanto, os resultados indicam uma convergência probabilística para os níveis mais elevados de proficiência para todas as classes sociais, considerando que nas classes mais baixas essa convergência é moderadamente menor.

As Regiões Norte e Nordeste, de fato, possuem os piores resultados e estes persistem ao longo das gerações. Não diferentemente, as demais regiões geográficas, apesar de apresentarem proporções significativas de alunos(as) nos níveis 2, 3 e 4 de proficiência, ainda detêm resultados não muito animadores ao longo do tempo. Conforme os resultados mostrados, há uma clara desigualdade regional em relação ao desempenho escolar.

Fica evidente, então, que o fator renda parece impactar melhor o desempenho escolar, e um resultado negativo e persistente na economia acarretaria prejuízos à qualidade do ensino no âmbito do acúmulo do capital humano (cognitivo) da população brasileira, refletindo-se claramente nos indicadores de avaliação educacional nacionais e estrangeiros.

O presente trabalho vem a contribuir com a literatura relacionada a fatores associados ao desempenho escolar na medida em que se utiliza de uma abordagem geracional fora dos moldes tradicionais de analisar ano a ano e apenas em um determinado local geográfico- estadual ou municipal. A modelagem econométrica também se faz diferenciada pelo fato de utilizar a regressão logística ordinal que se coloca como alternativa aos modelos

lineares que conforme alguns autores internacionais não favorecem uma análise apropriada a estudos desta natureza.

Ressalta-se também a importância deste trabalho em se tratando do tamanho amostral utilizado que possibilitou uma significância estatística bastante razoável na estimação dos parâmetros desejados e nas probabilidades previstas dos fatores em relação aos níveis de proficiências.

CAPÍTULO III

INFÂNCIA E ADOLESCÊNCIA NO ENSINO FUNDAMENTAL: EDUCAÇÃO E DESIGUALDADE SOCIOECONÔMICA NO BRASIL

RESUMO

Neste artigo, conduziu-se uma análise sobre as condições socioeconômicas de crianças e adolescentes do 5º ano do Ensino Fundamental que foram submetidas à avaliação do SAEB/Prova Brasil no ano de 2015. O propósito foi estabelecer, por meio de análise estatística e modelagem econométrica empírica usando regressão logística ordinal, elementos que apontassem possíveis disparidades entre crianças e adolescentes que estão distribuídas em diferentes classes econômicas (sociais) e suas relações familiares quanto a escolaridade dos pais, moradia com os pais, incentivo a estudar e ir à escola e trabalho fora de casa. Os resultados evidenciam que estudantes do gênero masculino apresentam tendência a níveis inferiores de proficiência em leitura; estudantes de diferentes raças não apresentaram efeito positivo em relação ao desempenho escolar (nível de proficiência superior), não se evidenciando desigualdade em relação aos fatores raça ou cor da pele. Percebe-se, no entanto, que os estudantes negros estão concentrados nas classes sociais C e D e são persistentes na busca por melhores resultados no desempenho escolar; as Regiões Norte e Nordeste ainda possuem resultados preocupantes em relação ao desempenho escolar; as crianças e os adolescentes, cujos pais se fazem mais presentes na moradia, incentivam os filhos a irem à escola, a estudarem e a não trabalharem fora de casa, apresentam maiores chances de serem observadas nos níveis superiores de proficiência. De modo geral, não se evidencia o fenômeno da desigualdade socioeconômica e da racial em relação ao desempenho escolar.

Palavras Chaves: Desigualdade racial. Desigualdade de classes. Desigualdade e educação.

3.1 INTRODUÇÃO

Recentemente, no Brasil, houve uma redução das desigualdades sócio econômica devido ao alcance de melhores condições de trabalho e de políticas de valorização do salário mínimo. Apesar desse quadro, o país apresenta uma alta desigualdade social, um baixo desempenho educacional, bem como uma expressiva desigualdade de renda.

Um outro fenômeno bastante desafiador em nosso país é o da desigualdade racial, sendo comum encontrar, nas literaturas das ciências econômicas e sociais, estudos que visam analisar os efeitos da condição racial nas sociedades de classes, destacando quase sempre as diferenças entre brancos e negros e a comparação de características socioeconômicas.

O que se tem na realidade são diversos estudos na perspectiva econômica, sociológica e antropológica como o de Santos e Silva (2018), que analisam o racismo científico no Brasil, que enfatiza a discriminação existente entre as raças nas relações econômicas e políticas, assumindo que a raiz dessa problemática decorre da desigualdade de classes. Também, o de Campos (2018), que demonstra a influência dos movimentos negros brasileiros na criação e na elaboração de programas sociais e políticos e leis no âmbito educacional; ou ainda o de Duarte e Santos (2017), que discute sobre desigualdade social, educação e demandas populares.

Cabe ressaltar, entre os trabalhos mais recentes que se utilizam de métodos quantitativos, as pesquisas de Filho e Kirschbaum (2015) e de Toste Daflon et al. (2017):

Numa avaliação sobre desigualdade de renda e educação no Brasil, Filho e Kirschbaum (2015) chegaram à conclusão de que brancos e amarelos têm renda mais elevada do que pretos e pardos, mesmo quando se consideram níveis de escolarização maiores; e que os pais brancos são mais escolarizados que os pais negros.

Ao analisarem as dimensões cor e posição econômica, Toste Daflon et al (2017), utilizando modelos de regressão logística com foco nas variáveis: sexo, região, idade, níveis educacionais e renda, chegaram à conclusão de que pretos e pardos apresentam índices idênticos em diversas variáveis socioeconômicas, principalmente relacionadas à escolaridade e à renda que se concentra entre aqueles de renda mais baixa (5,1% dos pretos e 5,0% dos pardos no 10º decil).

“Os pardos estão extremamente próximos dos pretos no que toca à desigualdade de oportunidades e de resultados e ao mesmo tempo longe destes quanto à percepção do preconceito e da discriminação.”
(TOSTE DAFLON, CARVALHAES e FERES JÚNIOR, 2017).

O resultado dos trabalhos de Filho e Kirschbaum (2015), Toste Daflon et al (2017), entre outros, sugere uma seleção que diferencia pretos e pardos à medida que estes avançam no nível de escolaridade (e renda), ocorrendo uma discriminação dos pretos. Esses resultados, além da desigualdade econômica, mostram também a desigualdade racial, que possui raízes históricas no Brasil.

No Brasil, o Instituto Brasileiro de Geografia e Estatística (IBGE) classifica a população conforme “cor ou raça” em cinco categorias: branca, preta, parda, amarela e indígena. As categorias preta e parda compartilham características de uma mesma raça, formando assim a raça negra. Segundo Osório (2003), a agregação de pretos e pardos e sua designação como negra justificam-se estatisticamente pela uniformidade de características socioeconômicas dos dois grupos e teoricamente pela origem comum das desigualdades dos pretos e pardos em relação aos brancos.

Arretche (2015) enfatiza três linhas de interpretação nos estudos sobre a situação social do negro e as relações sociais no Brasil: os negros possuem as piores condições sociais devido à origem escrava e não ao preconceito racial, mas sim o de classe; o preconceito racial tem origem na escravidão e é incompatível com uma sociedade de classes; a discriminação racial gera desigualdade por meio da desqualificação competitiva dos negros.

Numa nova perspectiva, o presente trabalho se propõe a contribuir com a temática desigualdade entre classes socioeconômicas (renda, raça e educação), sob a ótica das crianças e dos pré-adolescentes e suas relações em família, atendidos no sistema educacional brasileiro (Ensino Fundamental) e que foram avaliadas pelo Sistema de Avaliação da Educação Básica (SAEB) no ano de 2015, edição que estava disponível no momento de elaboração deste artigo.

A pergunta que se faz, a partir do contexto apresentado, é a respeito da realidade de crianças e adolescentes atendidos no primeiro ciclo do Ensino Fundamental e qual o efeito dos fatores socioeconômicos: gênero, raça, classe econômica, escolaridade dos pais, moradia com os pais, trabalho fora de casa, incentivo à escola e estudo, rede de ensino (pública, privada) e região geográfica para o desempenho escolar (nível de proficiência em leitura como *proxy* para capital intelectual ou humano).

A principal contribuição que se faz com esta investigação é tratar de questões referentes a desigualdade de renda, raça e desempenho educacional entre crianças e adolescentes no ciclo inicial do ensino fundamental abrangendo todas as regiões brasileiras, além é claro da adoção da regressão logística ordinal como método econométrico para análise dos dados.

O presente trabalho se compõe das seguintes seções: Introdução, Desigualdade socioeconômica e racial no Brasil, Desigualdade de renda, raça e educação, Estudos sobre desigualdade racial e de classes, Metodologia, Resultados e discussão e Conclusões finais.

3.2 DESIGUALDADE SOCIOECONÔMICA E RACIAL NO BRASIL

A contribuição dos aumentos no salário mínimo para uma melhor distribuição de renda no Brasil, no período compreendido entre 1995 e 2005, foi investigado por Saboia (2007) através de um ensaio do efeito desses aumentos sobre os rendimentos oriundos de trabalho, pensões e aposentadorias, confirmando a existência de uma redução das desigualdades no rendimento familiar *per capita*.

Estudos realizados no Brasil pelo Instituto Brasileiro de Geografia e Estatística (IBGE), no ano de 2017, com dados da Pesquisa Nacional por Amostra de Domicílios (PNAD), referente ao período de 2005 a 2015, reúne análises que retratam a existência de uma redução na desigualdade socioeconômica nos últimos anos.

O relatório Desenvolvimento Humano para Além das Médias publicado em 2017, pelo Programa das Nações Unidas para o Desenvolvimento (PNUD), mostrou que o Brasil se posiciona no 10º lugar dentre 143 países, como um dos mais desiguais do mundo. Mensuração feita com base no coeficiente de Gini, que mede as desigualdades sociais e o nível de concentração de renda da população, tal informação revela como o Brasil se encontra no cenário mundial em termos de desigualdade.

Conforme o contexto apresentado, parece haver uma contradição na literatura a respeito da desigualdade socioeconômica no Brasil.

No entanto, o importante papel da política de valorização do salário mínimo, aliada à reorganização do trabalho formal e suas melhores condições, é ressaltado por Dedecca (2014) como responsáveis pela redução do índice de Gini, observada justamente a partir de 2005. Contudo, chama atenção o fato de que no Brasil, embora tenha havido queda na desigualdade socioeconômica, seu nível ainda é alto e atinge principalmente a população que compõe a raça negra.

3.2.1 Desigualdade de renda, raça e educação

No Brasil, dados da Pesquisa Nacional por Amostra de Domicílio (PNAD) de 2011 mostram que cada ano adicional de estudo corresponde a um aumento médio de renda superior

a 10%. Essa dependência da renda com o número de anos de estudos atinge fortemente a renda familiar dos estudantes.

A renda das pessoas, segundo Helene (2013), depende fortemente da quantidade e da qualidade da educação formal que recebem. Assim sendo, a educação das crianças e dos jovens possui uma dependência forte de sua renda familiar.

O sistema educacional brasileiro, conforme Helene (2013), é um importante instrumento a perpetuar a desigualdade, não surpreendendo que a renda média dos 10% mais ricos seja 56 vezes maior que a renda média dos 10% mais pobres. Pensamento semelhante é compartilhado por Arretche (2015), que considera o acesso à educação formal um mecanismo produtor de desigualdades. Nesse contexto, a combinação renda-educação é fator de concentração de renda e de reprodução de desigualdades.

Ao analisarem a relação entre a oferta de trabalhadores com seus diversos níveis de escolaridade e os possíveis impactos nas diferenças de rendimentos, bem como das divergências de renda no mercado de trabalho brasileiro de 1960 a 2010, Filho e Kirschbaum (2015) descobriram que entre os anos de 2000 a 2010 ocorreu uma queda nas desigualdades de renda impactadas pelo trabalho, e que a partir de 1990, passou a crescer o número de acessos a todos os níveis escolares, destacadamente ao Ensino Fundamental e Médio completos.

Após um exame da relação entre as características socioeconômicas das famílias e as possibilidades de aumento de escolaridade dos jovens brasileiros, considerando o período de 1960 a 2010, Ribeiro et al. (2015) observaram que houve ampliação no acesso aos níveis escolares básicos, devido aos ganhos de acessibilidade no Ensino Fundamental, mas pontuam que houve ao mesmo tempo o deslocamento para os níveis imediatamente superiores da trajetória educacional do efeito de fatores limitadores da progressão educacional – escolaridade dos pais, renda familiar, sexo e cor – muito presente nas décadas de 1960 e 1970 no ensino fundamental completo.

Ainda sobre as possibilidades de progressão escolar dos jovens brasileiros, Ribeiro et al. (2015), ao observarem o efeito da escolaridade da mãe nessas transições, constataram que os filhos que possuem mãe com escolaridade básica, considerando 8 anos de estudo, ou seja, Ensino Fundamental completo, têm mais chances de prosseguir e completar a etapa escolar seguinte, o Ensino Médio. No entanto, nos demais níveis de escolarização da mãe, os efeitos foram imprecisos, a que eles atribuem a grande variação na escolarização materna no período considerado.

Em uma pesquisa centrada nos efeitos da educação da mãe sobre a escolaridade do jovem, Anazawa et al. (2016), usando dados longitudinais da corte de 1993 das crianças que

nasceram em Pelotas, no Rio Grande do Sul, Brasil, concluem que, de forma geral, quanto maior a escolaridade das mães, maiores e melhores são seus reflexos no desempenho educacional dos filhos.

Ao tratar das mudanças educacionais em relação à raça negra, Campos (2018) considera que o nível educacional é especialmente importante ao influenciar a condição econômica de uma pessoa, o que resulta numa mudança de vida. No entanto, ressalta que há uma ilusão no Brasil de que resolver as diferenças no campo econômico, de forma que não se comprometa o sistema atual, não haverá problemas raciais; e que é necessário lidar tanto com a discriminação quanto com a desigualdade em diferentes frentes.

A partir das contribuições desses autores, da relação entre escolaridade dos pais e a escolaridade dos filhos, observa-se o fenômeno da reprodução intergeracional das condições de vida. Assume-se, pois, a tamanha importância da qualidade e da quantidade da educação para melhoria constante das condições de vida dos indivíduos e uma consequente redução das desigualdades entre estes.

