

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Faculdade de Educação
Programa de Pós-Graduação: Conhecimento e Inclusão Social

Isis Arantes Maciel de Souza

**DESIGUALDADES DE TRAJETÓRIAS EDUCACIONAIS NA EDUCAÇÃO BÁSICA
BRASILEIRA**

Belo Horizonte
2023

Isis Arantes Maciel de Souza

**DESIGUALDADES DE TRAJETÓRIAS EDUCACIONAIS NA EDUCAÇÃO BÁSICA
BRASILEIRA**

Tese apresentada ao Programa de Pós-graduação: Conhecimento e Inclusão Social da Universidade Federal de Minas Gerais como requisito parcial para a obtenção do título de doutora em educação.

Orientadora: Profa. Dra. Flavia Xavier

Belo Horizonte
2023

S729d Souza, Isis Arantes Maciel de, 1988-
T Desigualdades de trajetórias educacionais na educação básica brasileira [manuscrito] / Isis Arantes Maciel de Souza. - Belo Horizonte, 2023. 156 f. : enc., il., color.

Tese -- (Doutorado) - Universidade Federal de Minas Gerais, Faculdade de Educação.
Orientadora: Flavia Xavier.
Bibliografia: f. 115-122.
Anexos: f. 138-156.
Apêndices: f. 123-138.

1. Educação -- Teses. 2. Educação básica -- Teses. 3. Censo escolar -- Teses. 4. Igualdade na educação -- Teses. 5. Políticas públicas -- Educação -- Teses.
I. Título. II. Xavier, Flavia. III. Universidade Federal de Minas Gerais, Faculdade de Educação.

CDD- 370.981

Catálogo da fonte: Biblioteca da FaE/UFMG (Setor de referência)

Bibliotecária: Raissa Michalsky Martins CRB6 3155/O



UNIVERSIDADE FEDERAL DE MINAS GERAIS
FACULDADE DE EDUCAÇÃO
PÓS-GRADUAÇÃO EM EDUCAÇÃO: CONHECIMENTO E INCLUSÃO SOCIAL

ATA DA DEFESA DE TESE DA ALUNA

ISIS ARANTES MACIEL DE SOUZA

Realizou-se, no dia 25 de agosto de 2023, às 14:00 horas, na sala 5102 da Faculdade de Educação, da Universidade Federal de Minas Gerais, a 935ª defesa de tese, intitulada *Desigualdades de trajetórias educacionais na educação básica brasileira*, apresentada por ISIS ARANTES MACIEL DE SOUZA, número de registro 2019650228, graduada no curso de CIÊNCIAS SOCIAIS, como requisito parcial para a obtenção do grau de Doutor em EDUCAÇÃO - CONHECIMENTO E INCLUSÃO SOCIAL, à seguinte Comissão Examinadora: Prof(a). Flávia Pereira Xavier - Orientador (UFMG), Prof(a). Mariane Campelo Koslinski (UFRJ), Prof(a). José Francisco Soares (UFMG), Prof(a). Maria Teresa Gonzaga Alves (UFMG), Prof(a). Raquel Rangel de Meireles Guimarães (UFPR).

A comissão considerou a tese: Aprovada, reconhecendo a sua contribuição para um campo de pesquisa ainda recente e recomenda a publicação de artigos a partir do trabalho.

Finalizados os trabalhos, lavrei a presente ata que, lida e aprovada, vai assinada por mim e pelos membros da Comissão.

Belo Horizonte, 25 de agosto de 2023.

Prof(a). Flávia Pereira Xavier - orientadora (Doutora)

Prof(a). Mariane Campelo Koslinski (Doutora)

Prof(a). José Francisco Soares (Doutor)

Prof(a). Maria Teresa Gonzaga Alves (Doutora)

Prof(a). Raquel Rangel de Meireles Guimaraes (Doutora)



Documento assinado eletronicamente por **Flavia Pereira Xavier, Professora do Magistério Superior**, em 28/08/2023, às 14:27, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **José Francisco Soares, Usuário Externo**, em 29/08/2023, às 19:19, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Mariane Campelo Koslinski, Usuária Externa**, em 05/09/2023, às 19:54, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Raquel Rangel de Meireles Guimarães, Usuária Externa**, em 11/09/2023, às 16:18, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Maria Teresa Gonzaga Alves, Professora do Magistério Superior**, em 11/09/2023, às 18:11, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **2577039** e o código CRC **4189452C**.



AGRADECIMENTOS

À minha orientadora Prof^a. Dr^a. Flávia Xavier por tanto aprendizado. Sou grata pelo entusiasmo e rigor com que conduziu esta tese. E também por sempre acreditar que eu iria conseguir e me lembrar que nós fazemos *a boa ciência*.

Ao Núcleo de Pesquisa em Desigualdades Escolares (Nupede - FaE/UFMG) e todos os seus integrantes que foram interlocutores fundamentais. Em especial à Prof^a. Dr^a. Teresa, por tantas contribuições e pela oportunidade do estágio docente.

Ao PPGE, por financiar o acesso à sala segura.

À Fulbright, pela temporada de doutorado sanduíche na Universidade da Califórnia–Riverside e ao professor Dr. Palardy que gentilmente me recebeu. Seus cursos e comentários permearam e desenvolveram esse trabalho.

À banca examinadora desta tese composta por pesquisadores tão fundamentais aos avanços educacionais do país. Além de grata me sinto honrada.

Às instituições Escola estadual Padre João de Matos, Escola estadual Pedro II, Secretaria de Educação - MG, Cefet - MG e Uemg. Experiências profissionais que se entrelaçam com minha pergunta de pesquisa.

Às vontades políticas e aos gestores públicos que viabilizaram as políticas educacionais da primeira década dos anos 2000. Agradeço com a certeza de que sem essas esse trabalho não existiria.

Aos meus pais e irmãos, por tanto.

Às minhas amigas. Especialmente: Natasha, Nath, Alexia, Luísa e Mari [por me receberem no Rio e em Brasília quando esse trabalho exigiu] e os Fulbrighters 2022-2023 [pelo olhar latino que juntos aprendemos a ter sobre a ciência, sobre nós e sobre o mundo – o que me fez uma pesquisadora melhor].

Ao Claudinho, por um sem-fim de motivos e por ter me ensinado a programar e a *construir scripts elegantes*, habilidades fundamentais para a execução desse trabalho.

RESUMO

A pergunta que esse trabalho se propõe a responder é: as características das escolas produzem efeitos capazes de reduzir as chances de intercorrências nas trajetórias escolares, quando controladas as variáveis individuais dos alunos? Fazemos isso através de uma investigação empírica que utiliza o Censo Escolar respondemos essa pergunta com duas contribuições. A primeira é um desenho do cenário das trajetórias escolares brasileiras na última década. Essa fotografia permite acompanhar e entender ao longo dos anos o funil que é o sistema educacional brasileiro. A segunda está nessas trajetórias colocadas sob três premissas teóricas através de modelos analíticos: o efeito das origens sociais, o efeito de pares e o efeito escola. O objetivo desses modelos foi mapear o objeto de pesquisa para nortear gestores na elaboração de políticas públicas que sejam capazes de aumentar as trajetórias regulares com equidade educacional. O trabalho também discute sobre a imputação de casos ausentes nas bases educacionais e como essa metodologia pode impactar sobre os alunos que deixam de frequentar a escola. Nosso principal resultado é o efeito da formação docente adequada sobre as trajetórias regulares.

Palavras-chave: trajetórias educacionais, efeito escola, Censo Escolar, bases de dados longitudinais.

ABSTRACT

This work aims to answer the question: do the characteristics of schools produce effects capable of reducing the chances of complications in school trajectories when the individual variables of students are controlled? Through an empirical and quantitative investigation, we answer this question with two contributions. The first is a sketch of the scenario of Brazilian school trajectories in the last decade. This photograph makes it possible to follow and understand the funnel that is the Brazilian educational system over the years. The second is in these trajectories placed under three theoretical premises through analytical models. The objective of the models was to map the research object to guide managers in the elaboration of public policies capable of increasing regular trajectories with educational equity. The main results are the effect of teacher training and regularity on students failing and dropping out of school. The work also discusses the imputation of missing cases in educational databases and how this methodology can impact students who stop attending school. Our main result is the effect of adequate teacher training on regular trajectories.

Keywords: educational trajectories, school effects, School Census, longitudinal databases.

LISTA DE GRÁFICOS

Gráfico 1- Fluxo dos estudantes na educação básica (2009 a 2020). Base Longitudinal II (coorte 02/03 - dados protegidos - abrangência 2009 - 2020).	45
Gráfico 2 - Alunos promovidos e em etapas adequadas em relação ao ano anterior (%). Base Longitudinal II (coorte 02/03 - dados protegidos - abrangência 2009 - 2020).....	46
Gráfico 3 - Trajetórias educacionais na Educação Básica Completa (2009 a 2020). Base Longitudinal II (coorte 02/03 - dados protegidos - abrangência 2009 - 2020).....	49
Gráfico 4 - Trajetórias escolares na educação básica por NSE das escolas (2009 a 2020). Base Longitudinal II (coorte 02/03 - dados protegidos - abrangência 2009 - 2020).....	50
Gráfico 5 - Trajetórias escolares na educação básica por dependência administrativa das escolas (2009 a 2020). Base Longitudinal II (coorte 02/03 - dados protegidos - abrangência 2009 - 2020).....	51
Gráfico 6 - Trajetórias educacionais no ensino fundamental I (2009 a 2013). Base Longitudinal II (coorte 02/03 - dados protegidos - abrangência 2009 - 2020).	52
Gráfico 7 - Trajetórias educacionais no Ensino Fundamental II (2009 a 2017). Base Longitudinal II (coorte 02/03 - dados protegidos - abrangência 2009 - 2020).....	52
Gráfico 8 - Trajetórias educacionais no ensino fundamental Completo (2009 a 2017). Base Longitudinal II (coorte 02/03 - dados protegidos - abrangência 2009 - 2020).....	53
Gráfico 9 - Trajetórias educacionais no Ensino Médio (2018 a 2020). Base Longitudinal II (coorte 02/03 - dados protegidos - abrangência 2009 - 2020).....	54
Gráfico 10 - Trajetórias escolares no ensino médio por NSE das escolas (2018 a 2020). Base Longitudinal II (coorte 02/03 - dados protegidos - abrangência 2009 - 2020).....	55
Gráfico 11 - Trajetórias escolares no ensino médio por dependência administrativa das escolas (2018 a 2020). Base Longitudinal II (coorte 02/03 - dados protegidos - abrangência 2009 - 2020).....	55
Gráfico 12 - Trajetórias escolares no ensino fundamental completo por NSE das escolas (2009 a 2017). Base Longitudinal II (coorte 02/03 - dados protegidos - abrangência 2009 - 2020)..	56
Gráfico 13 - Trajetórias escolares no ensino fundamental completo por dependência administrativa das escolas (2009 a 2017). Base Longitudinal II (coorte 02/03 - dados protegidos - abrangência 2009 - 2020).....	57
Gráfico 14 - Trajetórias escolares no ensino fundamental (2009 - 2016). Base Longitudinal I (coorte 99/00 - dados públicos – abrangência 2007 - 2017).	63
Gráfico 15 - Trajetórias escolares no ensino fundamental (2009 - 2016). Base Longitudinal II (coorte 02/03 - dados protegidos - abrangência 2009 - 2020).....	63
Gráfico 16 - Trajetórias escolares no ensino fundamental (2009 - 2016). Base Longitudinal III (coorte 02/03 - dados públicos – abrangência 2007 - 2017).	64
Gráfico 17 - Alunos promovidos e em etapas adequadas em relação ao ano anterior (%).	65
Gráfico 18 - Trajetórias escolares por sexo	89
Gráfico 19 - Trajetórias escolares por raça/cor.....	89
Gráfico 20 - Trajetórias escolares por dependência administrativa das escolas.....	89
Gráfico 21 - Trajetórias escolares por NSE das escolas	89
Gráfico 22 - Trajetórias escolares por regiões do Brasil	89
Gráfico 23 - Trajetórias escolares por estados da Federação.....	90

Gráfico 24 - Valores preditos marginais da trajetória regular (95% de intervalo de confiança)	94
Gráfico 25 - Probabilidades de uma reprovação, por perfil no Sudeste	94
Gráfico 26 - Probabilidades de mais de uma reprovação, por perfil no Sudeste	95
Gráfico 27 - Probabilidades para um ano sem frequentar a escola, por perfil no Sudeste	95
Gráfico 28 - Probabilidades de uma reprovação, por perfil no Nordeste	96
Gráfico 29 - Probabilidades de mais de uma reprovação, por perfil no Nordeste	96
Gráfico 30 - Probabilidades para um ano sem frequentar a escola, por perfil no Nordeste	97
Gráfico 31 - Probabilidades para mais de um ano sem frequentar a escola, por perfil no Nordeste	97

LISTA DE TABELAS

Tabela 1 - Coortes por etapas	37
Tabela 2 - Exemplo de deduplicação	38
Tabela 3 - Exemplo da base longitudinal em formato <i>empilhado</i>	39
Tabela 4 – Indivíduo na base de dados longitudinal em formato <i>espalhado</i>	40
Tabela 5 - Distribuição percentual da coorte por etapas de ensino. Base Longitudinal I (coorte 99/00 - dados públicos – abrangência 2007 - 2017).....	41
Tabela 6 - Variáveis Fluxo Escolar e Situação de Rendimento	43
Tabela 7 - Exemplo hipotético de estudantes sem registro.....	61
Tabela 8 - Exemplo hipotético de imputação	61
Tabela 9 - Casos imputados no Censo Escolar 2007-2015.....	61
Tabela 10 - Banco original com as etapas do aluno em cada ano	62
Tabela 11 – Banco imputado com as etapas do aluno em cada ano	62
Tabela 12 - Banco original com o fluxo do aluno em cada ano	62
Tabela 13 - Banco imputado com o fluxo do aluno em cada ano.....	62
Tabela 14 - Perfil socioeconômico dos casos imputados	64
Tabela 15 - Dependência administrativa das escolas dos casos imputados.....	65
Tabela 16 - Taxa de indivíduos em idade escolar sem frequentar a escola - ensino fundamental I - Belo Horizonte	67
Tabela 17 - Exemplo de atribuição de municípios	68
Tabela 18 - Taxa de flutuação entre escolas em Belo Horizonte em 2014	70
Tabela 19 - Anos escolares frequentando a mesma escola x Porcentagem de alunos	86
Tabela 20 - Perfil socioeconômico da coorte.....	88
Tabela 21 - Resultados da regressão logística multinomial - Modelo de desigualdades de origem sobre as trajetórias educacionais completas (2009-2020)	93
Tabela 22 - Anos frequentando a mesma escola (% de alunos).....	101
Tabela 23 - Estatísticas descritivas	106
Tabela 24 – Efeito das características socioeconômicas sobre as trajetórias escolares no ensino fundamental I.....	108
Tabela 25 – Correlação entre resíduos do modelo básico e os efeito das características das escolas.....	109
Tabela 26 – Efeito das características socioeconômicas sobre as trajetórias escolares no Ensino fundamental II.....	109
Tabela 27 – Correlação entre resíduos do modelo básico e os efeito das características das escolas.....	110
Tabela 28 – Efeito das características socioeconômicas sobre as trajetórias escolares no Ensino médio	110
Tabela 29 – Correlação entre resíduos do modelo básico e os efeito das características das escolas.....	111
Tabela 30 – Modelos dos efeitos das características das escolas sobre as trajetórias escolares	111
Tabela 31 - Regressão logística multinomial do modelo de desigualdades de origem sobre as trajetórias educacionais completas (2009-2020)	124

Tabela 32 - Modelo referente ao Gráfico 24 - Valores preditos marginais da trajetória regular (95% de intervalo de confiança).....	126
Tabela 33 - Regressão linear e correlações referentes ao modelo básico - Ensino fundamental I	128
Tabela 34 - Regressão linear e correlações referentes ao modelo básico - Ensino fundamental I I.....	131
Tabela 35 - Regressão linear e correlações referentes ao modelo básico - Ensino Médio	133
Tabela 36 - Regressão linear e correlações referentes ao modelo para características da escola - Ensino fundamental I.....	136
Tabela 37 - Regressão linear e correlações referentes ao modelo para características da escola - Ensino fundamental I I	137
Tabela 38 - Regressão linear e correlações referentes ao modelo para características da escola - Ensino Médio.....	137

LISTA DE FIGURAS

Figura 1 - Variáveis do Modelo Causal Principal.....	24
Figura 2 – Distribuição do indicador de adequação da formação docente das escolas frequentadas pela coorte por etapas da educação básica nos anos disponíveis.....	103
Figura 3 - Distribuição do indicador de regularidade do docente da Educação Básica das escolas frequentadas pela coorte por etapas da educação básica nos anos disponíveis.....	104
Figura 4 - Distribuição do indicador de complexidade da gestão das escolas frequentadas pela coorte por etapas da educação básica nos anos disponíveis.	105

LISTA DE QUADROS

Quadro 1 - Variáveis independentes do modelo logístico multinomial.....	92
---	----

LISTA DE ABREVIATURAS E SIGLAS

BH – Belo Horizonte

Cedeplar – Centro de Desenvolvimento e Planejamento Regional (UFMG)

EJA – Educação de Jovens e Adultos

EJA – Educação de Jovens e Adultos

IBGE – Instituto Brasileiro de Geografia e Estatística

Ideb – Índice de Desenvolvimento da Educação Básica

Inep – Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira

LDB – Lei de Diretrizes e Bases da Educação Nacional

MEC - Ministério da Educação

MG – Minas Gerais

NSE – Nível socioeconômico

Nupede – Núcleo de Pesquisa em Desigualdades Escolares

PISA – Programa Internacional de Avaliação de Alunos

Pisa – Programa Internacional de Avaliação de Estudantes

Pnad – Pesquisa Nacional por Amostra de Domicílios

Pnad-c – Pesquisa Nacional por Amostra de Domicílios Contínua

PNE – Plano Nacional de Educação

Saeb – Sistema Nacional de Avaliação da Educação Básica

Unesco – Organização das Nações Unidas para a Educação, a Ciência e a Cultura

SUMÁRIO

INTRODUÇÃO	18
Contribuições da tese.....	19
Premissas, hipótese, pergunta e modelo causal	20
a) Premissa 1 – As regularidades sociológicas	20
b) Premissa 2 – Melhor Misturar	21
c) Premissa 3 – Linha tênue.....	22
Hipótese.....	23
Pergunta da pesquisa.....	23
Percurso do trabalho	24
CAPÍTULO I - TRAJETÓRIAS ESCOLARES COMO VARIÁVEL RESPOSTA	26
1. Os dados	26
1.1. A importância do Censo Escolar como base de dados públicas.....	27
1.2. A Lei nº 12.527 e os dados fora do ar	28
1.3. O Censo Escolar protegido.....	30
1.3.1. O Serviço de Acesso a Dados Protegidos (Sedap).....	30
1.3.2. Banco de fluxo	31
1.4. Convênio	32
1.5. Núcleos.....	32
1.6. Proteger x restringir: a agenda de pesquisa em risco	33
1.7. O desenho de pesquisa	33
2. A construção das bases longitudinais	35
2.1. Base Longitudinal 1 - a partir dos dados públicos do Censo Escolar	35
2.1.1. Os filtros	35
2.1.1.1. Por idade e/ou por etapa?.....	35
2.1.1.2. Definições da coorte de data de nascimento	36
2.1.2. Deduplicação	37
2.2. O primeiro mapeamento.....	39
2.3. Base Longitudinal 2 - a partir dos dados seguros.....	42
2.4. Confirmando a magnitude do problema.....	44
3. Não deixar ninguém para trás: uma proposta	46
3.1. Trajetórias escolares na educação básica como variável resposta	48
4. RESULTADOS.....	49
4.1. Análises exploratórias	49
5. Considerações.....	57
CAPÍTULO II – INDIVÍDUOS EM IDADE ESCOLAR QUE NÃO FREQUENTARAM A ESCOLA	59
1. Um debate sobre imputação	59
1.1. Replicação	59
1.2. As imputações possuem um perfil específico?	64
2. Estar na escola como dimensão da qualidade educacional.....	65
2.1. Proposta.....	66
3. Considerações.....	70
CAPÍTULO III – REGULARIDADES SOCIOLÓGICAS, EFEITOS DESPROPORCIONAIS E AS TRAJETÓRIAS ESCOLARES	71
1. Acesso, reprovação, abandono, evasão e seus marcos legais.....	72
1.1. Anos 80, redemocratização e a democratização da escola.....	72
1.2. Anos 90 e o Consenso	72

1.2.1.	O Plano Decenal de Educação para Todos (1993-2003).....	72
1.2.2.	LDB	73
1.3.	Virada do milênio e algumas viradas no Brasil.....	76
1.3.1.	Planos, Pactos e a emenda 59	76
1.3.2.	O PNE 2001-2010.....	76
1.3.3.	O PDE, o PAR e o TODOS.....	77
1.3.4.	A emenda nº 59 de 2009	78
1.3.5.	O PNE 2014-2024.....	78
2.	Acesso, fluxo e desigualdades sociais	80
2.1.	Acesso e desigualdades de origem	80
2.2.	Reprovação, repetência e desigualdades de origem	81
2.3.	Abandono, evasão e desigualdades de origem	83
2.4.	Efeito de pares	83
2.5.	Ainda a pedagogia da repetência.....	84
3.	Trajatórias escolares e desigualdades de origem – o modelo	85
3.1.	Modelo de desigualdades de origem	85
3.2.	As variáveis socioeconômicas e demográficas	85
3.2.1.	Trajatória escolares em sistema aberto e as variações socioeconômicas e de escolas	86
3.2.2.	Regularidades sociológicas.....	88
4.	Efeitos desproporcionais.....	90
4.1.	Recategorizações.....	91
5.	Resultados.....	92
5.1.	O sexo, a cor, a escola e o lugar dos escolarizados	93
6.	Considerações.....	98
CAPÍTULO IV - LINHAS TÊNUES E AQUELA CHANCE		99
1.	Modelos	100
1.1.	Escolha do modelo	107
2.	RESULTADOS.....	108
2.1.	O efeito-professor.....	112
3.	Considerações.....	113
4.	Considerações finais	113
REFERÊNCIAS		115
APÊNDICE A.....		123
APÊNDICE B.....		124
ANEXO A		138

INTRODUÇÃO

No período entre 1º de abril de 2002 e 31 de março de 2003, nasceram pouco mais de 3 milhões¹ de crianças no Brasil (DATASUS). Em 2009, observou-se que 98,6% das crianças dessa coorte²³ frequentaram a escola e 84% delas foram promovidas⁴ de 2009 para 2010. Em 2020, ano em que deveria ocorrer o desfecho para os estudantes dessa coorte, apenas 26% haviam concluído o ensino médio. Em contraste, o Cadastro Geral de Empregados e Desempregados (CAGED, 2019) registrou, em 2022, que 77,5% das vagas de emprego exigiam ensino médio ou superior completo. Esses números mostram um quadro no qual não há uma correspondência entre exigências do mercado de trabalho e formação, onde apenas 50% da população⁵ brasileira adulta possui o nível de escolaridade exigido por 75% das oportunidades de trabalho, enquanto os 50% menos escolarizados competem por 25% das vagas disponíveis.

Os retornos econômicos, a redução da criminalidade e da mortalidade, a melhoria na saúde da população e o fortalecimento da democracia trazidos por trajetórias escolares longas e completas são conhecidos da literatura (Heckman, Lochner e Todd, 2008; Xavier e Neves, 2012; Lochner, 2011; Melo e Machado, 2018). Assim, o impacto da escolaridade sobre a qualidade da vida das pessoas, a depender de quão longe foram suas trajetórias escolares é a justificativa dessa tese sobre as trajetórias educacionais na educação básica, no Brasil. E sobre as desigualdades que permeiam essas trajetórias, as quais possuem taxas das irregularidades altas e concentradas em alguns grupos sociais.

As trajetórias educacionais, em uma definição ampla, são os caminhos percorridos pelos estudantes ao longo do sistema de ensino (Nogueira e Fortes, 2004). Possuem trajetórias regulares os estudantes que percorrem com sucesso todas as etapas da educação básica. E

¹ Referência a coorte de crianças nascidas entre 1º de abril de 2002 a 31 de março de 2003. O número exato de nascido é 3.059.402.

² Coorte é um conceito demográfico que define um conjunto de pessoas que experimentam um evento em comum em um mesmo período no tempo. Uma coorte pode ser de pessoas que nasceram em 1960, de pessoas que ingressaram na primeira série do primeiro grau em 1970, de mulheres casadas em 1974 e etc. (CARVALHO et. al.,1998).

³ Em 2009 a coorte tinha entre seis e sete anos de idade.

⁴ Aluno promovido é aquele matriculado em determinado ano escolar na etapa seguinte em relação a etapa em que ele estava matriculado no ano escolar anterior.

⁵ Dados referentes a população brasileira com 25 anos ou mais de idade.

possuem trajetórias irregulares os estudantes que experimentam uma ou mais intercorrências de repetência⁶, abandono⁷ e/ou evasão⁸ (Soares, Alves e Fonseca, 2021).

Contribuições da tese

Através de uma investigação empírica essa tese elaborou duas contribuições principais. A primeira é um desenho do cenário das trajetórias escolares brasileiras na última década. Essa *fotografia* permite acompanhar e entender ao longo dos anos o funil que é o sistema educacional brasileiro. A segunda está nessas trajetórias colocadas sob três premissas teóricas através de modelos analíticos. O objetivo desses modelos foi mapear o objeto de pesquisa para nortear gestores na elaboração de políticas públicas que sejam capazes de aumentar as trajetórias regulares com equidade educacional. Essa nossa proposta se entrelaça com desafios transnacionais com os quais o Brasil se propôs a contribuir no combate. Explicamos o porquê.

A Agenda 2030 das Nações Unidas – aprovada em 2015 por representantes dos Estados-membros da ONU, da qual o Brasil é signatário – comprometeu-se, por meio desse marco global que **ninguém será deixado para trás**, entre 2015 e 2030. A Agenda 2030 é pautada em escolhas necessárias para a melhoria da vida das pessoas agora e no futuro e estabeleceu 17 Objetivos de Desenvolvimento Sustentável para erradicar a pobreza e promover uma vida digna para todos (Unesco, 2016). A educação está contemplada no objetivo 4, que tem como mensagem principal que todo aluno é importante e tem igual significado (Alves, Oliveira, *et al.*, 2021). O objetivo 4 também espera “garantir uma educação de qualidade inclusiva e equitativa e promover oportunidades de aprendizagem ao longo da vida para todos”. E, até 2030, espera-se que “todas as meninas e meninos concluam a educação primária e secundária gratuita, equitativa e de qualidade, levando a resultados de aprendizagem relevantes e eficazes” (Unesco, 2016).

Para uma educação com mais equidade é fundamental que se conheça a população mais excluída (Unesco, 2014). Diante dessa jornada coletiva rumo a um 2030 sustentável, esta tese que investiga a trajetória educacional de estudantes brasileiros com foco nos marginalizados se alinha com a agenda 2030 e se dedica a produção de conhecimento sobre aqueles que foram

⁶ Os repetentes são os alunos matriculados na mesma etapa em relação ao ano anterior por reprovação ou por terem abandonado a escola

⁷ Abandonam a escola os alunos que deixam de frequentá-la durante determinado ano escolar.

⁸ São considerados evadidos os alunos que frequentam a escola em determinado ano escolar e não se matriculam no ano seguinte.

deixados para trás. Além disso, esse trabalho firma um debate com o enfrentamento das consequências educacionais da pandemia da Covid-19⁹, com o PNE 2020-2024 e com a importância das bases de dados públicas e abertas.

Premissas, hipótese, pergunta e modelo causal

Na década de 60, a sociologia da educação apontou o efeito das desigualdades socioeconômicas sobre as desigualdades escolares; nos anos 80 e 90 apontou o impacto da escola sobre os alunos; nos anos 2000 percebeu que sistemas escolares e escolas reproduzem desigualdades em diferentes níveis e questionou se haveria escolas e sistemas escolares mais justos e outros mais injustos. Houve nesse campo de estudo um adensamento dessas discussões e o surgimento de um emaranhado de variáveis sobre as quais os pesquisadores da área têm trabalhado em busca de explicação. Ancoradas nesse adensamento teórico partimos de três premissas fundamentais e construímos a hipótese, a pergunta de pesquisa e o modelo causal que estruturam essa tese.

a) Premissa 1 – As regularidades sociológicas

Não há discussão sociológica sobre educação capaz de se esquivar ao debate “a escola reproduz desigualdades sociais” (Bourdieu e Passeron, 1975). Essa perspectiva, iniciada na década de 60 do século XX, ficou conhecida na literatura como Teoria da Reprodução. Clássicos como “*A escola conservadora*”, “*Os três estados do capital cultural*” e “*A reprodução*” quebraram o modelo otimista e meritocrático dos funcionalistas liberais, que tinham a escola como fator importante para o desenvolvimento econômico e como instituição capaz de promover igualdade de oportunidades e mobilidade social (Parsons, 1968; Durkheim, 1973; Gomes, 1994; Petitat, 1994). Com essa quebra paradigmática estava posto o indiscutível achado da Sociologia da Educação – há efeito das desigualdades socioeconômicas sobre as desigualdades escolares. Aqui, referimos a essa tendência em que variáveis de origem social, como o nível socioeconômico e a raça, estão associados às desvantagens escolares dos indivíduos socialmente mais vulneráveis (Coleman e others, 1966; Bourdieu e Passeron, 1975; Patto, 1990; Alves, Soares e Xavier, 2016), como regularidades sociológicas.

⁹ Este trabalho foi iniciado antes da pandemia de Covid-19 e vislumbrava inicialmente trabalhar com dados de 2007 a 2017. Os anos do painel foram estendidos até 2020 em busca tentar lançar luz sobre as desigualdades escolares brasileiras diante da crise sanitária mundial.

b) Premissa 2 – Melhor Misturar

“*Todas as crianças podem aprender e suas oportunidades não são definidas por quem está sentado na mesa ao lado*”. Essa afirmação foi feita pelo procurador-geral adjunto do estado de Nova York ao responder um processo de 1998, sobre os nocivos efeitos de escolas de alta pobreza. Esse não era um debate novo para o governo americano que, em 1964, havia encomendado um relatório intitulado *Igualdade de oportunidades educacionais* a um grupo de pesquisadores. O documento, mais conhecido como Relatório Coleman, tinha mais de 700 páginas e foi criticado, posteriormente, por suas fragilidades metodológicas. No entanto, ele apresentou diversos resultados importantes, incluindo a constatação de que os estudantes negros seriam mais sensíveis ao efeito da escola em relação aos estudantes brancos, além de verificar que haveria um duplo desfavorecimento dos alunos negros que estudavam em escolas com elevadas taxas de alunos também negros (Rumberger e Palardy, 2005). Esses argumentos embasariam a famosa política norte-americana de dessegregação escolar que espalhou *yellow school buses* por todo o país (Jacobs, 1998). Apesar das polêmicas em volta de sua publicação, o Relatório Coleman inaugurou as discussões sobre os efeitos de pares¹⁰ e é um marco na literatura da Sociologia da Educação.

Era o começo de uma série de debates sobre a hipótese de que o desempenho dos alunos estaria mais relacionado ao efeito dos pares, ou seja, às características socioeconômicas de seus colegas de escola, do que a qualquer outra característica das próprias escolas. Para a teoria do efeito dos pares, os resultados educacionais dos alunos de uma escola estão associados com a média das origens socioeconômicas – em nível agregado – de todos os alunos que também frequentam essa escola. O central à teoria é que o baixo nível socioeconômico (NSE) de uma escola se torna mais uma variável a contribuir para que os alunos tenham desempenhos menores, se comparados com os resultados de alunos de origem socioeconômica semelhante que frequentam escolas de NSE mais alto - mesmo depois desses resultados serem controlados pelas características individuais dos estudantes.

O efeito de pares desfavorece os alunos pobres porque eles ficam duplamente em desvantagem, devido a soma dos efeitos da pobreza individual com os efeitos da pobreza escolar. Esse contexto promove um ambiente de privação tanto de infraestrutura e qualidade do corpo docente quanto de motivação e aspirações, ambos prejudiciais ao desempenho acadêmico

¹⁰ O conceito também foi denominado como composição social e como segregação social.

(Rumberger e Palardy, 2005). Enquanto isso, alunos com alto desempenho e motivação podem gerar um efeito de pares capaz de cultivar o sucesso escolar e ampliar as condições propícias para esse sucesso, condições essas que já são altas.

O Relatório Coleman também indicou limitações na capacidade do sistema educacional norte-americano em promover igualdade escolar. A diferença encontrada entre as escolas americanas havia sido pequena e isso apontava para a falta de relação entre o tipo e/ou a localização das escolas e o aprendizado escolar. Esse argumento deu consistência para a conclusão: *a escola não faz diferença* (Bressoux, 2003; Brooke e Soares, 2008).

c) Premissa 3 – Linha tênue¹¹

A reação ao argumento *a escola não faz diferença* ficou conhecida como **efeito-escola**. Conceito que também foi denominado pela literatura por **eficácia escolar**. Há uma variação entre esses dois conceitos: efeito escola se refere a quanto uma escola acrescenta no aprendizado dos alunos através de políticas e práticas internas e eficácia escolar sugere que existem escolas melhores do que outras (Brooke e Soares, 2008). Este trabalho tratará de efeito escola.

Diferentes pesquisas se empenharam em mostrar que traços sociais – individuais ou do grupo - não eram os únicos fatores capazes de gerar desvantagens educacionais. Discussões importantes para as escolas e sobre os impactos delas sobre os alunos tanto através de políticas públicas quanto das práticas escolares se abriram a partir do fim da década de 1970 (Rutter, Maughan, *et al.*, 2008; Madaus, Airasian e Kellaghan, 2008; Mortimore, Sammons, *et al.*, 1988). O momento iniciou um movimento disposto a abrir *a caixa preta da escola*. O objetivo era identificar escolas que por alguma diferença produzissem um certo efeito em alunos de baixo nível socioeconômico e gerassem resultados escolares diferente das expectativas. Essas diferenças podem estar no efeito-professor, no funcionamento e organização da escola, no papel do diretor, na distribuição dos alunos em turmas, em avaliações do aprendizado dos alunos, no tamanho da escola e das turmas, nas despesas das escolas e entre muitos outros fatores¹² (Bressoux, 2003).

¹¹ Aprendi a usar *linha tênue* para referir ao efeito escola no curso *Leituras sobre pesquisa em eficácia escolar/efeito escola: o debate atual* com a Profa. Maria Teresa Alves. A expressão intitula a seção pois nos parece a melhor definição do papel da escola na vida dos brasileiros – principalmente os pobres.

¹² O detalhamento dessa definição conceitual se encontra no capítulo 4.

O efeito das escolas brasileiras varia em média entre 14% e 22%, dependendo da série e da disciplina (Alves e Soares, 2007 *apud* Soares, Alves e Fonseca, 2021). Não é um efeito poderoso, porém é uma *linha tênue* com capacidade de separar os alunos de origens sociais desvantajosas entre aqueles com maiores e aqueles com menores chances de sucesso escolar.

Hipótese

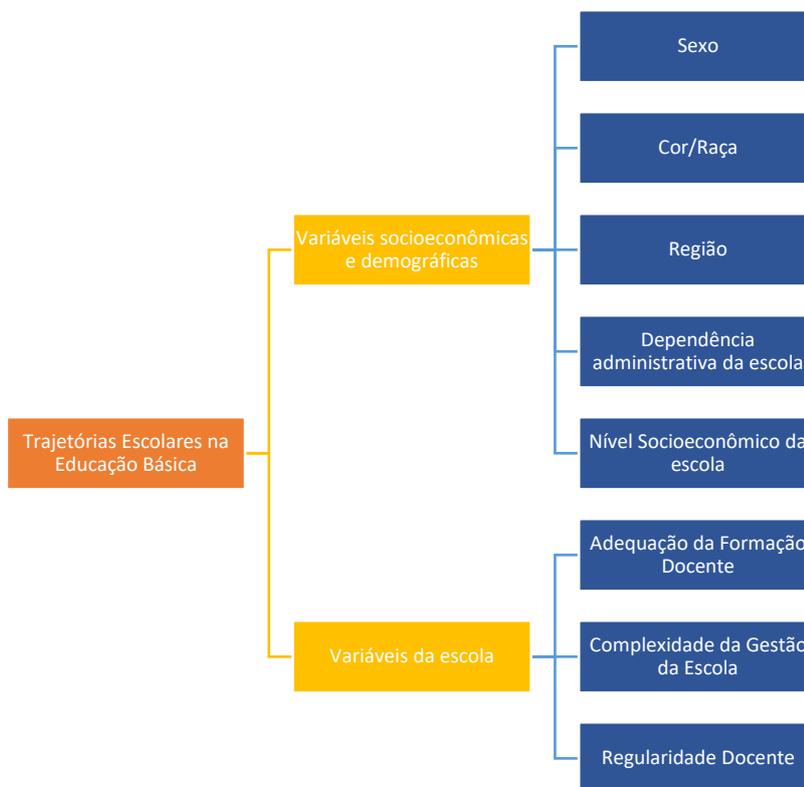
A desigualdade social brasileira concentra privilégios em determinados grupos sociais e faz com que parte dos estudantes estejam em desvantagens já no início de suas trajetórias escolares. Ao mesmo tempo, o sistema educacional oferece oportunidades escolares desiguais e é marcado por altas taxas de reprovação e evasão que são capazes de produzir efeitos cumulativos e negativos no curso de vida dos alunos (Barros e Mendonça, 1998). Os resultados nos arremessam para a máxima das regularidades sociológicas e dos efeitos de pares. Entretanto, se escolas podem ter efeitos (Lee, 2010) sobre as chances de sucesso escolar dos alunos, a hipótese deste trabalho é:

As trajetórias escolares estão associadas aos fatores intraescolares, independentemente das características individuais dos estudantes e do efeito dos pares.

Pergunta da pesquisa

Diante dessa hipótese, a pergunta que esse trabalho se propõe a responder é: as características das escolas produzem efeitos capazes de reduzir as chances de intercorrências nas trajetórias escolares, quando controladas as variáveis individuais dos alunos? A Figura 1 - Variáveis do Modelo Causal Principal apresenta as variáveis com as quais construímos o modelo analítico da tese.

Figura 1 - Variáveis do Modelo Causal Principal



Percurso do trabalho

Essa tese est organizada em 4 captulos. O captulo 1 faz uma proposta de varivel para as trajetrias escolares na educao bsica brasileira. O captulo 2 se dedica ao debate sobre a imputao de casos no Censo Escolar e aos impactos que essa metodologia pode ter sobre as taxas de indivduos que deixaram de frequentar a escola. O captulo 3 apresenta, analisa e discute um modelo estatstico que mensura os efeitos das origens sociais e dos efeitos de pares sobre as trajetrias escolares. O captulo 4 soma ao modelo anterior o efeito das escolas. Cada captulo tem estrutura prpria e pode ser lido separadamente. Entretanto, o desenho de pesquisa da tese usa diferentes bases de dados e essas diferenas ajudaram a construir os argumentos que apresentamos. Por isso, h tambm um argumento geral que s  compreendido na leitura do trabalho na ntegra.

Esse  um trabalho escrito na perspectiva extensionista da universidade. Ele traz, sim, modelagens estatsticas complexas e cdigos de programao, que daro ao leitor mais ntimo dessas tcnicas informaoes que podem no ser apreendidas por aqueles mais afastados desse

universo. Entretanto, essa tese foi pensada e desenhada para que sua parte metodológica não comprometa a possibilidade de ser acessada por toda a comunidade acadêmica e escolar.

Boa leitura!

CAPÍTULO I - TRAJETÓRIAS ESCOLARES COMO VARIÁVEL RESPOSTA

Esta tese propõe responder se características das escolas estão relacionadas com as trajetórias escolares na educação básica. Consideraremos como trajetórias escolares na educação básica as etapas ensino fundamental I, ensino fundamental II e ensino médio. Para responder a qualquer pergunta sobre trajetórias escolares em uma escala mais macro, é preciso ter uma variável que possa mensurá-las. Este é o objetivo deste capítulo. Apresentar uma proposta de mensuração para o conceito de trajetórias escolares na educação básica brasileira. Esse exercício é um desafio mão dupla, uma vez que ao mesmo tempo em que faz uma proposição de mensuração, ajuda a definir conceitualmente o que podemos entender como trajetórias escolares.

Para apresentar a proposta que construímos o capítulo está dividido em três seções. A primeira - Os dados - guia o leitor sobre a estratégia metodológica que traçamos para que essa pesquisa fosse possível e documenta sobre o acesso a importantes bases de dados educacionais no Brasil. A segunda seção - A construção das bases longitudinais - descreve como a partir dos microdados do Censo Escolar construímos três bases de dados longitudinais. A seção III - Não deixar ninguém para trás: uma proposta - apresenta nossa variável de trajetórias escolares e a metodologia utilizada para mensurar e operacionalizar a proposta.

1. Os dados

As bases de dados utilizadas neste trabalho são painéis longitudinais construídos através das bases do Censo Escolar. O Censo Escolar é o principal instrumento de coleta de informações da educação básica e a mais importante pesquisa estatística educacional brasileira (Inep, 2023). Essa coleta é realizada com edições anuais por uma parceria entre escolas, secretarias de educação municipais, dos estados e Distrito Federal e o Inep¹³. Uma característica importante dessa coleta é ela ser realizada em duas etapas nas quais um questionário padronizado é respondido por todas as escolas, das redes pública e privada. Esse levantamento é de abrangência nacional e contempla todas as etapas da educação básica compreendendo ensino regular¹⁴, Educação Especial, Educação de Jovens e Adultos (EJA) e Educação Profissional.

¹³ O Inep é um órgão federal vinculado ao Ministério da Educação do governo brasileiro.

¹⁴ O ensino regular brasileiro é composto por educação infantil, ensino fundamental e médio.

Os itens do Censo Escolar são preenchidos por um responsável indicado pelo estabelecimento de ensino com base nas informações escolares dos alunos e nas informações socioeconômicas fornecidas na ficha de matrícula dos alunos por suas famílias. Esse preenchimento dos dados é feito diretamente na Internet, por meio do sistema *Educacenso*. As secretarias municipais e estaduais têm por funções apoiar e fiscalizar o Censo Escolar enquanto o Inep coordena e recebe os dados coletados.

A primeira etapa do Censo Escolar é referente as matrícula iniciais e coleta dados das escolas, gestores, turmas, alunos e profissionais escolares em sala de aula. Essa coleta é feita com base na realidade das escolas na data de referência do Censo Escolar, última quarta-feira do mês de maio, instituída como o Dia Nacional do Censo Escolar na Portaria MEC nº 264/07. A segunda etapa é relativa ao final do ano letivo anterior e acontece entre fevereiro e março coletando informações de rendimento e movimento dos alunos que foram declarados na matrícula inicial. Para fornecer informações sobre rendimento do aluno a escola declara a condição de aprovado ou reprovado, transferido, deixou de frequentar ou falecido (Inep, 2023).

1.1. A importância do Censo Escolar como base de dados públicas

A cada ano o Censo Escolar coleta informações de em média 50 milhões de alunos, distribuídos em mais de 170 mil escolas que estão localizadas em 5.568 municípios. A realização desse censo tem o objetivo claro de subsidiar as políticas educacionais por meio do acompanhamento sistemático do desenvolvimento e evolução do sistema de ensino.

Apesar de seu uso ser prioritariamente administrativo, ou seja, ter como função subsidiar a administração pública e os governos em tomadas de decisão, o Censo Escolar foi ganhando a validade necessária para ser usado em pesquisas (Soares, Alves e Fonseca, 2021). Com informações antes desagregadas por alunos e atualmente por escolas¹⁵ e divulgadas em formato de microdados que estão disponíveis para *download online*¹⁶, tornou-se uma importante fonte de dados secundários com a qual pesquisadores contam para suas pesquisas educacionais.

Em 2017, o Censo Escolar ganhou mais um fator de relevância. O Inep tornou público em seu site, para as edições do Censo Escolar a partir de 2007, um código de identificação único

¹⁵ Discutiremos essa mudança na seção 1.2 desse capítulo.

¹⁶ Disponível em <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/censo-escolar>

dos estudantes matriculados na educação básica, os IDs. Utilizado no Censo Escolar anualmente no preenchimento das informações de cada estudante, o número dos IDs se mantém e acompanha cada aluno durante toda a educação básica, mesmo que ocorram mudanças de cidades ou escolas. Com a ligação desses IDs, entre edições do Censo Escolar, tornou-se possível construir bases longitudinais para analisar as trajetórias escolares dos alunos ao longo dos anos. Esse *link* entre os IDs, além de conectar a situação dos alunos em cada ano permite a conexão entre alunos, turmas, professores e escolas. Algumas metodologias para ligar os IDs foram traçadas nos últimos anos (Oliveira e Menezes Filho, 2018; Paula e Alves, 2019; Rangel Rigotti, Signorini e Hadad, 2020; Lima e Paes de Carvalho, 2021; Soares, Alves e Fonseca, 2021; Ferrão, 2022; Petrus, 2022; Malaguth, 2022; Fonseca, Rodrigues, *et al.*, 2022; Lunelli e Bruel, 2022). Todas elas utilizam o Censo Escolar.

Essa estrutura que o Censo Escolar adquiriu é de enorme importância para o Brasil, tanto para a elaboração de suas políticas públicas educacionais tais como o Fundeb, como para a possibilidade de ampliação da agenda de pesquisa sobre trajetórias educacionais. Não encontramos na literatura internacional¹⁷ referências a uma base de dados educacional censitária e em nível nacional como é o Censo da Educação que permita transformar dados administrativos em um painel longitudinal¹⁸.

1.2. A Lei nº 12.527 e os dados fora do ar

Em decorrência da Lei nº 12.527, de 18 de novembro de 2011, conhecida como *lei de acesso à informação*, a partir da edição de 2018 os IDs das bases do Censo Escolar disponibilizadas *online* foram criptografados e a conexão entre as edições de 2018, 2019 e 2020 com os anos anteriores ficou inviabilizada. Posteriormente, os IDs das edições de 2007 a 2017 também foram modificados e todas as ligações impossibilitadas.

Os microdados do Exame Nacional do Ensino Médio (Enem) 2020 e do Censo Escolar da Educação Básica 2021 estão disponíveis, desde esta sexta-feira, 18 de fevereiro, no portal do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). O formato de apresentação do conteúdo dos arquivos,

¹⁷ Na literatura há referências sobre bases de níveis municipais. Ver Roderick, 1994.

¹⁸ Bases educacionais longitudinais são construídas através de dados coletados ao longo do tempo com os mesmos alunos (Lee, 2010).

que reúnem um conjunto de informações detalhadas relacionadas à pesquisa estatística e ao exame, foram reestruturados para suprimir a possibilidade de identificação de pessoas, em atendimento às normas previstas na Lei n.º 13.709, de 14 de agosto de 2018 (Lei Geral de Proteção de Dados Pessoais – LGPD) (Inep, 2022).

Essa foi a comunicação oficial do Inep em 18 de fevereiro de 2022, data de retirada dos microdados de todas as edições do Censo Escolar do domínio *online* do instituto. As bases com microdados no nível dos alunos foram substituídas por bases com informações desagregadas no nível máximo das escolas. Essa reformulação foi amparada pelo Termo de Execução Descentralizada entre Inep e Universidade Federal de Minas Gerais, o TED 8750 – PRICE, intitulado *Privacidade nos Censos Educacionais*. Esse TED comparou as medidas de publicações dos Censos Escolares feitas pelo Inep com medidas de publicações internacionais. Os resultados encontrados foram menor rigidez e maior detalhadamente do Inep nas publicações e disponibilização dos dados quando comparados com as medidas internacionais as quais não publicam microdados individuais de estudantes, publicam amostras da população e disponibilizam permissão completa de acesso aos microdados sob autorização prévia e em salas seguras.

Para o TED, *desidentificação* (remoção de identificadores individuais óbvios tais como nome, CPF e RG) e *pseudonimização* (substituição de *IDs* óbvios por *IDs* artificiais) que são as técnicas de proteção de privacidade utilizadas pelo Inep na divulgação dos microdados são sujeitas a riscos de privacidade já identificados na literatura. Como exemplo, o estudo destacou o caso de *reidentificação* da AOL. Em contexto similar ao do Inep e com as mesmas técnicas de desidentificação e pseudonimização a *AOL Research* disponibilizou um arquivo de texto que continha milhões de palavras-chave usadas por 650.000 usuários.

Os dados divulgados mostravam que o indivíduo #4.417.749 havia realizado pesquisas no motor de busca por expressões como “dedos dormentes”, “homens solteiros com mais de 60 anos”, “cão que urina em tudo”, e “paisagistas em Lilburn, GA”. Cruzando os dados dessas pesquisas com informações de listas telefônicas, o jornal *The New York Times* foi capaz de identificar unicamente o indivíduo #4.417.749 como Thelma Arnold, uma viúva de 62 anos de Lilburn, estado da Geórgia, EUA (Alvim, Gonze e Graaf, 2020).

Os resultados do TED apontam que, se atacada, uma única base de dados com *quaseidentificadores* como data de nascimento ou escola em que o aluno está matriculado pode sofrer *reidentificação* de mais de 78% dos indivíduos. Se houver inferência de um atributo sensível como ser ou não portador de necessidades especiais ou deficiente a *reidentificação* pode ser em mais de 90% dos indivíduos. Diante desses resultados, o TED entendeu que a metodologia de divulgação dos Censos Educacionais feita pelo Inep oferecia riscos de violação de privacidade aos titulares dos dados e “poderia constituir em violação da Lei Geral de Proteção de Dados Pessoais.

1.3. O Censo Escolar protegido

A nova realidade restringiu o acesso a todos os censos educacionais. E outras possibilidades para as pesquisas sobre trajetórias escolares se tornaram necessárias.

1.3.1. O Serviço de Acesso a Dados Protegidos (Sedap)

O Serviço de Acesso a Dados Protegidos (Sedap) atende aos pesquisadores que requerem utilizar às bases de dados produzidas pelo Inep e protegidas pela lei de acesso à informação. As pesquisas são feitas na sala de acesso a dados protegidos, na sede do Inep, localizada nas instalações do Centro de Informação e Biblioteca em Educação (Cibec), Direção, SIG Quadra 04, Lote 327 – Edifício Villa Lobos – Térreo, Ala A – CEP: 70.610-908, Brasília – Distrito Federal.

Os microcomputadores não possuem acesso à internet ou à rede interna do Inep. Estão bloqueados para transferência de arquivos e suas entradas USB estão desabilitadas. As estações de trabalho possuem os seguintes softwares: SAS, STATA, R, SPSS e Pacote Office. [...] Não serão disponibilizados outros softwares além dos já mencionados, ou seja, o pesquisador não poderá instalar ou pedir para que sejam instalados outros softwares no ambiente seguro (Inep, 2021).

Realizar pesquisas na sala de acesso a dados protegidos é possível, como demonstra parte desse trabalho. Entretanto, há problemas estruturais que comprometem a qualidade das pesquisas e precisam ser discutidos. Três deles são: as desigualdades socioeconômicas e de gênero entre os pesquisadores, a atualização dos programas estatísticos dos computadores e a formação de jovens pesquisadores.

realizados nos dados para a construção desses indicadores e da base de dados *Banco de Fluxo* na qual os indicadores foram alocados. Apresenta também o processo de deduplicação e imputação de registros, as definições e a metodologia de cálculo dos indicadores de fluxo escolar. A nota destaca que os microdados públicos do Censo Escolar, os disponíveis *online*, não foram influenciados por esse processo que contemplou apenas as bases internas do Inep.

O banco de fluxo e seu dicionário não foram disponibilizados publicamente e há poucas informações sobre essa base disponíveis para a comunidade científica. Ao mesmo tempo, a variável *situação de rendimento ou movimento do aluno ao final do ano letivo* (TP_SITUAÇÃO) que tem como categorias *aprovado, reprovado, abandono, falecido* ou *sem informação de rendimento* e a variável *situação de conclusão* (IN_CONCLUINTE) que tem como categorias *não* e *sim* não foram incluídas nas bases públicas do Censo Escolar. Embora não esteja listado na relação de bases de dados disponíveis²¹, o banco de fluxo está disponível no Serviço de Acesso a Dados Protegidos (Sedap).

1.4. Convênio

Em junho de 2021, um acordo de cooperação técnica com vigência de 5 (cinco) anos foi realizado entre Faculdade de Ciências Econômicas (Face), da Universidade Federal de Minas Gerais (UFMG) e o Inep. Por esse acordo, os dados do banco de fluxo foram cedidos pelo Inep ao Cedeplar – Centro de Desenvolvimento e Planejamento Regional para fins de pesquisa²². Essa parceria permitiu que pesquisadores acessassem o banco de dados de fluxo escolar. Na publicação do diário oficial que se refere ao acordo entre UFMG e Inep ou no domínio do Instituto não há informações sobre a viabilidade desses convênios e a Lei Geral de Proteção de Dados Pessoais.

1.5. Núcleos

A portaria n.º 105/2022 criou os núcleos Sedap. Esses núcleos poderão ser implementados nas universidades federais, institutos federais e centros federais de educação

²¹<https://www.gov.br/inep/pt-br/areas-de-atuacao/gestao-do-conhecimento-e-estudos-educacionais/cibec/servico-de-acesso-a-dados-protegidos-sedap/base-de-dados>

²² Esses dados foram cedidos pelo Inep por meio de acordo de cooperação ao Cedeplar – Centro de Desenvolvimento e Planejamento Regional da Faculdade de Ciências Econômicas da Universidade Federal de Minas Gerais UFMG para fins de pesquisa.

tecnológica (CEFETs) através de cooperação técnica com o Inep²³. Esse novo formato permitirá que pesquisadores acessem os dados protegidos do Inep, via convênio, no âmbito das instituições conveniadas. Esses núcleos deverão possuir e manter uma sala segura em ambiente físico na instituição de ensino, com acesso controlado e seguro para o acesso a dados protegidos de acordo com as definições do Inep. Como os núcleos são espelhos da sala segura a atualização dos programas estatísticos disponíveis e a formação de jovens pesquisadores permanecem como desafios.

1.6. Proteger x restringir: a agenda de pesquisa em risco

É urgente que Ministério da Educação, Inep, comunidade acadêmica e organizações da sociedade civil discutam sobre as consequências que a importação de modelos mais restritivos de acesso aos dados educacionais podem gerar sobre as políticas públicas no Brasil. É preciso entender a série histórica do Censo Escolar como uma política pública brasileira bem-sucedida com êxito para ser modelo internacional e reconhecer os relevantes trabalhos produzidos com essas bases enquanto eram públicas como cruciais para os avanços educacionais que o país alcançou nas últimas décadas. Essas produções e as agendas de pesquisa que requerem o Censo da Educação estão colocadas em risco com o atual desenho de acesso a esses dados.

1.7. O desenho de pesquisa

Esse trabalho é mais um na literatura que enfrentou as dificuldades em realizar uma pesquisa longitudinal e tentou lidar com as circunstâncias através de um esforço de criatividade (Valbuena e Gil, 2020), percorrendo diferentes caminhos para trabalhar com o Censo Escolar em formato longitudinal. Apresentamos nosso desenho de pesquisa:

Em 2019, o projeto inicial desta tese tinha a expectativa de acessar os dados protegidos de modo remoto, o que não aconteceu. Iniciamos 2020 com o planejamento de utilizar o Sedap, opção que logo se tornou muito restrita e em alguns momentos inviável devido a pandemia da Covid-19. Diante das incertezas que permearam o período, construímos uma base de dados longitudinal utilizando os dados públicos do Censo Escolar que eram disponíveis *online* e com os IDs passíveis de ligação, entre os anos das edições de 2007 a 2017.

²³ Foram realizadas alterações na portaria para a inclusão de instituições privadas e de outras naturezas.

Em 2020, a UNESCO divulgou que, devido à pandemia da Covid-19, 1,5 bilhões de crianças e jovens foram afetadas pelo fechamento de escolas em 192 países, sendo que os mais vulneráveis foram os mais impactados. O Brasil também fez parte dessas estatísticas e suas desigualdades sociais e tecnológicas, agravadas pela pandemia, aumentou ainda mais o risco de grupos desfavorecidos terem aprendizado insuficiente ou abandonarem a escola. Todos os cenários indicaram que o fechamento de escolas aumentaria as desigualdades (Alves, Oliveira, *et al.*, 2021).

Com o objetivo de contribuir para a compreensão das consequências da crise sanitária global e lançar luz sobre as consequências do ano de 2020 nas desigualdades escolares, este trabalho foi realizado, entre fevereiro e setembro de 2022, no Serviço de Acesso a Dados Protegidos (Sedap) que, até o momento, é a única fonte de microdados no nível dos alunos disponível para o ano de 2020. Para pesquisas com o Censo Escolar, o Sedap disponibiliza duas bases de dados. Uma é o espelho dos microdados que anteriormente eram públicos, com a diferença de que os *ID's* dos alunos não estão criptografados. A outra é a base de Fluxo Escolar e foi com essa que construímos a base utilizada neste estudo.

Em 2021, nossa pesquisa foi contemplada²⁴ com um estágio de 9 meses na Universidade da Califórnia – Riverside, EUA, que teve início em setembro de 2022. Por isso, voltamos a trabalhar com a base de dados construída com os dados públicos. E em busca de comparações com os resultados extraídos dos dados protegidos na sala segura, construímos com os dados públicos do Censo Escolar mais uma base longitudinal que teve como filtro a mesma coorte que havia sido trabalhada nos dados seguros, porém ficamos limitadas ao ano de 2017.

Ao final, essa tese trabalhou com três bases de dados que são diferentes coortes e versões do Censo Escolar. Denominamos essas bases por ***Base Longitudinal 1***, ***Base Longitudinal 2*** e ***Base Longitudinal 3***. A primeira teve sua coorte definida com referência na literatura e pela disponibilidade dos dados. A segunda, por ser a única coorte que abrangia o ano de 2020 e atendia ao objetivo de investigação sobre tal ano. A terceira para replicar a coorte da segunda base com os dados públicos, que nos eram disponíveis. O leitor será avisado sobre qual versão estamos nos referindo.

²⁴ Esta pesquisa foi contemplada pelo Programa Fulbright de Doutorado Sanduíche nos EUA 2021-2022 no edital *Call For Applications – Cycle 2022-2023 Program For Brazilian Citizens - Doctoral Dissertation Research Award (DDRA)*.

Os filtros e coortes foram:

- **Base Longitudinal 1:** Abrangência de 2007 a 2017 e construída a partir das bases públicas do Censo da Educação (*disponíveis online até 2021*). Coorte: nascidos entre 1º de julho de 1999 e 30 de junho de 2000.
- **Base Longitudinal 2:** Abrangência de 2009 a 2020 e construída a partir da base de fluxo do Sedap. Coorte: nascidos entre 1º de abril de 2002 e 31 de março de 2003.
- **Base Longitudinal 3:** Abrangência de 2009 a 2017 e construída a partir das bases públicas do Censo da Educação (*disponíveis online até 2021*). Coorte: nascidos entre 1º de abril de 2002 e 31 de março de 2003.

2. A construção das bases longitudinais

2.1. Base Longitudinal 1 - a partir dos dados públicos do Censo Escolar

A base pública de matrículas da edição de 2007 do Censo Escolar possui 53.239.385 linhas, ou seja, mais de 50 milhões de alunos. E esses números, em média, se mantêm para todas as edições. Por isso, as pesquisas de nível nacional que constroem bases longitudinais através do Censo Escolar precisam lidar com o volume de dados dessas bases. Em um computador comum, mesmo se selecionadas um restrito número de variáveis, é inviável abrir ao mesmo tempo as edições de 2007 a 2017.

2.1.1. Os filtros

Para viabilizar a manipulação das bases, adotamos a estratégia de filtrar os dados por coorte.

2.1.1.1. Por idade e/ou por etapa?

A literatura nacional sobre trajetórias educacionais na educação básica vem definindo as populações de estudo com filtros por etapas de ensino ou por etapas de ensino e idade. Por etapas de ensino e idade foi a metodologia utilizada em estudos na perspectiva de avaliação de políticas públicas. Por exemplo, para entender as relações entre política de correção de fluxo e a transição para o ensino médio (Lima e Paes de Carvalho, 2021), para analisar a relação entre idade de ingresso na escola e progressão dos estudantes ao longo das séries do ensino

fundamental (Oliveira e Menezes Filho, 2018) ou para discutir a probabilidade de sucesso escolar especificamente em níveis de ensino mais tardios como o ensino fundamental II (Ferrão, 2022).

Os estudos com foco nas desigualdades entre trajetórias educacionais mais longas selecionaram a coorte de alunos pela data de nascimento, considerando a idade dos alunos e a prática vigente da legislação no primeiro ano das trajetórias analisadas (Soares, Alves e Fonseca, 2021; Rangel Rigotti e Hadad, 2018). Após teste e análises, adotamos esse último parâmetro e realizamos o filtro apenas por data de nascimento, principalmente porque esse filtro mantém na base de dados os estudantes que no início da trajetória já estão em desvantagem por não estarem na escola ou por frequentarem etapas anteriores ao ideal. Esses estudantes têm grandes chances de serem os mais vulneráveis de toda a população de estudos e por isso são também os que mais precisam ser analisados.

2.1.1.2. Definições da coorte de data de nascimento

Como os dados com IDs passíveis de conexão estavam disponíveis a partir da edição de 2007, não havia muito espaço para escolhas de coorte de nascimento. As decisões ficaram restritas às mudanças de legislação na oferta do ensino fundamental. A Lei nº 9394/96 (Brasil, 1996) estabelecia o ensino fundamental com etapas de oito anos de duração, obrigatórias e gratuitas para todas as crianças a partir de sete anos de idade. Em 2005, a Lei nº 11.114/05 (Brasil, 2005) estabeleceu a obrigatoriedade do ingresso no ensino fundamental aos seis anos de idade. E em 2006, a Lei nº 11.274, manteve a obrigatoriedade da matrícula aos seis anos, porém alterou a LDB e ampliou a oferta de etapas do ensino fundamental para nove anos. Estados e municípios tiveram até o ano de 2010 como prazo para a implementação da mudança.

Como o objetivo desta pesquisa é entender as trajetórias educacionais ao longo da educação básica, do início do 1º ano do ensino fundamental até o final do 3º do ensino médio, a edição mais recente disponível por data de nascimento seria a dos nascidos em 2001. Essa coorte completou 6 anos de idade ao longo de 2007 e, se regular, ingressou no ensino fundamental nesse mesmo ano. Então, consideramos a legislação vigente em 2007 (Oliveira e Menezes Filho, 2018), selecionamos essa coorte e montamos uma base de dados. Nos primeiros testes encontramos que, em 2007, 63% das crianças dessa coorte ainda frequentavam a pré-escola. A explicação para esse evento pode estar no prazo de implementação da nova legislação.

Esse resultado aponta para esse recorte ser interessante para análises sobre os municípios que aderiram rapidamente a legislação, porém, para a escala nacional as diferenças entre os estados e os municípios na implementação da nova legislação era significativa e fez com que tal recorte não fosse a melhor opção. Selecionamos então a coorte de crianças nascidas entre 1º de julho de 1999 e 30 de junho de 2000 (Soares, Alves e Fonseca, 2021).

Os alunos foram selecionados em cada edição do Censo Escolar, ano a ano, para assegurar que todos os alunos da coorte, matriculados em um ou em alguns anos na educação básica entre 2007 e 2017, fossem incluídos no estudo. Após a seleção, excluímos as matrículas em turmas de atividade complementar e atendimento educacional especializado²⁵ e realizamos um agrupamento por etapas de ensino²⁶ excluindo as matrículas em cursos concomitantes²⁷ e em cursos subsequentes²⁸. A comparação por etapas entre as duas coortes testadas pode ser vista na Tabela 1.

Tabela 1 - Coortes por etapas

Coorte – nascidos entre 1º de janeiro a 31 de dezembro de 2001			
2007	%	2008	%
Educação Infantil – Creche	0,8	Educação Infantil - Creche	0,1
Educação Infantil - Pré-escola	63	Educação Infantil - Pré-escola	4,6
EF de 9 anos - 1º Ano	30,7	EF de 9 anos - 1º Ano	36
EF de 8 anos - 1ª Série e EF de 9 anos - 2º Ano	5,2	EF de 8 anos - 1ª Série e EF de 9 anos - 2º Ano	53,7
Outras Etapas	0,2	EF de 8 anos - 2ª Série e EF de 9 anos - 3º Ano	3,3
		Outras Etapas	0,4
		NA	1,7
Coorte – nascidos entre 1º de julho de 1999 e 30 de junho de 2000			
2007	%	2008	%
Educação Infantil – Creche	0,1	Educação Infantil - Creche	0
Educação Infantil - Pré-escola	1,5	Educação Infantil - Pré-escola	0,4
EF de 9 anos - 1º Ano	10,2	EF de 9 anos - 1º Ano	1,4
EF de 8 anos - 1ª Série e EF de 9 anos - 2º Ano	64,1	EF de 8 anos - 1ª Série e EF de 9 anos - 2º Ano	18,1
EF de 8 anos - 2ª Série e EF de 9 anos - 3º Ano	23,3	EF de 8 anos - 2ª Série e EF de 9 anos - 3º Ano	54,8
Outras etapas	0,9	EF de 8 anos - 3ª Série e EF de 9 anos - 4º Ano	21,5
		Outras etapas	1,1
		NA	2,7

Elaboração própria a partir da Base de Dados Fluxo Escolar/Inep

2.1.2. Deduplicação

Os registros do Censo Escolar possuem muitos alunos duplicados. O caminho para lidar com essa duplicidade é a *deduplicação*. Realizamos esse procedimento em cada ano das edições do Censo Escolar entre 2007 e 2017, em dois formatos diferentes. No primeiro, combinamos

²⁵ Seguindo uma prática do Inep em suas sinopses estatísticas.

²⁶ O agrupamento está apresentado no anexo I.

²⁷ Esses são cursos de formação técnica frequentados pelos estudantes enquanto cursam o ensino médio em outra instituição ou após a conclusão do ensino médio.

²⁸ Esses são cursos de formação técnica frequentados por estudantes que já concluíram o ensino médio.

as variáveis *ano do censo*, *código de identificação único do aluno (ID)*, *etapa de ensino* e *ano de nascimento do aluno*. Se havia na base dois ou mais casos em que essas combinações eram iguais, um dos casos foi mantido e os demais excluídos²⁹. Em seguida, excluímos os casos de alunos com *IDs* duplicados que possuíam o mesmo ano escolar e de nascimento, porém etapas de ensino diferentes^{30, 31}. Essas duplicações são erros no cadastro ou alunos com mais de uma matrícula no ensino regular, por exemplo, os matriculados nos turnos da manhã e da tarde em diferentes escolas. Na impossibilidade de filtrar a correta linha duplicada e excluir a incorreta, todos os casos foram excluídos. Entretanto, esse segundo formato pode inflacionar³² as taxas de evasão. Exemplificamos na Tabela 2.

Tabela 2 - Exemplo de deduplicação

ID	2009	2010	Ano de nascimento	Situação
1	NA	2ºano	2001	Mantido
1	NA	2ºano	2001	Excluído
2	2ºano	3ºano	2001	Excluído
2	3ºano	3ºano	2001	Excluído

Fonte: Elaboração própria

O aluno de ID 1 foi tratado com a deduplicação por combinação de variáveis e uma das linhas foi mantida. No caso do aluno de ID 2, se um ou ambos os registros de 2009 fossem excluídos e os anos subsequentes mantidos, 2009 se tornaria sem registro e isso inflacionaria as taxas de evasão. Para lidar com essa superestimação, fizemos um controle para que os alunos duplicados retirados do banco não se misturassem aos ausentes. Para isso, criamos uma base de dados paralela com todos os alunos que - pelo menos uma vez - estiveram duplicados entre 2007 e 2017 e excluímos esses casos de todas as edições. O número total de alunos retirados do banco foi de 64.531.

²⁹ Para esse filtro utilizamos o comando *distinct* do pacote *dplyr* do programa R, o que garante a permanência de um dos casos entre os casos duplicados, triplicados etc.

³⁰ Esse procedimento é necessário para que um aluno não ganhe dois valores quando todas as edições se transformam em uma única base longitudinal. Se um aluno tem dois valores ele é transformado em *lista* e a base fica em formato de matriz.

³¹ Lima e Paes de Carvalho (2021) relatam o mesmo problema.

³² O Inep tem realizado procedimentos para lidar com essa sobre representação das taxas de evasão nas informações de fluxo escolar e estima uma sobre representação próxima de 3% quando os procedimentos duplicação, exclusão de casos duplicados e imputação de dados não são adotados. Porém, essas ações não se estenderam aos dados públicos.

2.2. O primeiro mapeamento

Após filtradas, as edições do Censo Escolar (de 2007 a 2017) foram empilhadas formando uma base única que tinha como variáveis *ano do censo*, *código de identificação único do aluno* (ID) e *etapa de ensino*. Em seguida, *pivotamos* os dados. *Pivotar* significa girar os dados das linhas para as colunas, o que transformou os *IDs* em casos (linhas) e os anos em variáveis (colunas) que passaram a ter as etapas de ensino como valores. Essa técnica gira a base do formato *empilhado* para o formato *espalhado*³³. Assim, construímos a **Base Longitudinal I** com um *n* de 4.319.625 estudantes. A Tabela 3 apresenta um exemplo da base em formato empilhado e a Tabela 4 mostra o mesmo exemplo em seu novo formato.

Tabela 3 - Exemplo da base longitudinal em formato *empilhado*

Edição	Código do aluno	Etapa de Ensino
2007	ID112XXXXXXXXX5	20
2008	ID112XXXXXXXXX5	30
2009	ID112XXXXXXXXX5	30
2010	ID112XXXXXXXXX5	40
2011	ID112XXXXXXXXX5	40
2012	ID112XXXXXXXXX5	40
2015	ID112XXXXXXXXX5	70
2016	ID112XXXXXXXXX5	80

Fonte: Elaboração própria a partir de dados do Censo Escolar/Inep

Construir uma base longitudinal com o Censo Escolar permite encontrar um importante resultado que é invisível quando as edições estão separadas. Note-se que na Tabela 3, que está em formato empilhado, o aluno não aparece nos registros para a edição de 2013. E por isso, como mostra a

Tabela 4, cada edição é uma coluna e com a *pivotagem* os anos de 2013 e 2014 recebem NA (*Not Available*), o que indica a ausência de registro do aluno em 2013 e 2014. Essa sinalização de casos ausentes a partir dos anos sem registros permite a mensuração das taxas de indivíduos que deixaram de frequentar a escola. Ou seja, a base em *formato espalhado* permite a visualização das trajetórias educacionais ano a ano, incluindo os anos nos quais os indivíduos

³³ Esse procedimento é feito no programa R pelo comando *pivot_wider*, do pacote *tidyverse*.

deixaram de frequentar a escola. Classificamos essa situação de trajetórias como **indivíduos em idade escolar que não frequentaram a escola**. Essa classificação foi dada porque a situação vai além da evasão, que é a mensurada de um ano em relação ao outro. E se difere do conceito de trajetória interrompida (Fonseca, Rodrigues, *et al.*, 2022) que filtra especificamente os indivíduos que não retornam para a escola.

Tabela 4 – Indivíduo na base de dados longitudinal em formato *espalhado*

Código do aluno	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
ID112XXXXXXXXX5	20	30	30	40	40	40	NA	NA	70	80

Fonte: Elaboração própria a partir de dados do Censo Escolar/Inep

De acordo com Inep:

[...] verificou-se a presença de inconsistências na trajetória de alguns alunos na base no período 2007-2015. Essas inconsistências estão basicamente relacionadas a irregularidades na trajetória dos alunos que são perceptíveis pelo acompanhamento da evolução das etapas de ensino no período.

Os dados inconsistentes foram submetidos a um procedimento de imputação denominado “hot deck”, que busca preservar a distribuição original e a coerência dos dados. De forma resumida, as informações ausentes de um aluno são substituídas por informações válidas de outros alunos (“doadores”) que possuem características similares a esse. [...]

O procedimento impõe restrições que garantem, por exemplo, que as informações imputadas de um determinado aluno sejam obtidas de outro aluno que em algum momento esteve na mesma escola, etapa de ensino e modalidade do primeiro – garantindo assim, que a trajetória imputada seja uma trajetória plausível para o aluno.

A segunda etapa consistiu na definição de uma segunda base de dados, denominada ELEGÍVEIS, contemplando os dados de alunos com trajetórias sem inconsistências e que são possíveis “doadores” de informações (Inep, 2017).