3.2.2 Estudos sobre desigualdade racial e de classes

No Brasil, são várias as pesquisas acerca da desigualdade racial e de classes. Algumas se destacam pelo fato de utilizarem métodos quantitativos para estabelecer relações entre os elementos considerados importantes nos aspectos socioeconômicos no que diz respeito a esta temática. Nesta seção, serão expostos alguns destes trabalhos que foram escolhas analíticas para a consecução deste trabalho de pesquisa.

Ao sintetizar um modelo de análise de realização econômica, objetivando caracterizar um processo de desigualdade entre pardos e pretos, Hasenbalg e Silva em 2005, segundo Arretche (2015), analisaram os efeitos da situação social familiar dos indivíduos e os recursos disponíveis, concluindo que a realização educacional, o ingresso e a posição no mercado de trabalho, bem como a composição da renda são elementos igualmente importantes em qualquer análise que se possa realizar nesse mesmo âmbito.

Dentre os problemas sociais no Brasil, segundo Santos (2016), as desigualdades são um entrave de longos anos e particularmente afetam as classes sociais e as raciais, impactando negativamente a educação, principalmente ao considerar que aproximadamente 54% da população brasileira é negra, ou seja, formada de pardos e pretos. O desafio na atualidade é o combate ao preconceito, sendo necessária a intervenção da sociedade, da escola e das instituições, reformulando o currículo escolar no eixo social e no seio familiar.

Ao estimar, através de técnicas de regressão logística, em que medida a escolaridade de mulheres negras contribui para desvantagens quanto à participação na classe social alta, Rocha (2017), baseando-se numa amostra do Censo Demográfico de 2010, conclui que as mulheres negras têm realizado progressos limitados nos níveis globais de educação e enfrentam um cenário crítico em relação à conversão destes níveis em oportunidade de riqueza. O autor também ressalta que estudos dessa natureza apontam para condições de vulnerabilidade e desigualdades por gênero e raça que podem ser atribuídas à discriminação.

“*Cor, raça e preconceito* no Brasil compõem a interconexão de concepções para o enfrentamento das questões raciais e de seus desdobramentos nocivos na formação de crianças e adolescentes.” (SANTOS e SILVA, 2018). Assim, os autores anunciam seu trabalho enfatizando que, sendo o Brasil de maioria negra, não se dá a essa característica a devida visibilidade, tendo em vista que o negro é visto como um problema social na condição que ocupa na escala social. Concluem, então, que há uma relação assimétrica entre as raças e que a discriminação racial e o preconceito são suas manifestações mais evidentes.

3.3 METODOLOGIA

Nesta seção, é apresentada uma descrição dos dados e das variáveis, o modelo teórico da regressão logística ordinal em seu aspecto teórico e a modelagem empírica econométrica adotada.

3.3.1 Descrição dos dados e das variáveis

Os dados secundários e amostrais utilizados neste trabalho são provenientes da base de Microdados do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) referente ao Sistema de Avaliação da Educação Básica (SAEB) do ano de 2015. A amostra é composta de 2.063.103 crianças e adolescentes do 5º ano do Ensino Fundamental com idade entre 8 e 15 anos.

O quadro 9 mostra as variáveis (explicativas) ou os fatores, conforme suas características, categoria/codificação e composição em variáveis *dummy*, que compõem este trabalho de pesquisa.

Quadro 9 – Demonstrativo de variáveis, categorias/codificação e *dummy* (continua)

Característica - Variável	Categoria do SAEB	<i>Dummy</i>
Variáveis independentes		
Sexo - SEX	1- Masculino 2 - Feminino	1 – Masculino, 0 - Feminino
Como você se considera? “cor ou raça” - COR	1- Branco(a) 2 - Pardo(a) 3 - Preto(a) 4 - Amarelo(a) 5 - Indígena	1 – Branco, 0 - Demais 1 – Pardo, 0 - Demais 1 – Preto, 0 - Demais 1 – Amarelo, 0 - Demais
Classe econômica - CLASS	1 - Classe A 2 - Classe B 3 - Classe C 4 - Classe D	1 – Classe A, 0 - Demais 1 – Classe B, 0 - Demais 1 – Classe C, 0 - Demais
Região geográfica- REG	1 – Norte 2 – Nordeste 3 – Sudeste 4 – Sul 5 – Centro-Oeste	1 – Norte, 0 - Demais 1 – Nordeste, 0 - Demais 1 – Sudeste, 0 – Demais 1 – Sul, 0 - Demais
Rede de Ensino - RED	1- Pública 2 - Privada	1- Pública, 0 - Privada

Quadro 9 – Demonstrativo de variáveis, categorias/codificação e *dummy* (continuação)

Escolaridade pai – EAP	Até Fund. Incompleto Ens. Fund. Completo Ensino Médio Ensino Superior	1 – Até Fund. Incompleto, 0 - Demais 1 – Fund. Completo, 0 - Demais 1 – Médio, 0 – Demais
Escolaridade mãe - EAM	Até Fund. Incompleto Ens. Fund. Completo Ensino Médio Ensino Superior	1 – Até Fund. Incompleto, 0 - Demais 1 – Fund. Completo, 0 - Demais 1 – Médio, 0 – Demais
Moradia com Pai – MP	Sim; Não	1 – Sim, 0 - Não
Moradia com Mãe – MM	Sim; Não	1 – Sim, 0 - Não
Incentivo a ir à escola - IESC	Sim; Não	1 – Sim, 0 - Não
Incentivo a estudar - IEST	Sim; Não	1 – Sim, 0 - Não
Trabalha fora - TF	Sim; Não	1 – Sim, 0 - Não

Fonte: Elaborado pelo autor, com base nos Microdados SAEB 2015

A variável COR se refere ao item do questionário aplicado aos alunos(as) submetidos(as) à avaliação do SAEB que expressava o seguinte questionamento: “Como você se considera? Branco(a), Pardo(a), Preto(a), Amarelo(a) ou Indígena”.

As variáveis EAP e EAM foram consideradas para esta investigação com o intuito de verificar o efeito da escolaridade dos pais no desempenho escolar (nível de proficiência em leitura).

As variáveis “Moradia com pai” (MP) e “Moradia com mãe” (MM) são relativas ao fato de os alunos(as) residirem com os responsáveis, pai ou mãe, e o objetivo da inclusão destas variáveis é averiguar o efeito da presença paterna e materna no desempenho escolar (proficiência em leitura).

Em relação às variáveis “Incentivo a ir à escola” (IESC) e “Incentivo a estudar” (IEST), estas se relacionam ao fato de os pais incentivarem, ou não, aos alunos(as) irem à escola e/ou a estudar.

O fato de “trabalhar fora de casa” é representado pela variável (TF) e sua inclusão tem o intuito de estimar o efeito de o estudante trabalhar, ou não, no desempenho escolar.

A variável independente classe econômica (social) é uma *proxy* para renda e representa o poder aquisitivo (calculado a partir do total de bens existentes nos domicílios) das famílias das crianças e dos adolescentes que foram submetidos à avaliação do SAEB no ano de 2015. Estes, ainda, foram classificados segundo o critério de classificação econômica da Associação Brasileira de Empresas de Pesquisa (ABEP), baseada na pesquisa de Orçamento Familiar (POF) do Instituto Brasileiro de Geografia e Estatística (IBGE), para o ano de 2015. O critério de corte para os extratos socioeconômicos pode ser verificado no apêndice G.

A variável resposta ou dependente (NIVPROC) foi utilizada para esta investigação e representa os níveis de proficiências em leitura dos estudantes na avaliação do SAEB no ano de 2015. Os níveis são os constantes no quadro 5, relativos à 4ª série/5º ano e que foram segmentados de 9 para 4, considerando adaptações metodológicas de ordem pedagógica em relação às habilidades cognitivas adotadas pelo INEP em 2013.

3.3.2 Modelo empírico econométrico – regressão logística ordinal

O modelo empírico para este trabalho de pesquisa, alicerçado nos aspectos teóricos do modelo de regressão logística ordinal explicitados nas seções 1.3.2 e 2.3.2, apresenta-se com a seguinte expressão:

$$P(Y \leq J) = \frac{1}{1 + e^{-(\alpha_j - \beta X)}} \quad (38)$$

Em que

$$\beta X = \beta_i SEX_{ik} + \beta_i COR_{ik} + \beta_i CLASS_{ik} + \beta_i REG_{ik} + \beta_i RED_{ik} + \beta_i EAP_{ik} + \beta_i EAM_{ik} + \beta_i MP_{ik} + \beta_i MM_{ik} + \beta_i IESC_{ik} + \beta_i IEST_{ik} + \beta_i TF_{ik}$$

α_j é a constante;

β_i – é o vetor de parâmetro de interesse que exprime as relações condicionais dos fatores sobre a ordem dos níveis de proficiência em leitura;

SEX_{ik} – Variável relacionada ao gênero;

COR_{ik} – Variável relacionada à cor da pele ou à raça;

$CLASS_{ik}$ – Variável relacionada à classe econômica;

REG_{ik} – Variável relacionada à região geográfica;

RED_{ik} – Variável relacionada à rede de ensino;

EAP_{ik} – Variável relacionada à escolaridade do pai;

EAM_{ik} – Variável relacionada à escolaridade da mãe;

MP_{ik} – Variável relacionada à moradia com o pai;

MM_{ik} – Variável relacionada à moradia com a mãe;

$IESC_{ik}$ – Variável relacionada a incentivo a ir à escola;

$IEST_{ik}$ – Variável relacionada a estudar;

TF_{ik} – Variável relacionada ao trabalho fora de casa.

Considerando $\alpha_1 < \alpha_2 < \dots < \alpha_j$ pontos de corte ou limites (níveis de proficiência em leitura), define-se como:

$$y = 1 \text{ if } y^* \leq \alpha_1 \quad (39)$$

$$y = 2 \text{ if } \alpha_1 < y^* \leq \alpha_2 \quad (40)$$

$$y = 3 \text{ if } \alpha_2 < y^* \leq \alpha_3 \quad (41)$$

$$y = 4 \text{ if } y^* > \alpha_3 \quad (42)$$

3.4 RESULTADOS E DISCUSSÃO

Nesta seção, são apresentados os resultados estatísticos descritivos dos dados amostrais, bem como as estimativas do modelo econométrico empírico.

3.4.1 Resultados estatísticos descritivos

A tabela 16 apresenta a proporção de alunos(as) na composição da amostra por variáveis (fatores), categorias.

Tabela 16 – Proporção de alunos(as) na amostra por categorias de cada variável (continua)

Variáveis	Categorias	%
Sexo - SEX	Masculino	50,9
	Feminino	49,1
Raça - COR	Branco	31,9
	Pardo	51,5
	Preto	10,9
	Amarelo	2,1
	Indígena	3,0
	Classe Econômica - CLASS	Classe A
	Classe B	6,8
	Classe C	34,3
	Classe D	58,2
Região - REG	Norte	12,0
	Nordeste	26,9
	Sudeste	38,5
	Sul	14,2
	Centro-Oeste	8,3
	Rede de Ensino - RED	Pública
Privada		1,5
Escolaridade do Pai - EAP	Fundamental incompleto	43,1
	Fundamental completo	15,7
	Ensino Médio	19,9
	Superior completo	21,3
Escolaridade da Mãe - EAM	Fundamental incompleto	40,9
	Fundamental completo	14,9
	Ensino Médio	23,9
	Superior completo	20,5
Moradia com Pai - MP	Sim	37,1
	Não	62,9

Tabela 16 – Proporção de alunos(as) na amostra por categorias de cada variável (continuação)

Variáveis	Categorias	%
Moradia com Mãe - MM	Sim	89,4
	Não	10,6
Incentivo a ir à Escola - IESC	Sim	95,3
	Não	4,7
Incentivo a Estudar - IEST	Sim	96,9
	Não	3,1
Trabalha Fora de Casa - TF	Sim	12,9
	Não	87,1

Fonte: Microdados SAEB/Prova Brasil (Elaborado pelo autor)

Em relação às proporções apresentadas na tabela 16, temos que a proporção de estudantes do sexo masculino (50,9%) é apenas 1% maior que de sexo feminino (49,1%).

A raça negra, composta de pardos(as) e pretos(as), representa 62,4% da população amostrada, portanto a maioria dos estudantes do 5º ano do Ensino Fundamental no ano de 2015 afirmaram serem de cor de pele parda ou preta, fenômeno este constatado por Santos (2016). O autor afirma que as fronteiras entre as autoafirmações entre categorias pardos e pretos são vagas e porosas, mas, nos últimos 20 anos, um número crescente de indivíduos optara por se identificar como negros, em resposta aos esforços de conscientização por parte dos movimentos negros. A raça branca e representada por 31,9% dos estudantes e as raças amarela e indígena (5,10%)

Ocorre uma grande concentração de estudantes nas classes econômicas baixas, posto que classe econômica D apresenta uma proporção de 58,2% e a classe C, 34,3%, totalizando 92,5%.

A rede pública de ensino atende a 98,5% dos estudantes do 5º ano do Ensino Fundamental e a rede privada apenas 1,5%. Já em relação as regiões geográficas, as Regiões Nordeste (26,9%) e Sudeste (38,2%) são as mais representadas na amostra, totalizando 65,1% dos alunos(as) avaliados pelo SAEB em 2015.

Em relação à escolaridade dos pais, a proporção de estudantes é bem semelhante tanto para a figura do pai quanto para da mãe. As proporções ficam em torno de 40% para o fundamental incompleto, 15% para o fundamental completo, 21% para o médio e superior.

Apesar de a maioria dos estudantes expressarem a ausência paterna na residência domiciliar (62,9%), a presença materna apresenta uma proporção expressiva (89,4%).

É bastante significativa a proporção dos pais que incentivam aos filhos(as) a irem à escola (95,3%) e também a estudarem (96,9%), como também a proporção de alunos que não trabalham fora de casa (87,1%).