A Tabela 5 apresenta a primeira visualização que tivemos da distribuição percentual das trajetórias escolares dos alunos, na educação básica e por etapas de ensino, ao longo de 11 anos.

Tabela 5 - Distribuição percentual da coorte por etapas de ensino. Base Longitudinal I (coorte 99/00 - dados públicos – abrangência 2007 - 2017).

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Idade	07/08	08/09	09/10	10/11	11/12	12/13	13-14	14-15	15-16	16-17	17-18
Ano escolar esperado	1S	2S	3S	4S	5S	6S	7S	8S	1EM	2EM	3EM
Pré-escola ³⁴	8,4	1,4	0,6	0,3	0,2	0,1	0,1	0,1	0,1	0	0
1ª Série	47,3	14,5	4,3	1,7	0,7	0,3	0,2	0,1	0	0	0
2ª Série	17,2	44,8	15,9	5,9	2,7	1,3	0,7	0,3	0,1	0,1	0
3ª Série	0,5	17,6	40,8	14,9	6,1	3	1,5	0,8	0,3	0,1	0,1
4ª Série	0,1	0,6	16,4	39	15	6,4	3,3	1,7	0,7	0,3	0,2
5ª Série	0	0,1	0,5	15,1	38,2	16,8	8,7	5,2	2,5	1	0,5
6ª Série	0	0	0,1	0,4	14,3	35,5	15,6	8,1	4,21	1,8	0,8
7ª Série	0	0	0	0	0,3	13,6	33,4	14,5	6,7	3	1,2
8ª Série	0	0,1	0,1	0,1	0,1	0,4	13,2	32	13,4	5,8	2,4
1º ano E. Médio	0	0	0	0	0	0	0,3	12,3	30,4	16,1	9,2
2º ano E. Médio	0	0	0	0	0	0	0	0,3	10,7	26,5	12,6
3º ano E. Médio	0	0	0	0	0	0	0	0	0,2	10,1	24,4
EJA	0	0,1	0,1	0	0,1	0,1	0,3	0,6	3,4	5,1	5,1
Indivíduos em idade escolar que não frequentaram a escola.	26,4	20,9	21,3	22,6	22,3	22,4	22,7	24	27,1	29,9	33,7 ³⁵

Fonte: Elaboração própria a partir de dados do Censo Escolar/Inep

Somando os estudantes da coorte que chegaram ao 3º ano do ensino médio em 2015, 2016 ou 2017 temos o total de 34,5%. Em 2014, a coorte deveria estar matriculada na 8ª Série/9º ano ou no 1º do EM e apenas 44,6% de alunos chegaram a essas etapas. Esse resultado é 8,4 pontos percentuais menor do resultado de trajetórias regulares encontradas para o ensino fundamental por (Soares, Alves e Fonseca, 2021). As maiores quedas de matrículas nas etapas adequadas em relação ao ano anterior aconteceram em 2008 e 2016. Em 2008, o segundo ano da trajetória, 16% dos estudantes não estavam nas etapas adequadas. Entre esses, em relação a 2007, 2,5% estavam e deixaram de estar; 9,6% estavam matriculados na mesma etapa ou em anteriores e 12,5% estavam na educação infantil ou não frequentavam a escola. Na comparação

³⁴ Somado alunos da pré-escola com os alunos do primeiro ano do ensino fundamental de 9 anos.

³⁵ Subtraímos do total 43,5 os 10,1 dos alunos que estavam no terceiro ano do ensino médio em 2016.

entre 2015 e 2016, houve um aumento de 2,7% nas etapas inadequadas. Isso mostra um gargalo na transição do 1º ano para o 2º do Ensino Médio. Outro resultado importante desta tabela é a média de 24,8% de indivíduos em idade escolar que não frequentaram a escola

Comparar o número de alunos presentes na base construída com os números do Sistema Nacional de Nascidos Vivos do Ministério da Saúde (SINASC/DATASUS) é uma forma de verificar a fragilidade e a robustez de base de dados construídas a partir dos dados públicos do Censo Escolar (Oliveira e Menezes Filho, 2018). De acordo com o Data SUS, entre 1º de julho de 1999 e 30 de junho de 2000 nasceram no Brasil 3.206.614 crianças. O banco que construímos possui 1.113.011 de alunos a mais. Essa diferença aponta para duplicidades que podem ser consequências de novas matrículas realizadas no sistema de ensino que geraram novos IDs e de outros fatores como erros de registro que fugiram ao controle da metodologia que utilizamos. Esta é uma fragilidade deste trabalho que pode ter inflacionado os resultados. Essa discussão será feita no Capítulo 2.

2.3. Base Longitudinal 2 - a partir dos dados seguros

Para pesquisas com o Censo Escolar, o Serviço de Acesso aos Dados Protegidos (Sedap) do Inep disponibiliza duas bases de dados. Uma das bases é o espelho dos microdados públicos, com a diferença de que os IDs dos alunos não estão criptografados. A outra base é a Fluxo Escolar e foi com essa que construímos a base longitudinal 2³⁶. Em 2022, a base Fluxo Escolar continha para os anos de 2007 a 2019 todas as variáveis presentes nas bases públicas do Censo Escolar e duas variáveis que não foram disponibilizadas *online* e são imprescindíveis para a definição de trajetórias escolares: a variável fluxo escolar e a variável situação de rendimento. Apresentamos essas variáveis na tabela 6. Em 2022, a variável fluxo escolar não estava disponível para o ano de 2020.

³⁶ A base longitudinal 2 foi construída com a mesma técnica de pivotagem utilizada para a base longitudinal 1.

Tabela 6 - Variáveis Fluxo Escolar e Situação de Rendimento

TP_FLUXO	Fluxo escolar - transição entre ano atual (considerando a etapa do aluno na 2ª fase do Censo Escolar) e ano seguinte (considerando a etapa do aluno na 1ª fase do Censo)	1. Promovido 2. Repetente 3. Evadido de escola 4. Migração para a EJA 6. Migração para o regular 7. Falecido 9. Não se aplica
TP_SITUAÇÃO	Situação de rendimento ou movimento do aluno ao final do ano letivo.	2. Abandono 3. Falecido 4. Reprovado 5. Aprovado 9. Sir (Matrícula sem informação de Rendimento ou Movimento)

Fonte: Inep

As categorias da variável fluxo escolar são referentes ao Modelo de Fluxo Escolar (Thonstad, 1980), proposto pela Unesco e adaptado³⁷ para o Brasil (Klein e Ribeiro, 1991; Klein, 1995; Rigotti e Cerqueira, 2004). A definição das categorias atualmente utilizadas pelo Inep são: “promovido” (aluno que se matricula na série seguinte àquela na qual estava matriculado no ano anterior), repetente (aluno que se matricula na mesma série que estava frequentando no ano anterior) e “evadido” (aluno que estava matriculado no início de um ano letivo, mas não se matriculou no ano seguinte). Já as categorias da variável situação de rendimento são “aprovado” (aluno que alcança notas e frequências satisfatórias, o que permite que ele seja promovido para a série seguinte no próximo ano), “abandono” (aluno que deixou de frequentar a escola durante o ano letivo), “falecido” (aluno que faleceu durante o ano letivo) e “reprovado” (aluno que não atinge notas e/ou frequências satisfatórias e por isso não é aprovado).

Na seleção sobre a coorte que seria analisada, levamos em consideração a Lei nº 11.274/2006, que estabelece o ensino fundamental com duração de 9 anos, iniciando-se aos seis anos de idade, e selecionamos as crianças nascidas entre 1º de abril de 2002 e 31 de março de 2003 que teriam suas trajetórias escolares, se regulares, no período de 2009 a 2020. Em 2009, a coorte tinha a idade de 6 anos e estava (ou deveria estar) no 1º ano do ensino fundamental I. Já em 2020, esses estudantes possuíam 17 anos e, para concluírem a educação básica com trajetória regular, era necessário serem aprovados no 3º ano do ensino médio.

³⁷ Essa adaptação foi feita para muitos países da América Latina.

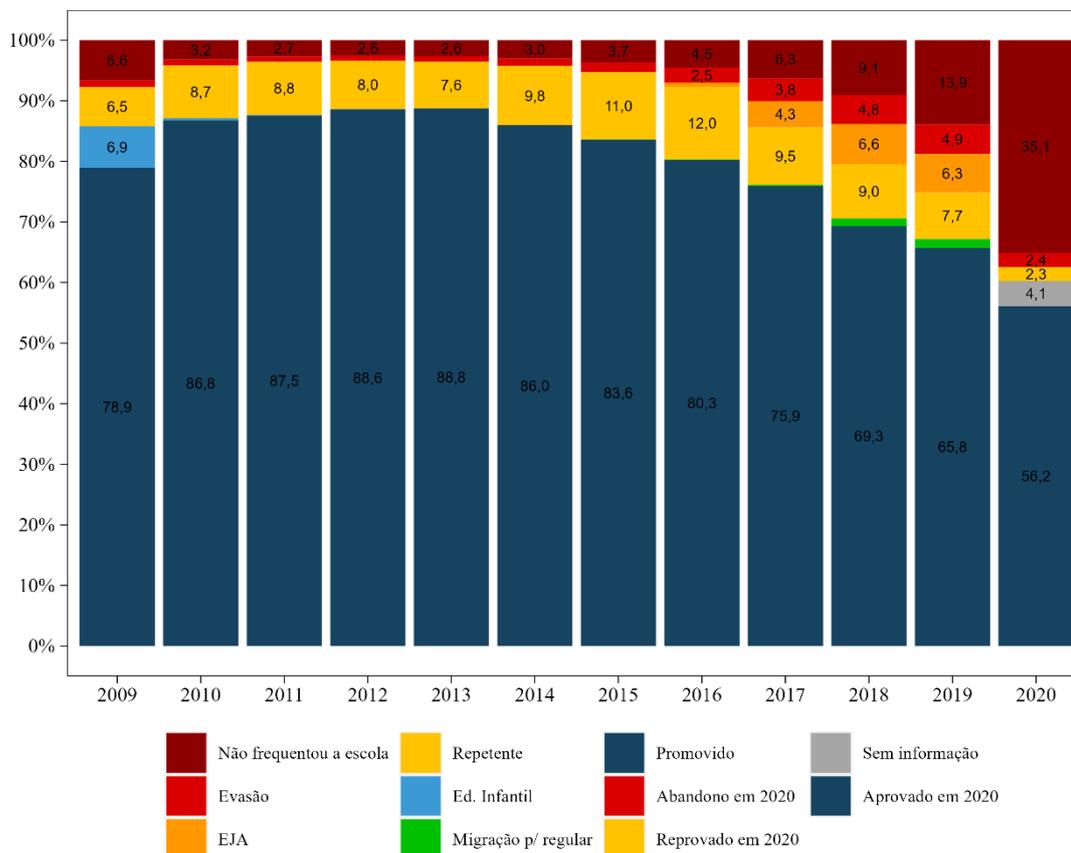
2.4. Confirmando a magnitude do problema

Nessa seção apresentamos resultados encontrados através da base longitudinal 2. O Gráfico 1 é referente ao fluxo dos estudantes entre 2009 e 2019 que estão distribuídos nas categorias repetente, promovido, evadido e alunos que se movimentaram entre a Educação de Jovens e Adultos (EJA) e o ensino regular. Para o ano de 2020, a referência é a situação dos estudantes ao final do ano letivo. As categorias são aprovado, reprovado, abandono e sem informação. Além dessas categorias, o gráfico permite a visualização dos indivíduos que deixaram de frequentar a escola³⁸. Um ponto importante da construção dessa categoria é que ela é referente ao ano atual. Por exemplo, em 2016, 80,3% dos alunos foram promovidos para 2017; 2,5% dos alunos frequentaram a escola em 2016 e não retornaram em 2017, ou seja, evadiram em 2016. Ao passo que em 2016 4,5% dos alunos da coorte não frequentaram a escola, os quais podem ter qualquer situação em 2017.

Entre 2009 e 2020, o número de registros de alunos dessa coorte que passaram pelo sistema escolar foi de 3.213.602 alunos. Analisando suas transições no Gráfico 1, notamos progressiva redução do percentual de estudantes promovidos a cada ano, o aumento da população que deixou de frequentar a escola e a migração para a EJA. Chama a atenção o resultado de 2020, primeiro ano da pandemia de Covid-19, com o percentual de indivíduos que deixaram de frequentar a escola mais elevado. Para o hiato em 2020, há duas explicações plausíveis. A primeira é que apenas uma reduzida parcela da coorte chega ao 3º ano do ensino médio, provocando a diferença entre as taxas. A segunda, é que esta diferença ainda mais acentuada entre as taxas ocorreu pelo contexto da pandemia, que levou muitos jovens a abandonarem a escola. Essa última, vai ao encontro de Artes, Pagaiame, *et al.* (2021) e será discutida na próxima seção.

³⁸ Esses alunos foram encontrados através da metodologia de pivotagem que considera os anos em que eles não aparecem registrados no Censo Escolar.

Gráfico 1- Fluxo dos estudantes na educação básica (2009 a 2020). Base Longitudinal II (coorte 02/03 - dados protegidos - abrangência 2009 - 2020).

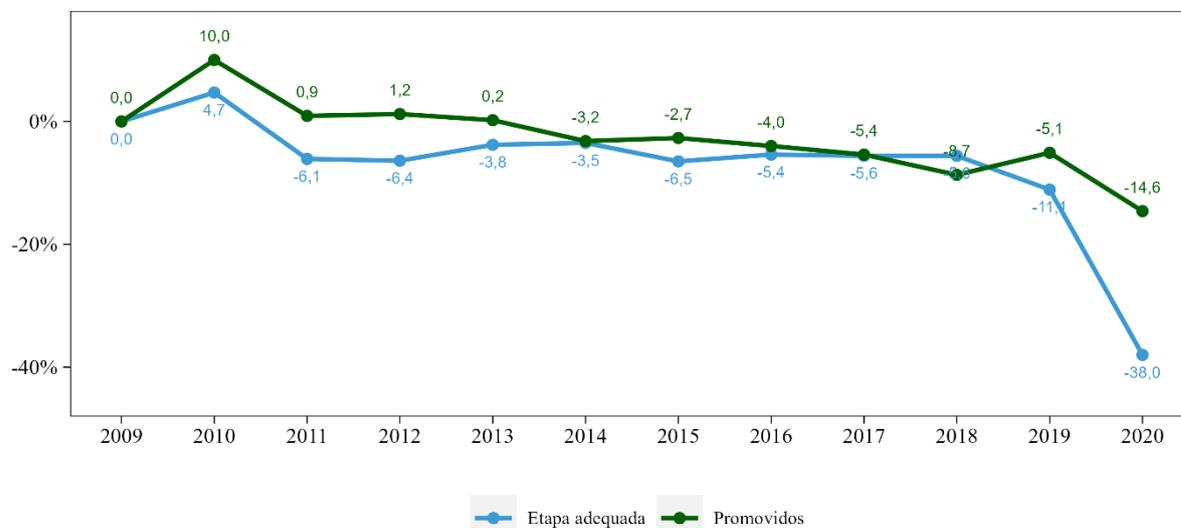


Fonte: elaboração própria a partir de dados do Fluxo Escolar/Inep

O Gráfico 2 apresenta uma comparação entre os alunos em etapa³⁹ adequada e os alunos promovidos (que independem da etapa adequada) ao longo do período de 2009 a 2020. Essas duas situações não são coincidentes, pois um aluno pode ser promovido em determinado ano, mas apresentar um atraso em relação ao ano escolar que deveria estar cursando. Ao longo do tempo, a distância entre o percentual de promovidos e o percentual em etapa adequada aumenta. Ou seja, verificamos que sucessivas reprovações, saída e retorno de alunos ao sistema, provocam o aumento dessa distância, que tem o seu ápice em 2020. A queda de 2018 para 2019 se refere a passagem do 1º ano do ensino médio para o 2º ano e revela o entrave neste momento da trajetória escolar. O movimento começa em 2017 quando as duas curvas são coincidentes e onde parte da população transita do 9º do ensino fundamental para o 1º ano do ensino médio. Uma análise interessante é perceber que a queda de 2018 é a mesma que acontece em 2015 na Base L1 (tabela 5, pag. 24). Isso porque ambos os anos se referem a passagem do 1º ano do ensino médio para o 2º ano.

³⁹ Aqui, utilizamos etapa como sinônimo de série.

Gráfico 2 - Alunos promovidos e em etapas adequadas em relação ao ano anterior (%). Base Longitudinal II (coorte 02/03 - dados protegidos - abrangência 2009 - 2020).



Fonte: elaboração própria a partir de dados do Fluxo Escolar/Inep

Gráfico 1 e Gráfico 2 mostram que o sistema educacional brasileiro possui gargalos em todas as categorias da variável de fluxo. Há elevadas taxas de repetência e de indivíduos que deixaram de frequentar a escola. As taxas de promoção caem a cada ano e a política de Ensino de Jovens e Adultos (EJA) tem pouca adesão frente ao elevado número de intercorrências. Dois pontos se destacam nos gráficos. O primeiro é a inversão que ocorre de 2016 para 2017 nos gráficos 1 e 2. Esses dois anos são respectivamente o 8º e 9º ano das trajetórias e o momento no tempo em que os estudantes têm entre 14 e 15 anos de idade. Até 2016, os estudantes foram mais reprovados do que estiveram sem frequentar a escola e em 2017 isso se inverte com mais alunos fora do sistema de ensino do que reprovados dentro dele.

3. Não deixar ninguém para trás: uma proposta

Trabalhos recentes elaboraram variáveis de trajetórias escolares a partir do indicador de fluxo escolar do Inep. O primeiro deles, Rangel Rigotti e Hadad (2018), sobre a relação entre migração e o desempenho educacional trabalhou uma variável binária para mensurar trajetórias educacionais. As categorias são trajetória regular (para os estudantes nascidos entre 1º de abril de 2000 e 31 de março de 2001 que concluíram o ensino fundamental em 2015) e trajetória irregular (para os estudantes da coorte que não concluíram o ensino fundamental em 2015). O trabalho encontrou 70% de trajetórias irregulares para o ensino fundamental completo (I + II).

Em Soares, Alves e Fonseca (2021), a variável de trajetórias escolares tem três categorias construídas a partir de uma soma denominada cinco critérios de qualidade para as trajetórias. Os critérios são: 1) presença de evasão, 2) número de anos de sucesso escolar, 3) número de anos de atraso escolar, 4) ano escolar atingido em 2015 e 5) presença de repetência. De acordo com cada critério, os estudantes receberam os valores 0, 1 ou 2 e somaram de 0 a 7 pontos. Aqueles com soma 0 ou 1 foram categorizados como trajetória regular, com soma 2 ou 3 como poucas irregularidades e com soma maiores que 3 como trajetória com grandes irregularidades. Os resultados encontrados foram 53,1% de alunos com trajetória regular, 16,5% com poucas irregularidades e 30,4% com grandes irregularidades para o ensino fundamental completo (I + II).

Para mensurar a probabilidade de progressão dos alunos do 5º para o 9º ano do Ensino fundamental (I + II), Ferrão (2022) também construiu uma variável binária de trajetória educacional. As categorias são sucesso escolar, para os estudantes que do 5º ao 9º [em 2011 e 2015] foram promovidos continuamente, e insucesso escolar para os estudantes que tiveram alguma intercorrência durante esse período. O resultado encontrado foi de 72,2% de sucesso escolar entre os alunos que frequentaram escolas públicas e realizaram as provas no final da 5ª série em 2011 enquanto 28% experimentaram alguma intercorrência, pelo menos uma vez.

Com uma proposta de modificações para o indicador desenhado por Soares, Alves e Fonseca (2021), Fonseca, Rodrigues, *et al.* (2022) apresentam uma variável de trajetórias em quatro categorias: regulares, com pouca irregularidade, com muita irregularidade e interrompidas. Tem trajetórias regulares os estudantes que entraram e concluíram o ensino fundamental e/ou médio em idades corretas. Os estudantes com pouca irregularidade tiveram uma ou duas intercorrências nas trajetórias e os com muita irregularidade três ou mais. Essas intercorrências podem ser repetência, abandono ou evasão. A principal mudança em relação ao indicador anterior é a distinção entre o grupo com trajetórias muito irregulares, porém que continuam a frequentar a escola e o grupo que evadiu da escola e não retornou, que são as trajetórias interrompidas. Essa proposta permite o monitoramento da permanência dos estudantes via indicador de regularidade de trajetórias educacionais. O estudo encontrou 52% com trajetórias regulares no ensino fundamental (I + II) e 41% para o ensino médio.

3.1. Trajetórias escolares na educação básica como variável resposta

Até o momento, está bem definido pela literatura o que se entende por trajetórias regulares, que incluem os estudantes que são promovidos em todas as séries da educação básica ou, no máximo, tiveram uma intercorrência (Oliveira e Menezes Filho, 2018; Rangel Rigotti e Hadad, 2018; Lima e Paes de Carvalho, 2021; Soares, Alves e Fonseca, 2021; Ferrão, 2022; Petrus, 2022; Fonseca, Rodrigues, *et al.*, 2022). Já a mensuração das trajetórias irregulares ainda está em debate. A literatura tem trabalhado para entender qual é a melhor configuração para medir as intercorrências de repetência, abandono e/ou evasão que os estudantes experimentam uma ou mais vezes.

É importante observar que os estudantes muitas vezes têm trajetórias de idas e vindas no sistema escolar. Isso já foi documentado pela literatura sociológica nos anos 90 (Zago, 2003). A contribuição que esta tese faz para esse debate é a possibilidade de mensurar esse movimento, incluindo no indicador de trajetória a categoria *indivíduos em idade escolar que não frequentaram a escola*. Assim, na esteira dos trabalhos de Soares, Alves e Fonseca (2021) e Fonseca, Rodrigues, *et al.* (2022), apresentamos uma proposta de mensuração para as trajetórias escolares na educação básica. Do ponto de vista analítico, argumentamos que ser reprovado e deixar de frequentar a escola são situações distintas e que o número de vezes que esses eventos acontecem podem diferenciar intensamente o próprio evento. Quer dizer, tratar essas situações dentro de uma única categoria de “irregulares” pode ser menos eficaz. Ao contrário, permitir a granularidade da informação pode nos aproximar de fatores explicativos que a literatura ainda desconhece e dar pistas para políticas públicas que ainda não foram elaboradas.

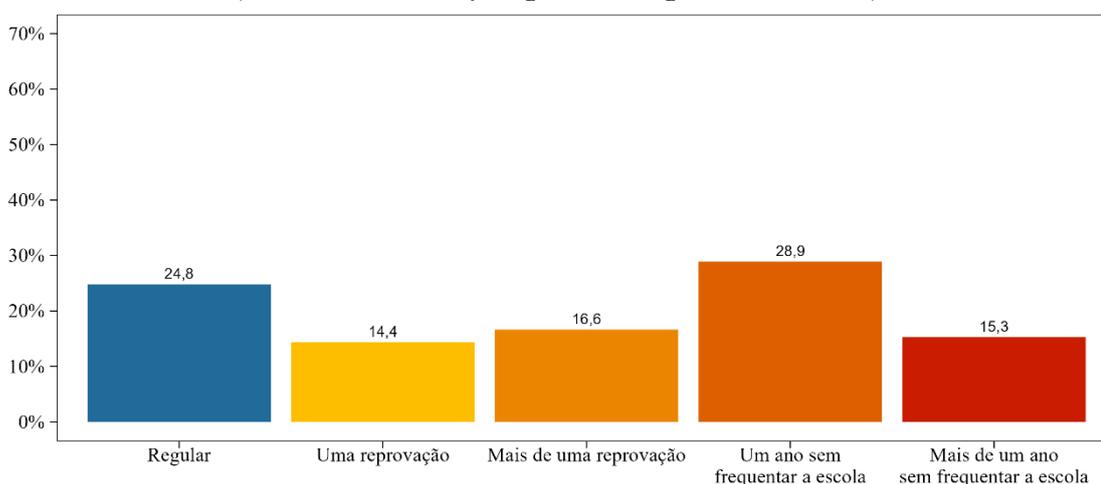
Nesse sentido, sugerimos uma variável de trajetórias escolares com cinco categorias: **trajetórias regulares** para os alunos sem intercorrências escolares, **trajetórias com uma reprovação** para os que experimentaram apenas uma reprovação durante toda a vida escolar, **trajetórias com mais de uma reprovação** para os que experimentam mais de uma reprovação, **trajetórias com um ano sem frequentar a escola** para os indivíduos que por apenas um ano deixaram de frequentar a escola em toda a vida escolar e **trajetórias com mais de um ano sem frequentar a escola** para aqueles que deixaram de frequentar a escola por mais de um ano. Nessa construção, deixar de frequentar a escola se sobrepõe a reprovação. Ou seja, indivíduos categorizados como deixaram de frequentar a escola podem ter experimentado reprovações, porém os reprovados não deixaram de frequentar a escola.

Chegamos à essa elaboração argumentativa através de uma análise exploratória, empírica dos dados do Censo Escolar. Apresentamos a seguir a descritiva da variável que construímos e em seguida as análises exploratórias que sustentam nosso argumento analítico.

4. RESULTADOS

Gráfico 3 mostra as porcentagens das trajetórias escolares da coorte analisada na educação básica (I + II + Médio): 24,8% de trajetórias regulares; 14,38% com uma reprovação; 16,64 % com mais de uma reprovação; 28,90% de alunos que estiveram um ano sem frequentar a escola e 15,28 % que estiveram mais de um ano sem frequentar a escola.

Gráfico 3 - Trajetórias educacionais na Educação Básica Completa (2009 a 2020). Base Longitudinal II (coorte 02/03 - dados protegidos - abrangência 2009 - 2020).



Fonte: elaboração própria a partir de dados do Fluxo Escolar/Inep

4.1. Análises exploratórias

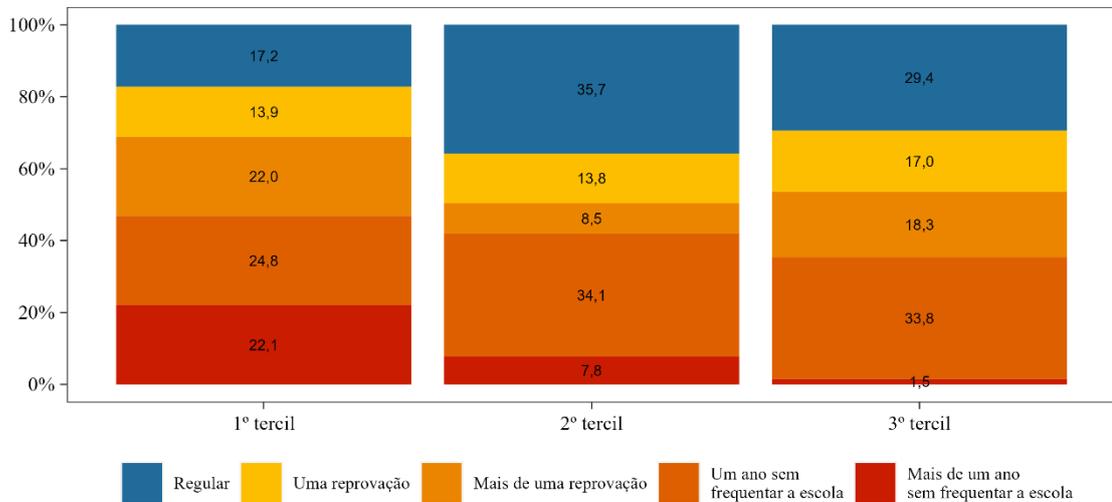
Os próximos gráficos apresentam os resultados das análises exploratórias que sustentam nosso argumento analítico.

Gráfico 4 mostra a relação entre o nível socioeconômico (NSE) das escolas⁴⁰ e as trajetórias na educação básica (I + II + Médio). Observamos que entre os estudantes mais ricos

⁴⁰ Esse indicador pertence a base de dados Nível Socioeconômico das Escolas desenvolvida no Núcleo de Pesquisa em Desigualdades Escolares sob a coordenação da Prof. Dr. Teresa Alves e tem como referência o ano de 2019. Ele foi dividido em três categorias através de cluster e sua distribuição está apresentada no apêndice I.

(3º tercil do NSE), 29,4% tiveram trajetórias regulares, 15% apresentaram uma reprovação, 15,3% tiveram mais de uma reprovação, 29,9% deixaram de frequentar a escola em algum momento e 1,5% deixaram de frequentar a escola mais de uma vez. Já no grupo mais pobre (1º tercil do NSE), 17,2% tiveram trajetórias regulares, 13,9% apresentaram uma reprovação, 22% tiveram mais de uma reprovação, 24,8% deixaram de frequentar a escola em algum momento e 22,1% deixaram de frequentar a escola mais de uma vez.

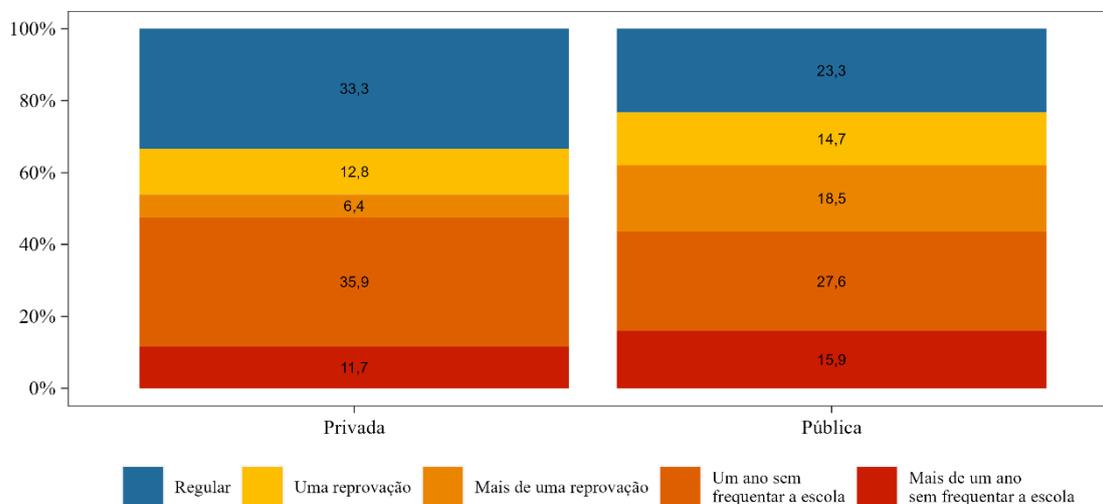
Gráfico 4 - Trajetórias escolares na educação básica por NSE das escolas (2009 a 2020). Base Longitudinal II (coorte 02/03 - dados protegidos - abrangência 2009 - 2020).



Fonte: elaboração própria a partir de dados do Fluxo Escolar/Inep

Gráfico 5 mostra que na educação básica (I + II + Médio) 33,3% dos estudantes de escolas privadas tiveram trajetórias regulares; 12,8% uma reprovação; 6,4% mais de uma reprovação; 35,9% deixaram de frequentar a escola uma vez e 11,7% mais de uma vez. Já os estudantes de escolas públicas 23,3% tiveram trajetórias regulares; 14,7% uma reprovação; 18,5% mais de uma reprovação; 27,6% deixaram uma vez fora de frequentar a escola e 15,9% mais de uma vez.

Gráfico 5 - Trajetórias escolares na educação básica por dependência administrativa das escolas (2009 a 2020). Base Longitudinal II (coorte 02/03 - dados protegidos - abrangência 2009 - 2020).



Fonte: elaboração própria a partir de dados do Fluxo Escolar/Inep

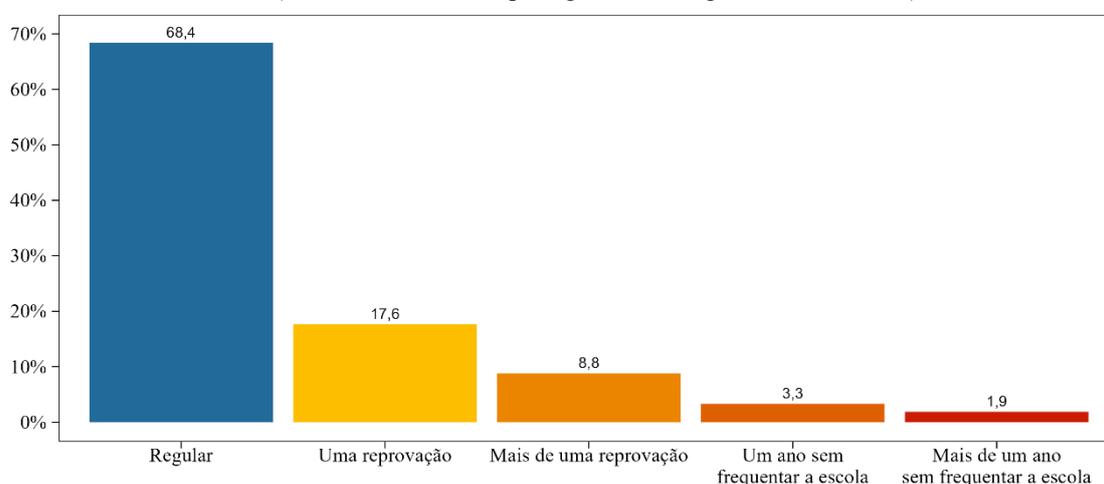
A predominância de estudantes de escolas com alto NSE e de escolas privadas na categoria um ano sem frequentar a escola são contraintuitivos diante da literatura. Dado a esses resultados divergentes da literatura optamos por realizar uma comparação com as trajetórias escolares separadas por etapas. Essa abordagem nos permitiu analisar os dados excluindo o ano de 2020, ano da pandemia da Covid-19, o que nos possibilitou uma visão mais consistente das tendências ao longo desse período.

Nossa proposta de variável para trajetórias é uma síntese do indicador fluxo escolar e permite análises por etapas da educação básica e/ou desse período por completo. Se por etapas, é importante compreender que essas estão temporalmente dentro da trajetória para se estabelecer a comparação e considera a regularidade. A seguir, os gráficos 6, 7, 8 e 9 mostram as trajetórias para o ensino fundamental I, ensino fundamental II, ensino fundamental completo (I + II) e ensino médio. Os resultados mostram que aos poucos a literatura tem caminhado para alguns consensos. A porcentagem de 67,6% de alunos regulares está próxima dos 72,2% encontrados por Ferrão (2022) e os 50,60% que encontramos para as trajetórias regulares ao longo do ensino fundamental também está próxima dos 53,1% de alunos com trajetória regular encontrados por Soares et al (2021). Vejamos as trajetórias por etapas:

No Gráfico 6, como esperado, vemos que os resultados para o ensino fundamental I têm melhores taxas: 68,4% dos alunos regulares, 17,6% com uma reprovação, 8,8% com mais de

uma reprovação, 3,3% com um ano sem frequentar a escola e 1,9% com mais de um ano sem frequentar a escola.

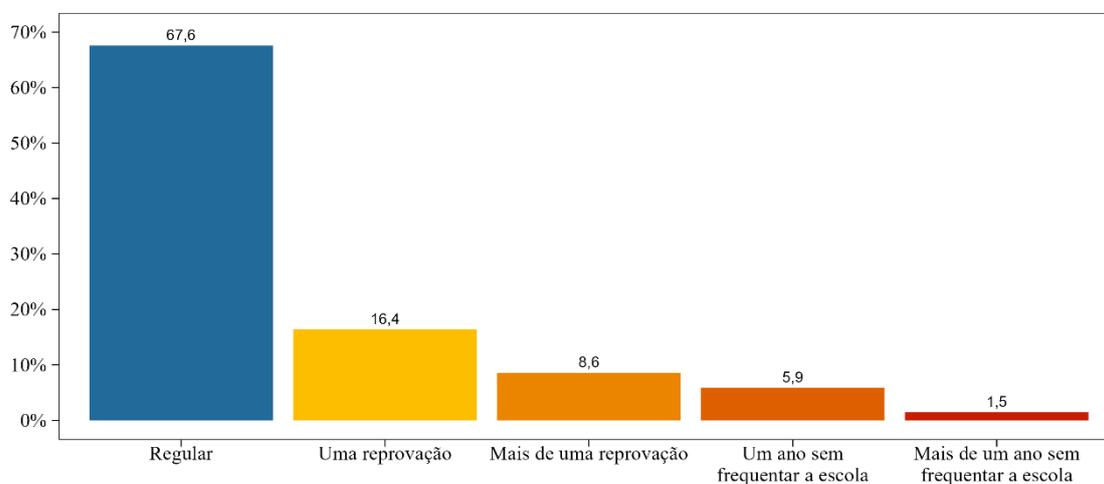
Gráfico 6 - Trajetórias educacionais no ensino fundamental I (2009 a 2013). Base Longitudinal II (coorte 02/03 - dados protegidos - abrangência 2009 - 2020).



Fonte: elaboração própria a partir de dados do Fluxo Escolar/Inep

Gráfico 7 é referente ao ensino fundamental II. Essa etapa tem um filtro para que somente os alunos regulares no primeiro ano da etapa fossem selecionados. Eles somam 2,841,938, que correspondem a 88,6% da coorte. Os resultados são 67,6% de trajetórias regulares, 16,4% de uma reprovação, 8,6% mais de uma reprovação, 5,9% um ano sem frequentar a escola e 1,5% mais de um ano sem frequentar a escola.

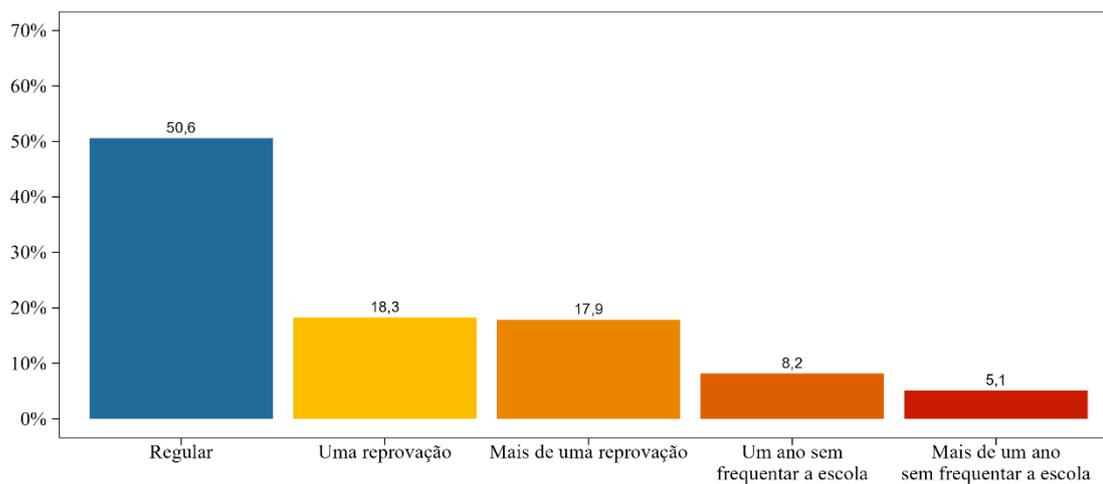
Gráfico 7 - Trajetórias educacionais no Ensino Fundamental II (2009 a 2017). Base Longitudinal II (coorte 02/03 - dados protegidos - abrangência 2009 - 2020).



Fonte: elaboração própria a partir de dados do Fluxo Escolar/Inep

Gráfico 8 são as trajetórias no ensino fundamental completo (I + II). Os resultados para esse período são 50,60% de alunos com trajetórias regulares; 18,26% com uma reprovação, 17,86% com mais de uma reprovação, 8,18% de alunos que estiveram um ano sem frequentar a escola e 5,1% que estiveram mais de um ano.

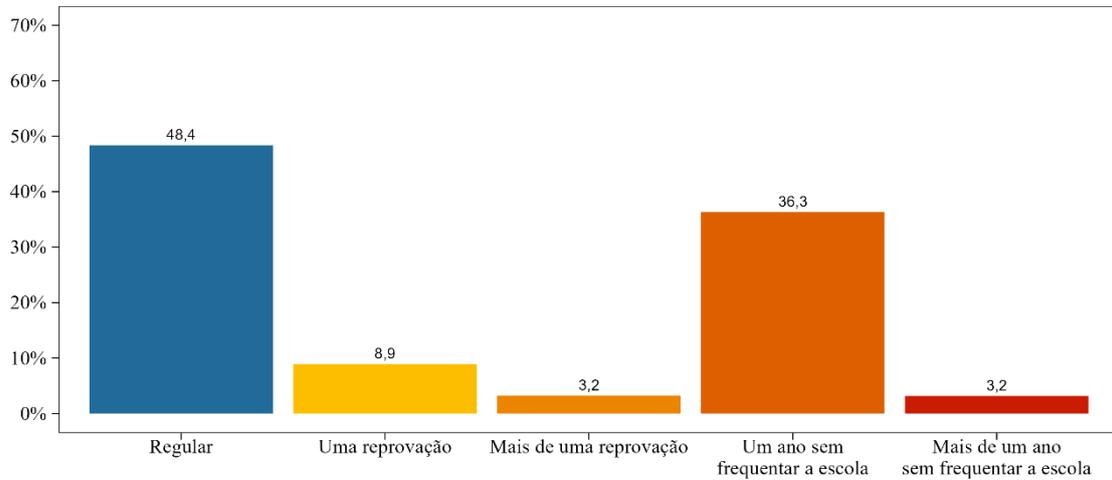
Gráfico 8 - Trajetórias educacionais no ensino fundamental Completo (2009 a 2017). Base Longitudinal II (coorte 02/03 - dados protegidos - abrangência 2009 - 2020).



Fonte: elaboração própria a partir de dados do Fluxo Escolar/Inep

O Gráfico 9 é referente ao ensino médio, também com filtro de seleção para os regulares no primeiro ano da etapa, 48,37% dos alunos tiveram trajetórias regulares; 8,93% uma reprovação; 3,25% mais de uma reprovação; 36,31% um ano sem frequentar a escola e 3,15% mais de um ano.

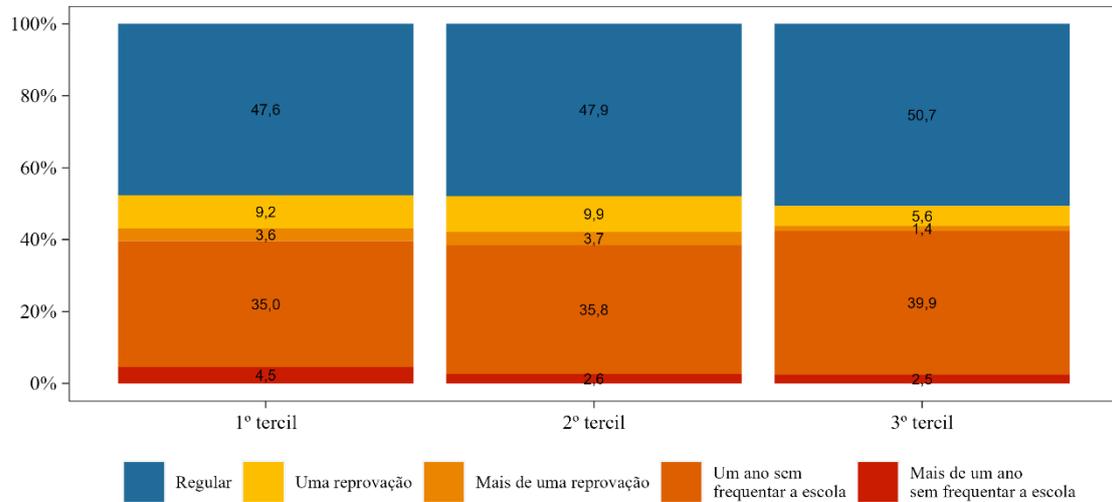
Gráfico 9 - Trajetórias educacionais no Ensino Médio (2018 a 2020). Base Longitudinal II (coorte 02/03 - dados protegidos - abrangência 2009 - 2020).



Fonte: elaboração própria a partir de dados do Fluxo Escolar/Inep

A primeira investigação que fizemos para entender a relação entre trajetórias e variáveis socioeconômicas foi com o ensino médio, pois a etapa abrangia o ano de 2020. Os gráficos 10 e 11 apresentam os resultados. No primeiro, os estudantes mais ricos (3º tercil) registraram 50,7% de trajetórias regulares, 5,6% de uma reprovação, 1,4% de mais de uma reprovação, 36,4% um ano sem frequentar a escola e 1,9% mais de um ano sem frequentar a escola. Já entre os estudantes mais pobres, 45,6% tiveram trajetórias regulares, 9,2% tiveram uma reprovação, 3,6% tiveram mais de uma reprovação, 39,9% estiveram sem frequentar a escola em algum momento e 2,5% estiveram sem frequentar a escola mais de uma vez.

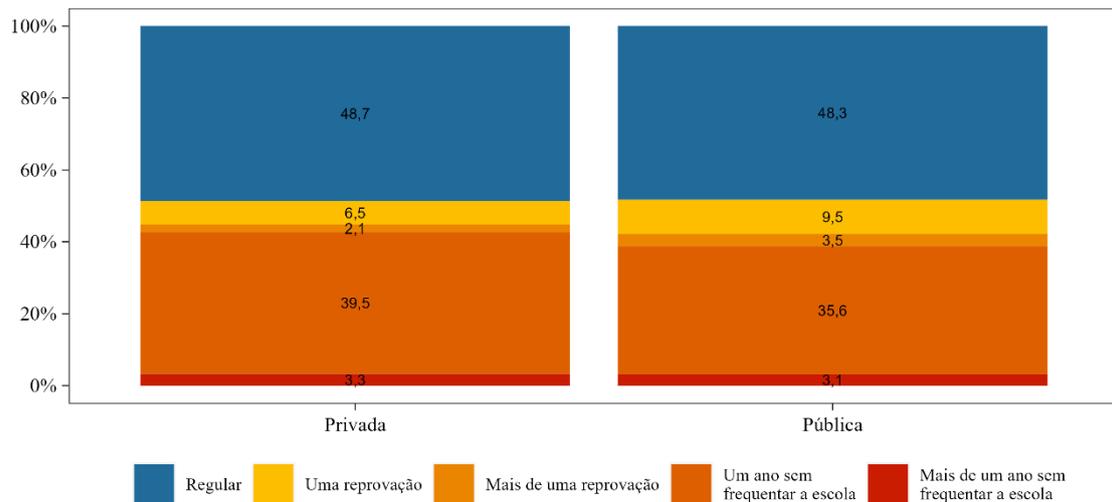
Gráfico 10 - Trajetórias escolares no ensino médio por NSE das escolas (2018 a 2020). Base Longitudinal II (coorte 02/03 - dados protegidos - abrangência 2009 - 2020).



Fonte: elaboração própria a partir de dados do Fluxo Escolar/Inep

Na análise por dependência administrativa o Gráfico 11 aponta estudantes de escolas privadas com 48,7% de trajetórias regulares; 6,5% com uma reprovação; 2,1% com mais de uma reprovação; 39,5% uma vez sem frequentar a escola e 3,3% mais de uma vez. Enquanto os de escolas públicas 48,3% tiveram trajetórias regulares; 9,5% uma reprovação; 3,5% mais de uma reprovação; 35,6% estiveram uma vez sem frequentar a escola e 3,1% mais de uma vez.

Gráfico 11 - Trajetórias escolares no ensino médio por dependência administrativa das escolas (2018 a 2020). Base Longitudinal II (coorte 02/03 - dados protegidos - abrangência 2009 - 2020).

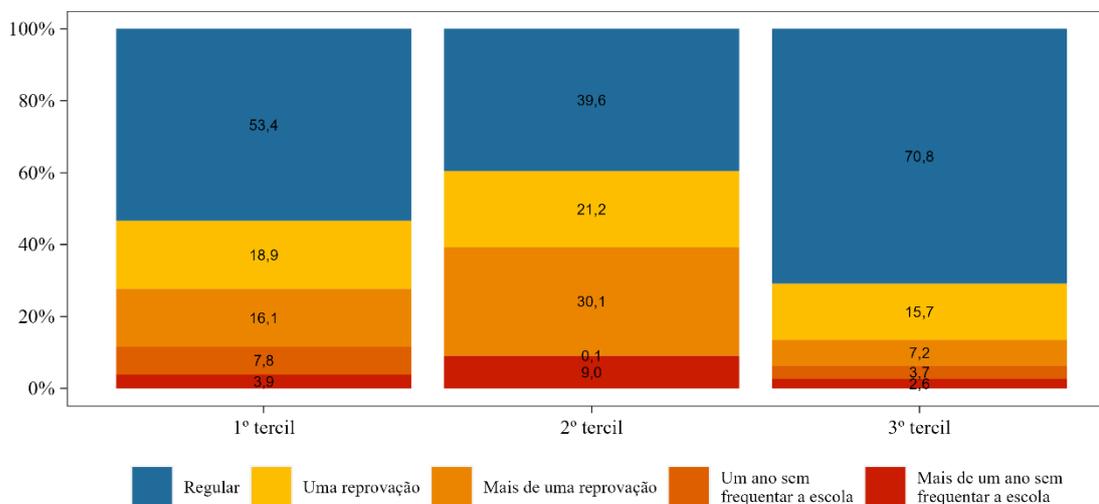


Fonte: elaboração própria a partir de dados do Fluxo Escolar/Inep

Os resultados do cruzamento entre as categorias de trajetória e as variáveis nível socioeconômico das escolas e dependência administrativa são contraintuitivos, com destaque para alunos que frequentam escolas de alto nível socioeconômico e escolas privadas terem deixado de frequentar a escola por um ano. Para investigar essa situação, realizamos uma comparação com as trajetórias escolares no ensino fundamental. Essa abordagem permite analisar os dados excluindo o ano de 2020 e possibilita uma visão mais consistente das tendências ao longo desse período.

O Gráfico 12 é o cruzamento entre as trajetórias escolares no ensino fundamental completo (I + II) e o NSE das escolas. Os resultados são os estudantes mais ricos com 70,8% de trajetórias regulares; 15,7% uma reprovação; 7,2% mais de uma reprovação; 3,7% estiveram uma vez sem frequentar a escola e 2,6% mais de uma vez. Já entre o grupo mais pobre 53,4% tiveram trajetórias regulares; 18,9% uma reprovação; 16,1% mais de uma reprovação; 7,8% estiveram uma vez sem frequentar a escola e 3,9% mais de uma vez.

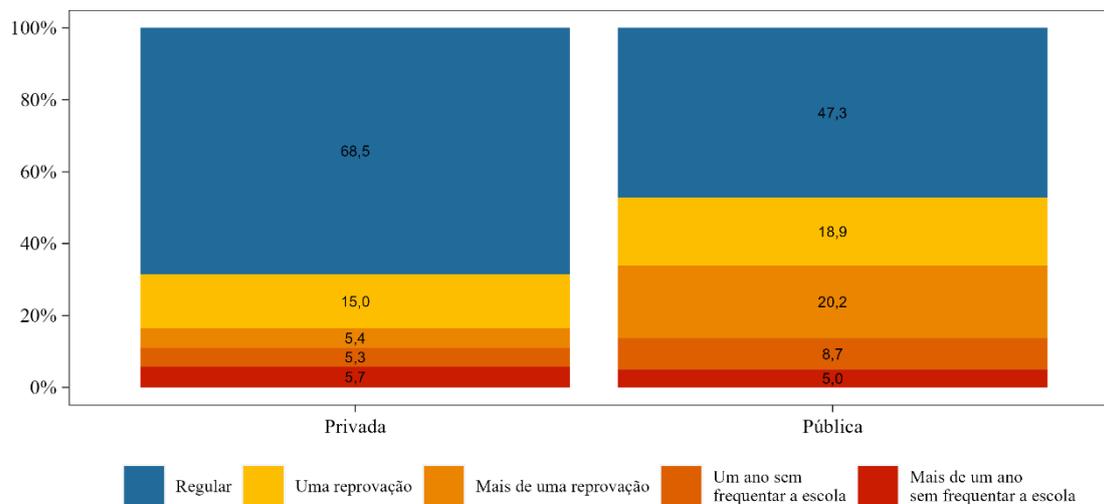
Gráfico 12 - Trajetórias escolares no ensino fundamental completo por NSE das escolas (2009 a 2017). Base Longitudinal II (coorte 02/03 - dados protegidos - abrangência 2009 - 2020).



Fonte: elaboração própria a partir de dados do Fluxo Escolar/Inep

No Gráfico 13 estão os resultados por dependência administrativa das escolas. Os estudantes de escolas privadas tiveram 68,5% de trajetórias regulares; 15% uma reprovação; 5,4% mais de uma reprovação; 5,3% estiveram uma vez sem frequentar a escola e 5,7% mais de uma vez. Já entre os estudantes de escolas públicas 47,3% tiveram trajetórias regulares; 18,9% uma reprovação; 20,2% mais de uma reprovação; 8,7% estiveram uma vez sem frequentar a escola e 5% mais de uma vez.

Gráfico 13 - Trajetórias escolares no ensino fundamental completo por dependência administrativa das escolas (2009 a 2017). Base Longitudinal II (coorte 02/03 - dados protegidos - abrangência 2009 - 2020).



Fonte: elaboração própria a partir de dados do Fluxo Escolar/Inep

A análise do ensino fundamental completo e sua comparação com a trajetória completa e no ensino médio permitem a inferência de que o ano em que os estudantes ricos estiveram sem frequentar a escola foi 2020, durante a pandemia da Covid-19. Esse achado é importante para pensarmos sobre o que significa ir ou não ir para a escola em momentos de crise.

Além disso, as diferenças dos resultados entre as categorias quando cruzadas com variáveis de nível socioeconômico nos levaram ao argumento de que ser reprovado e estar sem frequentar a escola são situações distintas e quantas vezes essas intercorrências acontecem impactam a própria intercorrência. Assim, entendemos que a variável de trajetórias escolares com cinco categorias é uma proposta com diferentes funções e pode ser uma interessante alternativa para a literatura.

5. Considerações

Esse capítulo traz dois pontos para o debate. Teria sido ficar um ano sem frequentar a escola uma estratégia privilegiada utilizada pelos alunos com nível socioeconômico mais alto? Além disso, há 25% entre esses mais ricos que se mantêm na escola sem intercorrências independente das circunstâncias?

No capítulo 3 retomaremos com a variável de trajetórias escolares e as variáveis socioeconômicas nos modelos analíticos. Antes disso, no próximo capítulo, discutiremos

metodologicamente outro importante entrave ao sistema educacional brasileiro que se estende e se espalha para toda a sociedade. Os indivíduos que deixam de frequentar a escola.

CAPÍTULO II – INDIVÍDUOS EM IDADE ESCOLAR QUE NÃO FREQUENTARAM A ESCOLA

A hipótese mais plausível para o pouco mais de 1 milhão de alunos encontrados a mais na base longitudinal 1 (coorte 99/00 – dados públicos) se comparados com os registros no DATASUS é a superestimação do número de alunos. Isso pode ter acontecido em razão dos alunos terem recebido IDs diferentes ao longo de suas trajetórias escolares. Entretanto, mesmo que inflacionadas, essa base ajuda no entendimento da trajetória de “indivíduos em idade escolar que deixaram de frequentar a escola”, que é uma dimensão com demanda de investigação, já que escapam aos atuais indicadores.

O capítulo está dividido em 2 partes. A parte I faz uma discussão sobre a imputação de casos ausentes nas bases educacionais e como essa metodologia pode impactar sobre os alunos que deixam de frequentar a escola. A parte II faz uma proposta de incorporação da dimensão “indivíduos em idade escolar que não frequentaram a escola” ao Índice de Desenvolvimento da Educação Básica (Ideb).

1. Um debate sobre imputação

Na constatação da disparidade entre as taxas de alunos que deixam de frequentar a escola na Base Longitudinal 1 em comparação com a Base Longitudinal 2 (conforme mencionado no capítulo 1, página 37) e possível interpretação de que as taxas da primeira estão acima do esperado realizamos um teste a fim de investigar sobre essa questão. Duas perguntas se desdobraram desse achado: o tratamento de imputação dado ao Censo da educação resulta em uma redução das taxas de “Indivíduos em idade escolar que não frequentaram a escola”? A imputação realizada afeta grupos com perfil socioeconômico específico? Na seção subsequente, apresentamos o desenvolvimento de nossa pesquisa sobre essa questão e as hipóteses que desenvolvemos diante dessas perguntas.

1.1. Replicação

Como mencionamos no capítulo anterior, construir uma base longitudinal com o Censo Escolar permite encontrar indivíduos que deixaram de frequentar a escola, que são invisíveis quando as edições estão separadas. Essa metodologia possibilitou a investigação sobre a

diferença entre as taxas de indivíduos sem frequentar a escola das bases Longitudinal 1 e 2. Para isso, construímos (através dos dados públicos *online*) uma terceira base de dados filtrando pela mesma coorte da Base Longitudinal 2. Essa terceira base foi denominada de Base Longitudinal 3.

Com o intuito de estabelecer uma comparação, procedemos à imputação na Base Longitudinal 3, adotando uma abordagem metodológica a mais próxima possível daquela utilizada pelo Inep⁴¹⁴². As trajetórias passíveis de imputação de acordo com a nota técnica do Inep são:

No ano t o aluno i está matriculado em uma das etapas de ensino de interesse, deixa de frequentar a escola por um ou mais anos e volta a frequentar (ano $t+n$) em etapa de ensino, tal que, comparada com a etapa inicial, indicaria uma trajetória regular do aluno na escola, isto é, sem evasão ou repetência.

No ano t o aluno i está matriculado em uma das etapas de ensino de interesse, deixa de frequentar a escola por dois ou mais anos e volta a frequentar (ano $t+n$) em etapa pelo menos duas séries acima da etapa inicial, de tal forma que se presume que o aluno estava matriculado em pelo menos um dos anos do período de ausência, mas que sua trajetória escolar não foi regular, isto é, há ao menos uma evasão ou repetência no período.

No ano t o aluno i está matriculado em uma das etapas de ensino de interesse, possui situação final de reprovado no mesmo ano, deixa de frequentar a escola por um ano e volta a frequentar em etapa superior.

O aluno i aparece pela primeira vez na base após 2007 em etapa igual ou superior ao terceiro ano do fundamental com evidência de inconsistência. Para alunos com idade adequada para a etapa, a situação apresentada foi considerada como inconsistente e para aqueles com distorção idade-série, foi avaliado se o período em que o aluno permaneceu não matriculado correspondia ou não a uma situação de inconsistência (comparando esse período com o número de anos de distorção idade-série);

No ano t o aluno i está matriculado na creche ou na pré-escola deixa de frequentar a escola por um ou mais anos e volta a frequentar (ano $t+n$; $n>1$) em etapa igual ou superior ao terceiro ano do ensino fundamental.

Com o objetivo de facilitar a compreensão das trajetórias passíveis de imputação, elaboramos exemplos hipotéticos para cada uma dessas possibilidades. Apresentamos os exemplos em Tabela 7 e Tabela 8. Consideramos que, em 2009, a situação ideal para os cinco

⁴¹ Disponível em

https://download.inep.gov.br/informacoes_estatisticas/indicadores_educacionais/2007_2016/nota_tecnica_taxas_transicao_2007_2016.pdf

⁴² Houve algumas limitações pela ausência de variáveis nos dados públicos.

alunos era estarem matriculados no 2ºano do ensino fundamental, no 3º ano em 2010, no 4ºano em 2011 e no 5ºano em 2012. Na tabela 11, apresentamos a quantidade de alunos que foram imputados nos dados oficiais do Inep. Esses números mostram que trajetórias passíveis de imputação nas situações 1 e 4 têm maiores impactos.

Tabela 7 - Exemplo hipotético de estudantes sem registro

Situação	2009	2010	2011	2012
1	Matrícula no 2ºano com aprovação	Sem registro	Sem registro	5ºano
2	Matrícula no 2ºano com aprovação	Sem registro	Sem registro	Sem registro
3	Matrícula no 2ºano com reprovação	Sem registro	4ºano	
4	Sem registro	Sem registro	2ºano	
5	Matrícula na Pré-escola	Sem registro	Sem registro	5ºano

Fonte: Elaboração própria

Tabela 8 - Exemplo hipotético de imputação

Situação	2009	2010	2011	2012
1	2ºano	Imputação	Imputação	5ºano
2	2ºano	Imputação	Imputação	Sem registro
3	2ºano/reprovação	Imputação	4ºano	
4	Imputação	Imputação	2ºano	
5	Pré-escola	Imputação	Imputação	5ºano

Fonte: Elaboração própria

Tabela 9 - Casos imputados no Censo Escolar 2007-2015

Situação	n
1	2.678.170
2	159.747
3	262.937
4	2.326.464
5	191.537

Fonte: Elaboração própria a partir de dados da Nota Técnica /Inep

Para replicar a imputação, fixamos as etapas ideais para cada ano escolar de acordo com a idade da coorte, a cada ano. Se no ano $t+n$ o aluno estava na etapa adequada e ausente no ano t , a etapa adequada foi imputada para o ano t . Essa metodologia se aproxima para todas as situações acima, exceto a situação 3. Em seguida, Tabela 10, Tabela 11, Tabela 12, e Tabela 13 são exemplos reais de quatro alunos imputados em nossa Base Longitudinal 3. As informações em negrito são as imputadas.

Tabela 10 - Banco original com as etapas do aluno em cada ano

ID	2009	2010	2011	2012	2013	2014	2015	2016	2017
1	NA	2ºano	3ºano	4ºano	5ºano	6ºano	7ºano	8ºano	9ºano
2	2ºano	3ºano	4ºano	5ºano	NA	NA	NA	9ºano	1ºEM
3	1ºano	NA	NA	NA	NA	6ºano	7ºano	8ºano	9ºano
4	2ºano	3ºano	NA	NA	NA	NA	NA	NA	1ºEM

Fonte: Elaboração própria a partir de dados do Censo Escolar/Inep

Tabela 11 – Banco imputado com as etapas do aluno em cada ano

ID	2009	2010	2011	2012	2013	2014	2015	2016	2017
1	1ºano	2ºano	3ºano	4ºano	5ºano	6ºano	7ºano	8ºano	9ºano
2	2ºano	3ºano	4ºano	5ºano	6ºano	7ºano	8ºano	9ºano	1ºEM
3	1ºano	2ºano	3ºano	4ºano	5ºano	6ºano	7ºano	8ºano	9ºano
4	2ºano	3ºano	4ºano	5ºano	6ºano	7ºano	8ºano	9ºano	1ºEM

Fonte: Elaboração própria a partir de dados do Censo Escolar/INEP

Tabela 12 - Banco original com o fluxo do aluno em cada ano

ID	2009	2010	2011	2012	2013	2014	2015	2016
1	Não frequentaram a escola	Promovido						
2	Promovido	Promovido	Promovido	Evadido	Não frequentaram a escola	Não frequentaram a escola	Não frequentaram a escola	Promovido
3	Evadido	Não frequentaram a escola	Promovido	Promovido	Promovido			
4	Promovido	Evadido	Não frequentaram a escola					

Fonte: Elaboração própria a partir de dados do Censo Escolar/INEP

Tabela 13 - Banco imputado com o fluxo do aluno em cada ano

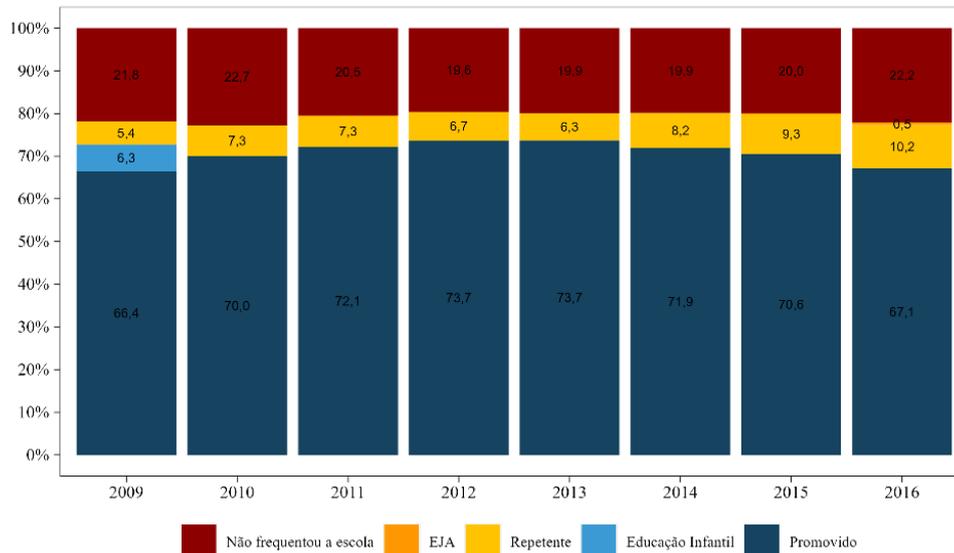
ID	2009	2010	2011	2012	2013	2014	2015	2016
1	Promovido							
2	Promovido							
3	Promovido							
4	Promovido							

Fonte: Elaboração própria a partir de dados do Censo Escolar/INEP

A seguir apresentamos os gráficos 14, 15 e 16 construídos com as bases longitudinais 1, 2 e 3. Filtramos a base longitudinal 1 (que não foi imputada) pela coorte das bases 2 e 3, ou seja, os três gráficos são referentes a mesma coorte. Ao compararmos esses gráficos, observamos que os resultados a partir dos dados oficiais (L2) apresentam taxas mais elevadas de alunos promovidos e maiores taxas de reprovação. Enquanto mesmo com imputação

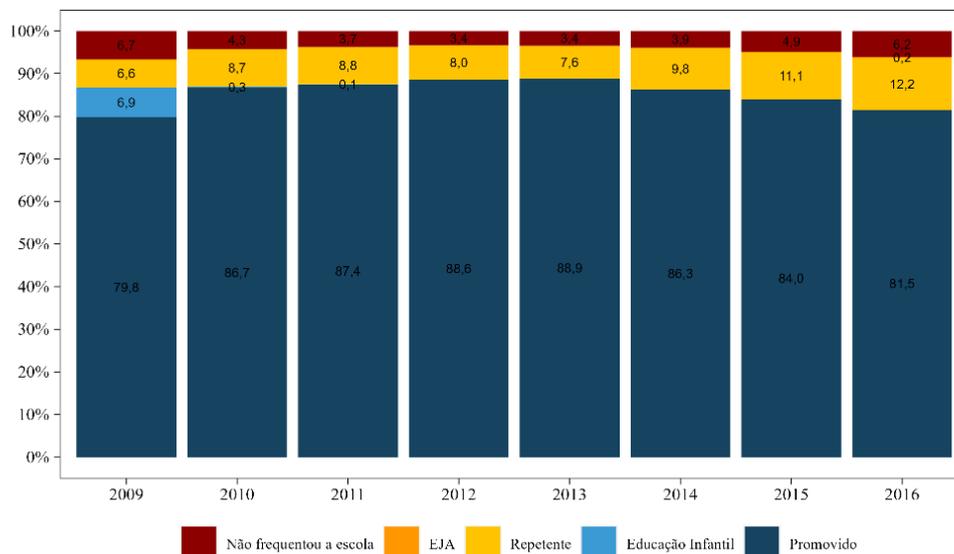
realizada a base 3 (com os dados públicos) apresentam maiores taxas de indivíduos sem frequentar a escola do que de reprovados.

Gráfico 14 - Trajetórias escolares no ensino fundamental (2009 - 2016). Base Longitudinal I (coorte 99/00 - dados públicos – abrangência 2007 - 2017).



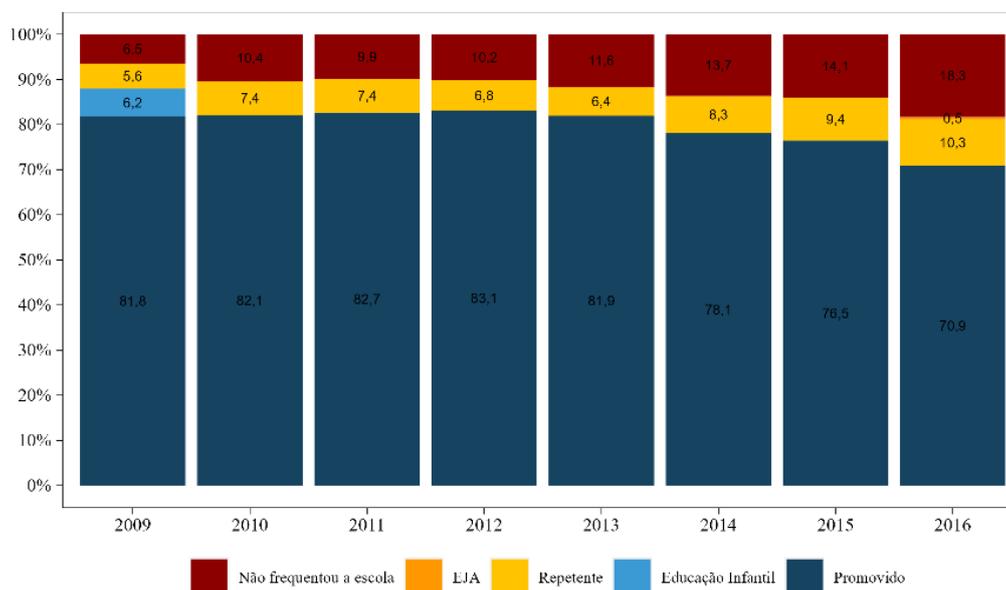
Fonte: elaboração própria a partir de dados do Censo Escolar/Inep

Gráfico 15 - Trajetórias escolares no ensino fundamental (2009 - 2016). Base Longitudinal II (coorte 02/03 - dados protegidos - abrangência 2009 - 2020).



Fonte: elaboração própria a partir de dados do Fluxo Escolar/Inep

Gráfico 16 - Trajetórias escolares no ensino fundamental (2009 - 2016). Base Longitudinal III (coorte 02/03 - dados públicos – abrangência 2007 - 2017).



Fonte: elaboração própria a partir de dados do Censo Escolar/Inep

1.2. As imputações possuem um perfil específico?

Traçamos o perfil dos alunos que foram imputados na base longitudinal 3 para verificar se existe algum padrão socioeconômico dos casos imputados ou se a distribuição é aleatória. Calculamos as porcentagens imputadas em cada grupo para as variáveis raça/cor, dependência administrativa da escola e sexo. A variável sexo não apresentou variação. De acordo com a Tabela 14 constatamos que 27% dos estudantes com imputação não tinham declaradas também as informações de raça/cor. Já a Tabela 15 revela que a maioria dos estudantes com imputação (59,9% do total dos imputados) pertence às escolas federais. Esses resultados evidenciam a existência de um perfil distinto entre os alunos que passaram pelo processo de imputação de informações.

Tabela 14 - Perfil socioeconômico dos casos imputados

Raça/cor	Base original	Base imputada	% de casos imputados
Não declarada	757.273	204.643	27,0
Branca	1.093.611	191.162	17,5
Preta	119.785	26.223	21,9
Parda	1.595.189	358.803	22,5
Amarela	21.832	4.942	22,6
Indígena	29.919	6.917	23,1

Fonte: Elaboração própria a partir de dados do Censo Escolar/Inep

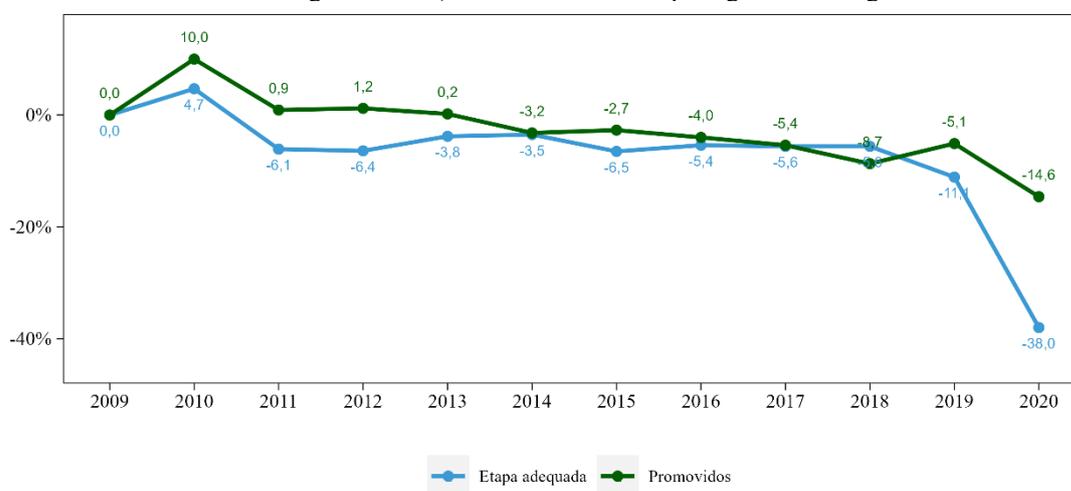
Tabela 15 - Dependência administrativa das escolas dos casos imputados

Dependência Administrativa da escola	Base original	Base imputada	% de casos imputados
Federais	2.689	1.612	59,9
Estaduais	484.605	172.509	35,6
Municipais	2.353.149	376.804	16,0
Privadas	777.166	241.765	31,0

Fonte: Elaboração própria a partir de dados do Censo Escolar/Inep

O gráfico 17 foi apresentado no capítulo anterior. Ele é retomado aqui para apresentarmos a hipótese de que é preciso testar se os tratamentos de imputação dados ao Censo da educação resulta em uma redução das taxas de “indivíduos em idade escolar que não frequentaram a escola”. Isso porque os anos em que a etapa adequada menos reduz são anos de finalização do Ensino Fundamental I e II, nos quais as reclassificações para etapas adequadas são mais prováveis. Ou seja, é preciso testar se as imputações estão ocorrendo sobre alunos de determinado perfil socioeconômico que não estavam frequentando a escola.