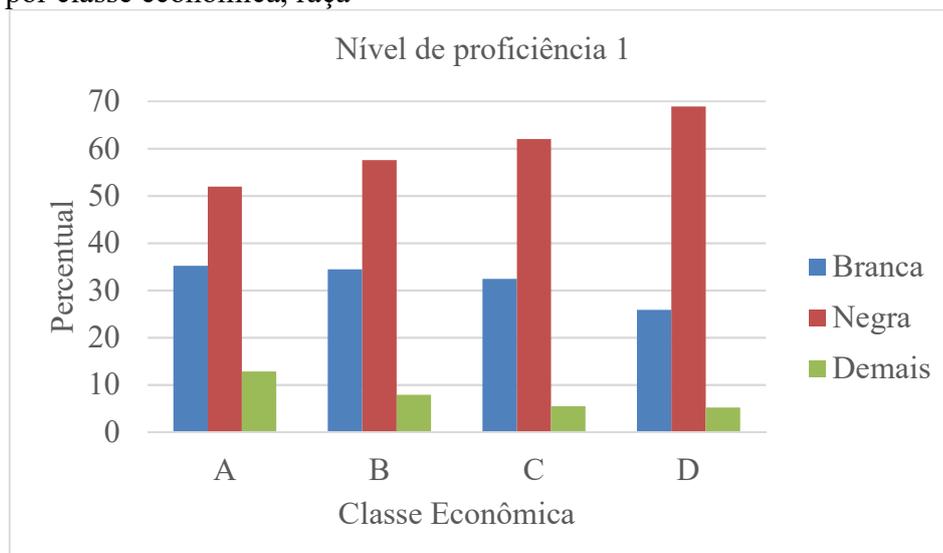
A tabela 17 e os gráficos de 20 a 23 mostram a proporção de estudantes por classe econômica, nível de proficiência e raça.

Tabela 17 – Proporção de estudantes por classe econômica por nível de proficiência e raça

Nível de Proficiência	Raça	Classe Econômica %			
		A	B	C	D
1	Branca	35,23	34,48	32,46	25,87
	Negra	51,93	57,57	62,01	68,87
	Demais	12,85	7,95	5,53	5,27
2	Branca	40,44	40,38	35,41	26,09
	Negra	53,12	53,70	59,16	68,38
	Demais	6,44	5,92	5,43	5,53
3	Branca	47,32	46,23	40,15	29,19
	Negra	47,15	47,95	54,16	64,86
	Demais	5,54	5,82	5,69	5,95
4	Branca	55,85	53,18	46,12	33,96
	Negra	37,99	41,15	48,37	60,26
	Demais	6,16	5,67	5,51	5,79

Fonte: Microdados SAEB/Prova Brasil (Elaborado pelo autor)

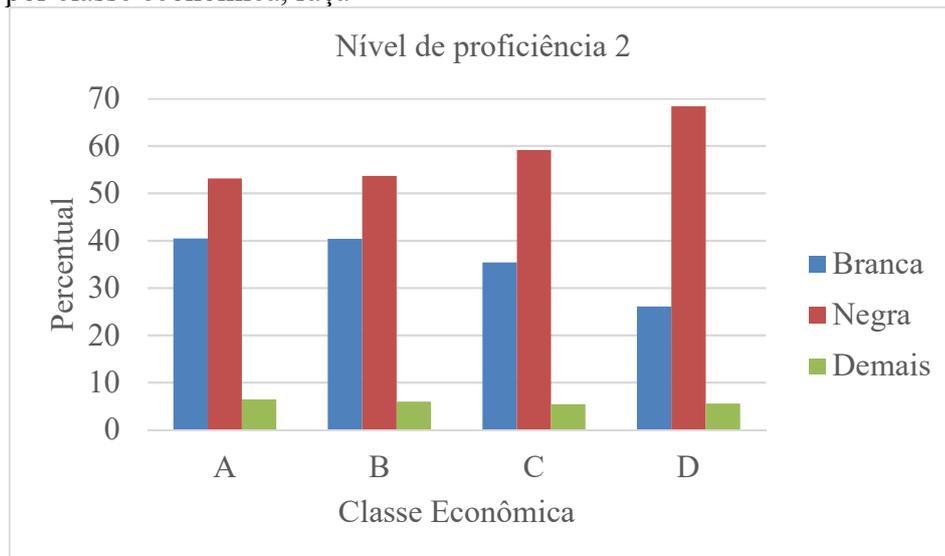
Gráfico 20 – Proporção de estudantes no nível de proficiência em leitura 1 por classe econômica, raça



Fonte: Microdados SAEB/Prova Brasil – 2015 (Elaborado pelo autor)

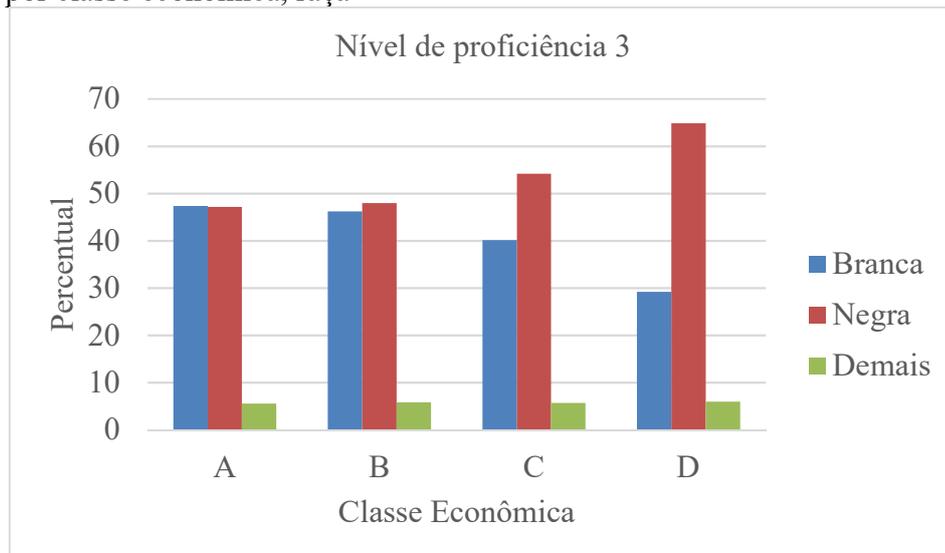
O gráfico 20 demonstra que a proporção de estudantes da raça negra predomina no menor nível de proficiência em leitura em todas as classes econômicas, observa-se que quanto menor a classe econômica maior a proporção de negros e menor a de brancos e as demais raças. Semelhante fenômeno é observado no nível 2 de proficiência, conforme o gráfico 21.

Gráfico 21 – Proporção de estudantes no nível de proficiência em leitura 2 por classe econômica, raça



Fonte: Microdados SAEB/Prova Brasil – 2015 (Elaborado pelo autor)

Gráfico 22 – Proporção de estudantes no nível de proficiência em leitura 3 por classe econômica, raça

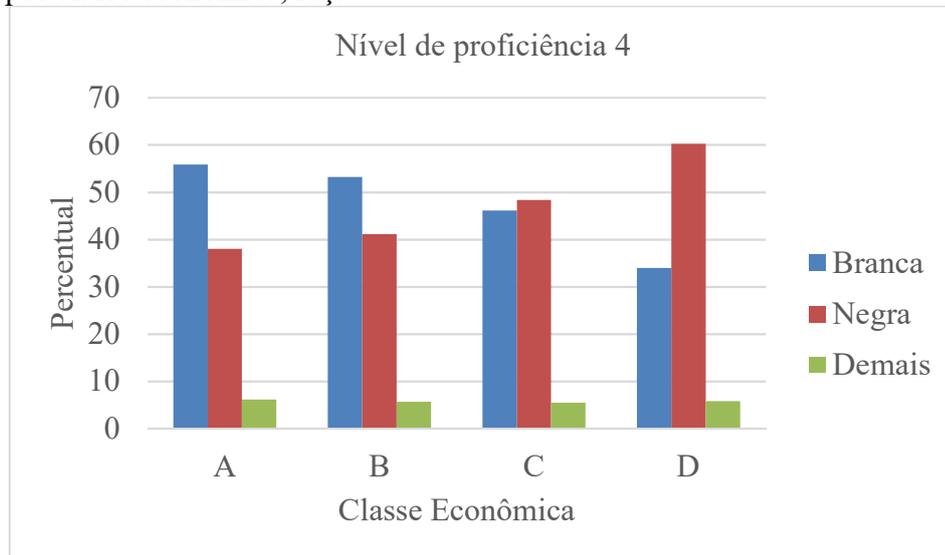


Fonte: Microdados SAEB/Prova Brasil – 2015 (Elaborado pelo autor)

No nível 3 de proficiência, conforme demonstrado no gráfico 22, ocorre uma predominância da raça negra apenas nas classes econômica C e D. Um certo equilíbrio é

observado na proporção de estudantes brancos e negros nas classes A e B. Ainda assim quanto menor a classe econômica maior a proporção de estudantes negros e menor a de brancos.

Gráfico 23 – Proporção de estudantes no nível de proficiência em leitura 4 por classe econômica, raça



Fonte: Microdados SAEB/Prova Brasil – 2015 (Elaborado pelo autor)

O gráfico 23 evidencia, no nível 4 de proficiência, que a proporção de estudantes negros aumenta conforme decresce a classe econômica, de forma inversa ocorre com a proporção de estudantes brancos.

Conclui-se, partir da leitura da tabela 16 e dos gráficos 20 a 23, os seguintes pontos:

- No maior nível de proficiência, ocorre um decréscimo na proporção de estudantes brancos e um crescimento na proporção de estudantes negros em relação às classes econômicas. Este fenômeno evidencia que os estudantes brancos, cujas famílias possuem maior poder aquisitivo, são mais presentes no maior nível de proficiência. Um outro elemento evidenciado é em relação aos estudantes da raça negra que compõem as classes econômicas C e D, os quais se apresentam com proporções maiores no nível 4 de proficiência;
- Os estudantes da raça negra de todas as classes sociais apresentam proporções significativas nos níveis 1 e 2 de proficiência;
- Estudantes brancos e negros, que compõem as classes econômicas A e B, apresentam proporções semelhantes no nível 3. Contudo, em relação às classes C e D, há uma predominância de estudantes negros em relação aos estudantes brancos de mesma classe social;

- De forma geral, evidencia-se uma certa resistência e persistência dos estudantes da raça negra nas menores classe sociais e em busca por melhores resultados.

3.4.2 Resultados do modelo econométrico

A tabela 18 apresenta os coeficientes estimados dos fatores associados ao desempenho escolar (nível de proficiência em leitura) a partir do modelo de regressão logística ordinal.

Tabela 18 – Estimativa dos efeitos esperados pelos fatores do modelo econométrico empírico.

Variáveis	Estimativas	Erro Padrão	z	p-value	OR
SEX (Masculino)	- 0,355	0,005	- 71,98	0,000	1,43
COR (Branca)	- 0,047	0,014	- 3,26	0,001	1,05
COR (Parda)	- 0,100	0,014	- 7,13	0,000	1,11
COR (Preta)	- 0,566	0,016	- 36,03	0,000	1,76
COR (Amarela)	- 0,137	0,020	- 6,74	0,000	1,15
CLASS A	- 0,251	0,026	- 9,60	0,000	1,29
CLASS B	0,098	0,009	10,52	0,000	0,91
CLASS C	0,199	0,005	36,16	0,000	0,82
REG (Norte)	- 0,562	0,011	- 51,17	0,000	1,75
REG (Nordeste)	- 0,578	0,009	- 59,19	0,000	1,78
REG (Sudeste)	0,276	0,009	30,51	0,000	0,76
REG (Sul)	0,211	0,103	20,55	0,000	0,81
REDE (Pública)	- 0,981	0,017	- 56,26	0,000	2,67
EAP (Fund. Incompleto)	- 0,160	0,008	- 20,57	0,000	1,17
EAP (Fund. Completo)	- 0,071	0,009	- 8,12	0,000	1,07
EAP (Médio)	0,261	0,008	33,11	0,000	0,77
EAM (Fund. Incompleto)	- 0,498	0,008	- 62,67	0,000	1,65
EAM (Fund. Completo)	- 0,234	0,009	- 25,94	0,000	1,26
EAM (Médio)	0,214	0,008	25,42	0,000	0,81
MP (Não)	- 0,136	0,005	- 25,14	0,000	1,15
MM (Não)	- 0,216	0,009	- 24,68	0,000	1,24
IESC (Não)	- 1,026	0,014	- 71,37	0,000	2,79
IEST (Não)	- 0,806	0,018	- 42,83	0,000	2,24
TF (Não)	1,090	0,008	140,48	0,000	0,34

Fonte: SAEB/2015 - Elaborado pelo autor

Conforme os resultados dos coeficientes estimados por meio do modelo proposto, temos as seguintes assertivas, conforme tabela 17:

- Ser do sexo masculino apresentou efeito negativo para níveis de proficiência com chances maiores (1,43) do aluno ser observado nos menores níveis de proficiência. Este resultado parece confirmar uma tendência diagnosticada no relatório da UNESCO (2017) que ressalta um avanço maior do sexo feminino em proficiência em leitura que o sexo masculino no período compreendido de 2007 a 2013;
- Todas as raças apresentaram efeito negativo e razão de chances maiores para nível de proficiência baixo, das quais a raça negra tem maiores chances (1,76) de se encontrar nos menores níveis de proficiência que as demais. O relatório da UNESCO (2017) evidencia que embora tenha havido redução de alunos de diferentes raças nos níveis menores de proficiência em leitura no período de 2007 a 2013, os alunos autodeclarados pretos permanecem com um percentual elevado nos níveis de proficiência menores;
- A classe econômica A tem efeito negativo em relação a desempenho escolar com chances maiores (1,29) de se encontrar nos menores níveis de proficiência. Já as classes econômicas B e C apresentam efeito positivo com razão de chances menores 0,91 e 0,82, respectivamente, de se encontrarem nos níveis baixos de proficiência;
- Podemos perceber que estudantes cujos pais têm baixo nível de escolaridade apresentam chances maiores de serem observados nos níveis baixos de proficiência; para a figura paterna que tem apenas Ensino Fundamental incompleto, a chance é de 1,17 e com fundamental completo é de 1,07. Para a figura materna com fundamental incompleto, a razão de chances é de 1,65 e com fundamental completo, 1,26;
- A ausência do pai no domicílio aumenta em 1,15 a chance do estudante alcançar nível de baixo de proficiência e sendo a mãe a ausente a chance aumenta para 1,24 para níveis menores de proficiência;
- Não incentivar o estudante a ir à escola aumenta em 2,79 sua chance de apresentar baixo nível de proficiência. O mesmo ocorre com estudantes não incentivados a estudar, posto que apresentam uma razão de chance de 2,24;
- A proporção de alunos do 5º ano do Ensino Fundamental nas edições de 2007 a 2013 segundo percentis do fator “envolvimento dos pais” no aprendizado em leitura apresentou redução de 36,4% para 31,6% nos níveis menores de proficiência, conforme o relatório da

UNESCO (2017), resultado este que confirma a presença dos pais e o incentivo destes para os estudos como fator positivo em relação ao desempenho escolar;

- Não trabalhar fora de casa é um fator que apresenta efeito positivo em relação a desempenho escolar, pois diminui em 0,34 a chance para níveis de proficiência menores;
- O fato de a criança ou o adolescente estudar na escola pública diminui a chance de alcançar níveis de proficiência elevados, já que estudar na rede pública de ensino apresenta efeito negativo para um melhor desempenho escolar com chance maior em 2,67 de ser observado com nível de proficiência baixo;
- As Regiões Norte e Nordeste apresentam efeito negativo e razão de chance de 1,75 e 1,78, respectivamente, de se encontrar com baixo nível de proficiência em leitura. O mesmo não acontece com as Regiões Sudeste e Sul, pois apresentam efeito positivo com chances menores para nível baixo de proficiência.

Tabela 19 – Efeito marginal dos fatores associados por nível de proficiência. (continua)

Variáveis	Probabilidade			
	(1)	(2)	(3)	(4)
	0,207	0,421	0,295	0,077
SEX (Masculino)	0,058	0,025	- 0,057	- 0,025
COR (Branca)	0,008	0,003	- 0,008	- 0,003
COR (Parda)	0,016	0,007	- 0,016	- 0,007
COR (Preta)	0,105	0,018	- 0,089	- 0,033
COR (Amarela)	0,023	0,008	- 0,022	- 0,009
CLASS A	0,044	0,012	- 0,040	- 0,016
CLASS B	- 0,016	- 0,007	0,015	0,007
CLASS C	- 0,032	- 0,014	0,032	0,014
REG (Norte)	0,104	0,019	- 0,089	- 0,033
REG (Nordeste)	0,103	0,025	- 0,092	- 0,036
REG (Sudeste)	- 0,045	- 0,020	0,045	0,020
REG (Sul)	- 0,033	- 0,016	0,034	0,016
REDE (Pública)	0,119	0,120	- 0,137	- 0,103
EAP (Fund. Incompleto)	0,027	0,011	- 0,026	- 0,011
EAP (Fund. Completo)	0,012	0,005	- 0,011	- 0,005

Tabela 19 – Efeito marginal dos fatores associados por nível de proficiência.
(continuação)

Variáveis	Probabilidade			
	(1)	(2)	(3)	(4)
	0,207	0,421	0,295	0,077
EAP (Médio)	- 0,041	- 0,021	0,042	0,019
EAM (Fund. Incompleto)	0,085	0,029	- 0,080	- 0,034
EAM (Fund. Completo)	0,040	0,013	- 0,038	- 0,016
EAM (Médio)	- 0,031	-0,015	0,031	0,014
MP (Não)	0,023	0,009	- 0,022	- 0,009
MM (Não)	0,038	0,011	- 0,035	- 0,014
IESC (Não)	0,211	- 0,011	- 0,151	- 0,049
IEST (Não)	0,161	0,003	- 0,122	- 0,041
TF (Não)	- 0,218	- 0,001*	0,163	0,056

Fonte: SAEB/2015 - Elaborado pelo Autor

Nota: Não estatisticamente significativa (*), todas as demais significantes a 1%.

Na tabela 19, são apresentados os efeitos marginais dos fatores em cada nível de proficiência, verifica-se as probabilidades das variáveis se encontrarem no ponto médio de cada nível de proficiência em leitura, as seguintes probabilidades são dadas em ordem crescente dos níveis: 20,7%, 42,1%, 29,5% e 7,7%. O nível 2 apresenta a maior probabilidade, seguido do nível 3.

Diante dos resultados apresentados na tabela 19 para os fatores associados, temos as asserções:

- Redução nas probabilidades para níveis superiores (especificamente para os níveis 3 e 4) de proficiência para as variáveis (fatores): sexo masculino, raça ou cor, classe econômica A, Regiões (norte e nordeste), rede de ensino pública, escolaridade dos pais (pai e mãe) até o Ensino Fundamental, ausência de moradia com os pais (pai e mãe) e ausência de incentivos aos (motivação) estudantes a irem à escola e a estudarem pelo pais;
- Aumento nas probabilidades para níveis superiores de proficiência em leitura para os fatores: classe econômica B e C, Regiões (Sudeste e Sul), escolaridade para os pais que têm Ensino Médio e para estudantes que não trabalham fora de casa.
- A redução na probabilidade para os níveis superiores de proficiência em leitura dos (as) alunos (as) da classe econômica A pode ser explicado pela pequena proporção amostral de

apenas 0,7% relativo a esta classe na composição da amostra ou a outros fenômenos com efeito na educação infantil relacionados a crise econômica de 2015.

3.5 CONSIDERAÇÕES FINAIS

A partir dos resultados expostos, afirma-se que, embora as desigualdades em vários aspectos socioeconômicos venham apresentando diminuição no Brasil nos últimos anos, ainda existem grandes obstáculos a serem transpassados.

Em relação ao fator classe econômica dos estudantes, o que se percebe é que as classes B e C apresentaram resultados positivos em relação ao desempenho escolar (nível de proficiência em leitura), sendo mais provável que sejam observadas nos níveis superiores de proficiência. Fenômeno inverso ocorre em relação aos estudantes da classe econômica A. Cabe ressaltar que 2015 não foi um ano fácil para economia brasileira, pois conforme o Relatório da Pesquisa Nacional por Amostra de Domicílios (PNAD) 2015 verificou-se uma diminuição do rendimento médio do trabalho em todas as Grandes Regiões e conseqüentemente uma possível migração de alunos(as) da rede privada para a rede pública.

O fator raça (cor da pele) não apresentou diferenças significativas acerca do desempenho escolar entre as raças, considerando que todas elas apresentaram efeito negativo em relação a nível de proficiência, ou seja, probabilidades maiores de serem observadas nos menores níveis de proficiência.

Os resultados dos fatores classe econômica (*proxy* para renda) e raça (cor da pele) parecem ser um contraponto aos resultados do estudo de Santos (2016), que enfatiza a desigualdade social (econômica) e a racial como fonte de impacto negativo em relação à educação.

O modelo logístico ordinal como modelo empírico para este trabalho infere que estudantes cujos pais têm baixa escolaridade apresentam probabilidades maiores de serem observados nos níveis de proficiências inferiores. O mesmo não ocorre a estudantes com pais mais escolarizados, visto que apresentam chances menores de serem observados nos níveis inferiores. Esse fenômeno é concordante com os trabalhos de Ribeiro et al (2015) e de Anazawa et al. (2016), que chegaram à mesma conclusão em relação à escolarização dos pais no desempenho educacional dos filhos

O fato dos(as) alunos(as) residirem no mesmo domicílio dos pais, serem incentivados pelos pais a irem à escola e a estudarem, bem como a não trabalharem fora de casa possibilitam a ascensão para níveis superiores de proficiência.

No contexto apresentado neste trabalho de pesquisa segundo o objetivo a que se propôs, as expectativas foram atendidas, de modo que foi lançada luz sobre aspectos ainda obscuros, relacionando desigualdade de renda, educação e raça no âmbito da educação básica

no Brasil, principalmente ao fato de não confirmar diferenças de classes (econômica ou social) e de raça como elementos ou fatores inibidores de melhor desempenho escolar entre crianças e adolescentes no primeiro ciclo do ensino fundamental.

CONCLUSÃO GERAL

A universalização do Ensino Fundamental no Brasil é uma realidade já comprovada pelos indicadores nacionais de educação, como a taxa líquida de matrícula, que atingiu em 2018 o patamar de 98%. Segundo dados do Anuário da Educação 2019, 128,4 mil escolas atendem a 15,2 milhões de estudantes no Ensino Fundamental nos anos iniciais e 12 milhões nos anos finais. Só a rede pública atende a 27,2 milhões no ensino básico fundamental. Ainda assim, o sistema escolar brasileiro é reprodutor de diversas desigualdades sociais e, segundo o referido anuário, é necessário evidenciá-las para possibilitar o planejamento de políticas públicas mais equitativas, eficientes e com maior cobertura. Sendo importante recortes por região, renda, raça/cor e outros fatores que permitam atingir este objetivo.

A meta 3 do movimento Todos Pela Educação 2019 prevê que, até 2022, pelo menos 70% dos estudantes estejam aprendendo o que é adequado para o seu ano. Isso significa que em Língua Portuguesa os alunos do 5º ano do Ensino Fundamental superem os 59,4% alcançados em 2017 e os alunos do 9º ano os 39,5%. Cabe enfatizar que, no Programa Internacional de Avaliação de Estudantes (PISA), menos da metade dos estudantes brasileiros alcançam o nível adequado de proficiência, o nível 2, que é considerado o mínimo adequado em Leitura, segundo a Organização para a Cooperação e do Desenvolvimento Econômico (OCDE). Estudantes até este nível de proficiência apresentam deficiências cognitivas que comprometem as habilidades necessárias para o ingresso no mercado de trabalho e suas expectativas sociais.

É nesse contexto da realidade da educação básica no Brasil que se apresentam esses três ensaios sobre a avaliação do Ensino Fundamental no Brasil.

No primeiro ensaio, avaliou-se a relação e os efeitos dos fatores extraescolares: gênero, raça, classe econômica, região geográfica e rede administrativa (pública e privada); no desempenho de estudantes do Ensino Fundamental segmentados nos ciclos iniciais (4ª série/5º ano) e finais (8ª série/9º ano), compreendendo o período de 1999 a 2017. Considerando uma modelagem empírica baseada num modelo de regressão logística ordinal, os resultados evidenciam ao longo de todo o período uma aglomeração nos níveis mais baixos de proficiência bastante persistente em termos probabilísticos, mas também aponta posição positiva em relação ao fator classe econômica (renda), que avançou timidamente em relação aos maiores níveis de proficiência.

No segundo ensaio, investigou-se também a relação e os efeitos dos fatores extraescolares: gênero, raça, classe econômica e região geográfica no desempenho escolar,

numa perspectiva de gerações sobrepostas e ainda utilizando uma modelagem empírica regressiva logística ordinal. Os resultados apontaram possibilidades maiores de observar os níveis mais baixos de proficiência ao longo das gerações. No entanto, o fator classe econômica foi o que melhor realçou tendência para níveis mais altos de proficiência.

No terceiro ensaio, conduziu-se uma análise sobre as condições socioeconômicas de crianças e adolescentes do primeiro ciclo (5º ano) do Ensino Fundamental que participaram da avaliação do SAEB, edição 2015. O propósito foi investigar elementos que apontassem possíveis disparidades entre estudantes que estão distribuídas em diferentes classes socioeconômicas. Os resultados apresentaram aspectos conciliadores em relação aos fatores raça, classe econômica e desempenho escolar entre os estudantes, bem como elementos relacionados a moradia com os pais, escolaridade dos pais e motivações familiares para estudar e irem à escola.

As conclusões a que se chega com estes ensaios e com a metodologia proposta conduzem as seguintes evidências:

- A persistência de tendência para os níveis inferiores de proficiência em leitura dos fatores investigados: gênero masculino, raças (negra, parda e amarela), regiões (Norte e Nordeste), rede de ensino pública; com exceção do fator classe econômica, ainda que a convergência das classes econômicas mais abastadas seja mais rápida para os níveis maiores de proficiência;
- Evidencia-se ainda uma possível ineficiência das políticas públicas educacionais de caráter nacional em alavancar a qualidade do Ensino Fundamental, de modo a impactar positivamente os indicadores educacionais em leitura, principalmente nos anos finais do Ensino Fundamental;
- Parece não ocorrer ao longo das gerações uma melhoria nos níveis de proficiência em leitura. Há uma persistente desigualdade educacional, racial, de classe e regional no Brasil.

Em proposição às conclusões apresentadas, sugere-se a utilização do método aplicado nestes ensaios para outros tipos de avaliação e o desenvolvimento de novas pesquisas de mesma natureza no tema em estudo.

REFERÊNCIAS

- AGRESTI, Alan. **Analysis of ordinal categorical data**. 2. ed. New Jersey: John Wiley e Son, 2010.
- ALMEIDA, Rita K.; PACKARD, Truman G. **Competências e Empregos: uma agenda para a juventude. Relatório Síntese de constatações, conclusões e recomendações políticas**. Brasília, 2018. Disponível em: <<http://documents.worldbank.org/curated/en/953891520403854615/Síntese-de-constatações-conclusões-e-recomendações-de-políticas>> Acesso em: 27 jun. 2019
- ALVES, M. T. G.; SOARES, J. F. **O efeito das escolas no aprendizado dos alunos: um estudo com dados longitudinais no ensino fundamental**. **Educação e Pesquisa.**, São Paulo, v. 34, n. 3, p. 527-544, set./dez. 2008.
- ANAZAWA, Leandro et al. **A loteria da vida: Examinando a Relação entre a Educação da Mãe e a Escolaridade do Jovem com Dados Longitudinais do Brasil**. **Centro de Políticas Públicas.**, São Paulo, n. 22, nov. 2016.
- ARRETCHE, Marta (Org.). **Trajetórias das desigualdades: como o Brasil mudou nos últimos cinquenta anos**. 1.ed. São Paulo: Editora Unesp, 2015.
- ARTIGE, Lionel. Endogenous growth and parental funding of education in na OLG model with a fixed fator. **CREPP**, Working Paper 1003. HEC-Management School, University of Liège. 2010. Disponível em: <https://www.researchgate.net/publication/46452303_Endogenous_Growth_and_Parental_Funding_of_Education_in_an_OLG_Model_with_a_Fixed_Factor/references>. Acesso em: 04 jun. 2019.
- BARBETTA, P. A.; ANDRADE, D. F.; TAVARES, H. R. Estudos de fatores associados através de Regressão Logística Quantílica Hierárquica. **Estudos de Avaliação Educacional**, São Paulo, v. 29, n. 71, p. 320-349, maio/agosto. 2018.
- BARRO, Robert J.; MARTIN, Xavier Sala-i. **Economic growth**. 2. ed. Massachusets: MIT, 2003.
- BLANCHARD, Olivier. **Macroeconomia**. Tradução: Luciana do Amaral Teixeira. 5. ed. Pearson Prentice Hall, 2011.
- BRASIL. Programa das Nações Unidas para o Desenvolvimento/Instituto de Pesquisa Econômica Aplicada/ Fundação João Pinheiro. **Desenvolvimento humano para além das médias: 2017**. Brasília, PNUD: IPEIA: FJP, 2017.127 p.
- CAMPOS, Fernando Barbosa. Resistência e luta: a influência dos movimentos negros na criação de leis e políticas públicas educacionais no Brasil. **Anais do Congresso Brasileiro de Processo Coletivo e Cidadania**, n. 6, p. 732-754, out./2018.
- CARVALHO, Francisca Eliane Dias de. **Fatores socioeconômicos associados ao desempenho dos estudantes na prova de redação do exame nacional do ensino médio**. 2017. Tese (doutorado em Educação) - Universidade Federal do Ceará, Programa de Pós-Graduação em Educação, Fortaleza, 2017.