Gráfico 17 - Alunos promovidos e em etapas adequadas em relação ao ano anterior (%).
Base Longitudinal II (coorte 02/03 - dados protegidos - abrangência 2009 - 2020).



Fonte: elaboração própria a partir de dados do Fluxo Escolar/Inep

2. Estar na escola como dimensão da qualidade educacional

O Índice de Desenvolvimento da Educação Básica (Ideb) é um indicador de qualidade educacional que combina informações sobre o desempenho dos estudantes em testes

padronizados⁴³ e rendimento escolar⁴⁴. No entanto, a literatura tem discutido sobre a limitação do indicador em abranger todos os componentes da qualidade educacional e em ser insensível às trajetórias irregulares dos estudantes, deixando a dimensão da evasão escolar escapar. Como resultado, pesquisadores têm buscado caminhos para que essa dimensão seja incorporada ao indicador (Alves, Oliveira, *et al.*, 2021; Fonseca, Rodrigues, *et al.*, 2022). A proposta do monitoramento da permanência dos estudantes de Fonseca, Rodrigues, *et al.* (2022) identificando a categoria *trajetória interrompida* é um avanço nessa direção. Para os autores, “A permanência dos estudantes nas escolas é tanto um processo como um resultado abrangente da escolarização. Pode ser entendida como uma proxy para a ‘oportunidade de aprender’ [...] na medida em que, no contexto da educação escolar, não há aprendizado para aqueles que estão fora da escola” (p. 05).

A proposta a seguir concorda com a perspectiva da literatura e pretende ser mais um passo na inclusão dos alunos ausentes da escola no Ideb. O que propomos é uma estratégia de mensuração da dimensão *indivíduos sem frequentar a escola* através dos IDs dos alunos no Censo Escolar que pode ser incorporada (em nível municipal) ao Ideb. As taxas podem ser calculadas para os municípios através das coortes de nascimento, como exemplificamos na Tabela 16. E para isso, a discussão sobre a imputação, argumentada na seção anterior, importa.

2.1. Proposta

Nossa proposta consiste em mensurar as taxas de indivíduos em idade escolar que estão sem frequentar a escola, por ano, para cada município. Para ilustrar a metodologia da nossa proposta utilizamos a Base Longitudinal 1. Medimos a taxa indivíduos em idade escolar que estão sem frequentar a escola no município de Belo Horizonte – MG considerando o ensino fundamental I e o ano de 2014. Essa análise exige cinco coortes, sendo selecionados os nascidos entre 01/04/2003 e 31/03/2004; 01/04/2004 e 31/03/2005; 01/04/2005 e 31/03/2006; 01/04/2006 e 31/03/2007; 01/04/2007 e 31/03/2008. A Tabela 16 mostra a idade dos estudantes ao longo dos anos e as etapas nas quais eles deveriam estar, se regulares.

⁴³ Prova Brasil ou Saeb.

⁴⁴ O rendimento escolar é calculado considerando se os estudantes foram ou não aprovados ao final do ano letivo.

Tabela 16 - Taxa de indivíduos em idade escolar sem frequentar a escola - ensino fundamental I - Belo Horizonte

Ano	2010	2011	2012	2013	2014	2015	2016	2017
Etapa Ideal	1º ano EF	2º ano EF	3º ano EF	4º ano EF	5º ano EF	6º ano EF	7º ano EF	8º ano EF
Idade da Coorte 1	6 anos	7 anos	8 anos	9 anos	10 anos	11 anos	12 anos	13 anos
Idade da Coorte 2		6 anos	7 anos	8 anos	9 anos	10 anos	11 anos	12 anos
Idade da Coorte 3			6 anos	7 anos	8 anos	9 anos	10 anos	11 anos
Idade da Coorte 4				6 anos	7 anos	8 anos	9 anos	10 anos
Idade da Coorte 5					6 anos	7 anos	8 anos	9 anos

Fonte: Elaboração própria.

Via *pivotagem*, construímos uma base de dados para cada coorte considerando as etapas de ensino nas quais os estudantes estavam em cada ano e o código da cidade das escolas frequentadas pelos estudantes. Os alunos que em 2014 estavam matriculados em escolas de cidades que não eram Belo Horizonte foram excluídos da base. Para fazer a agregação municipal nos termos de Fonseca, Rodrigues, *et al.* (2022), consideramos o último município em que o aluno esteve matriculado e retiramos da base 0.9% de estudantes que estavam matriculados em outras cidades.

A Tabela 17 apresenta exemplos da coorte 1. Os casos 3, 5 e 11 são exemplos desse filtro. Os estudantes que não tiveram cidades vinculadas até 2014 como nos casos 13 e 14 também foram excluídos. Esse filtro retirou da base 4,6% dos estudantes. Isso acontece porque eles foram matriculados na escola, pela primeira vez, após 2014. No final, a base passou de 35,911 para 33,924 alunos.

Tabela 17 - Exemplo de atribuição de municípios

Aluno	2010	2011	2012	2013	2014	2015	2016	2017	Município de matrícula - 2010	Município de matrícula - 2011	Município de matrícula - 2012	Município de matrícula - 2013	Município considerado para 2014
1	Matriculado	Sem frequentar a escola	Sem frequentar a escola	Sem frequentar a escola	Matriculado	Matriculado	Matriculado	Matriculado	BH	Sem Informação	Sem Informação	Sem Informação	BH
2	Matriculado	Sem frequentar a escola	BH	Sem Informação	Sem Informação	Sem Informação	BH						
3	Matriculado	Matriculado	Matriculado	Matriculado	Sem frequentar a escola	Matriculado	Matriculado	Sem frequentar a escola	BH	Santa Bárbara	Santa Bárbara	Santa Bárbara	Santa Bárbara
4	Matriculado	Matriculado	Sem frequentar a escola	BH	BH	Sem Informação	Sem Informação	BH					
5	Matriculado	Matriculado	Sem frequentar a escola	Matriculado	Matriculado	BH	Contagem	Sem Informação	Sem Informação	Contagem			
6	Matriculado	Matriculado	Matriculado	Sem frequentar a escola	BH	BH	BH	Sem Informação	BH				
7	Matriculado	Matriculado	Matriculado	Matriculado	Sem frequentar a escola	Sem frequentar a escola	Matriculado	Matriculado	BH	BH	BH	BH	BH
8	Matriculado	Matriculado	Matriculado	Matriculado	Sem frequentar a escola	Matriculado	Matriculado	Sem frequentar a escola	BH	BH	BH	BH	BH
9	Matriculado	Sem frequentar a escola	Matriculado	Matriculado	BH	Sem Informação	Sem Informação	Sem Informação	BH				
10	Matriculado	Matriculado	Matriculado	Matriculado	Sem frequentar a escola	Sem frequentar a escola	Matriculado	Sem frequentar a escola	BH	BH	BH	BH	BH
11	Sem frequentar a escola	Sem frequentar a escola	Matriculado	Matriculado	Sem frequentar a escola	Sem Informação	Sem Informação	BH	Ourém	Ourém			
12	Matriculado	Marituba	Marituba	Marituba	BH	BH							

13	Sem frequentar a escola	Matriculado	Matriculado	Matriculado	Sem Informação								
14	Sem frequentar a escola	Matriculado	Matriculado	Matriculado	Sem Informação								

Fonte: Elaboração própria a partir de dados do Censo Escolar/INEP

A Tabela 18 mostra as taxas de flutuação entre escolas, em 2014, no município de Belo Horizonte. É possível analisar os resultados por coorte ou pela taxa total. Por exemplo, a taxa de alunos sem frequentar a escola para a coorte 2 é de 6.5% enquanto a taxa total é de 6,2%. Essa taxa é uma possibilidade para o Ideb.

Tabela 18 - Taxa de flutuação entre escolas em Belo Horizonte em 2014

Coorte	1	2	3	4	5	Total
Matrículas	33.925	32.273	32.354	31.480	29.352	159.384
Sem registro	2.467	2.083	2.272	2.131	915	9.868
Taxa	7,3	6,5	7	6,8	3,1	6,2

Fonte: Elaboração própria a partir de dados do Censo Escolar/INEP

3. Considerações

Esse capítulo elabora duas hipóteses: 1) a imputação de casos ausentes em bases de dados educacionais pode ter um efeito sobre as taxas de alunos que não frequentaram a escola; 2) os casos que recebem a imputação tem um perfil específico. É importante que as instituições responsáveis por bases de dados educacionais façam divulgação das porcentagens de casos imputados como também de testes sobre a aleatoriedades dessas imputações.

CAPÍTULO III – REGULARIDADES SOCIOLÓGICAS, EFEITOS DESPROPORCIONAIS E AS TRAJETÓRIAS ESCOLARES⁴⁵

O objetivo deste capítulo é entender quais são os efeitos das desvantagens de origem socioeconômica sobre as trajetórias escolares no Brasil. Essa pergunta parte da premissa que há uma forte associação, amplamente comprovada pela literatura, entre desvantagens de origem socioeconômica e desigualdades educacionais. Assim, esperamos encontrar que o nível socioeconômico, a raça, o gênero do estudante, a rede de ensino e a região da escola frequentada estão associados às desvantagens escolares (Bourdieu e Passeron, 1975; Willms, 1992; Coleman e others, 1966; Patto, 1990; Alves, Soares e Xavier, 2016).

Para entender tal relação construímos um modelo logístico multinomial que tem como variável resposta a variável de trajetórias escolares na educação básica (I + II + ensino médio) - construída no capítulo 1 - e como explicativas cinco variáveis sociodemográficas. Conforme apresentamos nos capítulos anteriores, a variável de trajetórias escolares é uma síntese do fluxo escolar⁴⁶. E esse, por sua vez, é a condição do aluno em determinado ano escolar considerando sua transição entre esse ano escolar e o ano escolar seguinte.

Tanto na literatura internacional quanto na nacional há muitos estudos sobre esses componentes das trajetórias escolares (Oliveira e Araújo, 2015; Oliveira, 2007; Klein e Ribeiro, 1991; Ferrão, 2015; Barros e Mendonça, 1998; Faria, 2011; Almeida e Xavier, 2021; Roderick, 1994; Patto, 1990). Além de diversas discussões sobre o acesso à escola, que é condição elementar para que a trajetória escolar aconteça. Todos esses estudos são consensuais com a premissa das regularidades sociológicas e foram, em parte, o referencial teórico para a construção de nosso modelo analítico. Essa construção também foi pensada a partir dos marcos legais que moldaram essas questões no contexto brasileiro ao longo do tempo e variaram entre abordagens mais tímidas e outras mais audaciosas.

O capítulo está dividido em quatro seções. A primeira apresenta os marcos legais da legislação brasileira que atravessam a temática enquanto a segunda apresenta nosso referencial teórico. A intenção é demonstrar onde amparamos nossos argumentos. A terceira seção reporta nosso modelo, seus resultados e análises. Por último, na quarta seção, analisamos esses resultados considerando os pactos internacionais acordados na Agenda 2030, na qual o Brasil é

⁴⁵ A base de dados utilizada neste capítulo foi a base longitudinal 2.

⁴⁶ Definição no capítulo I, pag.38.

signatário. Os resultados mostram que as trajetórias educacionais brasileiras estão na contramão dos marcos legais educacionais e dos pactos internacionais e corroborando a literatura especializada.

1. Acesso, reprovação, abandono, evasão e seus marcos legais

1.1. Anos 80, redemocratização e a democratização da escola

As Constituições brasileiras de 1824, 1891, 1934, 1937, 1946 e 1967 mencionavam o ensino gratuito, embora as escolas primárias públicas fossem escassas e milhões de brasileiros fossem analfabetos (Ferraro, A. R., 2008). Entretanto, foi somente com a promulgação da Constituição Federal Brasileira de 1988 (CF88) que a educação foi reconhecida como um direito social (Art. 6º), universal e responsabilidade do Estado e da família (Art. 205º), baseada nos princípios de igualdade de oportunidades para acesso e permanência, assegurando o ensino público gratuito e de qualidade (Art. 206º).

O texto original da Constituição Federal de 1988 estabeleceu a prioridade que o país daria ao ensino fundamental, tornando-o obrigatório e gratuito. Para o ensino médio, nessa primeira redação, ficou prevista uma progressiva extensão da obrigatoriedade e gratuidade (Art. 208º). O §1º do art. 208º garante o acesso ao ensino obrigatório e gratuito como um direito público subjetivo, seguido pelo §2º, que coloca a responsabilidade da oferta regular do ensino obrigatório sob a autoridade competente. Esses incisos do artigo 208º não foram alterados desde o texto inicial, mantendo o ensino obrigatório como um direito e sua oferta como dever do Estado desde 1988 até os dias atuais.

1.2. Anos 90 e o Consenso

1.2.1. O Plano Decenal de Educação para Todos (1993-2003)

As tendências políticas educacionais da década de 90 foram marcadas pelos problemas de fluxo iniciados na década anterior e pelo contexto internacional das relações constituídas após o Consenso de Washington. Em **1990**, diversos países foram convocados por instituições como UNESCO, UNICEF e Banco Mundial para participarem da Conferência Mundial de Educação para Todos, realizada na cidade de Jomtien - Tailândia. Esse encontro resultou na Declaração de Jomtien, na qual o Brasil integrou o grupo dos signatários comprometendo-se a

assegurar à sua população o direito à educação. Três anos depois, com referência nos acordos de Jomtien, o Brasil elaborou o Plano Decenal de Educação para Todos (PDET) **1993-2003**.

O PDET colocou a educação fundamental no centro de seu debate. O plano trouxe discussões sobre falta de acesso à escola básica; desigualdades sociais com impacto nas condições de acesso à escola e na expansão da escolaridade; a pobreza das famílias e a dificuldade dos estudantes na conclusão das séries iniciais; a heterogeneidade dos padrões de oferta escolar que geravam acúmulo de repetências, a evasão e dificuldades de conclusão do ensino obrigatório; a incapacidade do sistema educacional em associar o acesso, permanência, qualidade e equidade entre os alunos. O Plano também apresentou estratégias para a universalização da educação fundamental e metas para em dez anos elevar em 94% a cobertura da população em idade escolar, assegurar a melhoria do fluxo escolar e reduzir as repetências, sobretudo na 1ª e 5ª séries, para que 80% das gerações escolares, do final do período concluíssem o ensino fundamental com bom aproveitamento.

O plano traçou linhas de ação estratégica para melhorias do acesso e da permanência reconhecendo que os alunos permanecerem na escola com sucesso escolar era o maior desafio do final do século XX, e ressaltou que o aperfeiçoamento dos processos de avaliação escolar contribuiria para estimular o progresso do aluno e a superação da *cultura da repetência*.

1.2.2. LDB

Em 1996⁴⁷ foi estabelecida a Lei de Diretrizes e Bases da Educação (LDB). A conexão entre essa lei e acesso escolar está no alargamento das etapas de escolarização. A LDB estabeleceu nova redação relativas ao acesso à escola nos incisos I e II do art. 208 da CF88. O inciso I passou a ser “*garantia de ensino fundamental, obrigatório e gratuito, passar a ser assegurada, inclusive, para os que a ele não tiveram acesso na idade própria*” enquanto o texto original era “*garantia de ensino fundamental, obrigatório e gratuito, inclusive para os que a ele não tiveram acesso na idade própria*”. O inciso II passa de “*garantia de progressiva extensão da obrigatoriedade e gratuidade ao ensino médio*” para “*garantia de progressiva universalização do ensino médio gratuito*”.

No contexto de um sistema educacional congestionado e influenciado pela agenda internacional, a partir de **1995** foram desenvolvidas estratégias em escala nacional, sob a

⁴⁷ LDB já haviam sido estabelecidas em 1961 e 1971, porém tiveram pouco ou nenhum impacto.

coordenação do Ministério da Educação, para a expansão das ações de **correção de fluxo escolar** em estados e municípios brasileiros (Parente e Lück, 2004).

Os **ciclos escolares** organizam o tempo escolar em unidades maiores e mais flexíveis do que o ano letivo do sistema seriado enquanto os **regimes de progressão** dos alunos são **regular, parcial e continuada**. Embora na LDB não tenha as definições do que consistem as progressões, através da legislação educacional dos Estados é possível entender que na **progressão regular** o aluno é promovido de um ano/série para o seguinte em sequência, com possibilidade de reprovação ao final de cada ano letivo se houver baixo desempenho escolar ou abandono; **progressão parcial** permite que o aluno avance para o ano/série seguinte mesmo que tenha havido retenção em alguns¹⁵ componentes curriculares, desde que cumpram dependência; na **progressão continuada** a prática de reprovação dos alunos segundo o critério de desempenho é impedida ao final de determinados anos/séries escolares. É comum a progressão continuada nos anos/séries escolares ser adotada em conjunto com o sistema de ciclo (Machado, 2020).

Essas medidas resultaram em significativa redução nos altos percentuais de matrículas na 1ª série do ensino fundamental, em distribuição mais equitativa nas séries seguintes e em melhora do fluxo escolar (Oliveira e Araújo, 2015; Machado, 2020). Porém, como já vimos e ainda veremos nas próximas seções o Brasil mantém altas taxas de reprovação.

Em 27 de outubro de 2009, a lei nº 12.061 alterou a própria LDB nos incisos II do art. 4º e VI do art. 10. O intuito era assegurar o acesso de todos os interessados ao ensino médio público. A nova redação do inciso II passou a ser “universalização *do ensino médio gratuito*” substituindo “*a progressiva extensão da obrigatoriedade e gratuidade ao ensino médio*”. E o inciso VI foi substituído de “*assegurar o ensino fundamental e oferecer, com prioridade, o ensino médio*” por “*assegurar o ensino fundamental e oferecer, com prioridade, o ensino médio a todos que o demandarem*”.

A preocupação com a regularização do fluxo permeia o texto da LDB/1996 com políticas de não repetência como adoção de ciclos de escolarização, promoção continuada, programas de aceleração da aprendizagem, recuperação paralela, reclassificação, entre outras medidas. As referências aos ciclos estão nos artigos 23, 24 e 32. É importante notar que todos eles permanecem em sua redação original de 1996. O art. 23º define:

Art. 23. A educação básica poderá organizar-se em séries anuais, períodos semestrais, ciclos, alternância regular de períodos de

estudos, grupos não-seriados, com base na idade, na competência e em outros critérios, ou por forma diversa de organização, sempre que o interesse do processo de aprendizagem assim o recomendar.

§ 1º A escola poderá reclassificar os alunos, inclusive quando se tratar de transferências entre estabelecimentos situados no País e no exterior, tendo como base as normas curriculares gerais.

Já o inciso II do art. 24º estabelece

II - a classificação em qualquer série ou etapa, exceto a primeira do ensino fundamental, pode ser feita:

a) por promoção, para alunos que cursaram, com aproveitamento, a série ou fase anterior, na própria escola;

b) por transferência, para candidatos procedentes de outras escolas;

c) independentemente de escolarização anterior, mediante avaliação feita pela escola, que defina o grau de desenvolvimento e experiência do candidato e permita sua inscrição na série ou etapa adequada, conforme regulamentação do respectivo sistema de ensino;

III - nos estabelecimentos que adotam a progressão regular por série, o regimento escolar pode admitir formas de progressão parcial, desde que preservada a sequência do currículo, observadas as normas do respectivo sistema de ensino;

IV - Poderão organizar-se classes, ou turmas, com alunos de séries distintas, com níveis equivalentes de adiantamento na matéria, para o ensino de línguas estrangeiras, artes, ou outros componentes curriculares;

V - a verificação do rendimento escolar observará os seguintes critérios:

a) avaliação contínua e cumulativa do desempenho do aluno, com prevalência dos aspectos qualitativos sobre os quantitativos e dos resultados ao longo do período sobre os de eventuais provas finais;

b) possibilidade de aceleração de estudos para alunos com atraso escolar;

c) possibilidade de avanço nos cursos e nas séries mediante verificação do aprendizado;

d) aproveitamento de estudos concluídos com êxito; e) obrigatoriedade de estudos de recuperação, de preferência paralelos ao período letivo, para os casos de baixo rendimento escolar, a serem disciplinados pelas instituições de ensino em seus regimentos;

Enquanto o art. 32º estabelece que “é facultado aos sistemas de ensino desdobrar o ensino fundamental em ciclos” como também “[...] estabelecimentos que utilizam progressão regular por série podem adotar no ensino fundamental o regime de progressão continuada [...]” (Brasil, 1996).

1.3. Virada do milênio e algumas viradas no Brasil

1.3.1. Planos, Pactos e a emenda 59

A partir dos anos 2000 teve início o surgimento de planos e pactos federativos que apresentaram políticas educacionais mais amplas. Os documentos mostram que o país começa a entender que acesso, reprovação, abandono, evasão, qualidade e equidade precisam ser tratadas em conjunto. Nesses marcos legais do novo milênio não faltam menções sobre acesso ao ensino gratuito, educação como direito social de todos, igualdade de condições de acesso, permanência e a conclusão na idade adequada. Também são inúmeras as referências ao combate das repetências, da evasão e da melhoria do fluxo escolar acompanhada da qualidade e da equidade. Vejamos:

1.3.2. O PNE 2001-2010

Previsto e obrigatório nos artigos 214º da CF88 e 9º da LDB, o Plano Nacional de Educação (PNE) **2001-2010** foi regulamentado pela Lei nº 10.172/2001. Seus objetivos e prioridades foram: aumentar a escolaridade da população; melhorar a qualidade do ensino em todos os níveis; reduzir as desigualdades sociais e regionais em acesso e permanência com sucesso; garantir o ensino fundamental obrigatório para todos com idade entre 7 e 14 anos, assegurando ingresso, permanência, conclusão (Brasil, 2001). O plano apresentou diagnósticos e metas para o ensino fundamental e médio. Os diagnósticos para o ensino fundamental foram a distorção idade-série causada por altos índices de reprovação e evasão. As diretrizes eram atingir a universalização com permanência e qualidade para que houvesse conclusão.

O plano reconhecia a necessidade de políticas educacionais destinadas à correção das distorções idade-série e tinha entre as metas a regularização do fluxo escolar. O objetivo era reduzir em 50% as taxas de repetência e evasão através de programas de aceleração da aprendizagem e recuperação paralela que deveria garantir a aprendizagem. Para o ensino médio,

o diagnóstico ressaltou o reduzido acesso à etapa causado principalmente pelas baixas taxas de conclusão do ensino fundamental. As metas eram expansão da oferta, melhoria da qualidade, correção do fluxo, infraestrutura que assegurasse o oferecimento de vagas para 100% da demanda e redução de repetência e evasão. Outro ponto importante a ser notado no plano é a preocupação com o estabelecimento de um sistema de avaliação para ensino fundamental e médio (Brasil, 2001).

1.3.3. O PDE, o PAR e o TODOS

Durante a vigência do PNE 2001-2010, em 2007 foi criado o Plano de Desenvolvimento da Educação (PDE). O novo plano pertencia à esfera executiva federal e foi constituído por programas e divididos em eixos e um desses eixos foi a educação básica. O texto do PDE propõe um enlace entre educação, território e desenvolvimento com qualidade, equidade e potencialidade. O objetivo era reduzir as desigualdades sociais e regionais através da equalização das oportunidades de acesso à educação de qualidade.

Partindo do pressuposto de que todo processo educacional deve ser avaliado e do diagnóstico de que a “indústria da aprovação automática” e a “indústria da repetência” não fazem o sistema educacional avançar o PDE propôs a criação de um indicador de qualidade. Dessa proposta originou-se o Índice de Desenvolvimento da Educação Básica – o Ideb⁴⁸. O índice, como vimos no capítulo II, ainda utilizado e com previsões de melhorias, combina resultados de desempenho escolar fornecidos pela Prova Brasil com resultados de rendimento escolar apurados pelo Censo Escolar.

O plano também apresentou as enormes desigualdades regionais encontradas através da Prova Brasil, muitas vezes no interior do sistema de ensino, e discutiu as boas práticas de escolas e redes de ensino que resultam em aprendizagem satisfatória. O documento se preocupou com os números relativos à aprendizagem e com as altas taxas de evasão e repetência apresentando um forte entendimento de que o problema da qualidade só será resolvido se o problema da equidade também for enfrentado.

O PDE foi articulado pelo Governo Federal a outros dois planos de sustentação: o Plano de Metas Compromisso Todos pela Educação [o Todos] e o Plano De Ação Articulada [o PAR]. Esse acoplamento foi realizado pelo decreto nº 6.094/2007 (Brasil, 2007). O Plano de Metas

⁴⁸ Calculado pelo Instituto Nacional de Estudos e Pesquisas Nacionais (INEP), órgão vinculado ao MEC.

estabelecia pactos entre União, Estados, Distrito Federal e municípios em busca da melhoria da qualidade da educação básica através de 28 diretrizes dentre as quais se destacavam: a aprendizagem; o acompanhamento da frequência e do desempenho dos alunos, o combate à repetência e à evasão, o acompanhamento das metas de evolução do Ideb. No art. 3º do capítulo II há o detalhamento sobre como a qualidade da educação básica passaria a ser aferida, objetivamente, com base no Ideb. E em parágrafo único definiu que o Ideb seria o indicador objetivo para a verificação do cumprimento de metas fixadas no termo de adesão ao Compromisso.

O PAR foi um conjunto articulado de ações, que seria apoiado pelo MEC e visava o cumprimento das metas do Compromisso e a observância das suas diretrizes. Os municípios assumiriam o compromisso de melhorar a qualidade do ensino ofertada e o Ideb seria o indicador verificador do cumprimento de metas fixadas através de um termo de adesão – o Compromisso Todos pela Educação – regulado no Decreto nº 6.094/2007. O PAR recebeu adesão de todos os 26 estados, do Distrito Federal e de todos os municípios brasileiros.

1.3.4. A emenda nº 59 de 2009

Ainda em 2009, através da emenda constitucional nº 59, o inciso I do Art. 208 foi alterado para a redação “*garantia de educação básica obrigatória e gratuita dos 4 (quatro) aos 17 (dezessete) anos de idade, assegurada inclusive sua oferta gratuita para todos os que a ela não tiveram acesso na idade própria*”. E em 2013 os incisos I e IV do artigo 4 da LDB ganham novas redações, respectivamente: “*garantia de educação básica obrigatória e gratuita dos 4 (quatro) aos 17 (dezessete) anos de idade, organizada da seguinte forma: a) pré-escola; b) ensino fundamental; c) ensino médio*” e “*garantia de acesso público e gratuito aos ensinos fundamental e médio para todos os que não os concluíram na idade própria*”. Nessas condições, atualmente o estado brasileiro deve garantir à toda a população acesso à toda população o ensino obrigatório de 4 a 17 anos, ou seja, a educação infantil e o ensino médio passam a ser contemplados na educação básica em instituições públicas e de modo gratuito.

1.3.5. O PNE 2014-2024

O mais recente texto legal brasileiro que atravessa as discussões sobre trajetórias educacionais é o Plano Nacional de Educação (PNE) 2014-2024. Aprovado pela lei nº 13.005 de 25 de junho de 2014, o texto apresenta 20 metas para consolidar um sistema educacional capaz de concretizar o direito à educação em sua integralidade, resolver as dificuldades de

acesso e a permanência e reduzir as desigualdades. O texto retoma a CF88 e a LDB para argumentar que a concretização do direito à educação perpassa pela garantia de que todos tenham oportunidades de acesso à escola e condições propícias para a conclusão na idade certa das etapas de ensino com níveis satisfatórios de aprendizagem. Há um destaque para a importância do Inep no monitoramento e a avaliação do Plano. Três das 20 metas propostas estão diretamente ligadas as trajetórias educacionais. Vejamos:

A meta 2 possui dois objetivos: *universalizar o ensino fundamental de 9 anos para toda a população de 6 a 14 anos e garantir que até 2024 pelo menos 95% dos alunos concluam essa etapa na idade recomendada*. Em 2013, o percentual da população de 6 a 14 anos que frequentava a escola era de 98,4%, números que se mantiveram em 2020. Para o plano, a inclusão dos 1,6% restantes exigiria esforços mais específicos. Já o percentual de conclusão do ensino fundamental na idade recomendada (16 anos) foi de 66,7% em 2013 (Brasil, 2015) e de 82,4% em 2020 (OPNE, 2020). 2023

A meta 3 também possui dois objetivos: universalizar, até 2016, o atendimento escolar para toda a população de 15 a 17 anos e elevar, até 2024, a taxa líquida de matrículas do ensino médio para **85%**. Em 2013, o percentual da população de 15 a 17 anos que frequentava a escola era de 84,3%. Para atingir a meta era necessário o crescimento de 15 pontos percentuais, o que ainda não aconteceu. Em 2016, o atendimento escolar para toda a população de 15 a 17 anos foi de 90,8% e 94,5% em 2020. Enquanto o percentual da população de 15 a 17 anos que frequenta o ensino médio foi de 55,3% em 2013, 30 pontos distante da meta. Em 2020, com 75,4% dos jovens de 15 a 17 anos cursando o ensino médio estávamos a 9,9% da meta (BRASIL, 2015; OPNE, 2020).

A meta 7 tem como objetivo fomentar a qualidade da educação básica em todas as etapas e modalidades, com melhoria do fluxo escolar e da aprendizagem até 2021. A intenção era atingir as seguintes médias nacionais para o Ideb: a) para os anos iniciais do ensino fundamental 5,2 em 2015; 5,5 em 2017; 5,7 em 2019 e 6 em 2021. b) para os anos finais do ensino fundamental 4,7 em 2015; 5 em 2017; 5,2 em 2019 e 5,5 em 2021 c) para o ensino médio 4,3 em 2015; 4,7 em 2017; 5 em 2019 e 5,2 em 2021. A meta dos anos iniciais foi alcançada com média de 5,9 em 2019. Com média de 4,9 em 2019 para os anos finais e 4,2 em 2019 para o ensino médio essas etapas não alcançaram a meta (Brasil, 2014; OPNE, 2020).

2. Acesso, fluxo e desigualdades sociais

2.1. Acesso e desigualdades de origem

A expansão do sistema educacional no Brasil⁴⁹ possibilitou, ao longo do tempo, um acesso gradual da população à escola. As matrículas no ensino fundamental aumentaram gradativamente. Os números foram de 11,6 milhões em **1965**, 15,9 milhões em **1970**, 22,5 milhões em **1980**, 32,1 milhões em **1994**, 36 milhões em **1999** e como era de se esperar reduziram e estabilizaram entre **28,5** e **26,5** milhões em 2003 e 2021, respectivamente. Esses números foram consoantes com a nova constituição, com o caminho em direção à democratização de acesso à escola e com a aproximação da universalização do atendimento do ensino fundamental (Oliveira e Araújo, 2015; IBGE, 2023), o que é um avanço indiscutível.

De **1960 a 2005** a taxa de pessoas com mais de 25 anos de idade e menos de 4 anos de estudos reduziu de **75%** para **28%**, enquanto as taxas referentes as pessoas com 25 anos ou mais e ensino médio completo aumentaram de **1,1%** para **12,3%**. Em **2020**, o ensino fundamental se aproximou da universalização com **98%** das crianças e jovens de 6 a 14 anos frequentando as escolas e a taxa de matrícula no ensino médio foi de **75,4%** entre os jovens de 15 a 17 anos. O problema da oferta limitada do ensino fundamental foi corrigido através da construção de escolas, uma política educacional relativamente simples, e o acesso ao sistema escolar por diferentes grupos sociais aconteceu (OLIVEIRA, 2007).

Como a transição para uma etapa é condicionada a ter completado a etapa anterior, poderíamos esperar uma ampliação da conclusão do ensino fundamental e médio ao passo que o acesso escolar foi ampliado. Entretanto, não foi o que aconteceu para o Brasil. Apesar dos avanços, em números absolutos, o contingente que ainda não tem acesso à educação formal na idade adequada é bastante expressivo (Ribeiro, 2011, Todos Pela Educação, 2021) e as desigualdades sociais de origem marcam essa exclusão educacional.

Ao longo do tempo houve uma ampliação no número de pessoas que completaram os anos iniciais do ensino fundamental, e dentro desse grupo o percentual de conclusão do ensino fundamental completo também aumentou. Porém, as taxas de finalização do ensino médio não acompanharam as médias de aumento das etapas anteriores. A origem social e o tipo de escola frequentada são importantes para explicar quais grupos acessaram primeiro essas expansões das

⁴⁹ Em 1961, 1971 e 1982 foram realizadas as primeiras reformas na educação fundamental.

oportunidades escolares e as desigualdades de oportunidades geradas por essas diferenças (Ribeiro, 2011).

O acesso ao ensino fundamental aconteceu inicialmente para as populações com nível socioeconômico mais alto e posteriormente foi se expandindo para os mais baixos. Quando a parcela mais pobre da população acessou o ensino fundamental a estratificação social se deslocou para o próximo nível educacional, o ensino médio. Esse fenômeno é conceituado como desigualdade maximamente mantida. O ensino fundamental brasileiro é marcado por um aumento na acessibilidade com diminuição limitada das desigualdades enquanto o ensino médio, mesmo com a expansão da oferta, possui uma persistência das desigualdades na conclusão da etapa, o que favorece os jovens das classes mais altas (Brito, 2017).

2.2. Reprovação, repetência e desigualdades de origem

Se o final da década de 80 constitucionalizou o acesso, o início dos anos 90 foi marcado por outro fenômeno. O aumento do número de crianças ingressas no sistema escolar na idade adequada somado com a volta à escola daquelas anteriormente excluídas intensificou as desigualdades intraescolares (Oliveira, 2007) e teve início uma dificuldade de permanência que se tornou um entrave na escolarização dos brasileiros. A discussão sobre repetência ganha relevância no Brasil no início dos anos 90 e é acompanhada pela literatura nacional, com destaque para os clássicos *A pedagogia da repetência* e *O censo educacional e o modelo de fluxo: o problema da repetência*.

Enquanto o Ministério da Educação afirmava que os alunos brasileiros não concluíam os estudos porque abandonavam a escola Ribeiro (1991) e Klein e Ribeiro (1991) defenderam que a causa central da não conclusão era a excessiva taxa de repetência escolar, principalmente nas primeiras séries, o que foi conceituado pela literatura como problema de fluxo e não como uma questão de evasão. Havia um esforço de escolarização, entretanto a repetência era o fator de sua própria causa, em consequência disso, os alunos evadiam após tentarem permanecer na escola e acumularem várias repetências. As taxas de repetência se concentravam na primeira série, e essa etapa inicial era intransponível para muitos alunos pobres e gerava as evasões (Klein e Ribeiro, 1991; Ribeiro, 1991).

Os dados mostram que a probabilidade de reprovação, em geral, é maior para alunos que já foram retidos, e essa probabilidade aumenta com mais de uma reprovação. A cada repetência, o grupo de alunos repetentes é mais selecionado, o que por conseguinte amplia o

universo de estudantes de baixo desempenho ao redor daqueles que já são repetentes. Essa seleção impacta na motivação e autoestima dos estudantes e nas expectativas de sucesso escolar da escola sobre os alunos, a qual será menor para os repetentes agrupados. Todas essas condições juntas fazem com que esses alunos apresentem menor probabilidade de sucesso escolar e maiores chances de evasão, mesmo quando o background familiar é controlado (Barros e Mendonça, 1998; Faria, 2011; Ferrão, 2022).