CASTELO-CLIMENT, Amparo; DOMÉNECH, Rafael. Human capital inequality, life expectancy and economic growth. **The Economic Journal**, Vol.118, abr. 2008. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.132.9735&rep=rep1&type=pdf>>. Acesso em: 3 jul. 2019.

CHANDRA, Ritu; AZIMUDDIN, Shaikh. Influence of socio economic status on academic achievement of secondary school students Lucknow city. **International Journal of scientific & Engineering research**, v. 4, Issue 11, nov. 2013.

COLL, César; MARCHESI, Álvaro; PALÁCIOS, Jesús. **Desenvolvimento psicológico e educação**. Porto Alegre: Artmed, 2004.

DEDECCA, C. S.; TROVÃO, C. J. B. M.; SOUZA, L. F. Desenvolvimento e equidade. **Novos estud. - CEBRAP**, São Paulo, v. 33, n. 1, p. 23-41, 2014.

DIAS, Fernanda. O desenvolvimento cognitivo no processo de aquisição de linguagem. **Letrônica**, Porto Alegre, n. 2, p. 107-119, dez. 2010.

FÁVERO, Luiz Paulo. et al. **Análise de dado: modelagem multivariada para a tomada de decisão**. Rio de Janeiro: Elsevier, 2009.

FERNANDES, Luana de Mendonça et al. **Preditores do desempenho escolar ao final do ensino fundamental: histórico de reprovação, habilidades sociais e apoio social**. Temas em Psicologia. Volume 26, nº 1, 215-228. Março, 2018.

FILHO, M. N.; KIRSCHBAUM, C. Educação e desigualdade no Brasil. **In: Trajetória das desigualdades: como o Brasil mudou nos últimos cinquenta anos**.1. ed. São Paulo: Editora Unesp, 2015. 489p.

FRANCO, A. M.; MENEZES FILHO, N. Os determinantes do aprendizado com dados em painel de escolas do SAEB. **Revista Economia Aplicada**, v. 21, n. 3, p. 525-548, 16 abr. 2018.

FULLERTON, Andrew S.; XU, Jun. **Ordered Regression Models: parallel, partial, and non-parallel alternatives**. Chapman & Hall/CRC Statistic in the Social and Behavioral Sciences Series, 2016.

GARCIA, Maria de Fátima; KATZ, Frederico Jayme. **Emprego e trabalho: uma visão multidisciplinar**. PME-UEM. Maringá, 2004.

GONÇALVES, Debora Souza Neves. **A importância da leitura nos anos iniciais escolares**. São Gonçalo, 2013. Disponível em:< <http://www.ffp.uerj.br/arquivos/dedu/monografias/dsng.pdf>>. Acesso em: 26 set. 2019.

GLOMM, Gerhard; RAVIKUMAN, B. Public versus private Investment in human capital: endogenous growth and Income inequality. **The Journal of Political Economy**. Vol. 100, Nº 4 (Aug. 1992), pp. 818-834, The University of Chicago Press. Disponível em :<<<https://www.isid.ac.in/~tridip/Teaching/DevEco/Readings/11Education/Glomm&Ravikumar-JPE1992.pdf>>>. Acesso em: 04 jun. 2019.

GREENE, William H. **Econometric analysis**. 7. ed. Pearson, 2012.

GUBA, Egon S.; LINCOLN, Yvonna. **Avaliação de quarta geração**. Tradução Beth Honorato. Campinas, SP: Editora da Unicamp, 2011.

GUJARATI, Damodar N. **Econometria básica**. Tradução: Maria José Cyhlar Monteiro. Rio de Janeiro: Elsevier, 2006.

GYLFASON, Thorvaldur; ZOEGA, Gylfi. **Education, social equality and economic growth: a view of the landscape**. CESinfo, Working paper 876, 2003. Disponível em: <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=385780>. Acesso em: 04 jun. 2019.

HECKMAN, James J. *The economic, technology, and neuroscience of human capability formation*. PNAS. V. 104. n. 33. Los Angeles, 2007. Disponível em: <http://jenni.uchicago.edu/papers/Heckman_2007_PNAS_v104_n33.pdf>. Acesso em: 04 jun. 2019.

HELENE, Otaviano. **Um diagnóstico da educação brasileira e de seu financiamento**. Campinas, São Paulo: Autores associados, 2013.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Síntese de indicadores sociais: uma análise das condições de vida da população brasileira: 2016 / IBGE**, Coordenação de População e Indicadores Sociais. - Rio de Janeiro: IBGE, 2016, n. 36, 146 p.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Pesquisa nacional por amostra de domicílios: síntese de indicadores 2015**. Rio de Janeiro: IBGE, 2015, 108 p.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Pesquisa nacional por amostra de domicílios: síntese de indicadores 2011**. Rio de Janeiro: IBGE, 2011.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. Portaria nº 1100, de 28 de dezembro de 2018. Brasília, 2018. Disponível em: <http://download.inep.gov.br/educacao_basica/saeb/2018/legislacao/portaria_n1100_26122018_saeb2019.pdf>. Acesso em: 02 jul. 2019.

KLEINKE, Maurício Urban. **O Enem e o impacto de variáveis socioeconômicas e educacionais**. In: Inep/MEC — Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **BOLETIM NA MEDIDA**. Ano 5. n. 11. Brasília-DF, fevereiro, 2017.

KUMARI, Santoshi; Shivgotra, V. K.; KUMAR, Pawan. Application of ordinal logistic Regression analysis in determining the risk factors for intelligence quotient of school going children. **International Education e Research Journal**. E-ISSN N° 2454-9916. V. 3, Issue 5. 2017.

LA FUENTE, Angel. **Métodos e modelos matemáticos para economistas**. 1. ed. Lisboa, Portugal: Instituto Piaget, 2007.

LAROS, A. J.; MARCIANO, J. L.; ANDRADE, M. J. **Fatores que afetam o desempenho na prova de Matemática do SAEB: um estudo multinível**. *Avaliação Psicológica*, Porto Alegre, v.9. n. 2, p. 113-189, ago. 2010.

LEÃO, N.; CANDIDO, M. R.; CAMPOS, L. A.; JÚNIOR, J. F. Relatório das desigualdades de raça, gênero e classe. GEMMA. Rio de Janeiro, n. 1, p. 1-21, 2017.

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

MARIANO, Zilania Francisca. **Ensaio sobre o Ensino profissionalizante no Ceará.** 2016. Tese (doutorado em Economia) – Faculdade de Economia, Administração, Atuária e Contabilidade. Universidade Federal do Ceará, 2016.

MENARD, Scott W. **Applied logistic regression analysis.** 2. ed. Thousand Oaks: Sage University Papers Series on Quantitative Applications in the Social Sciences, 2001.

MINISTÉRIO DA EDUCAÇÃO. Portaria nº 931, de 21 de março de 2005. Brasília, 2005. Disponível em: < http://download.inep.gov.br/educacao_basica/prova_brasil_saeb/legislacao/Portaria931_NovoSaeb.pdf>. Acesso em: 02 jul. 2019.

O'CONNELL, A. A. **Logistic Regression Models for Ordinal Response Variables.** Quantitative Applications in the Social Sciences. Thousand Oaks, CA: Sage Publications, 2006.

OECD (2016), **PISA 2015 Results (Volume I): Excellence and Equity in Education,** PISA, OECD Publishing, Paris.

OSÓRIO, R. G. O sistema classificatório de “cor ou raça” do IBGE. **Texto para Discussão n. 996.** Brasília: IPEA, nov., 2003.

PESSÔA, Samuel de Abreu. **Um modelo de acumulação de capital físico e humano: um diálogo com a economia do trabalho.** ANPEC 2, 2001. Disponível em: < <http://www.anpec.org.br/encontro2001/artigos/200102173.pdf>>. Acesso em: 02 jul. 2019.

RIBEIRO, C. C.; CENEVIVA, R.; BRITO, M. M. A. Estratificação educacional entre jovens no Brasil: 1960 a 2010. **In: Trajetória das desigualdades: como o Brasil mudou nos últimos cinquenta anos.** 1. ed. São Paulo: Editora Unesp, 2015.

ROCHA, Emerson Ferreira. Riqueza e status entre mulheres negras no Brasil. **Revista Sociedade e Estado,** Volume 32, n. 1, jan. 2017.

ROCHA, Pedro Fernando Damasceno. **Avaliação de fatores socioeconômico e comportamentais nos resultados dos alunos na Prova Brasil de 2011 nos municípios cearenses.** 2014. Dissertação (mestrado profissional) - Faculdade de Economia, Administração, Atuária e Contabilidade. Universidade Federal do Ceará, 2014.

SABOIA, João. O salário mínimo e seu potencial para a melhoria de distribuição de renda no Brasil. **In: Desigualdade de renda no Brasil: uma análise da queda recente.** Brasília: IPEA, 2007, v. 2. 553p.

SANDRONI, Paulo. (Org.). **Novíssimo dicionário de economia**. 1. ed. São Paulo: Best Seller, 1999. Disponível em: < <http://sinus.org.br/2014/wp-content/uploads/2013/11/FMI.BMNov%C3%ADssimo-Dicion%C3%A1rio-de-Economia.pdf> > Acesso em: 28 jun. 2019.

SANTOS, Bruno Freitas. **Discriminação racial: uma breve discussão**. **Revista Opara – Ciências Contemporâneas Aplicadas**. Petrolina: FACAPE, v. 6, n. 1, p. 68-79, jan./dez., 2016.

SANTOS, R. A. dos; SILVA, R. M. N. B. e. **Racismo científico no Brasil: um retrato racial do Brasil**. **Educar em Revista**, Curitiba, Brasil, v. 34, n. 68, p. 253-268, mar./abr. 2018

SCHMIDT, Cristiane Alkmin Junqueira. et. al. **Macroeconomia**. 3. ed. Rio de Janeiro: Elsevier, 2010.

SEN, Amartya. **Desenvolvimento como liberdade**. Tradução Laura Teixeira Mota. et. al. São Paulo: Companhia das Letras, 2010.

SILVA, Vera Lucia. **Ensaio em economia da educação: efeitos da escola, programa mais educação em background familiar sobre o desempenho dos estudantes**. Tese (doutorado em Economia) - Faculdade de Economia, Administração, Atuária e Contabilidade. Universidade Federal do Ceará, 2016.

TODOS PELA EDUCAÇÃO. **Anuário brasileiro da Educação Básica: 2018**. Disponível em: <https://todospelaeducacao.org.br/_uploads/20180824-Anuario_Educacao_2018_atualizado_WEB.pdf?utm_source=conteudoSite>. Acesso em: 04 jul. 2019.

TOSTE DAFLON, Verônica; CARVALHAES, Flávio; FERES JÚNIOR, João. **Sentindo na pele: percepções de discriminação cotidiana de pretos e pardos no Brasil**. **Dados - Revista de Ciências Sociais**, vol. 60, n. 2, pp. 293-330, 2017.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS PARA A EDUCAÇÃO A CIÊNCIA E A CULTURA. **Desigualdades de aprendizado entre alunos das escolas públicas brasileiras: evidências da prova Brasil (2007 a 2013)**. Série: Debates ED, n. 5, abr. 2017.

VIANA, Iara Azevedo Vitelli. **Investimento produtivo em educação no Brasil: educação infantil para crianças pobres**. 2012. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) - Escola Nacional de Ciências Estatísticas. Rio de Janeiro, 2012.

WOOLDRIDGE, Jeffrey M. **Econometric analysis of cross section and panel data**. 2nd ed. Massachusetts: MIT Press. Cambridge, 2010.

WOOLDRIDGE, Jeffrey M. **Introdução à econometria: uma abordagem moderna**. Tradução José Antônio Ferreira. Revisão técnica de Galo Carlos Lopez Noriega. São Paulo. 2013.

**APÊNDICE A – TABELAS DEMONSTRATIVAS DA PROPORÇÃO DE
ALUNOS(AS) NOS NÍVEIS DE PROFICIÊNCIA CONFORME ESCALA SAEB
PARA A 4ª SÉRIE/5º ANO E 8ª SÉRIE/9º ANO.**

Tabela 2 – Proporção nos níveis de proficiência (4ª Série e 5º Ano)

Ano	Nível de proficiência								
	1	2	3	4	5	6	7	8	9
1999	35,6	21,1	17,1	12,2	7,6	4,1	1,6	0,5	0,2
2001	35,9	18,4	17,2	13,3	8,4	4,4	1,8	0,6	0,1
2003	29,9	19,1	18,5	15,0	10,0	4,9	1,9	0,5	0,2
2005	29,4	19,6	18,8	14,9	10,1	5,1	1,7	0,4	0,1
2007	30,1	23,4	20,9	14,1	7,4	2,9	0,9	0,2	0,0
2009	26,1	22,2	20,0	14,6	9,4	4,9	2,0	0,5	0,2
2011	22,4	19,8	19,9	16,3	11,0	6,3	2,9	1,1	0,2
2013	23,1	18,4	17,5	15,6	12,1	7,8	4,0	1,4	0,3
2015	13,6	15,6	18,4	18,5	15,7	10,6	5,0	1,8	0,9
2017	12,2	14,0	17,0	18,4	16,6	11,8	6,5	2,8	0,7
Média	25,8	19,2	18,5	15,3	10,8	6,3	2,8	1,0	0,3
Desvio Padrão	8,1	2,8	1,4	2,0	3,2	2,9	1,8	0,8	0,3

Fonte: Microdados SAEB/Prova Brasil (Elaborado pelo autor)

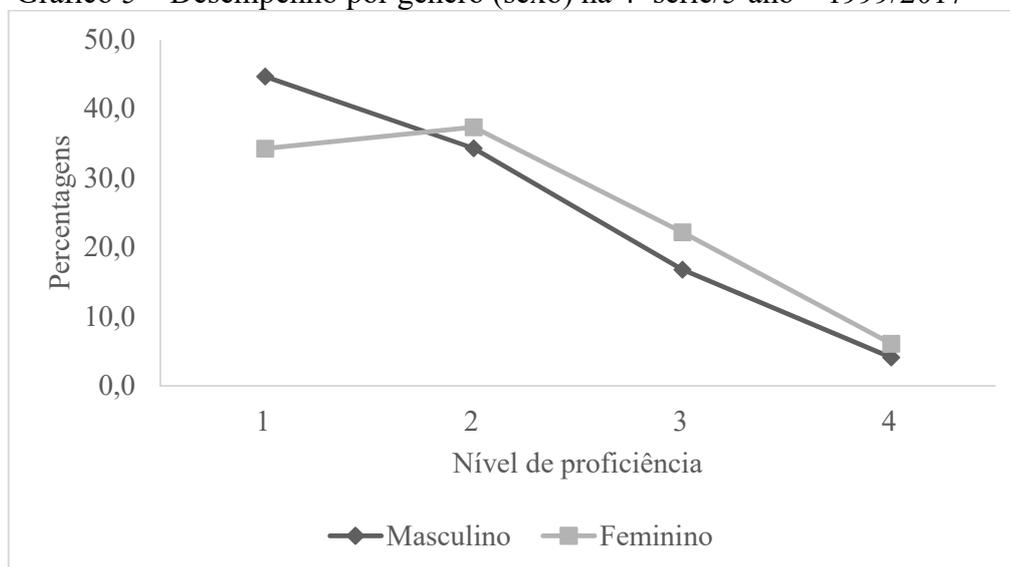
Tabela 3 – Proporção dos níveis de proficiência (8ª Série e 9º Ano)

Ano	Nível de Proficiência							
	1	2	3	4	5	6	7	8
1999	22,5	25,5	22,0	15,7	9,1	3,9	1,1	0,2
2001	18,9	23,0	21,7	17,3	11,8	5,4	1,8	0,2
2003	20,2	22,7	21,7	17,0	11,1	5,6	1,5	0,2
2005	21,0	23,2	21,3	16,3	11,0	5,4	1,5	0,2
2007	27,2	28,3	22,6	13,8	6,0	1,7	0,3	0,0
2009	21,7	25,6	23,6	16,7	8,7	3,0	0,6	0,0
2011	21,5	25,0	23,1	16,9	9,2	3,5	0,8	0,0
2013	20,5	23,5	22,7	17,7	10,4	4,2	1,0	0,0
2015	17,5	22,2	23,2	19,1	11,8	4,8	1,2	0,1
2017	16,1	20,3	22,7	20,5	13,0	5,4	2,0	0,0
Média	20,5	23,8	22,5	17,3	10,3	4,3	1,2	0,1
Desvio Padrão	3,1	2,3	0,8	1,9	2,1	1,4	0,6	0,1

Fonte: Microdados SAEB/Prova Brasil (Elaborado pelo autor)

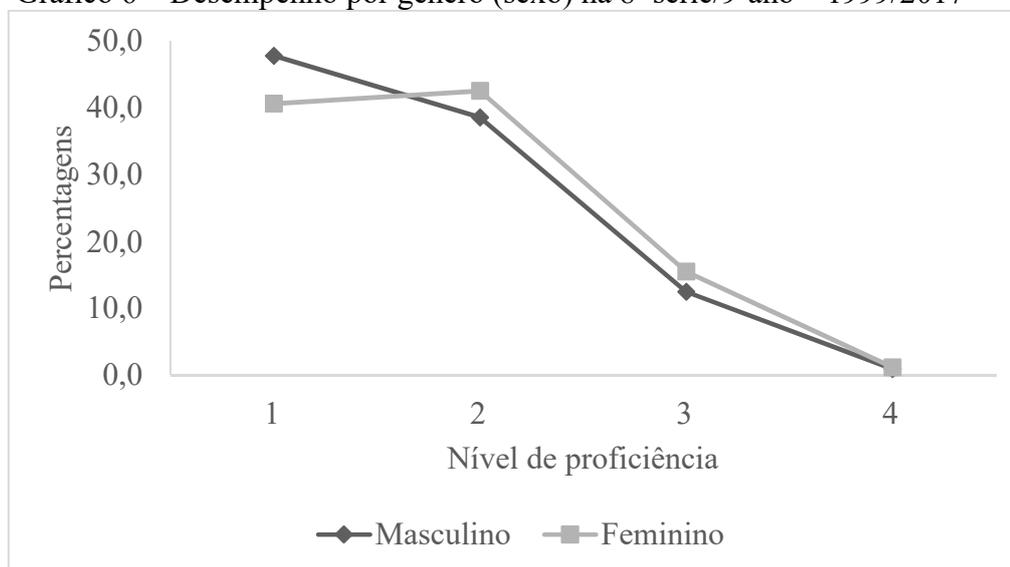
APÊNDICE B – GRÁFICOS DEMONSTRATIVOS DE DESEMPENHO POR GÊNERO (SEXO)

Gráfico 5 – Desempenho por gênero (sexo) na 4ª série/5ºano – 1999/2017



Fonte: Microdados SAEB/Prova Brasil (Elaborado pelo autor)

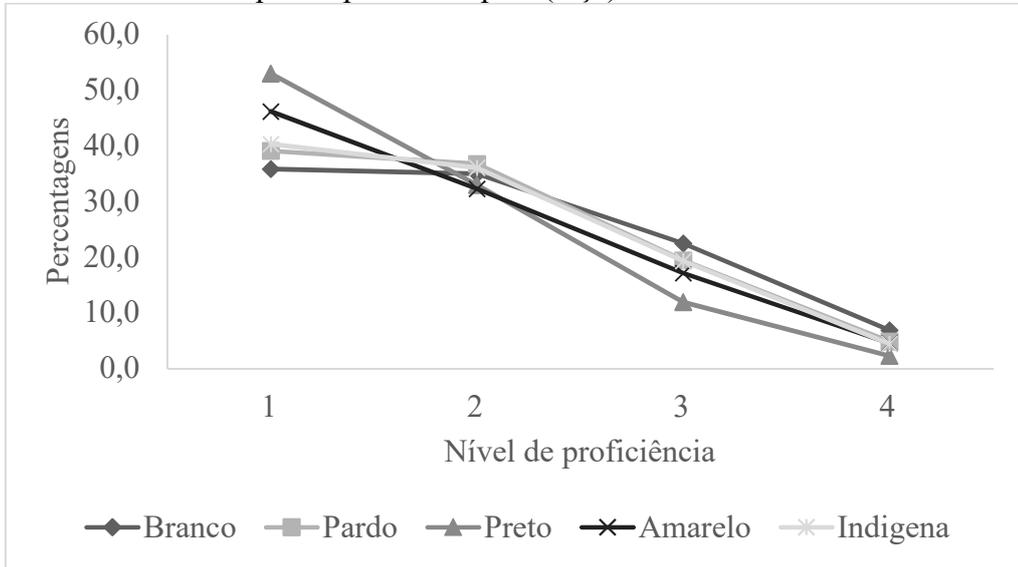
Gráfico 6 – Desempenho por gênero (sexo) na 8ª série/9ºano – 1999/2017



Fonte: Microdados SAEB/Prova Brasil (Elaborado pelo autor)

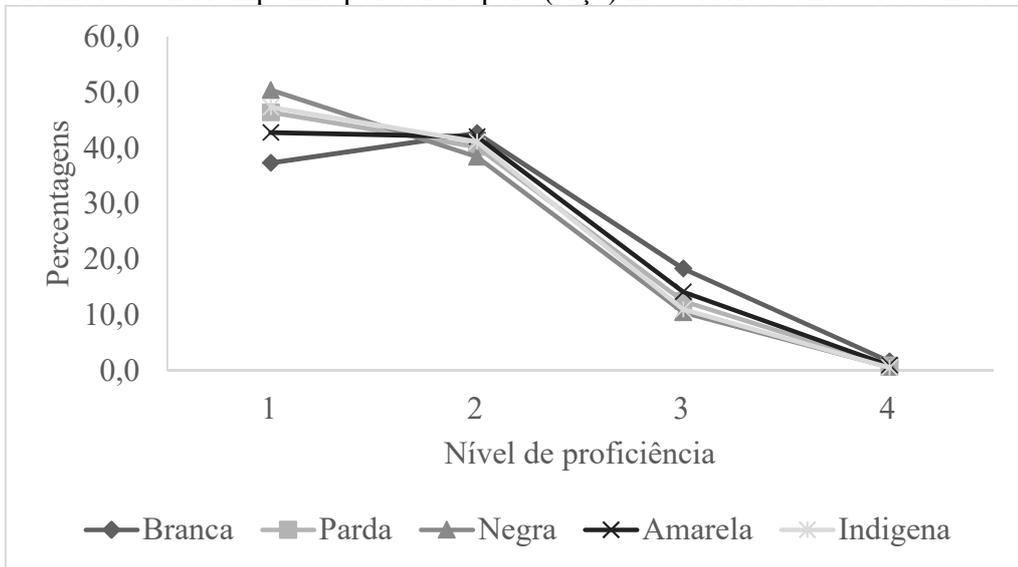
**APÊNDICE C – GRÁFICOS DEMONSTRATIVOS DE DESEMPENHO
POR COR DA PELE OU RAÇA**

Gráfico 7 – Desempenho por cor da pele (raça) na 4ª série/5ºano – 1999/2017



Fonte: Microdados SAEB/Prova Brasil (Elaborado pelo autor)

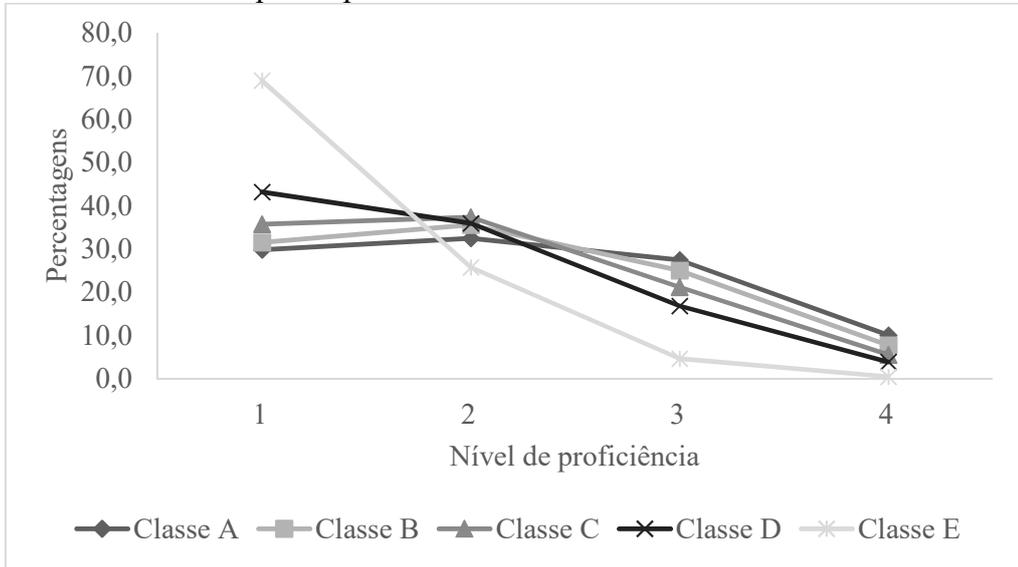
Gráfico 8 – Desempenho por cor da pele (raça) na 8ª série/9ºano – 1999/2017



Fonte: Microdados SAEB/Prova Brasil (Elaborado pelo autor)

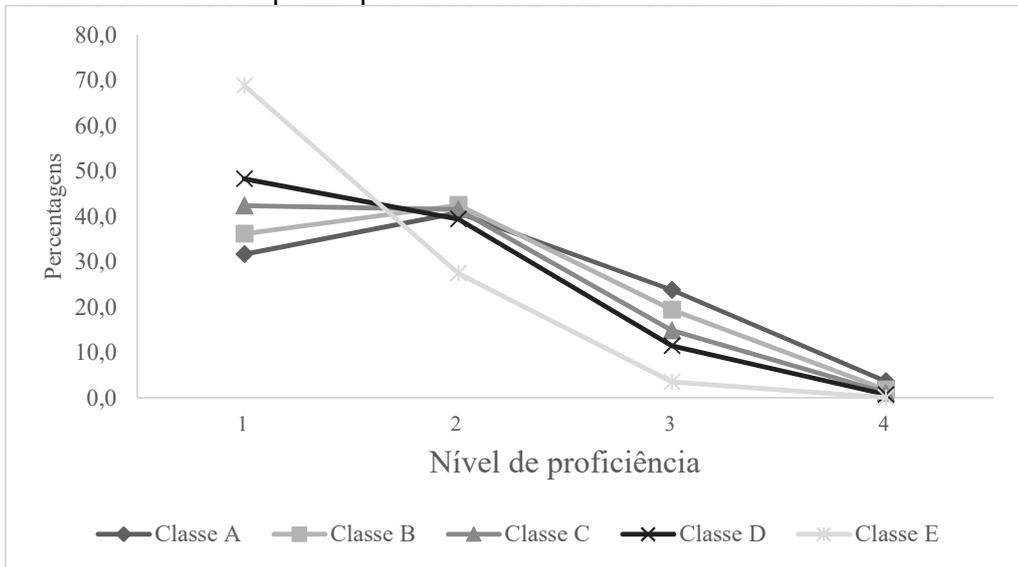
**APÊNDICE D – GRÁFICOS DEMONSTRATIVOS DE DESEMPENHO
POR CLASSE SOCIAL**

Gráfico 9 – Desempenho por classe social na 4ª série/5ºano – 1999/2017



Fonte: Microdados SAEB/Prova Brasil (Elaborado pelo autor)