A retenção o congestionamento do sistema educacional, é uma política injusta, ineficaz e cara para os recursos públicos (Klein e Ribeiro, 1991; Gomes-Neto e Hanushek, 1994; Ferrão, Beltrão e Santos, 2002). As comprovações desses argumentos contra a retenção são diversos: aumenta a probabilidade dos alunos repetentes abandonarem a escola em até 30%; a retenção precoce é 35 vezes maior entre os 10% mais pobres e é um forte preditor da retenção tardia; a maioria (51%) dos alunos reprovados no 1º ou 2º ciclos do ensino básico são reprovados no 3º e essa retenção precoce está associada ao nível socioeconômico; os ganhos dos alunos retidos durante o ano de retenção não são suficientes para nivelá-los aos promovidos e se dissipam em dois ou três anos; alunos retidos tem maior probabilidade de trabalharem em empregos com baixos salários; a experiência durante o ano retido não é diferente da experiência anterior à retenção (Roderick, 1994; Roderick e Nagaoka, 2005; Jacob e Lefgren, 2009; Correa, Bonamino e Soares, 2014; Ferrão, 2015; Ribeiro, Kasmirski, *et al.*, 2018).

Somadas com essas desvantagens estão as enturmações associada à repetência. Esse mecanismo que é enturmar os alunos reprovados nas mesmas salas de aula gera mais repetência e afeta injustamente os alunos mais desfavorecidos socialmente limitando suas oportunidades educacionais. Além disso, concentrar esses alunos com baixo nível socioeconômico e desempenho insuficiente em salas específicas é uma decisão da escola (Almeida e Xavier, 2021). A tendência são os estereótipos em torno dos alunos pobres serem usados pelos professores como justificativa para o fracasso escolar e os estereótipos do que é ser pobre e da pobreza circularem livremente pela escola, marcando as justificativas da reprovação (Patto, 1990). Somando nesse contexto, temos as turmas com alta concentração de alunos repetentes destinadas aos professores menos experientes e qualificados enquanto os mais qualificados e experientes são destinados para escolherem as turmas com baixas taxas de repetência e/ou bons padrões de comportamento (Paiva, 1998).

Noruega, Japão, Coreia, Islândia, Eslovênia, Reino Unido, Finlândia, República Eslováquia, República Tcheca, Dinamarca, Suécia, Estados Unidos, Nova Zelândia e Polônia

estão próximos de resolverem o problema de atraso escolar. No PISA de 2009, esses países tiveram, no máximo, 5% dos alunos declarando terem experimentado reprovações pelo menos uma vez. Ao passo que os países da América do Sul apresentam taxas muito elevadas. O Chile tem o menor índice (23%) que em comparação é alto, e Brasil⁵⁰ o maior (40%) o que nos localizou em segundo lugar, atrás apenas da Bélgica francesa (Crahay e Baye, 2013).

2.3. Abandono, evasão e desigualdades de origem

O abandono escolar acontece quando o aluno deixa de frequentar a escola durante o ano letivo. As consequências são a reprovação se o aluno retorna para a escola no ano seguinte ou a evasão, se não retorna. Ambas as situações são intercorrências nas trajetórias escolares. Em “*Grade Retention and School Dropout: Investigating the Association*” (Roderick, 1994) investiga a relação entre altas taxas de evasão e as altas taxas de repetência questionando como a retenção pode influenciar a conclusão do *high school*. Os achados demonstram que retenções entre o período do jardim de infância e a sexta série estão associadas com um aumento substancial nas chances de desistência. Uma variável explicativa importante dessa correlação é a idade inadequada na coorte porque estar inadequado na série aumenta as chances de desengajamento durante o *middle school* e posteriormente as chances da desistência.

Durante a década de **1980**, os alunos faziam tentativas de permanecerem no sistema escolar, porém deixavam de frequentar as escolas após acumularem retenções. No início de **1990**, a literatura apontaria que os anos iniciais eram intransponíveis para muitos alunos pobres. O sistema educacional estava permeado por uma naturalizada pedagogia da repetência que responsabilizava os alunos por seus **fracassos escolares** previamente determinados por essa pedagogia fortemente marcada por variáveis socioeconômicas (Klein e Ribeiro, 1991; Ribeiro, 1991; Patto, 1990; Oliveira e Soares, 2012).

2.4. Efeito de pares

Além dessas desigualdades socioeconômicas e demográficas em nível individual a literatura aponta mais uma desigualdade que é formada por essas individuais no formato agregado, o efeito de pares. Essa variável também é conhecida como composição social ou segregação, ou ainda NSE da escola e está relacionada ao contexto da escola. Se os alunos estão

⁵⁰ Análise comparativa entre os países da Organização para a Cooperação e Desenvolvimento Econômico (OCDE) e países convidados para o Programa Internacional de Avaliação de Estudantes (PISA) de 2009.

racial e/ou economicamente segregados por escolas, há desigualdades socioeconômicas entre escolas e isso amplia as desigualdades entre os estudantes.

Escolas com muitas matrículas de nível socioeconômico elevado têm maior composição social e com isso tendem a receber mais recursos. Essas escolas também têm um clima acadêmico mais adequado, atraem e retêm professores bem formados e com boas expectativas sobre os alunos, as famílias apoiam mais as questões escolares, têm menos problemas disciplinares e uma atmosfera mais propícia ao aprendizado. Há, então, para os alunos com alto desempenho e motivação, um efeito de pares que faz com que eles criem uma cultura de sucesso na escola e ampliem as condições propícias ao sucesso escolar, que já são altas. Já as escolas que possuem um nível socioeconômico mais baixo, desfavorecem duplamente os alunos pobres porque o efeito da pobreza individual se soma ao efeito da pobreza escolar promovendo um ambiente de privação prejudicial ao desempenho acadêmico (Rumberger e Palardy, 2005; Lee e Burkam, 2002; Mortimore, Sammons, *et al.*, 1988).

2.5. Ainda a pedagogia da repetência

Após 30 anos das pesquisas de Ribeiro (1991) e várias tentativas de correção de fluxo que marcaram a década de 90, o sistema educacional brasileiro ainda é permeado por altos índices de **reprovação, abandono e evasão** – taxas profundamente marcadas pelas desigualdades socioeconômicas – e está distante de resolver seus problemas educacionais. O acesso não eliminou as desigualdades e elas foram intensificadas dentro das escolas. Atualmente, existem desigualdades entre regiões, sistemas de ensino e escolas, bem como entre grupos sociais definidos por gênero, raça e nível socioeconômico. Muitos alunos não concluem o ensino fundamental na idade esperada, fato que é tendência predominante entre os alunos rurais, negros e pobres que têm em média menos de 12 anos de estudo (Alves, Oliveira, *et al.*, 2021). O Brasil possui uma das maiores taxas de reprovação do mundo, que geram evasão escolar, não garante melhores condições de aprendizagem (Correa, Bonamino e Soares, 2014) e são injustas com os estudantes mais pobres e/ou que estudam em escolas com nível socioeconômico mais baixo porque eles têm maiores chances de reprovação apenas por serem pobres (Ferrão, Costa e Matos, 2017).

O sistema de ensino brasileiro apresenta graves problemas quanto ao nível de desempenho e de equidade. Existe um perfil socioeconômico dos problemas do aprendizado no ensino fundamental. Para os alunos com nível socioeconômico mais baixo, pretos e do sexo

masculino, as proficiências mais baixas os colocam em situação muito desvantajosa para seguirem suas trajetórias escolares e esse hiato tende a aumentar se a média geral do desempenho aumenta (Alves, Soares e Xavier, 2016). Meninas tendem a ter mais sucesso no ambiente escolar (Senkevics e Carvalho, 2016; Rosemberg e Madsen, 2011). O mesmo acontece para os estudantes brancos, os que frequentam escolas privadas, para aqueles que estudam em escolas das regiões Sudeste e Centro- Oeste (Alves, Soares e Xavier, 2016; Unesco, 2019).

3. Trajetórias escolares e desigualdades de origem – o modelo

Os trabalhos que analisaram as trajetórias escolares brasileiras encontraram relações entre desigualdades educacionais e as desigualdades socioeconômicas e demográficas. Os meninos, os estudantes da cor preta, os mais pobres, os que frequentam as escolas municipais e as localizadas no Norte e Nordeste do país tem trajetórias mais irregulares enquanto as meninas, estudantes de cor branca, os mais ricos, os que frequentam as escolas privadas e as localizadas no Sudeste possuem trajetórias mais regulares (Soares, Alves e Fonseca, 2021; Ferrão, 2022; Fonseca, Rodrigues, *et al.*, 2022).

3.1. Modelo de desigualdades de origem

Para mensurar os efeitos socioeconômicos e demográficos sobre as trajetórias escolares utilizamos a base longitudinal 2 (coorte 2003/2004 - dados protegidos). Construímos um modelo logístico multinomial que tem como variável resposta a variável de trajetórias escolares (construída no capítulo 1) e como explicativas variáveis socioeconômicas e demográficas escolhidas com base na literatura sociológica e na disponibilidade das bases de dados do Inep. Denominamos esse modelo de modelo de desigualdades de origem.

3.2. As variáveis socioeconômicas e demográficas

Nossas variáveis explicativas são sexo, raça, região do país, dependência administrativa da escola e nível socioeconômico das escolas (NSE)⁵¹. As quatro primeiras foram extraídas do Censo Escolar e o indicador de NSE das escolas foi construído a partir de informações do

⁵¹ Para sua inclusão no modelo, as três categorias do indicador foram binarizadas

Sistema de Avaliação da Educação Básica (Saeb) e do Exame Nacional do Ensino Médio (ENEM).

3.2.1. Trajetória escolares em sistema aberto e as variações socioeconômicas e de escolas

É possível que em estudos de painel ocorram mudanças nas variáveis socioeconômicas ao longo do tempo. Essas mudanças podem acontecer no Censo Escolar em função de alteração da situação socioeconômica dos alunos, na declaração racial ou, embora sejam menos frequentes, no gênero que as famílias e os alunos fornecem na ficha de matrícula. Mas também podem ocorrer pelo equívoco no preenchimento do formulário do Censo Escolar ou pelo não lançamento da informação requerida no sistema do MEC (Unesco, 2022). Alterações no nível socioeconômico médio das escolas é um evento mais raro (Paula, 2020). Além disso, os alunos tendem a permanecer em escolas com o mesmo NSE, apesar das muitas trocas de estabelecimento de ensino que são realizadas durante a vida escolar. A tabela 21 mostra as mudanças de escola da corte selecionada para a nossa pesquisa. Podemos ver que 45% dos alunos permanecem 5 anos ou menos em uma mesma escola. Enquanto 55 % permanecem 6 anos ou mais. Considerando a estrutura da oferta do ensino fundamental I e II e Ensino Médio, as mudanças de dependências administrativas, de municípios⁵² e de regiões do país a tendência é que os alunos mudem de escola no mínimo três vezes.

Tabela 19 - Anos escolares frequentando a mesma escola x Porcentagem de alunos

Anos	% de alunos
1	0
2	0,6
3	5,8
4	18,1
5	20,5
6	14
7	15,1
8	8,7
9	8,1
10	3,4
11	3,5
12	2,2

Fonte: Elaboração própria com dados do Banco de fluxo/Inep

⁵² A partir de uma pesquisa que acompanhou as coortes de estudantes que nasceram entre 1993 e 2009, Rigotti e Hadad (2018) encontraram que, em 3 anos, 14,85% dos estudantes migraram pelo menos uma vez entre municípios.

Essas condições tornam-se um desafio para a definição de qual escola os alunos devem ser vinculados para fins de análise. Outra frequente mudança é na variável raça/cor. Os alunos e/ou suas famílias foram, ao longo dos anos, se declarando com maior frequência como pretos ou pardos em consequência da adoção de políticas públicas, tais como as cotas raciais nas universidades, pela conscientização racial feita pelo Movimento Negro e elevação das taxas de escolaridades e renda (Telles e Paschel, 2014; Francis e Tannuri-Pianto, 2012). Os resultados da Pesquisa Nacional por Amostra em Domicílio Contínua (Pnad-c) do IBGE apontaram essa tendência com o avanço do percentual de pessoas autodeclaradas pretas de 7,4%, em 2012, para 9,1% em 2021.

Para lidar com esse desafio metodológico, optamos por atribuir a moda⁵³ para as variáveis socioeconômicas sexo, raça, dependência administrativa e região ao longo dos 12 anos⁵⁴. Por exemplo, se o aluno frequentou por 9 anos escolas municipais e por 3 anos escolas estaduais, será atribuído a ele escola municipal. Ou ainda, se nos primeiros 4 anos da trajetória o estudante foi considerado pardo, nos 5 anos finais preto e em dois anos não tem informação ele será considerado preto. As porcentagens para sexo, raça e região da coorte se manteve o mesmo da população brasileira do Censo 2022, composta por mais mulheres, mais pardos, e com concentração no Sudeste.

Nós reconhecemos os limites dessa opção para as variáveis que são influenciadas pela mudança de rede nos iniciais e finais do ensino fundamental. Entretanto a variável dependência administrativa da escola se aproximou das estatísticas nacionais com mais alunos frequentando escolas municipais (Unesco, 2019). A Tabela 20 mostra os números absolutos e as frequências das cinco variáveis socioeconômicas investigadas: sexo, raça região e dependência administrativa da escola.

⁵³ Moda é o valor que ocorre com maior frequência num conjunto de dados (SPIEGEL, 1976, p. 74).

⁵⁴ Agradeço ao professor Ricardo Tavares pelo auxílio na programação em R para definir a função da moda.

Tabela 20 - Perfil socioeconômico da coorte

Perfil	N	%
Sexo		
Feminino	1,634,559	51
Masculino	1,571,959	49
Total	3,206,518	100%
Cor/raça		
Não declarada	809,961	25
Branca e amarela	986,811	31
Preta	125,568	4
Parda	1,249,680	39
Amarela	11,967	0
Indigna	22,531	1
Total	3,206,518	100%
Região		
Norte	357,281	11
Nordeste	998,113	31
Sudeste	1,205,661	38
Sul	406,103	13
Centro-Oeste	239,360	8
Total	3,206,518	100%
Nível socioeconômico das escolas		
Nse 1º tercil	533,612	17
Nse 2º tercil	1,046,770	33
Nse 3º tercil	1,104,681	35
Sem Informação de Nse	521,455	16
Total	3,206,518	100%
Dependência Administrativa da escola		
Federal	5,299	0
Estadual	1,230,738	38
Municipal	1,487,116	46
Privada	483,365	15
Total	3,206,518	100%

Fonte: Elaboração própria com dados do Banco de fluxo/Inep

3.2.2. Regularidades sociológicas

A seguir, apresentamos nos gráficos 18 a 22 os cruzamentos da variável de trajetória escolar na educação básica (I + II + ensino médio) com as variáveis socioeconômicas.

Gráfico 18 - Trajetórias escolares por sexo

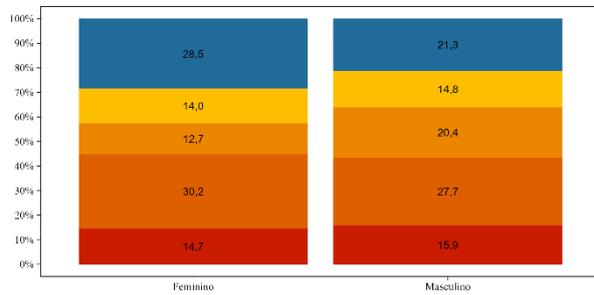


Gráfico 19 - Trajetórias escolares por raça/cor

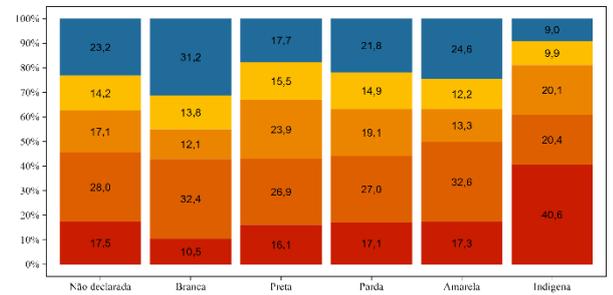


Gráfico 20 - Trajetórias escolares por dependência administrativa das escolas

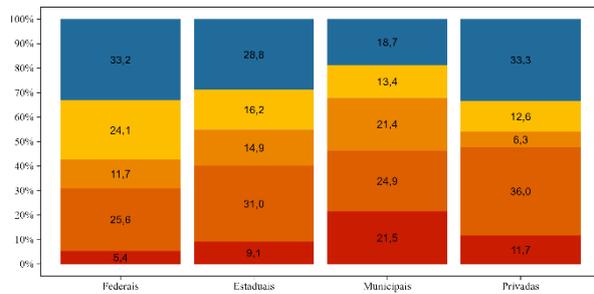


Gráfico 21 - Trajetórias escolares por NSE das escolas

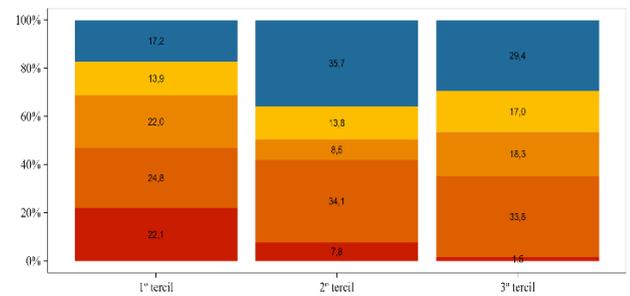
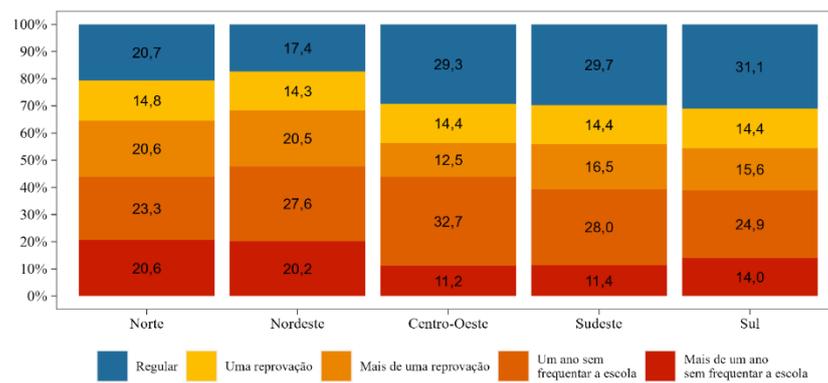
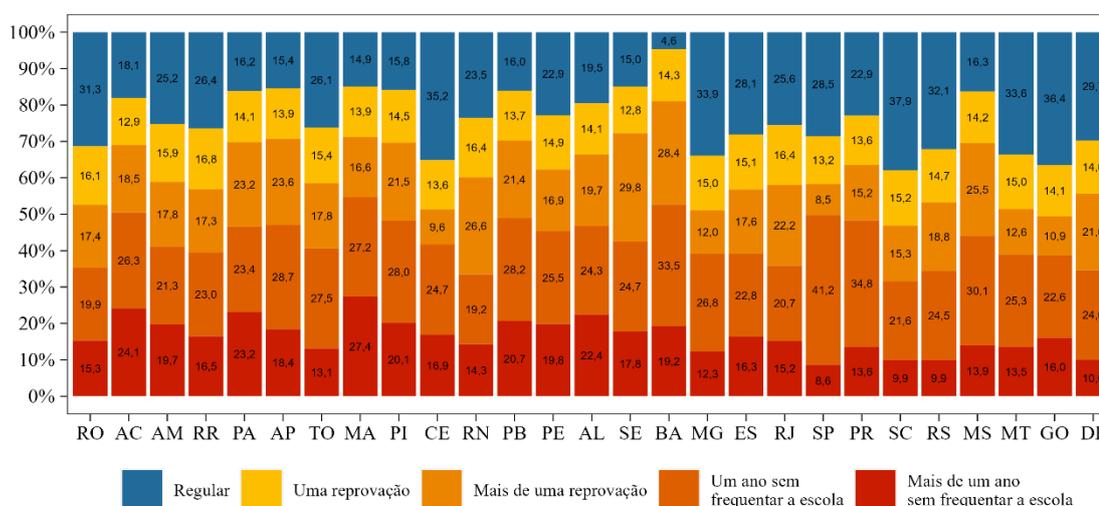


Gráfico 22 - Trajetórias escolares por regiões do Brasil



Fonte: elaboração própria a partir de dados do Fluxo Escolar/Inep

Gráfico 23 - Trajetórias escolares por estados da Federação



Fonte: elaboração própria a partir de dados do Fluxo Escolar/Inep

Nossos resultados vão de encontro com a literatura. Os gráficos mostram as meninas, os brancos, os estudantes de escolas privadas, aqueles que frequentam escolas de NSE mais alto e são moradores do Sudeste e Centro-Oeste com trajetórias mais regulares. O Gráfico 23 decompõe a variável região por estados e mostra resultados importantes como a Bahia com apenas 5% de estudantes com trajetórias regulares enquanto tem 33% de indivíduos que não frequentaram a escola. Santa Catarina é o estado com a maior porcentagem de estudantes regulares e São Paulo apresenta a menor taxa de estudantes que não frequentaram a escola mais de uma vez.

4. Efeitos desproporcionais

Modelos logísticos multinomiais é um dos muitos métodos utilizados para analisar variáveis de resposta categórica não ordenada (nominal) nas pesquisas de ciências sociais (Long e Freese, 2003; Agresti, 2002). Esses modelos são capazes de mensurar simultaneamente variáveis resposta categóricas estimando as probabilidades de um evento ocorrer de acordo com as variáveis explicativas que estão associadas com tal evento. A equação para esse modelo é:

$$Pr (Y_i = k | X_i) = 1 / (1 + \exp (X_i \beta_k))$$

Onde:

$P (Y = k | X)$ é a probabilidade da variável resposta Y estar na categoria k , consideradas as variáveis explicativas X .

K representa as categorias possíveis de Y .

β_k são os coeficientes associados com as categorias k da variável dependente.

X são os valores das variáveis independentes.

Mensuramos as razões de chance e os incrementos percentuais de um estudante experimentar uma reprovação, duas reprovações, ter deixado de frequentar a escola por um ano ou por mais de um ano em relação a ter trajetórias regulares, dependendo de suas características socioeconômicas e demográficas. Então,

1 se a trajetória regular (categoria de referência)

2 se uma reprovação

$K = 3$ se mais de uma reprovação

4 um ano sem frequentar a escola

5 mais de um ano sem frequentar a escola

4.1. Recategorizações

Como o objetivo final deste capítulo perpassa a construção de modelos explicativos, as categorias de variáveis com números absolutos muito baixos performam melhor se reagrupadas. Fizemos duas recategorizações considerando a semelhança das categorias. Na variável dependência administrativa da escola agrupamos as escolas federais com as privadas e a variável cor/raça foi recategorizada agrupando os brancos com os amarelos. Além disso, modelos logísticos exigem que cada categoria das variáveis explicativas categóricas sejam binarizadas. O quadro 1 apresenta essa recategorização.

Quadro 1 - Variáveis independentes do modelo logístico multinomial

Variável	Tipo	Descrição
Sexo	Binária	1 = menino e 0 = menina
Preta ou indígena	Binária	1 = Preta ou indígena e 0 = branca ou amarela
Parda	Binária	1 = parda e 0 = branca ou amarela
Cor não declarada	Binária	1 = Cor não declarada e 0 = branca ou amarela
Estadual	Binária	1 = Estadual e 0 = Privada ou Federal
Municipal	Binária	1 = Municipal e 0 = Privada ou Federal
Norte	Binária	1 = Norte e 0 = Sudeste
Nordeste	Binária	1 = Nordeste e 0 = Sudeste
Sul	Binária	1 = Sul e 0 = Sudeste
Centro-Oeste	Binária	1 = Centro-Oeste e 0 = Sudeste
NSE da escola no 1º tercil	Binária	1 = NSE da escola no 1º tercil e 0 = NSE da escola no 3º tercil
NSE da escola no 2º tercil	Binária	1 = NSE da escola no 2º tercil e 0 = NSE da escola no 3º tercil
NSE da escola sem informação	Binária	1 = NSE da escola sem informação e 0 = NSE da escola no 3º tercil

Fonte: elaboração própria.

5. Resultados

A Tabela 21 apresenta os resultados do modelo⁵⁵ estimado e indica que as variáveis socioeconômicas possuem efeitos significativos nas trajetórias dos estudantes. É possível observar que os grupos em desvantagens sociais são aqueles que tendem a ter mais de uma intercorrência. Em outras palavras, os estudantes pretos ou indígenas, os estudantes que frequentam escolas municipais, estudantes matriculados no Nordeste ou com nível socioeconômico mais baixo são mais prováveis de terem experimentado uma ou mais reprovações ou deixado de frequentar a escola por mais de um ano.

Algumas probabilidades se destacam: os meninos comparados às meninas têm 43,3% a mais de chances de terem uma reprovação em relação a possuírem trajetórias regulares. Percentual que aumenta para 120,4% quando se trata de duas reprovações. Outro resultado importante é o aumento percentual das chances que estudantes pretos ou indígenas tem nas categorias mais de uma reprovação e mais de um ano sem frequentar a escola.

Na comparação com estudar em escolas privadas ou federais, estudar em escolas municipais aumenta em 277,9% as chances de ter mais de uma reprovação na comparação com frequentar escolas privadas e federais enquanto as chances daquelas que frequentam escolas estaduais aumenta em 123,2%. Esses resultados inferiores que as redes municipais apresentam em relação as redes estaduais aparecem em outros indicadores tais como Unesco (2019) e

⁵⁵ Ver Apêndice II para detalhes do modelo.

QEDu⁵⁶ (2019). Enquanto frequentar escolas de baixo nível socioeconômico aumenta as chances de deixar de frequentar a escola por mais de um ano em 193,5% enquanto escolas de médio aumenta essas chances 117,9%.

Tabela 21 - Resultados da regressão logística multinomial - Modelo de desigualdades de origem sobre as trajetórias educacionais completas (2009-2020)

	Uma reprovação		Mais de uma reprovação		Um ano sem frequentar a escola		Mais de um ano sem frequentar a escola		Em relação a
	Razão de chance	Inc. %	Razão de chance	Inc. %	Razão de chance	Inc. %	Razão de chance	Inc. %	
Se é menino	1,43**	43,3	2,20**	120,4	1,24**	23,8	1,47**	47,1	Ser menina
Se cor não declarada	1,25**	25,3	1,51**	50,7	1,09**	8,6	1,55**	54,8	Ser branco ou amarelo
Se é preto ou indígena	1,71**	70,9	2,62**	161,9	1,36**	35,5	2,33**	132,6	
Se é pardo	1,25**	24,7	1,47**	47,2	1,03**	2,7	1,37**	36,7	
Se mora no Norte	1,29**	28,6	1,64**	64	-0,96**	-3,9	1,47**	47,4	Morar no Sudeste
Se mora no Nordeste	1,52**	52	1,88**	88,4	1,33**	33	1,46**	45,9	
Se mora no Sul	1,04**	3,5	1,51**	50,8	0,87**	-13,3	1,26**	-186,7	
Se mora no Centro-oeste	-0,89**	-189,1	1,07**	7,3	-0,70**	-30,2	1,05**	5,2	
Se estuda em escola estadual	1,42**	42,2	2,23**	123,2	0,99*	-1,4	-0,61**	-161,3	Estudar em escola privada ou federal
Se estuda em escola municipal	1,55**	54,8	3,78**	277,9	1,07**	7,4	1,63**	63,4	
Se estuda em escola de baixo NSE	1,12**	12,2	1,65**	65,1	1,15**	15,2	2,93**	193,5	Se estuda em escola de alto NSE
Se estuda em escola de médio NSE	1,18**	17,5	1,58**	58,2	1,17**	16,9	2,18**	117,9	
Sem informação de NSE	-0,99	-199,3	2,50**	150,3	1,14**	13,8	18,68**	1767,8	

Fonte: Elaboração própria com dados do Banco de fluxo/Inep

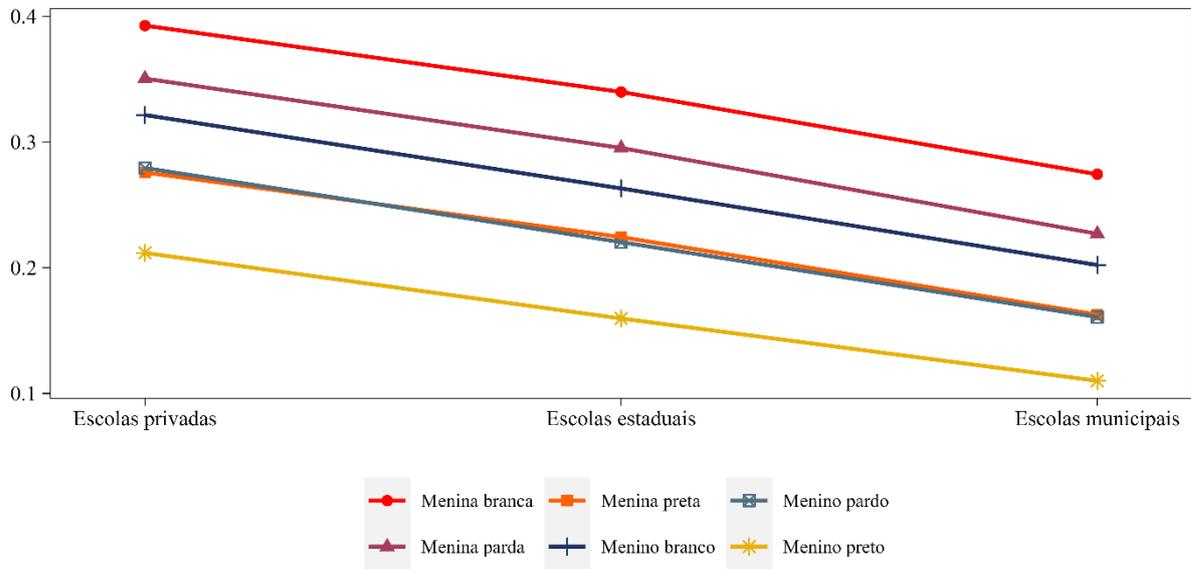
Nota: * p-valor < 0.05 e ** p-valor < 0.00.

5.1. O sexo, a cor, a escola e o lugar dos escolarizados

Existem muitas camadas de desigualdade na sociedade brasileira – muitas vezes relacionadas e sobrepostas – o que gera acúmulo de desvantagens para alguns grupos sociais. Com um sistema educacional excludente, uma das camadas da desigualdade brasileira é a das desigualdades escolares. Nessas condições, o Brasil nega para alguns grupos sociais o direito à educação - reconhecido como universal pela Declaração Universal dos Direitos Humanos (Unesco, 2019). É o que vemos no gráfico 24. Elaboramos este gráfico usando o valor predito dos coeficientes das variáveis explicativas do modelo de desigualdades de origem (Tabela 21) e mapeamos os perfis mais prováveis de estarem em cada categoria das trajetórias escolares.

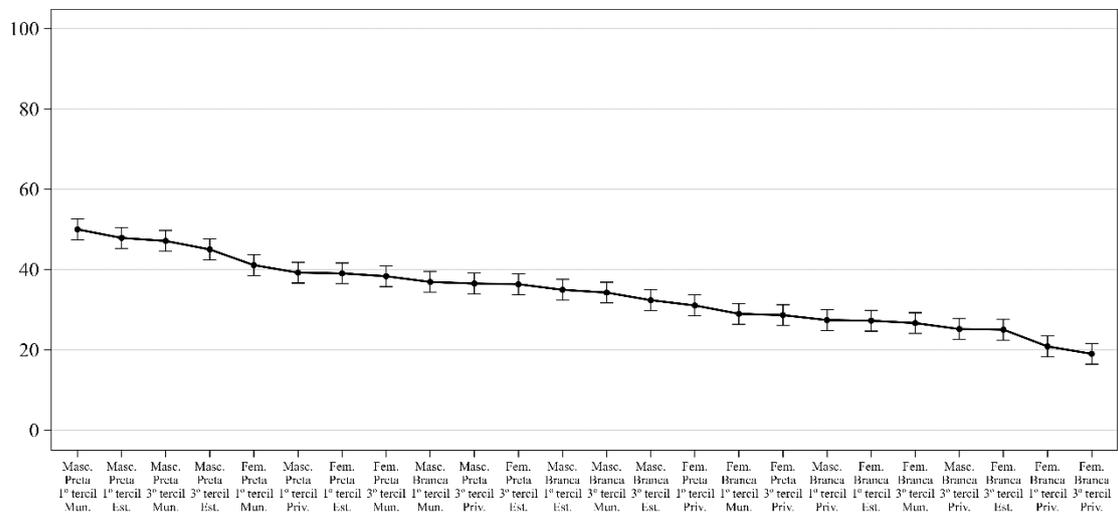
⁵⁶ Disponível em <https://qedu.org.br/>

Gráfico 24 - Valores preditos marginais da trajetória regular (95% de intervalo de confiança)



Fixados pelas regiões do país, os gráficos 25 a 31 mostram os perfis com maiores e menores chances de finalizarem a educação básica com diferentes intercorrências. Os gráficos apresentam intervalos de confiança para as quatro categorias das trajetórias escolares das regiões Sudeste e nordeste: uma reprovação, duas reprovações, um ano sem frequentar a escola e mais de um ano sem frequentar a escola em relação a trajetórias regulares. Os intervalos de confiança dos quatro grupos com chances mais altas de uma reprovação se sobrepõem. A sobreposição dos intervalos indica que não há diferenças significativa entre os grupos e se não se sobrepõe indica que os grupos são diferentes entre si e da categoria de referência (meninas, brancas, que frequentam escolas privadas e de alto NSE).

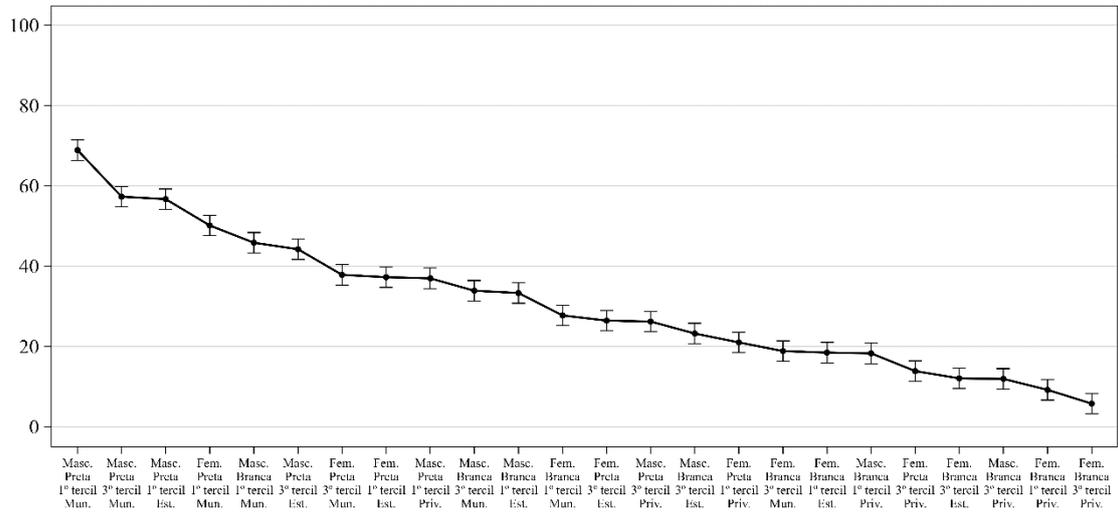
Gráfico 25 - Probabilidades de uma reprovação, por perfil no Sudeste



Fonte: elaboração própria a partir de dados do Fluxo Escolar/Inep

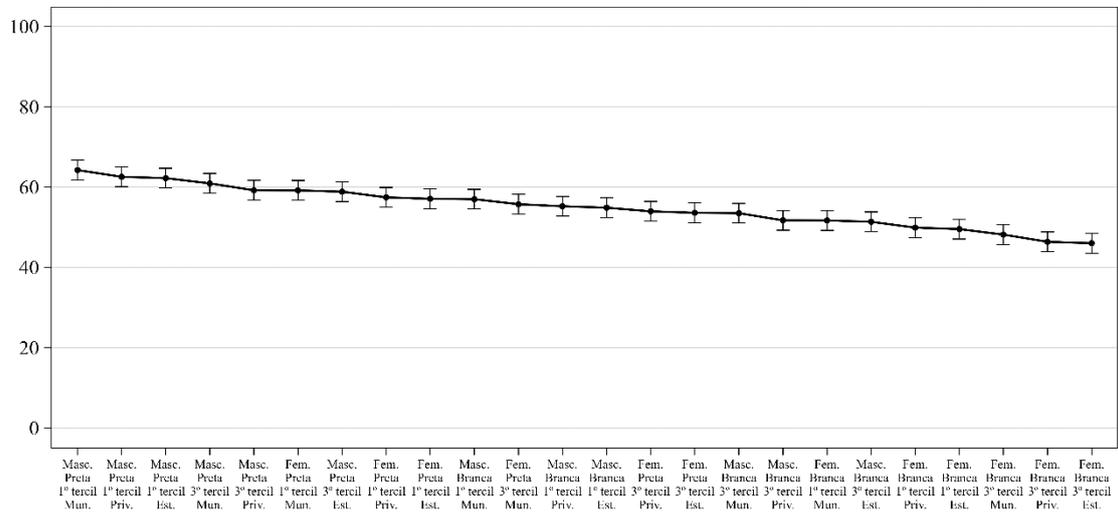
As pontas das linhas revelam de diferentes maneiras a desigualdade nas trajetórias escolares entre os grupos sociais. Uma possibilidade de análise é olhar para a localização dos perfis nos eixos X dos gráficos.