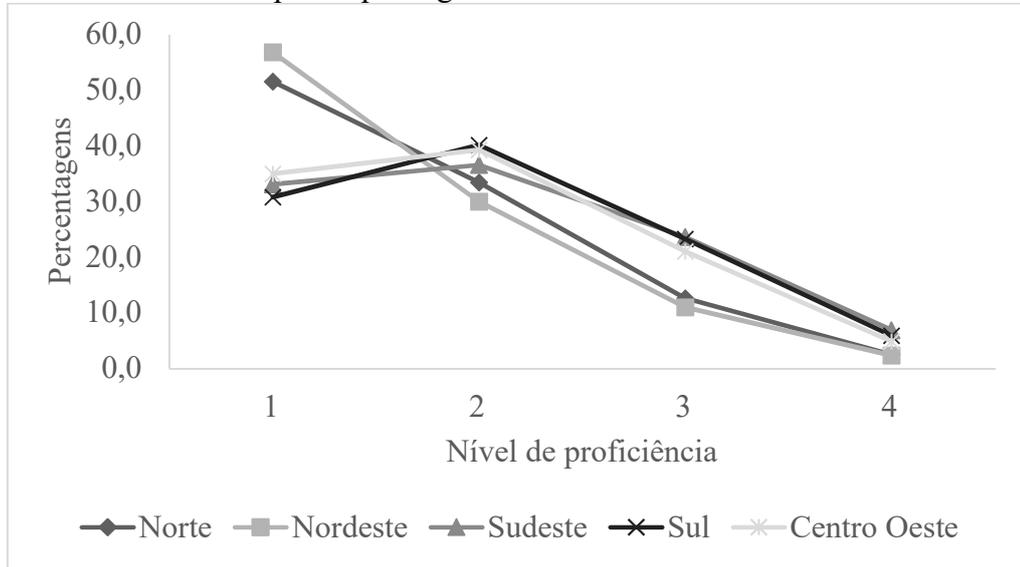
Gráfico 10 – Desempenho por classe social na 8ª série/9ºano – 1999/2017



Fonte: Microdados SAEB/Prova Brasil (Elaborado pelo autor)

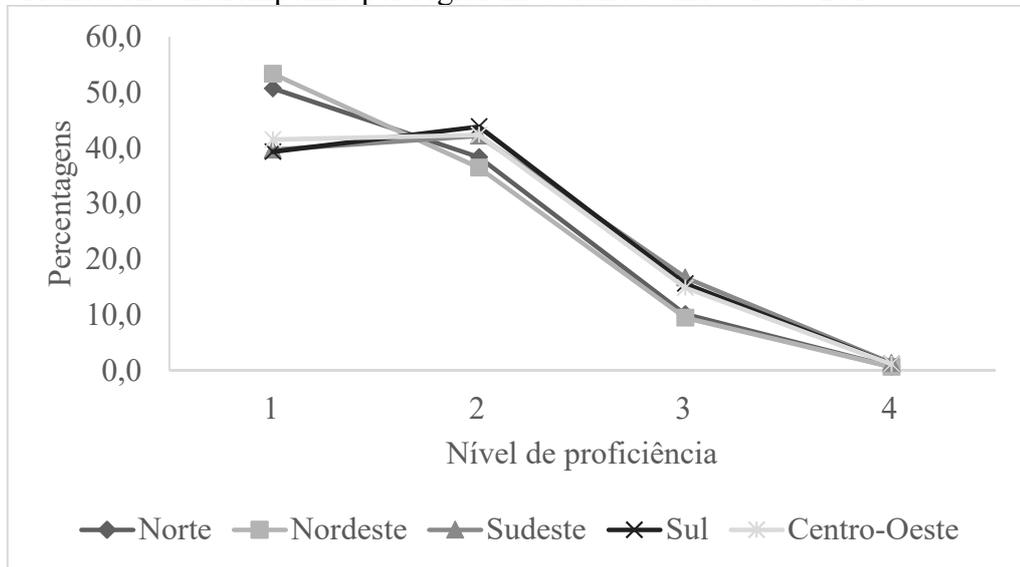
**APÊNDICE E – GRÁFICOS DEMONSTRATIVOS DE DESEMPENHO
POR REGIÃO GEOGRÁFICA**

Gráfico 11 – Desempenho por região na 4ª série/5ºano – 1999/2017



Fonte: Microdados SAEB/Prova Brasil (Elaborado pelo autor)

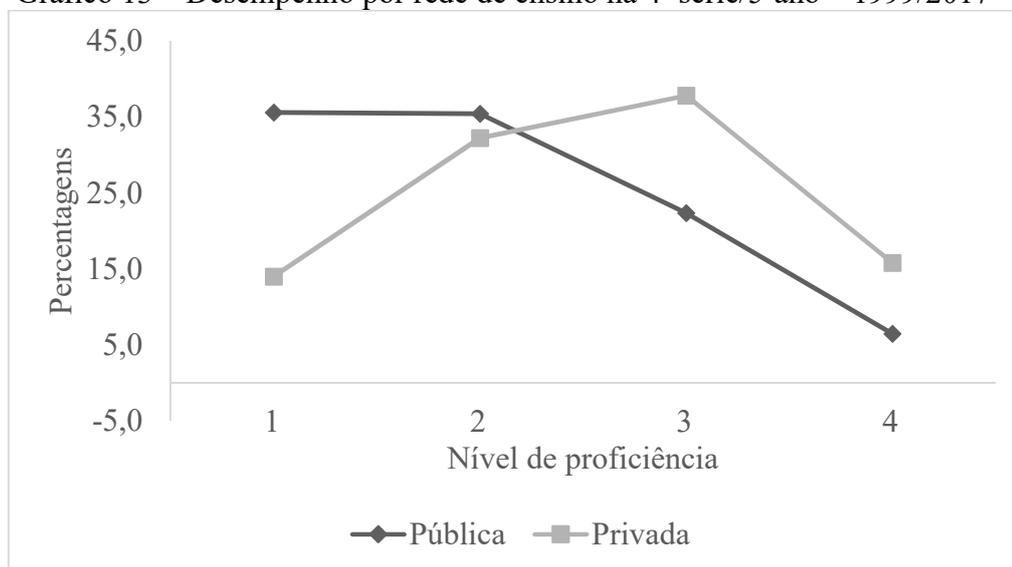
Gráfico 12 – Desempenho por região na 8ª série/9ºano – 1999/2017



Fonte: Microdados SAEB/Prova Brasil (Elaborado pelo autor)

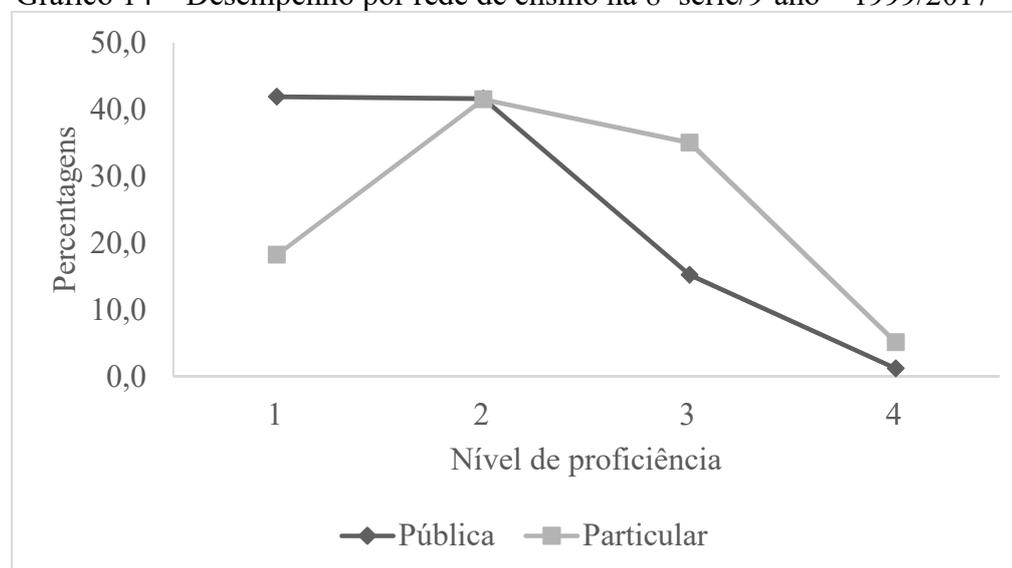
APÊNDICE F – GRÁFICOS DEMONSTRATIVOS DE DESEMPENHO POR REDE DE ENSINO

Gráfico 13 – Desempenho por rede de ensino na 4ª série/5ºano – 1999/2017



Fonte: Microdados SAEB/Prova Brasil (Elaborado pelo autor)

Gráfico 14 – Desempenho por rede de ensino na 8ª série/9ºano – 1999/2017



Fonte: Microdados SAEB/Prova Brasil (Elaborado pelo autor)

APÊNDICE G – CRITÉRIO DE CLASSIFICAÇÃO ECONÔMICA DO BRASIL

Quadro 10 - Estimativa para a renda média domiciliar para os estratos socioeconômicos

Estratos Socioeconômicos	Cortes do Critério Brasil de 1999 a 2017		
	1999 a 2004	2005 a 2013	2014 a 2017
A	25 - 34	35 - 46	45 - 100
B	17 - 24	23 - 34	29 - 44
C	11 - 16	14 - 22	17 - 28
D	6 - 10	8 - 13	0 - 16
E	0 - 5	0 - 7	

Fonte: Associação Brasileira de Empresas de Pesquisa (ABEP) – Adaptado pelo autor

APÊNDICE H – TABELAS DE PONDERAÇÕES DE AMOSTRAS E SUBAMOSTRAS

Tabela 20 – Ponderações de amostras e subamostras 4ª série/5º ano

Variáveis	Categorias	1999		2001		2003		2005		2007		2009		2011		2013		2015		2017	
		am.	sub.																		
Sexo	Masculino	50,4	51,8	50,9	51,7	50,2	49,4	49,8	50,2	50,3	50,3	51,1	51,1	51,2	51,3	50,9	50,8	50,9	50,9	50,8	50,9
	Feminino	49,2	48,2	49,1	48,3	49,8	50,6	50,2	49,8	49,7	49,7	48,9	48,9	48,8	48,7	49,1	49,2	49,1	49,1	49,2	49,1
Raça	Branco	43,3	43,0	43,8	43,1	40,6	39,4	37,6	37,9	34,8	34,7	35,1	35,2	34,0	34,0	34,3	34,2	31,9	32	33,0	33,0
	Pardo	37,3	37,1	37,1	37,4	43,1	44,1	44,4	43,8	46,4	46,5	46,1	46	50,5	50,5	48,8	48,9	51,5	51,4	49,3	49,3
	Preto	10,6	10,9	11,1	11,7	9,5	9,8	11,7	11,7	11,7	11,7	12,5	12,5	10,7	10,7	11,1	11,1	10,9	10,9	11,6	11,6
	Amarelo	4,9	4,9	4,4	4,1	2,7	2,6	3,2	3,4	3,2	3,2	2,5	2,5	2,1	2,1	2,7	2,7	2,7	2,7	3,0	3,1
	Indígena	3,6	4,1	3,6	3,7	4,1	4,1	3,1	3,3	3,9	3,9	3,8	3,8	2,7	2,7	3,1	3,1	3,0	3,0	3,0	3,0
Classes	A	17,5	17,6	2,7	3,1	2,2	2,2	0,2	0,2	0,1	0,1	0,1	0,1	0,1	0,1	0,2	0,2	0,7	0,7	1,0	1,0
	B	81,7	81,3	21,6	21,3	20,2	19,8	22,0	21,9	12,4	12,3	15,5	15,5	12,1	12	16,1	16	6,8	6,8	13,3	13,2
	C	0,7	1,1	31,6	30,6	31,7	31,8	41,8	41,7	58,2	58,3	59,4	59,4	62,9	63,1	61,8	61,9	34,4	34,5	43,6	43,5
	D	-	-	44,1	44,9	36,3	36,1	29,1	29,6	25,9	26,0	20,3	20,3	22,1	22,1	20,8	20,8	58,2	58,1	42,1	42,3
	E	-	-	-	-	9,6	10,2	7,0	6,9	3,4	3,4	4,7	4,7	2,9	2,9	1,3	1,2	-	-	-	-
Região	Norte	15,6	15,5	16,2	16,4	14,3	14,1	20,4	20,1	9,2	9,1	10,1	10,1	10,5	10,5	11,8	11,8	12,0	12,0	12,0	12,0
	Nordeste	45,7	45,3	41,3	41,8	33,9	34,3	36,0	35,4	22,8	22,9	26,5	26,5	26,9	26,9	28,2	28,1	26,9	26,8	28,5	28,5
	Sudeste	14,9	14,8	15,6	15,1	21,5	21,5	16,1	16,7	45,6	45,8	41,5	41,6	43,1	43,1	37,2	37,4	38,5	38,7	37,9	38,0
	Sul	10,2	10,2	12,8	12,4	16,5	15,7	13,6	13,9	14,8	14,8	14,8	14,7	12,2	12,2	14,3	14,3	14,2	14,2	13,7	13,7
	Centro Oeste	13,6	14,2	14,1	14,3	13,8	14,4	13,9	13,9	7,6	7,5	7,1	7,1	7,3	7,3	8,5	8,5	8,3	8,3	7,9	7,9
Rede	Pública	71,6	71,0	68,4	69,0	72,1	72,5	70,3	70,3	-	-	-	-	98,6	98,7	98,1	98,1	98,5	98,5	99,0	99,0
	Privada	28,4	29,0	31,6	31,0	27,9	27,5	29,7	29,7	-	-	-	-	1,4	1,3	1,9	1,9	1,5	1,5	1,0	1,0

Fonte: SAEB 1999/2017 (Elaborado pelo autor)

Tabela 21– Ponderações de amostras e subamostras 8ª série/9º ano

Variáveis	Categorias	1999		2001		2003		2005		2007		2009		2011		2013		2015		2017	
		am.	sub.																		
Sexo	Masculino	46,1	46	46,3	46	46,5	46,2	45,3	46,2	46,3	46,3	46,2	46,4	47	46,9	47,9	48,2	48,4	48,4	48,9	48,9
	Feminino	53,9	54	53,7	54	53,5	53,8	53,5	53,8	53,7	53,7	53,8	53,6	53	53,1	52,1	51,8	51,6	51,6	51,1	51,1
Raça	Branco	46,9	46,2	45,3	45,7	43,1	42,9	37,7	38,1	36	36,1	34	33,9	35,4	35,3	34	34	29,7	29,5	31,1	31,2
	Pardo	37,5	37,4	37,8	37,7	42	42,8	45,5	46,6	45,8	45,7	48,1	48,1	47,9	48,1	48,5	48,4	51	51	49,7	49,5
	Preto	6,4	6,8	7,1	7,3	7,4	7	7,9	7,8	10,9	10,9	11,2	11,1	11	10,9	11,5	11,5	12,7	12,8	12,7	12,8
	Amarelo	6,6	6,9	7,4	7	4,1	3,9	5,3	5,3	3,8	3,8	3,2	3,2	3,4	3,4	3,7	3,7	4,3	4,3	3,9	3,9
	Indigena	2,7	2,7	2,5	2,3	3,4	3,4	2,1	2,2	3,4	3,4	3,6	3,7	2,3	2,3	2,3	2,3	2,3	2,3	2,6	2,6
Classes	A	27,6	27,6	5,1	5,5	3,4	3,3	0,7	0,8	0,1	0,1	0,1	0,1	0,1	0,1	0,2	0,2	0,5	0,5	1	1
	B	72,2	72,1	25,8	26	26,8	26,5	28,5	28,9	13,2	13,2	16,5	16,6	15,1	15,1	23,7	23,9	6,4	6,4	12,3	12,3
	C	0,2	0,2	31,2	29,9	33	33,5	41,9	41,2	57,4	57,4	58,8	58,8	61,5	61,5	62	62	36	36	44,5	44,6
	D	-	-	32,7	33,4	31,3	31,4	24,3	24,5	26,4	26,4	20,8	20,6	21,2	21,1	13,2	13,2	57,1	57,1	42,2	42,1
	E	-	-	5,3	5,3	5,4	5,3	4,6	4,6	3	3	3,8	3,8	2,3	2,2	0,9	0,9	-	-	-	-
Região	Norte	15,3	15,3	15,8	16,2	15,1	14,9	17,4	17,5	7,4	7,4	8,4	8,4	8,2	8,2	9,2	9,1	9,8	9,7	10,8	10,8
	Nordeste	38,5	38,5	40,2	40	34,2	34,9	38,6	39,1	22,8	22,7	26,4	26,2	25,9	26	25,9	25,9	27,3	27,2	29,7	29,7
	Sudeste	17,6	17,4	17,2	17,3	22,3	22	16,9	16,5	48,2	48,4	43,2	43,4	43,9	43,8	42,9	42,9	43,6	43,6	37,2	37,3
	Sul	13,1	13,1	12,4	12,1	15,8	15,2	13,1	13,2	14,3	14,3	14,2	15	14,4	14,5	14,5	14,6	11,8	11,9	13,7	13,8
	Centro Oeste	15,5	15,8	14,4	14,4	12,6	12,9	13,9	13,7	7,3	7,3	7,1	7	7,5	7,5	7,5	7,5	7,5	7,5	8,5	8,5
Rede	Pública	64,9	64,7	64,3	63,9	69,5	70,3	66,4	66,1	-	-	-	-	98,1	98,1	97,9	97,9	98,7	98,7	98,8	98,8
	Privada	35,4	35,3	35,7	36,1	30,5	29,7	33,6	33,9	-	-	-	-	1,9	1,9	2,1	2,1	1,3	1,3	1,2	1,2

Fonte: SAEB 1999/2017 (Elaborado pelo autor)

**ANEXO A – DESCRITORES DO NÍVEL DE PROFICIÊNCIA DO SAEB PARA A 4ª
SÉRIE/5º ANO**

Quadro 3 – Descritores da escala de Proficiência em LP – 4ª Série/5º Ano e 8ª Série/9º Ano. (continua)

Nível	Descrição do Nível
Nível 1 Desempenho menor que 150	Localizar informações explícitas em textos narrativos curtos, informativos e anúncios; Identificar o tema de um texto; Localizar elementos como o personagem principal; Estabelecer relação entre partes do texto: personagem e ação; ação e tempo; ação e lugar.
Nível 2 Desempenho maior ou igual a 150 e menor que 175	Localizar informações explícitas em contos; Identificar o assunto principal e a personagem principal em reportagem e em fábulas; Reconhecer a finalidade de receitas, manuais e regulamentos; Inferir características de personagens em fábulas; Interpretar linguagem verbal e não verbal em tirinhas.
Nível 3 Desempenho maior ou igual a 175 e menor que 200	Localizar informação explícita em contos e reportagens; Localizar informação explícita em propagandas com ou sem apoio de recursos gráficos; Reconhecer relação de causa e consequência em poemas, contos e tirinhas; Inferir o sentido de palavra, o sentido de expressão ou o assunto em cartas, contos, tirinhas e histórias em quadrinhos com o apoio de linguagem verbal e não verbal.
Nível 4 Desempenho maior ou igual a 200 e menor que 225	Identificar informação explícita em sinopses e receitas culinárias; Identificar assunto principal e personagem em contos e letras de música; Identificar formas de representação de medida de tempo em reportagens; Identificar assuntos comuns a duas reportagens; Identificar o efeito de humor em piadas; Reconhecer sentido de expressão, elementos da narrativa e opinião em reportagens, contos e poemas; Reconhecer relação de causa e consequência e relação entre pronomes e seus referentes em fábulas, poemas, contos e tirinhas; Inferir sentido decorrente da utilização de sinais de pontuação e sentido de expressões em poemas, fábulas e contos; Inferir efeito de humor em tirinhas e histórias em quadrinhos.

Fonte: Microdados SAEB 1999/2011. (continua)

Quadro 3 – Descritores da escala de Proficiência em LP – 4ª Série/5º Ano e 8ª Série/9º Ano. (continuação)

<p>Nível 5 Desempenho maior ou igual a 225 e menor que 250</p>	<p>Identificar assunto e opinião em reportagens e contos; Identificar assunto comum a cartas e poemas; Identificar informação explícita em letras de música e contos; Reconhecer assunto em poemas e tirinhas; Reconhecer sentido de conjunções e de locuções adverbiais em verbetes, lendas e contos; Reconhecer finalidade de reportagens e cartazes; Reconhecer relação de causa e consequência e relação entre pronome e seu referente em tirinhas, contos e reportagens; Inferir elementos da narrativa em fábulas, contos e cartas; Inferir finalidade e efeito de sentido decorrente do uso de pontuação e assunto em fábulas; Inferir informação em poemas, reportagens e cartas; Diferenciar opinião de fato em reportagens; Interpretar efeito de humor e sentido de palavra em piadas e tirinhas.</p>
<p>Nível 6 Desempenho maior ou igual a 225 e menor que 250</p>	<p>Identificar opinião e informação explícita em fábulas, contos, crônicas e reportagens; Identificar informação explícita em reportagens com ou sem o auxílio de recursos gráficos; Reconhecer a finalidade de verbetes, fábulas, charges e reportagens; Reconhecer relação de causa e consequência e relação entre pronomes e seus referentes em poemas, fábulas e contos; Inferir assunto principal e sentido de expressão em poemas, fábulas, contos, crônicas, reportagens e tirinhas; Inferir informação em contos e reportagens; Inferir efeito de humor e moral em piadas e fábulas.</p>
<p>Nível 7 Desempenho maior ou igual a 275 e menor que 300</p>	<p>Identificar assunto principal e informações explícitas em poemas, fábulas e letras de música; Identificar opinião em poemas e crônicas; Reconhecer o gênero textual a partir da comparação entre textos e assunto comum a duas reportagens; Reconhecer elementos da narrativa em fábulas; Reconhecer relação de causa e consequência e relação entre pronomes e seus referentes em fábulas, contos e crônicas; Inferir informação e efeito de sentido decorrente do uso de sinais gráficos em reportagens e em letras de música; Interpretar efeito de humor em piadas e contos; Interpretar linguagem verbal e não verbal em histórias em quadrinhos.</p>
<p>Nível 8 Desempenho maior ou igual a 300 e menor que 325</p>	<p>Identificar assunto principal e opinião em contos e cartas do leitor; Reconhecer sentido de locução adverbial e elementos da narrativa em fábulas e contos; Reconhecer relação de causa e consequência e relação entre pronomes e seus referentes em fábulas e reportagens; Reconhecer assunto comum entre textos de gêneros diferentes; Inferir informações e efeito de sentido decorrente do uso de pontuação em fábulas e piadas.</p>
<p>Nível 9 Desempenho maior ou igual a 300 e menor que 325</p>	<p>Identificar opinião em fábulas e reconhecer sentido de advérbios em cartas do leitor.</p>

Fonte: Microdados SAEB 1999/2011.

**ANEXO B – DESCRITORES DO NÍVEL DE PROFICIÊNCIA DO SAEB PARA A 8ª
SÉRIE/9º ANO**

Quadro 4 – Descritores da escala de Proficiência em LP – 8ª Série/9º Ano. (continua)

Nível	Descrição do Nível
Nível 1 Desempenho maior ou igual a 200 e menor que 225	Reconhecer expressões características da linguagem (científica, jornalística etc.) e a relação entre expressão e seu referente em reportagens e artigos de opinião; Inferir o efeito de sentido de expressão e opinião em crônicas e reportagens.
Nível 2 Desempenho maior ou igual a 225 e menor que 250	Localizar informações explícitas em fragmentos de romances e crônicas; Identificar tema e assunto em poemas e charges, relacionando elementos verbais e não verbais; Reconhecer o sentido estabelecido pelo uso de expressões, de pontuação, de conjunções em poemas, charges e fragmentos de romances; Reconhecer relações de causa e consequência e características de personagens em lendas e fábulas; Reconhecer recurso argumentativo em artigos de opinião; Inferir efeito de sentido de repetição de expressões em crônicas.
Nível 3 Desempenho maior ou igual a 250 e menor que 275	Localizar informações explícitas em crônicas e fábulas; Identificar os elementos da narrativa em letras de música e fábulas; Reconhecer a finalidade de abaixo-assinado e verbetes; Reconhecer relação entre pronomes e seus referentes e relações de causa e consequência em fragmentos de romances, diários, crônicas, reportagens e máximas (provérbios); Interpretar o sentido de conjunções, de advérbios, e as relações entre elementos verbais e não verbais em tirinhas, fragmentos de romances, reportagens e crônicas; Comparar textos de gêneros diferentes que abordem o mesmo tema; Inferir tema e ideia principal em notícias, crônicas e poemas; Inferir o sentido de palavra ou expressão em história em quadrinhos, poemas e fragmentos de romances.
Nível 4 Desempenho maior ou igual a 275 e menor que 300	Localizar informações explícitas em artigos de opinião e crônicas; Identificar finalidade e elementos da narrativa em fábulas e contos; Reconhecer opiniões distintas sobre o mesmo assunto em reportagens, contos e enquetes; Reconhecer relações de causa e consequência e relações entre pronomes e seus referentes em fragmentos de romances, fábulas, crônicas, artigos de opinião e reportagens; Reconhecer o sentido de expressão e de variantes linguísticas em letras de música, tirinhas, poemas e fragmentos de romances; Inferir tema, tese e ideia principal em contos, letras de música, editoriais, reportagens, crônicas e artigos; Inferir o efeito de sentido de linguagem verbal e não verbal em charges e história em quadrinhos; Inferir informações em fragmentos de romance; Inferir o efeito de sentido da pontuação e da polissemia como recurso para estabelecer humor ou ironia em tirinhas, anedotas e contos.

Fonte: Microdados SAEB 2013/2017. (continua)

Quadro 4 – Descritores da escala de Proficiência em LP – 8ª Série/9º Ano.(continuação)

<p>Nível 5 Desempenho maior ou igual a 300 e menor que 325</p>	<p>Localizar a informação principal em reportagens; Identificar ideia principal e finalidade em notícias, reportagens e resenhas; Reconhecer características da linguagem (científica, jornalística etc.) em reportagens; Reconhecer elementos da narrativa em crônicas; Reconhecer argumentos e opiniões em notícias, artigos de opinião e fragmentos de romances; Diferenciar abordagem do mesmo tema em textos de gêneros distintos; Inferir informação em contos, crônicas, notícias e charges; Inferir sentido de palavras, da repetição de palavras, de expressões, de linguagem verbal e não verbal e de pontuação em charges, tirinhas, contos, crônicas e fragmentos de romances.</p>
<p>Nível 6 Desempenho maior ou igual a 325 e menor que 350</p>	<p>Identificar ideia principal e elementos da narrativa em reportagens e crônicas; Identificar argumento em reportagens e crônicas; Reconhecer o efeito de sentido da repetição de expressões e palavras, do uso de pontuação, de variantes linguísticas e de figuras de linguagem em poemas, contos e fragmentos de romances; Reconhecer a relação de causa e consequência em contos; Reconhecer diferentes opiniões entre cartas de leitor que abordam o mesmo tema; Reconhecer a relação de sentido estabelecida por conjunções em crônicas, contos e cordéis; Reconhecer o tema comum entre textos de gêneros distintos; Reconhecer o efeito de sentido decorrente do uso de figuras de linguagem e de recursos gráficos em poemas e fragmentos de romances; Diferenciar fato de opinião em artigos e reportagens; Inferir o efeito de sentido de linguagem verbal e não verbal em tirinhas.</p>
<p>Nível 7 Desempenho maior ou igual a 350 e menor que 375</p>	<p>Localizar informações explícitas, ideia principal e expressão que causa humor em contos, crônicas e artigos de opinião; Identificar variantes linguísticas em letras de música; Reconhecer a finalidade e a relação de sentido estabelecida por conjunções em lendas e crônicas.</p>
<p>Nível 8 Desempenho maior ou igual 375</p>	<p>Localizar ideia principal em manuais, reportagens, artigos e teses; Identificar os elementos da narrativa em contos e crônicas; Diferenciar fatos de opiniões e opiniões diferentes em artigos e notícias; Inferir o sentido de palavras em poemas.</p>

Fonte: Microdados SAEB 2013/2017.