Gráfico 26 - Probabilidades de mais de uma reprovação, por perfil no Sudeste



Fonte: elaboração própria a partir de dados do Fluxo Escolar/Inep

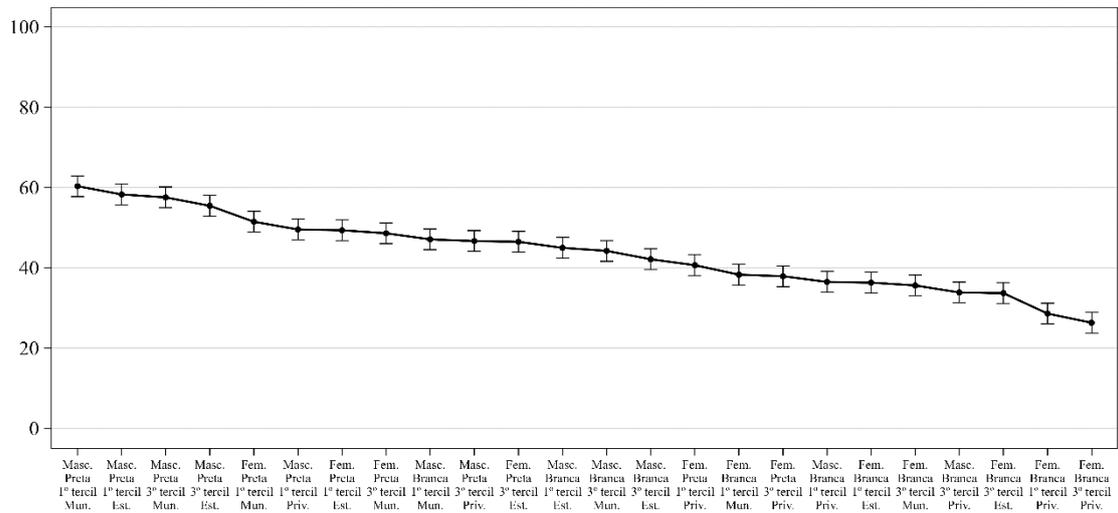
Gráfico 27 - Probabilidades para um ano sem frequentar a escola, por perfil no Sudeste



Fonte: elaboração própria a partir de dados do Fluxo Escolar/Inep

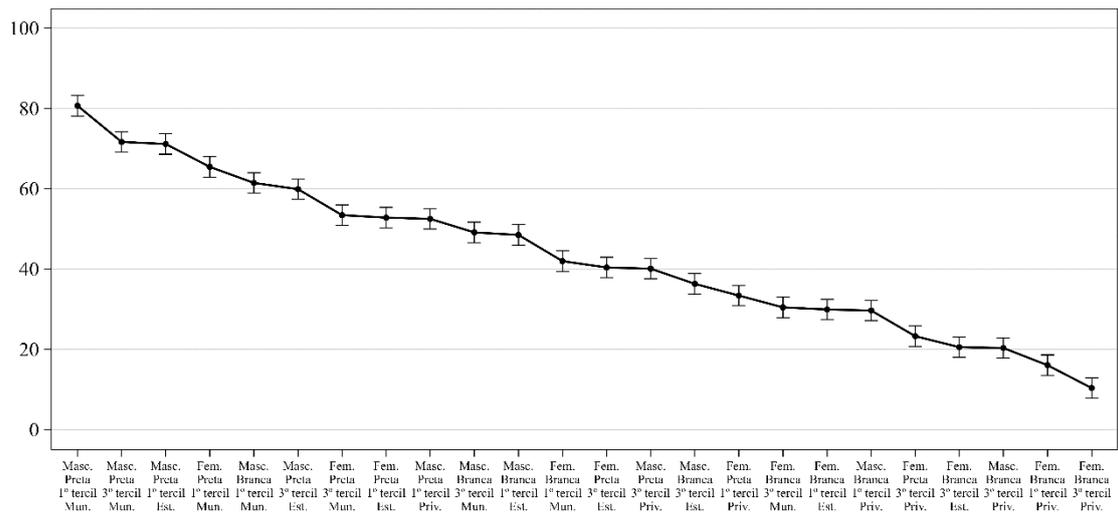
Vemos que os meninos pretos que frequentam escolas municipais de baixo NSE estão quase sempre localizados na extremidade esquerda, com altas probabilidades de intercorrências e as meninas brancas que com alto NSE estão sempre localizadas na extremidade direita com chances mais baixas para uma intercorrência e chances próximas de zero para mais de uma intercorrência.

Gráfico 28 - Probabilidades de uma reprovação, por perfil no Nordeste



Fonte: elaboração própria a partir de dados do Fluxo Escolar/Inep

Gráfico 29 - Probabilidades de mais de uma reprovação, por perfil no Nordeste

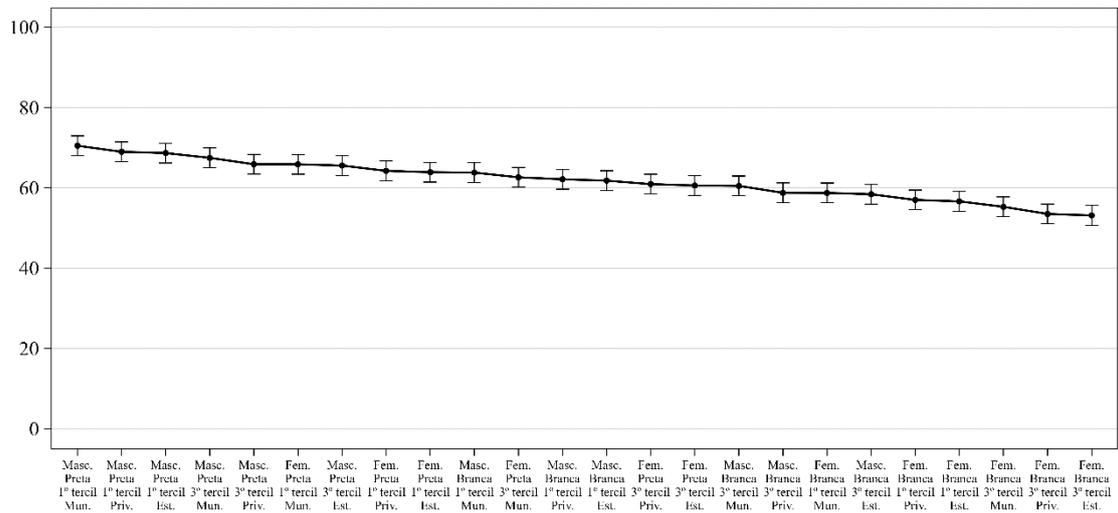


Fonte: elaboração própria a partir de dados do Fluxo Escolar/Inep

Outra maneira de análise é através da amplitude das variações entre as pontas dos gráficos. O Sudeste tem para uma reprovação a variação de 31 pontos percentuais, para mais de uma reprovação 40 pontos percentuais, para um ano sem frequentar a escola 18 pontos percentuais e para mais de um ano sem frequentar a escola 53 pontos percentuais. Enquanto o Nordeste ao mesmo tempo que apresenta chances mais altas para as intercorrências tem respectivamente 34, 66, 17,4 e 66 pontos percentuais de amplitude das variações.

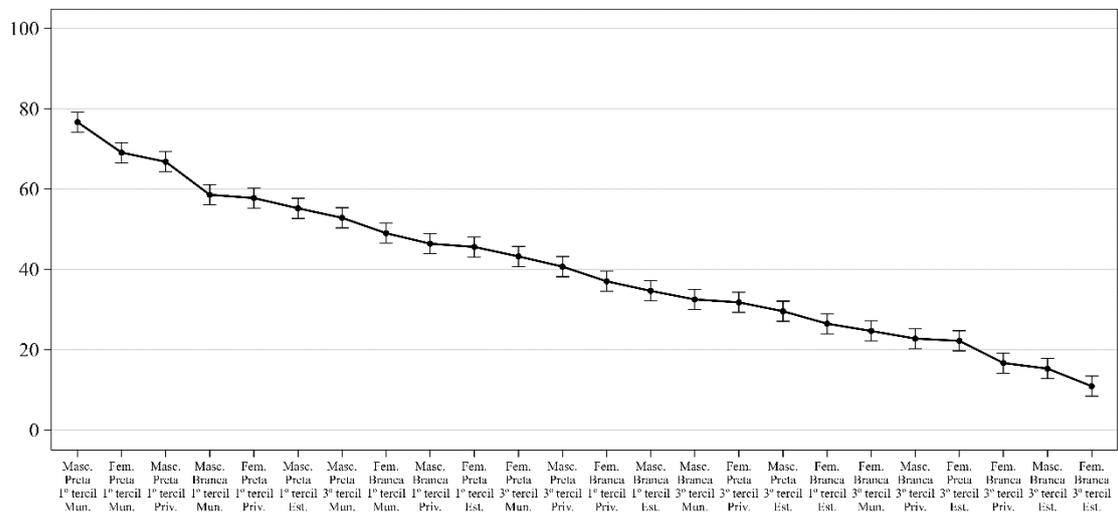
Esses números apontam que a região possui menos alunos com trajetórias escolares regulares e essas regularidades estão concentradas em alguns grupos. Por exemplo, as probabilidades de mais de uma reprovação são maiores para os grupos mais vulneráveis, ou seja, estão concentradas nos alunos pretos, meninos e mais pobres.

Gráfico 30 - Probabilidades para um ano sem frequentar a escola, por perfil no Nordeste



Fonte: elaboração própria a partir de dados do Fluxo Escolar/Inep

Gráfico 31 - Probabilidades para mais de um ano sem frequentar a escola, por perfil no Nordeste



Fonte: elaboração própria a partir de dados do Fluxo Escolar/Inep

O Nordeste apresenta as piores situações enquanto o Sudeste tem as melhores situações seguido do Centro-Oeste com probabilidades muito semelhantes. As diferenças entre as

probabilidades variam conforme as categorias das trajetórias. Isso sustenta a decisão de tratar separadamente as categorias reprovação e sem frequentar a escola.

Para visualizar outras regiões e combinações acesse o gráfico dinâmico em https://bit.ly/tese_isis_arantes.

6. Considerações

O que podemos entender da relação entre a origem socioeconômica dos estudantes e as etapas educacionais da educação básica brasileira é que a origem não perde seu efeito ao longo do tempo e impacta de diferentes maneiras sobre o alcance educacional. E na medida em que o acesso se expandiu, marcado por desigualdades, o sistema educacional passou a gerar desigualdades escolares que vão interagir com as de origem.

Os números são injustos e reiteram o argumento que apresentamos, de que no Brasil o direito à educação – de acesso, permanência, aprendizado e conclusão na idade adequada - é negado para alguns grupos sociais. O contexto de injustiça se intensifica com o papel utilitário que os diplomas escolares desempenham na definição das oportunidades de emprego e, conseqüentemente, no futuro dos indivíduos. Esse cenário exige escolas justas, capazes de valorizar todos os alunos, independentemente de sua origem social (Dubet, 2004).

Os resultados são uma sobreposição das desigualdades sociais brasileiras. Por um lado, uma concentração de trajetórias regulares nos grupos privilegiados e concentração de trajetórias mais irregulares nos grupos de estudantes mais vulneráveis. Ao mesmo tempo, o sistema educacional brasileiro oferece oportunidades desiguais e é marcado por altas taxas de reprovação e evasão capazes de produzir efeitos cumulativos negativos no curso de vida dos alunos, causando muitas desvantagens (Barros e Mendonça, 1998).

Entretanto, considerando que a qualidade das escolas pode variar e algumas são mais eficazes do que outras (Lee, 2008) e que a eficácia escolar é mais importante para as chances de melhora de vida dos alunos com desvantagens sociais – por impactarem mais sobre eles do que sobre os alunos privilegiados (Coleman e others, 1966; Rutter, Maughan, *et al.*, 2008; Mortimore, Sammons, *et al.*, 1988 *apud* Brooke e Soares, 2008) existiriam chances de compensação das desvantagens de largada através de variáveis de efeito escola? Esse é o tema do próximo capítulo.

CAPÍTULO IV - LINHAS TÊNUES⁵⁷ E AQUELA CHANCE⁵⁸

As evidências mostraram que as desvantagens de origem social reduzem as chances de trajetórias escolares regulares. Entretanto, os traços sociais não são os únicos fatores capazes de impactar os resultados escolares. Há também o *efeito escola*.

Eficácia escolar ou efeito escola⁵⁹ é a contribuição da escola sobre os resultados escolares. Essa contribuição já foi mensurada para o Brasil e explica em torno de 20% dos resultados dos alunos (Alves e Soares, 2007). O conceito é utilizado pela literatura em vários sentidos. A referência pode aludir aos efeitos de determinadas políticas ou práticas de uma escola, uma determinada intervenção, as práticas dos professores, o tamanho da escola e das turmas ou as despesas das escolas. Pode também ser utilizado para indicar a variação dos resultados entre as escolas ou os efeitos gerais de frequentar ou não determinada escola (Willms, 1992).

As variáveis associadas com as variações entre escolas estão classificadas em cinco categorias: 1) recursos escolares; 2) organização e gestão da escola; 3) clima acadêmico; 4) formação e salário docente; 5) ênfase pedagógica” (Alves e Franco, 2008). Recursos escolares são a existência, a conservação e o uso dos equipamentos e do prédio escolar, ou seja, a infraestrutura das escolas (Unesco, 2019). Organização e gestão da escola são fatores de liderança da direção escolar e o comprometimento junto ao empenho dos docentes. Clima acadêmico está associado com percepções em relação a escola como confiança, respeito, segurança, justiça, bom processo de ensino e aprendizagem e diálogo ou esses fatores em risco (Castellini, 2019; Vinha, Nunes, *et al.*, 2017). Formação e salário docente são o nível de formação acadêmica somada à experiência dos professores e os salários que eles recebem. Ênfase pedagógica são as metodologias de ensino.

Como existem diferenças tanto entre sistemas escolares quanto entre escolas (Alves e Franco, 2008) e algumas escolas são mais eficazes do que outras dentro de um mesmo sistema de ensino, a depender da qualidade (Lee, 2008), entendemos que existem características dos sistemas de ensino e das escolas (Bressoux, 2003; Darvene, 2011; Barbosa e Fernandes, 2001) que afetam as trajetórias dos alunos. Isso porque as trajetórias regulares estão condicionadas ao

⁵⁷ Definição apresentada na introdução, página 19.

⁵⁸ A base de dados utilizada neste capítulo foi a base longitudinal 2.

⁵⁹ Definição apresentada na introdução, página 19.

aprendizado e conclusão das etapas de ensino na idade adequada e esses são resultados que dependem das contribuições das escolas, e essas contribuições são variadas. Assim, argumentamos que apesar de o efeito das escolas ser menor do que os efeitos das características individuais sobre as trajetórias, frequentar uma escola com determinadas características pode ser *aquela chance* necessária para que alunos em desvantagem social tenham trajetórias escolares regulares ou, ao menos, poucas intercorrências.

Neste capítulo testamos a hipótese de que há efeito das características das escolas sobre as trajetórias escolares e demonstramos que, controladas as variáveis individuais dos alunos nossa hipótese é corroborada. Apresentamos nas próximas seções os modelos analíticos com os quais testamos essa hipótese e em seguida seus resultados.

1. Modelos

Os modelos que desenvolvemos para compreender a relação entre trajetórias escolares e os efeitos das escolas buscam seguir a sugestão do modelo integrado de avaliação da eficácia escolar (Scheerens, 1990), o qual se propõe a integrar e servir como estrutura de referência para as variáveis que influenciam o desempenho escolar. Nossa abordagem considera categorias analíticas contextuais, de *input* e de processo.

Para mensurar o efeito das escolas utilizamos como variáveis explicativas alguns dos indicadores educacionais desenvolvidos pelo Inep. Os modelos que construímos são compostos por adequação da formação docente (variável de *input*); pela regularidade do docente na escola e pela complexidade da gestão das escolas (variáveis de processo). As variáveis de controle são Nível Socioeconômico das Escolas (NSE) (variável contextual) e as de características individuais dos estudantes⁶⁰. Como explicamos no capítulo 1, a escolha das variáveis que compõe nossos modelos foi feita diante dos limites metodológicos da pesquisa e de acordo com a literatura.

A adequação da formação docente é a taxa de professores em exercício na educação básica considerando a formação acadêmica e a(s) disciplina(s) que leciona. O indicador é composto por cinco perfis:

1. Docentes com formação superior de licenciatura na mesma disciplina que lecionam, ou bacharelado na mesma disciplina com curso de complementação pedagógica concluído.

⁶⁰ Fontes e descrição apresentadas no Capítulo 3 na pag.82.

2. Docentes com formação superior de bacharelado na disciplina correspondente, mas sem licenciatura ou complementação pedagógica.
3. Docentes com licenciatura em área diferente daquela que leciona, ou com bacharelado nas disciplinas da base curricular comum e complementação pedagógica concluída em área diferente daquela que leciona.
4. Docentes com outra formação superior não considerada nas categorias anteriores.
5. Docentes que não possuem curso superior completo (Inep, 2014).

Para empregar esse indicador, calculamos a média do grupo 1 para cada escola e utilizamos uma variável contínua que representa a taxa de professores com licenciatura na mesma disciplina que lecionam nas escolas. O indicador de regularidade do docente da Educação Básica mensura a regularidade do corpo docente nas escolas de educação básica considerando a permanência dos professores nas escolas nos últimos cinco anos (Inep, 2015). A complexidade da gestão nas escolas é composta por (1) porte da escola; (2) número de turnos de funcionamento; (3) complexidade das etapas ofertadas pela escola e (4) número de etapas/modalidades oferecidas (Inep, 2014).

Um fator importante na relação entre o efeito das escolas e as trajetórias escolares são as mudanças de escolas que os alunos fazem durante a trajetória. A Tabela 22 mostra as mudanças de escola realizadas pelos alunos da coorte a cada etapa da educação básica e durante a trajetória completa. Como sabemos, se o aluno frequenta ensino fundamental I e II e ensino médio a tendência é que ele mude de escola no mínimo três vezes. Vemos que cerca de 40% dos alunos frequentaram a mesma escola durante toda uma etapa.

Tabela 22 - Anos frequentando a mesma escola (% de alunos)

Anos	Ensino fundamental I (2009-2013)	Ensino fundamental II (2004-2017)	Ensino fundamental completo (2009-2017)	Ensino médio (2017-2020)	Trajetoária escolar completa (2009 - 2020)
1	0,6	2,0	-	9,7	-
2	10,9	21,2	1,9	50,5	0,6
3	21,8	30,5	9,2	39,8	5,8
4	28,1	46,3	23,3	-	18,1
5	38,6	-	27,4	-	20,5
6	-	-	10,1	-	14,0
7	-	-	7,1	-	15,1
8	-	-	9,3	-	8,7
9	-	-	11,6	-	8,1
10	-	-	-	-	3,4
11	-	-	-	-	3,5
12	-	-	-	-	2,2

Fonte: Elaboração própria com dados do Banco de fluxo/Inep

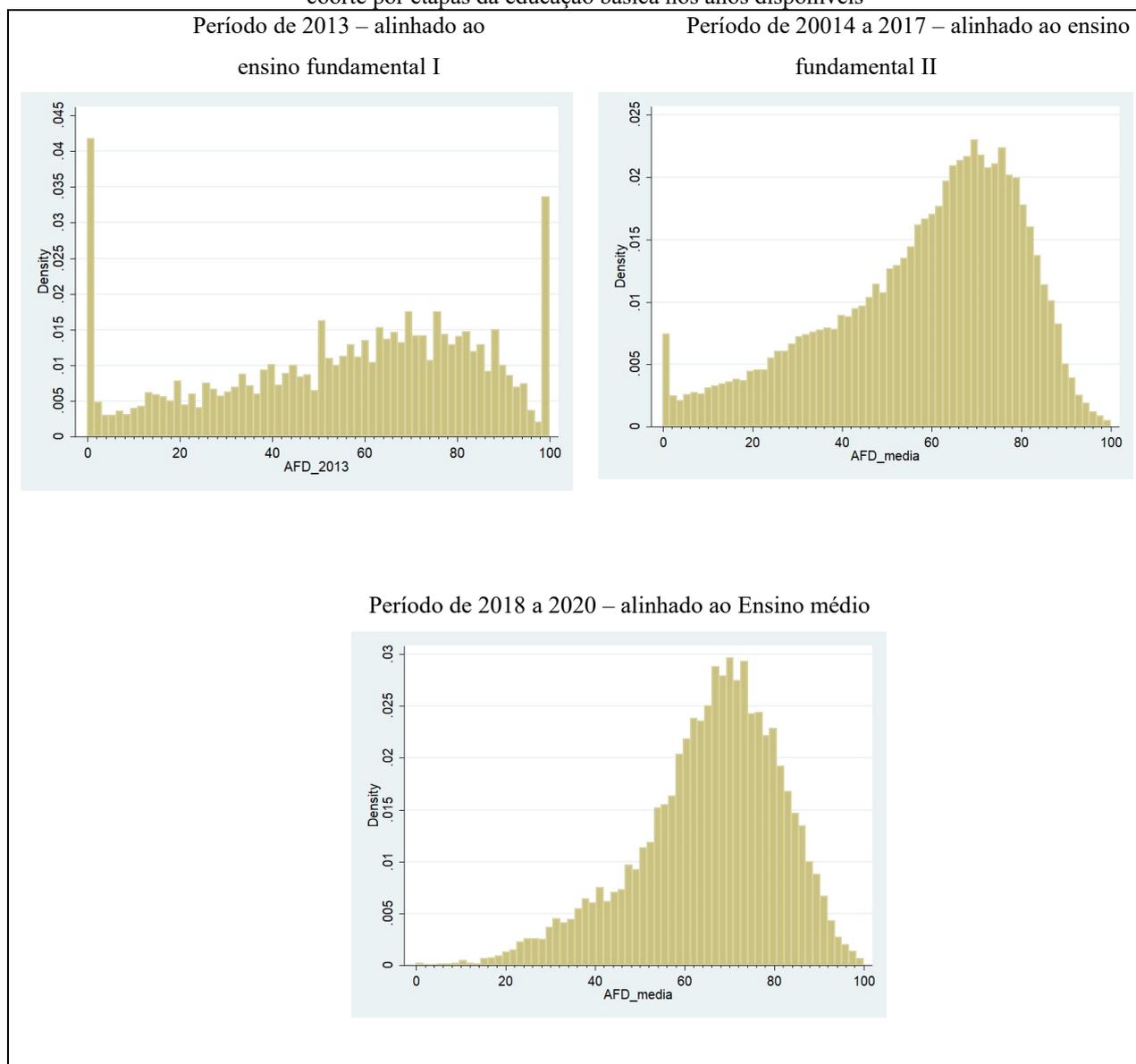
Para lidar com isso, entendemos que a melhor escolha metodológica é considerar na construção dos modelos a estrutura da oferta de ensino, ou seja, ensino fundamental I ensino fundamental II e ensino médio. Ao todo construímos seis modelos, dois para cada etapa de ensino. Um modelo básico com variáveis de origem socioeconômica e um modelo que inclui características da escola para as etapas ensino fundamental I, ensino fundamental II e ensino médio.

Para definir a escola que teria suas características atribuídas ao aluno fizemos a moda⁶¹ em cada etapa. E para alinhar a disponibilidade dos indicadores com os períodos analisados fizemos a média dos indicadores coincidindo os anos para os quais eles estavam disponíveis e os anos do período analisado. Se para o período analisado havia mais de um ano com indicador disponível, fizemos a média. Por exemplo, a trajetória no ensino fundamental II foi de 2014 a 2017 e o indicador de formação docente está disponível para os anos de 2013 até 2020. Então, na análise do ensino fundamental completo fizemos a média do indicador de 2014 a 2017. Se só havia um ano disponível para o período, utilizamos apenas um ano. Foi o caso de regularidade docente para o ensino fundamental II. Ao final consideramos a média de 2014 a 2017 do indicador de formação docente da escola que o aluno mais frequentou no ensino fundamental completo.

A Figura 2 apresenta a média da distribuição do indicador de adequação da formação docente das escolas mais frequentadas pela coorte. Na comparação entre as etapas as discrepâncias chamam a atenção. Alta taxa de escolas com a adequação da formação docente em 0 no ensino fundamental I. Enquanto no ensino fundamental II e ensino médio as escolas apresentam taxas deslocadas à direita, bem acima da média. Considerando o fato de que nessas etapas estão os alunos que chegam às respectivas etapas sem intercorrências, podemos inferir que a aqueles com trajetórias regulares frequentam escolas com taxas mais altas de adequação docente.

⁶¹ Explicação da metodologia na pag.79.

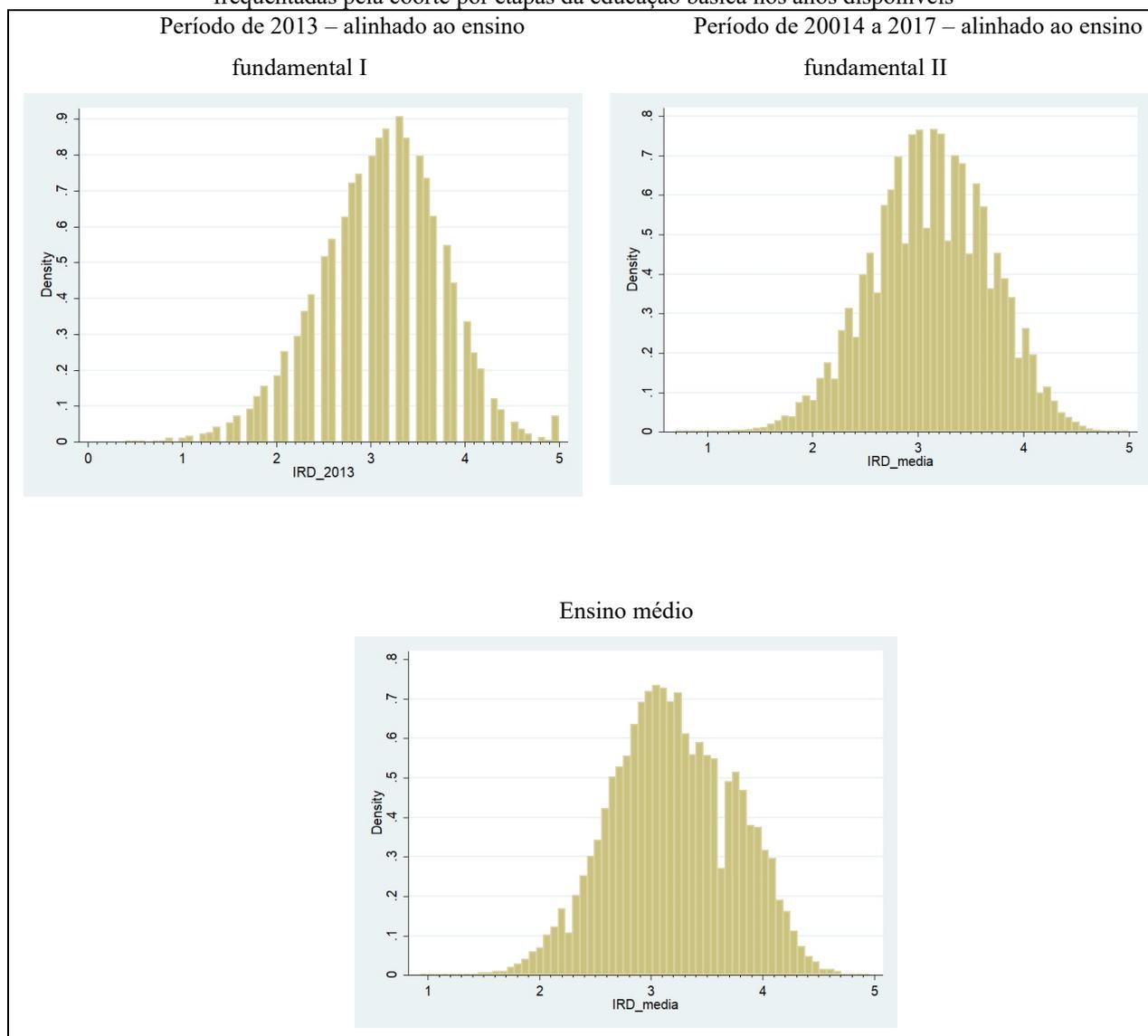
Figura 2 – Distribuição do indicador de adequação da formação docente das escolas frequentadas pela coorte por etapas da educação básica nos anos disponíveis



Fonte: Elaboração própria com dados do Banco de fluxo/Inep

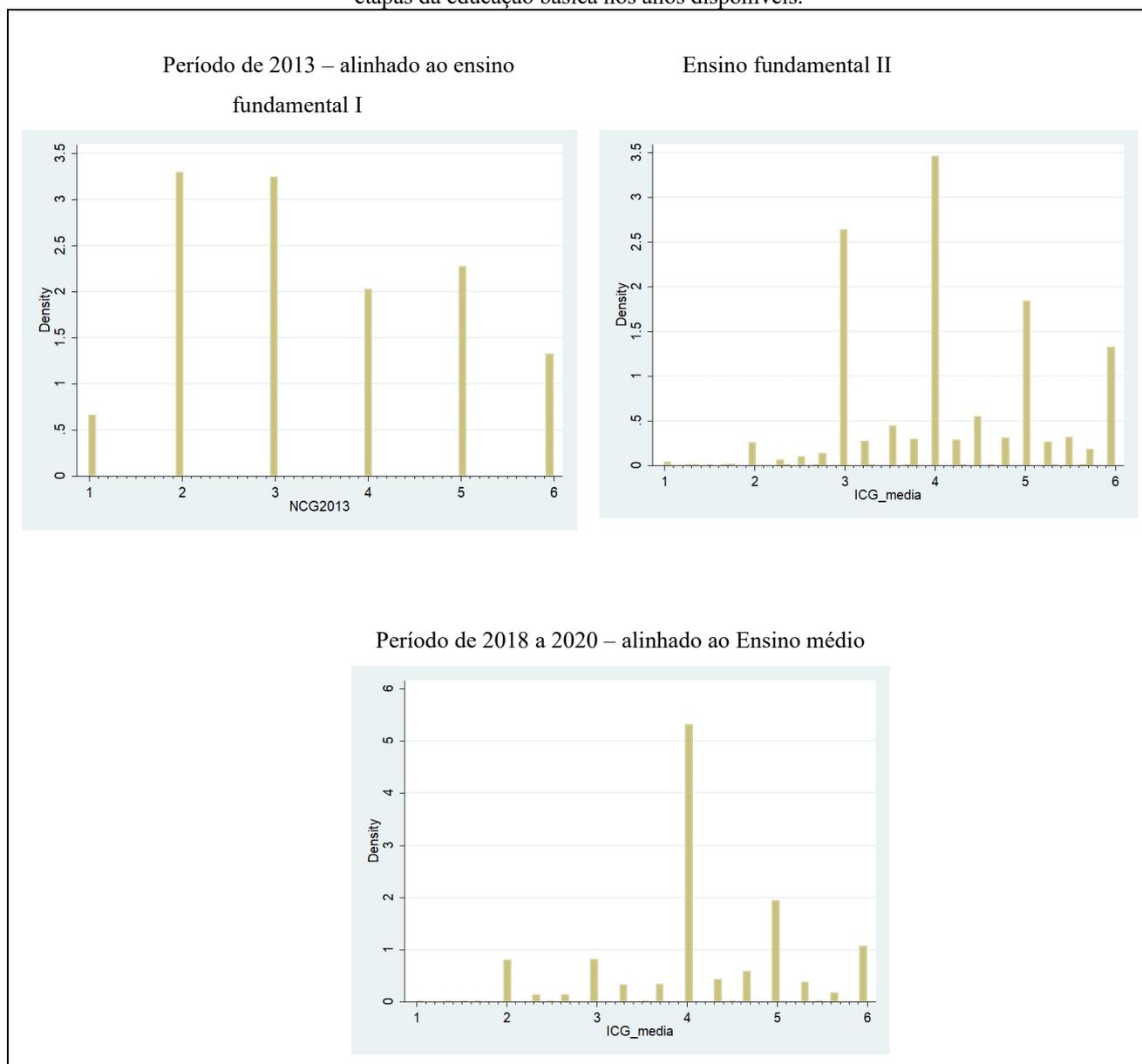
A figura 3 se refere à regularidade do docente. Podemos ver uma distribuição normal para todas as etapas, exceto para o ensino fundamental I. A Figura 4 apresenta a complexidade da gestão das escolas.

Figura 3 - Distribuição do indicador de regularidade do docente da Educação Básica das escolas frequentadas pela coorte por etapas da educação básica nos anos disponíveis



Fonte: Elaboração própria com dados do Banco de fluxo/Inep

Figura 4 - Distribuição do indicador de complexidade da gestão das escolas frequentadas pela coorte por etapas da educação básica nos anos disponíveis.



Fonte: Elaboração própria com dados do Banco de fluxo/Inep

A Tabela 23 apresenta as estatísticas descritivas das variáveis referentes ao efeito das escolas. Essas variáveis não possuem escalas comparáveis: adequação da formação docente varia entre 0 e 100, complexidade da gestão escolar entre 1 e 6 e regularidade docente entre 0 e 5.

Tabela 23 - Estatísticas descritivas

	Adequação da formação docente	Complexidade da gestão escolar	Regularidade docente
Ensino Fundamental I			
Média	55,9	-	3,1
N	3.029.133,00	-	3.083.597,00
Máximo	100	-	5
Mínimo	0	-	0,4
Desvio Padrão	28,6	-	0,7
Variância	816,4	-	0,4
Coefficiente de variação	0,5	-	0,2
Mediana	60,6	-	3,2
Ensino Fundamental II			
Média	58,3	4,1	3,1
N	2.808.955	2.827.979	2.754.867
Máximo	100	6	5
Mínimo	0	1	0,7
Desvio Padrão	21,5	1	0,6
Variância	464,3	1,1	0,3
Coefficiente de variação	0,4	0,3	0,2
Mediana	62,9	4	3,1
N	2.476.143	3.182.019	3.110.059
Máximo	100	6	5
Mínimo	0	1	0,4
Desvio Padrão	26,9	1,3	0,6
Variância	724,2	1,6	0,3
Coefficiente de variação	0,5	0,3	0,2
Mediana	55,6	3,8	3,1
Ensino Médio			
Média	65,5	4,2	3,2
N	1.880.355	1.882.003	1.828.035
Máximo	100	6	5
Mínimo	0	1	0,9
Desvio Padrão	15,9	1	0,5
Variância	252	1	0,3
Coefficiente de variação	0,2	0,2	0,2
Mediana	67,7	4	3,2
N	3.178.677	3.183.411	3.113.367
Máximo	100	6	5
Mínimo	0	1	0,4
Desvio Padrão	21,6	1,2	0,5
Variância	466,1	1,4	0,3
Coefficiente de variação	0,4	0,3	0,2
Mediana	63,8	4	3,2

Fonte: elaboração própria com dados do Banco de fluxo/Inep

1.1. Escolha do modelo

As pesquisas sobre o efeito escola tem por definição seu desenho fundamentado no fato de que os alunos estão aninhados em salas de aula e escolas. Tal desenho requer como metodologia uma modelagem hierárquica que permite

(1) obtenção de melhores estimativas para os parâmetros relativos a unidades específicas, pois por meio do modelo hierárquico é possível obter uma equação para cada escola, fazendo assim uso de toda a informação presente na amostra de forma eficiente; (2) possibilidade de formular e testar hipóteses relativas a efeitos entre níveis; (3) partição da variância em componentes, ou seja, variância do nível 1 (alunos) e variância do nível 2 (escolas) (Jesus e Laros, 2004).

Essa classe de modelos é especialmente adequada para investigações sobre interações entre variáveis de contexto das escolas e variáveis individuais. Por exemplo, o modelo hierárquico é capaz de responder se as desvantagens de ter baixo nível socioeconômico são intensificadas em determinadas escolas (Rumberger e Palardy, 2005).

Apesar da pergunta desta tese ser sobre efeito escola, entendemos inicialmente que para respondê-la seria suficiente estimar o efeito das variáveis contextuais das escola sobre as trajetórias a partir de modelos clássicos de regressão e que o uso desse modelo contemplaria nossa pergunta sobre as características das escolas e seus efeitos sobre as trajetórias. Entretanto, a variável de trajetórias em nível individual, referente aos alunos, fez com que os efeitos da características escolares se dissipassem. O que nos levou a agregar as variáveis de trajetória, de sexo e de raça no nível das escolas e elaborar um modelo linear que contém apenas variáveis agregadas.

Os resíduos dos modelos estatísticos são valores equivalentes a parte do modelo a qual suas variáveis não foram capazes de explicar e em resíduos dos modelos estatísticos educacionais esses resíduos podem conter valores referentes ao efeito escola (Lee, 2008). Assim, a partir dos modelos com variáveis agregadas primeiramente elaboramos um **modelo básico** com as variáveis percentual de meninos nas escolas, percentual de pretos ou indígenas nas escolas, região, dependência administrativa das escolas e nível socioeconômico das escolas. Para entender quais características da escola continham nos resíduos do modelo básico correlacionamos esses resíduos do modelo básico com as variáveis adequação da formação

docente, complexidade da gestão escolar e regularidade docente. Apresentamos os resultados a seguir.

2. RESULTADOS

A Tabela 24 apresenta o modelo básico para o ensino fundamental I. Esse modelo mostra como escolas de nível socioeconômico baixo tendem a ter taxas de trajetórias regulares mais baixas. O beta calculado para o modelo aponta para a tamanho do efeito dessa variável. A Tabela 25 apresenta a correlação entre os resíduos do modelo básico e as variáveis a adequação da formação docente, complexidade da gestão escolar e regularidade docente. A variável adequação da formação docente apresentou correlação enquanto complexidade da gestão escolar e regularidade docente tiveram correlações baixas com resultados próximos de zero.

Tabela 24 – Efeito das características socioeconômicas sobre as trajetórias escolares no ensino fundamental I

Modelo básico	R2 = 0,364	
	Coefficiente	Beta
Percentual de meninos nas escolas	-0,14	-0,06
Percentual de pretos ou indígenas nas escolas	-0,39	-0,15
Norte	-1,71	-0,24
Nordeste	-1,28	-0,26
Sul	-6,23	-0,09
Centro-oeste	-6,54	-0,08
Escola estadual	-1,05	-0,02
Escola municipal	-8,30	-0,17
Escola de NSE no 1º tercil	-17,39	-0,36
Escola com NSE no 2º tercil	-7,85	-0,17
Sem informação de NSE	-35,98	-0,10
Constante	99,52	

Fonte: Elaboração própria com dados do Banco de fluxo/Inep

Nota:* P-valor < 0.00.

A correlação próxima de zero entre os resíduos e as variáveis complexidade da gestão escolar e regularidade docente mostram que essas características da escola não estão relacionadas com as taxas de trajetórias regulares. Enquanto a correlação mais alta entre adequação da formação docente e os resíduos do modelo básico mostram associação positiva entre a variável e as trajetórias regulares, o que aponta para que escolas com maiores taxas de adequação da formação docente tendem a maiores taxas de trajetórias regulares.

Tabela 25 – Correlação entre resíduos do modelo básico e os efeito das características das escolas

	Resíduos	Adequação da formação docente	Regularidade docente	Complexidade da gestão escolar
Resíduos	1			
Adequação da formação docente	0,1510 0,0000	1		
Complexidade da gestão escolar	0,0224 0,0000	0,0243 0,0000		1
Regularidade docente	0,0542 0,0000	0,0098 0,0000	0,0853 0,0000	1

Fonte: Elaboração própria com dados do Banco de fluxo/Inep

Nota:* P-valor < 0.00.

Na Tabela 26 está o modelo básico para as trajetórias escolares no ensino fundamental II. Nesse modelo estão apenas os alunos que chegaram regularmente no primeiro ano do ensino fundamental. Os coeficientes reduzem seus valores, entretanto a tendência se mantém. A Tabela 27 apresenta os resíduos desse modelo, apontando que entre as características das escolas apenas a complexidade da gestão escolar está correlacionada com as trajetórias.

Tabela 26 – Efeito das características socioeconômicas sobre as trajetórias escolares no Ensino fundamental II

Modelo básico	R2 = 0,2781	
	Coefficiente	Beta
Percentual de meninos nas escolas	-0,15	-0,07
Percentual de pretos ou indígenas nas escolas	-0,41	-0,17
Norte	-7,05	-0,11
Nordeste	-8,34	-0,21
Sul	-6,95	-0,13
Centro-oeste	-3,09	-0,04
Escola estadual	-7,82	-0,21
Escola municipal	-12,74	-0,34
Escola de NSE no 1º tercil	-10,53	-0,27
Escola com NSE no 2º tercil	-6,41	-0,17
Sem informação de NSE	-19,06	-0,02
Constante	95,95	

Fonte: Elaboração própria com dados do Banco de fluxo/Inep

Nota:* P-valor < 0.00.

Tabela 27 – Correlação entre resíduos do modelo básico e os efeito das características das escolas

	Resíduos	Adequação da formação docente	Regularidade docente	Complexidade da gestão escolar
Resíduos	1			
Adequação da formação docente	-0,0114 0,0000	1,0000		
Complexidade da gestão escolar	-0,1047 0,0000	0,0566 0,0000	1,0000	
Regularidade docente	-0,0460 0,0000	-0,0035 0,0000	0,1313 0,0000	1,0000

Na Tabela 28 vemos que o modelo do ensino médio apresenta um ajuste baixo, o que é mostrado por seu R2 próximo de zero. Esse resultado pode ter relação com o ano 2020 e com a Covid-19. Como pode ser visto em Gráfico 10 e Gráfico 11 do capítulo 3, a variabilidade entre as variáveis é pequena o que pode explicar o baixo ajuste do modelo. Quanto as correlações, apresentadas na Tabela 29, tanto adequação da formação docente quanto complexidade da gestão escolar apresentam correlações. A primeira positiva a segunda negativa. Então elaboramos o segundo modelo com essas duas variáveis incorporadas.

Tabela 28 – Efeito das características socioeconômicas sobre as trajetórias escolares no Ensino médio

Modelo básico	R2 = 0,0681	
	Coefficiente	Beta
Percentual de meninos nas escolas	-0,08	-0,03
Percentual de pretos ou indígenas nas escolas	-0,4	-0,12
Norte	4,17	0,06
Nordeste	-7,46	-0,17
Sul	4,89	0,08
Centro-oeste	5,98	0,08
Escola estadual	0,81	0,02
Escola municipal	10,59	0,04
Escola de NSE no 1º tercil	2,12	0,05
Escola com NSE no 2º tercil	-2,27	-0,06
Sem informação de NSE	20,57	0,04
Constante	53,86	

Fonte: Elaboração própria com dados do Banco de fluxo/Inep
Nota:* P-valor < 0.00.

Tabela 29 – Correlação entre resíduos do modelo básico e os efeitos das características das escolas

	Resíduos	Adequação da formação docente	Regularidade docente	Complexidade da gestão escolar
Resíduos	1			
Adequação da formação docente	0,1492	1		
	0			
Complexidade da gestão escolar	-0,1036	-0,0002	1	
	0	0,8171		
Regularidade docente	0,0243	-0,07	0,0236	1
	0	0	0	

Para mensurar as associações das variáveis que apresentaram correlações incorporamos as variáveis adequação da formação docente, complexidade da gestão escolar e regularidade docente⁶² aos modelos básicos. Os novos modelos⁶³ estão apresentados na Tabela 30 e confirmam a existência do efeito das escolas.

Tabela 30 – Modelos dos efeitos das características das escolas sobre as trajetórias escolares

	Ensino fundamental I		Ensino fundamental II		Ensino médio	
R2	0,389		0,315		0,114	
	Coefficiente	Beta	Coefficiente	Beta	Coefficiente	Beta
Percentual de meninos nas escolas	-0,14	-0,06	-0,14	-0,07	-0,08	-0,03
Percentual de pretos ou indígenas nas escolas	-0,36	-0,14	-0,40	-0,16	-0,33	-0,10
Norte	-16,55	-0,23	-6,60	-0,11	3,88	0,06
Nordeste	-11,71	-0,24	-7,79	-0,19	-5,86	-0,13
Sul	-6,38	-0,09	-7,39	-0,14	5,04	0,09
Centro-oeste	-7,36	-0,08	-3,16	-0,05	10,19	0,14
Escola estadual	-3,11	-0,05	-6,86	-0,19	2,13	0,04
Escola municipal	-10,00	-0,21	-12,10	-0,32	11,72	0,04
Escola de NSE no 1º tercil	-13,94	-0,29	-10,00	-0,25	4,89	0,11
Escola com NSE no 2º tercil	-6,51	-0,14	-6,22	-0,17	-1,18	-0,03
Sem informação de NSE	-28,17	-0,08	-18,11	-0,02	11,10	0,02
Baixa adequação da formação docente	-7,93	-0,15	0,78	0,02	-10,27	-0,19
Média Adequação da formação docente	-2,09	-0,04	0,17	0,00	-0,40	-0,01
Sem Informação sobre adequação da formação docente	-12,21	-0,12	-34,65	-0,20	-46,46	-0,13
Baixa Regularidade docente	-1,11	-0,02	1,55	0,04	2,31	0,05
Média Regularidade docente	0,26	0,01	1,47	0,04	-2,37	-0,09
Sem Informação sobre Regularidade docente	-6,28	-0,05	0,06	0,00	2,17	0,02
Baixa complexidade da gestão escolar	0,28	0,01	2,42	0,02	7,68	0,10
Média complexidade da gestão escolar	1,40	0,03	4,24	0,11	2,56	0,07
Sem informação sobre complexidade da gestão escolar	6,52	0,03	40,12	0,15	56,51	0,13
Constante	101,25		90,98		51,62	101,25

Fonte: Elaboração própria com dados do Banco de fluxo/Inep

Nota: * P-valor < 0.00.

⁶² Para a incorporação das variáveis adequação da formação docente, complexidade da gestão escolar e regularidade docente ao modelo elas foram divididas em três categorias através de cluster e em seguida binarizadas.

⁶³ Ver apêndice II para detalhes dos modelos.

Os modelos que estimam os efeitos das características das escolas sobre as trajetórias escolares reiteram o que vimos no capítulo 3. O nível socioeconômico baixo tem forte efeito sobre as trajetórias regulares, assim como frequentar escolas municipais e escolas do Norte e Nordeste do país. Sobre as características das escolas, vemos que no ensino fundamental I e no ensino médio escolas com menores taxas de professores com formação docente adequada possuem menores taxas de trajetórias regulares. Enquanto escolas com gestão de baixa e média complexidade parecem contribuir para trajetórias mais regulares no ensino fundamental II e no médio.

2.1. O efeito-professor

O principal achado desse capítulo está associado ao efeito-professor. Vemos que a formação adequada dos docentes reduz intercorrências e a regularidade docente está associada com o aumento das chances de reprovação. Esse achado é de grande importância, pois revela uma associação entre os professores, o contexto escolar e as práticas pedagógicas. A primeira associação é relevante, pois escolas com níveis socioeconômicos mais elevados têm maior probabilidade de atrair e reter professores mais qualificados. Essa tendência, aliada ao fato de que os alunos mais vulneráveis são mais sensíveis à qualidade e à experiência dos professores (Willms, 1992; Darvene, 2011; Madaus, Airasian e Kellaghan, 2008), os quais são menos propensos a adotar práticas meritocráticas, como a reprovação (Ribeiro, Kasmirski, *et al.*, 2018), pode estar relacionada aos resultados encontrados em nossa pesquisa.

Além disso, no Brasil, devido à composição homogênea das salas de aula (Cruz, Xavier e Oliveira, 2021) os professores mais qualificados e experientes têm preferências por turmas com baixas taxas de repetência e/ou bons padrões de comportamento, enquanto os professores iniciantes e/ou menos qualificados ministram turmas com alta concentração de alunos repetentes. Esse contexto acarreta uma configuração apressada de estereótipos que é difícil de ser desfeita. A tendência são os estereótipos em torno dos alunos pobres serem usados pelos professores como justificativa para o fracasso escolar desses alunos (Patto, 1990).

Acreditamos que os estereótipos e a estigmatização (Goffman, 1982) em torno da pobreza e da repetência estão associados com a relação entre regularidade docente e reprovação.

3. Considerações

Diante de nossos achados entendemos que políticas compensatórias (Downey e Condrón, 2016; Simielli, 2017) de formação de professores são urgentes para que o Brasil possa enfrentar suas desigualdades de trajetórias educacionais. Além disso, essas políticas precisam atuar sobre as crenças que professores (principalmente o sem formação docente adequada) têm na reprovação (Ribeiro, Kasmirski, *et al.*, 2018). Caso exista a necessidade de focalização, entendemos que as políticas devem ter início nas escolas com a maioria de alunos de origem social vulnerável, conforme mostramos nos modelos das desigualdades de origem. Isso porque a literatura já sabe que a desigualdade na distribuição de professores mais qualificados afeta as crianças mais pobres, que moram na zona rural, em estados do Norte e Nordeste (Simielli, 2017) e que a eficácia escolar é mais importante para as chances de melhora de vida dos alunos com desvantagens sociais – por impactarem mais sobre eles do que sobre os alunos privilegiados (Coleman e others, 1966; Rutter, Maughan, *et al.*, 2008; Mortimore, Sammons, *et al.*, 1988 *apud* Brooke e Soares, 2008).

Essa sugestão se encontra com o que alguns autores já vêm propondo que é diferenciar o ensino de acordo com princípios da justiça corretiva garantindo a igualdade de aquisição (Crahay, 2013). Nossos resultados sugerem que uma **política compensatória** de professores pode ser um bom ponto de partida para um sistema educacional mais justo e capaz de limitar o peso social da origem no desempenho dos alunos (Crahay e Baye, 2013).

4. Considerações finais

Essa tese é uma pequena parte de uma agenda de pesquisa que se consolida no país. Acreditamos que em breve, juntos, seremos capazes de traçar trajetórias completas das vidas brasileiras, da educação infantil aos destinos no mercado de trabalho formal e informal passando pela educação básica, a profissionalizante e ensino superior e discutir sobre suas desigualdades e as variáveis que são sensíveis as políticas públicas. Seus principais objetivos foram traçar e entender as desigualdades entre as trajetórias escolares na educação básica. Fizemos isso colocando essas trajetórias sob três premissas teóricas através de modelos analíticos: o efeito das origens sociais, o efeito de pares e o efeito escola. Seus principais resultados foram o efeito-professor. O que nos permite reiterar sobre a importância da formação adequada dos docentes. As vantagens desse trabalho são os resultados inéditos que ele apresenta, principalmente sobre o ano de 2020 e sua possibilidade de replicação.

Como já sabemos, há inúmeras dificuldades para os estudos das trajetórias escolares e em consequência disso pouca literatura sobre a temática. Essa escassez se soma a resultados que ainda se desencontram. Uma das formas de reduzir esses desencontros é o compartilhamento dos códigos utilizados nos programas estatísticos para a elaboração dos trabalhos. Neste link <https://github.com/Isis-Arantes> compartilhamos os códigos da construção de nossas bases de dados longitudinais e das análises que fizemos nesta tese. Esperamos que pesquisadores e institutos de pesquisa publiquem seus códigos para que a literatura possa avançar nos estudos das trajetórias educacionais com dados de larga escala.

As limitações deste trabalho estão associadas ao uso de dados sigilosos que fizemos, o que não permitiu a elaboração de modelos mais robustos que poderiam responder com maiores detalhes as perguntas que perseguimos. As pesquisas que se desdobrarão desta tese estão relacionadas com a sua replicação para outras coortes, principalmente as que estarão na escola nos próximos anos.

REFERÊNCIAS

- Agresti, A. **Categorical Data Analysis**. Second edition. ed. New York: Wiley, 2002.
- Almeida, F. A.; Xavier, F. P. Desigualdades no interior da escola: a formação da turma dos repetentes. **Revista Brasileira De Política E Administração Da Educação**, 2021. 109–131.
- Alves, M. T. G. et al. **Inclusão, equidade e desigualdade entre estudantes das escolas públicas de ensino fundamental - versão 1.3 (Relatório de pesquisa)**. [S.l.]. 2021.
- Alves, M. T. G.; Franco, C. A pesquisa em eficácia escolar no Brasil: evidências sobre o efeito das escolas e fatores associados à eficácia escolar. In: Brooke, N.; Soares, F. J. **Pesquisa em eficácia escolar: origem e trajetórias**. Belo Horizonte: Editora UFMG, 2008. p. 482-500.
- Alves, M. T. G.; Soares, J. F. As pesquisas sobre o efeito das escolas: contribuições metodológicas para a sociologia da educação. **Sociedade e Estado**, 22, 2007. 435-473.
- Alves, M. T. G.; Soares, J. F. Efeito-escola e estratificação escolar: o impacto da composição de turmas por nível de habilidade dos alunos. **Educação em Revista**, Belo Horizonte, 45, Junho 2007. 25-58.
- Alves, M. T. G.; Soares, J. F. As pesquisas sobre o efeito das escolas: contribuições metodológicas para a sociologia da educação. **Sociedade e Estado**, Brasília, 22, n. 2. 435-473.
- Alves, M. T. G.; Soares, J. F.; Xavier, F. P. Desigualdades educacionais no ensino fundamental de 2005 a 2013: hiato entre grupos sociais. **Revista Brasileira de Sociologia**, 4, 2016. 49-82.
- Alvim, M. S.; Gonze, R. G.; Graaf, J. V. **TED 8750 - PRICE: Privacidade nos censos educacionais**. UFMG. Belo Horizonte. 2020.
- Artes, A. et al. Estudo Desigualdades na educação brasileira: ressignificação do abandono escolar no contexto de pandemia. **Educação escolar em tempos de pandemia**, 2021. Disponível em: <<https://www.fcc.org.br/fcc/educacao-pesquisa/abandono-escolar-e-a-pandemia-no-brasil-efeitos-nas-desigualdades-escolares/>>. Acesso em: 2023.
- Barbosa, M. E. F.; Fernandes, C. A escola brasileira faz diferença? Uma investigação dos efeitos da escola na proficiência em matemática dos alunos da 4ª série. In: Franco, C. **Avaliação, ciclos e promoção na educação**. Porto Alegre: Artmed, 2001.
- Barros, R. P.; Mendonça, R. **Consequências da repetência sobre o desempenho educacional**. Brasília: Ministério da Educação e Desporto, 1998.
- Bourdieu, P.; Passeron, J. **A reprodução: elementos para uma teoria do sistema de ensino**. Rio de Janeiro: Francisco Alves, 1975.
- Brasil. **Constituição [de 1988] da República Federativa do Brasil**. Diário Oficial da União [da] República Federativa do Brasil. Brasília. 1988.
- Brasil. Constituição [de 1988] da República Federativa do Brasil, Brasília, DF, 05 Outubro 1988.
- Brasil. Lei n. 9.394. Diretrizes e Bases da Educação Nacional, Brasília, 23 Dezembro 1996.
- Brasil. Lei nº 10.172. Plano Nacional de Educação (PNE). Diário Oficial da União [da] República Federativa do Brasil, Brasília, DF, 10 Janeiro 2001.

Brasil. Lei nº. 11.114. Altera os arts. 6º , 30, 32 e 87 da Lei nº 9.394, de 20 de dezembro de 1996, com o objetivo de tornar obrigatório o início do ensino fundamental aos seis anos de idade, 2005.

Brasil. Decreto nº 6.094. Plano de Metas Compromisso Todos pela Educação, 24 Abril 2007.

Brasil. Lei nº 13.005. Plano Nacional de Metas Compromisso Todos pela Educação, pela União Federal, em regime de colaboração com municípios, Distrito Federal e Estados, e a participação das famílias e da comunidade, 25 Junho 2014.

Brasil. Lei nº. 13.005. Plano Nacional de Educação - PNE, 26 Junho 2014.

Brasil, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Serviço de acesso a dados protegidos (Sedap):guia do usuário, Brasília, v. 2, 2020.

Brasil, Ministério da Educação e do Desporto. Plano Decenal de Educação Para Todos, Brasília, DF, 1993.

Bressoux, P. As pesquisas sobre o efeito-escola e o efeito-professor. **Educação em Revista**, Belo Horizonte, 38, Dezembro 2003. 17-88.

Brito, M. M. A. d. ovas tendências ou velhas persistências? Modernização e expansão educacional no Brasil. **Cadernos De Pesquisa**, 2017. 224–263.

Brooke, N.; Soares, F. J. **Pesquisa em Eficácia Escolar: origem e trajetórias**. Belo Horizonte: UFMG, 2008.

CAGED. **Ministério do Trabalho e Emprego**, Brasília, 2019.

Carvalho, J. A. M.; Sawyer, D. O.; Rodrigues, R. N. **Introdução a alguns conceitos básicos e medidas em demografia**. São Paulo: ABEP, 1994.

Castellini, T. S. J. **Clima escolar e desempenho em avaliação externa. Adaptação de um instrumento de avaliação do clima e relação com os resultados obtidos em provas externas municipais**. Universidade Estadual de Marília. Marília. 2019.

Coleman, J. S.; others, a. **Equality of educational opportunity**. U.S. Department of Health, Education, and Welfare. Washington. 1966.

Correa, E. V.; Bonamino, A.; Soares, T. M. Evidências do efeito da repetência nos primeiros anos escolares. **Estudos em Avaliação Educacional**, São Paulo, 25, n. 59, 2014. 242-269.

Crahay, M. Como a escola pode ser mais justa e mais eficaz? **Cadernos Cenpec**, 3, 2013. 9-40.

Crahay, M.; Baye, A. Existem escolas justas e eficazes? Esboço de resposta baseado no PISA 2009. **Cadernos de Pesquisa**, 43, n. 150, 2013. 858–883.

Cruz, T. M.; Xavier, F. P.; Oliveira, V. C. d. Desigualdades no interior das escolas: uma análise da composição social das turmas. **Revista De Gestão E Avaliação Educacional**, Santa Maria, 10, n. 19, 2021. 1–19.

Darvene, C. Trajetórias Escolares. In: VAN ZANTÉN, A. **Dicionário de Educação**. Petrópolis-RJ: Vozes, 2011. p. 768-770.

DATASUS. **Ministério da Saúde**. Disponível em: <<https://datasus.saude.gov.br/>>. Acesso em: 2023.

Downey, D. B.; Condrón, D. J. Fifty Years since the Coleman Report: Rethinking the Relationship between Schools and Inequality. **Sociology of Education**, 89, 2016. 207-220.

Dubet, F. O que é uma escola justa? **Cadernos de pesquisa**, 34, 2004. 539-555.

Durkheim, É. A educação – sua natureza e função. In: Durkheim, É. **Educação e Sociologia**. Rio de Janeiro: Melhoramentos, 1973.

Faria, E. M. Os alunos reprovados no Brasil: uma análise das proficiências e das taxas de abandono por meio das avaliações Prova Brasil e Pisa. **Estudando Educação: portal de pesquisas em educação.**, São Paulo, 01, 2011.

Ferrão, M. E. Retenção escolar e desenvolvimento cognitivo no ensino básico. In: Nunes, L. **A escola e o desempenho dos alunos**. Lisboa: Fundação Francisco Manuel dos Santos, 2015. p. 87-104.

Ferrão, M. E. Retenção escolar e desenvolvimento cognitivo no ensino básico. In: Nunes, L. **A escola e o desempenho dos alunos**. Lisboa: Fundação Francisco Manuel dos Santos, 2015. p. 87-104.

Ferrão, M. E. The evaluation of students' progression in lower secondary education in Brazil: Exploring the path for equity. **Stud Educ Eval**, 2022.

Ferrão, M. E.; Beltrão, K. I.; Santos, D. P. O impacto de políticas de não repetência sobre o aprendizado dos alunos da 4ª série. **Pesquisa e Planejamento Econômico**, 32, n. 3, 2002. 495-514.

Ferrão, M. E.; Costa, P. M.; Matos, D. A. S. The relevance of the school socioeconomic composition and school proportion of repeaters on grade repetition in Brazil: a multilevel logistic model of PISA 2012. **Large-scale Assessments in Education**, 5, n. 1, 2017.

Ferraro, A. R. Direito à Educação no Brasil e dívida educacional: e se o povo cobrasse? **Educação e Pesquisa**, São Paulo, maio/agosto 2008. 273-289.

Fonseca, I. et al. Monitoramento da Permanência na Educação Básica com o Indicador de Regularidade de Trajetórias Educacionais, 2022. Manuscrito submetido para publicação. Texto disponível com os autores.

Francis, A.; Tannuri-Pianto, M. Using Brazil's Racial Continuum to Examine the Short-Term Effects of Affirmative Action in Higher Education. **The Journal of Human Resources**, 47, 2012. 754-784.

Goffman, E. **Estigma**: notas sobre a manipulação da identidade deteriorada. Rio de Janeiro: Zahar, 1982.

Gomes, C. A. O funcionalismo e a educação. In: Gomes, C. A. **A educação em perspectiva sociológica**. São Paulo: EPU, 1994. p. 31-44.

Gomes-Neto; Hanushek, E. A. Causes and consequences of grade repetition: Evidence from Brazil. **Econ. Dev. Cult. Change**, 43, n. 1, 1994.

Heckman, J.; Lochner, L.; Todd, P. Earnings functions and rates of return. **Journal of Human Capital**, 1, 2008. 1-31.

IBGE, 2023. Disponível em: <<https://www.ibge.gov.br/estatisticas/sociais/educacao.html>>.

Inep. **Indicador de adequação da formação do docente da educação básica**. Brasília. 2014. (20).

Inep. **Indicador para mensurar a complexidade da gestão nas escolas a partir dos dados do Censo Escolar da Educação Básica**. Brasília. 2014. (40).

Inep. **Indicador de regularidade do docente da Educação Básica**. Brasília. 2015. (11).

Inep. **Estimativas de fluxo escolar a partir do acompanhamento longitudinal dos registros de aluno do Censo Escolar do período 2007-2016**. Brasília. 2017. (8).

Inep. **Guia do usuário**. Brasília. 2021.

Inep, 2022. Disponível em: <<https://www.gov.br/inep/pt-br/assuntos/noticias/institucional/inep-publica-microdados-do-enem-2020-e-do-censo-escolar-2021>>.

Inep, 2023. Disponível em: <<https://www.gov.br/inep/pt-br/areas-de-atuacao/pesquisas-estatisticas-e-indicadores/censo-escolar>>.

Inep, 2023. Disponível em: <<https://www.gov.br/inep/pt-br/areas-de-atuacao/pesquisas-estatisticas-e-indicadores/censo-escolar>>.

Jacob, B. A.; Lefgren, L. The Effect of Grade Retention on High School Completion. **American Economic Journal: Applied Economics**, 1, n. 3, 2009. 33-58.

Jacobs, G. S. **Getting Around Brown: Desegregation, Development, and the Columbus Public Schools (Urban Life and Urban Landscape)**. Columbus: Ohio State University Press, 1998.

Jesus, G. R. d.; Laros, J. A. Eficácia escolar: Regressão Multinível com Dados de Avaliação em Larga Escala. **Avaliação Psicológica**, Campinas - SP, 2004.

Klein, R. **Produção e utilização de indicadores educacionais**. LNCC/CNPq. Rio de Janeiro. 1995.

Klein, R.; Ribeiro, S. C. O censo educacional e o modelo de fluxo: o problema da repetência. **Revista Brasileira de Estatística**, Rio de Janeiro, 52, 1991. 05-45.

Leão, N. e. a. Trabalho e vida das mulheres na pandemia. In: Oliveira, D.; Pochmann, M. **Devastação do trabalho : a classe do labor na crise da pandemia**. [S.l.]: Positiva, 2020. p. 289–310.

Lee, V. E. Utilização de modelos hierárquicos lineares para estudar contextos sociais: o caso dos efeitos da escola. In: BROOKE, N.; SOARES, J. F. **Pesquisa em eficácia escolar: origem e trajetória**. Belo Horizonte: UFMG, 2008.

Lee, V. E. Dados longitudinais em educação: um componente essencial da abordagem de valor agregado no que se refere à avaliação de desempenho escolar. **Est. Aval. Educ.**, São Paulo, 21, set/dez 2010. 531-542.

Lee, V. E.; Burkam, D. T. **Inequality at the starting gate: Social background differences in achievement as children begin school**. Washington: Economic Policy Institute, 2002.

Lima, M. F. M.; Paes de Carvalho, C. Transição para o ensino médio: Gestão de política de correção de fluxo e trajetórias escolares. **Arquivos Analíticos de Políticas Educativas**, 29, n. 51, 2021.

- Lochner, L. Non-production benefits of education: Crime, health, and good citizenship. **Handbook of the Economics of Education**, 4, 2011.
- Long, J. S.; Freese, J. **Regression Models for Categorical Dependent Variables Using Stata**. College Station, TX: US: Stata Press, 2003.
- Lunelli, P.; Bruel, A. L. Trajetórias escolares e segregação socioespacial em Curitiba: uma análise sobre os estudantes vinculados ao Programa Bolsa Família. **Jornal de Políticas Educacionais**, Curitiba, Maio 2022. 01-27.
- Machado, M. R. C. Políticas de não-repetência na educação básica brasileira: impactos na trajetória escolar e na proficiência dos alunos. **Tese de doutorado**, Brasília, 2020.
- Madaus, G.; Airasian, P.; Kellaghan, T. Insumos escolares, processos e recursos. In: Brooke, N.; Soares, J. F. **Pesquisa em eficácia escolar: origem e trajetórias**. Belo Horizonte: Editora UFMG, 2008. p. 273-296.
- Malaguth, T. Z. **Migrações e fluxo escolar da coorte de estudantes de 2008 a 2019, em Minas Gerais (dissertação de mestrado)**. Universidade Federal de Minas Gerais. Belo Horizonte. 2022.
- Melo, G.; Machado, A. Educational trajectories. Evidence from Uruguay. **International Journal of Educational Research**, 92, 2018. 110-134.
- Monticelli, T. Divisão sexual do trabalho, classe e pandemia: novas percepções? **Sociedade e Estado**, 2021. 83–107.
- Mortimore, P. et al. **School matters**. [S.l.]: University of California Press, 1988.
- Nogueira, C. M. M.; Fortes, M. F. A Importância dos Estudos sobre Trajetórias Escolares na Sociologia da Educação Contemporânea. **Paidéia**, Belo Horizonte, 3, n. 2, 2004. 57-74.
- Oliveira, A. P.; Menezes Filho, N. **Idade de ingresso escolar, repetência e evasão escolar no Brasil: uma abordagem para estimação de efeitos causais**. 46º Encontro Nacional de Economia. Rio de Janeiro: Apec. 2018.
- Oliveira, L. F. B. d.; Soares, S. S. D. Determinantes da Repetência Escolar no Brasil: Uma Análise de Painel dos Censos Escolares entre 2007 e 2010. **Texto para Discussão**, Brasília, n. 1706, fevereiro 2012.
- Oliveira, R. L. P.; Araújo, G. C. Qualidade do ensino: uma nova dimensão da luta pelo direito à educação. In: Coutinho, S. A. **Qualidade da educação: democratização do acesso, permanência, avaliação, condições de participação e aprendizagem**. Curitiba: APPRIS, v. 4, 2015. p. 29-50.
- Oliveira, R. P. Da universalização do ensino fundamental ao desafio da qualidade: uma análise histórica. **Educação & Sociedade**, 28, 2007.
- OPNE. OPNE, 2020. Disponível em: <<https://www.observatoriodopne.org.br/>>.
- Paiva, V. Revolução educacional e contradições da massificação do ensino. **Contemporaneidade e Educação**, São Paulo, 3, n. 3, 1998. 44-99.
- Parente, M.; Lück, H. Mecanismos e experiências de correção do fluxo escolar no ensino fundamental. **Texto para Discussão**, Brasília, n. 1032, 2004.

Parsons, T. A classe como sistema social. In: Brito, S. **Sociologia da juventude 3: a vida coletiva juvenil**. [S.l.]: Jorge Zahar, 1968. p. 47-74.

Patto, M. H. S. **A produção do fracasso escolar**. São Paulo: T.A. Editor, 1990.

Paula, T. **Eficácia escolar sobre o aprendizado em matemática: um estudo longitudinal sobre como a desigualdade prejudica a estimação do efeito das escolas para estudantes da rede municipal de Belo Horizonte (tese de doutorado)**. Universidade Federal de Minas Gerais. Belo Horizonte. 2020.

Paula, T.; Alves, M. T. G. **Apontamentos sobre as taxas de fluxo escolar produzidas pelo Inep: um conselho aos navegantes**. X Reunião Da Abave – Avaliação Educacional No Brasil: “O Desafio Da Qualidade”. Anais. São Paulo: Associação Brasileira de Avaliação Educacional. 2019.

Pedraza, D. F. Qualidade do Sistema de Informações sobre Nascidos Vivos (Sinasc): análise crítica da literatura. **Cienc. Saúde Coletiva**, 17, n. 10, out 2012. 2729-37.

Petit, A. As teorias gerais – II A corrente funcionalista. In: Petit, A. **Produção da escola – produção da sociedade**. Porto Alegre: Artes Médicas, 1994. p. 11-21.

Petrus, J. **A tipologia da pré-escola cursada e sua relação com a trajetória de crianças nos primeiros anos do Ensino Fundamental (dissertação de mestrado)**. Universidade Federal de Minas Gerais. Belo Horizonte. 2022.

Ramos Simielli, L. E. Equidade e oportunidades educacionais: O acesso a professores no Brasil. **Arquivos Analíticos de Políticas Educativas**, 25, 8 maio 2017. 1-30.

Rangel Rigotti, J. I.; Hadad, R. M. **An analysis of the relationship between internal migration and education in Brazil**. Unesco. Paris. 2018.

Rangel Rigotti, J. I.; Signorini, B. A.; Hadad, R. M. Migração intermunicipal de estudantes do ensino básico do Brasil entre 2007 e 2015. **Perspectiva**, 38, n. 4, 2020. 1-24.

Ribeiro, C. A. C. Desigualdade de oportunidades e resultados educacionais no Brasil. **Dados**, Rio de Janeiro, 54, n. 1, 2011. 41-87.

Ribeiro, S. C. A Pedagogia da Repetência. **Estudos Avançados**, Rio de Janeiro, 1991.

Ribeiro, V. M. et al. Crença de professores sobre reprovação escolar. **Educação em Revista**, Belo Horizonte, 2018.

Rigotti, J. I. R.; Cerqueira, C. .
As bases de dados do INEP e os indicadores educacionais: conceitos e aplicações. In: Rios-Neto, E. L. G.; Riani, J. L. R. **Introdução à demografia da educação**. Campinas: Associação Brasileira de Estudos Populacionais - ABEP, 2004.

Roderick, M. .; Nagaoka, J. Retention Under Chicago's High-Stakes Testing Program: Helpful, Harmful, or Harmless? **Educational Evaluation and Policy**, 2005.

Roderick, M. Grade Retention and School Dropout: Investigating the Association. **American Educational Research Journal**, 4, 1994. 729-59.

Rosemberg, F.; Madsen, N. Educação formal, mulheres e gênero no Brasil contemporâneo. In: Barsted, L. L. a. P. J. **O Progresso das Mulheres no Brasil 2003-2010**. Rio de Janeiro: CEPIA, 2011.

Rumberger, R. W.; Palardy, G. J. Does segregation still matter? The impact of student composition on academic achievement in high school. **Teachers College Record**, 107, n. 9, 2005. 1999-2045.

Rutter, M. et al. Conclusões, especulações e implicações. In: Brooke, N.; Soares, J. F. **Pesquisa em eficácia escolar: origem e trajetórias**. Belo Horizonte: Editoria UFMG, 2008. p. 225-251.

Scheerens, J. School Effectiveness Research and the Development of Process Indicators of School Functioning, School Effectiveness and School Improvement. **An International Journal of Research, Policy and Practice**, 1, n. 1, 1990. 61-80.

Senkevics, A. S.; Carvalho, M. P. 'O que você quer ser quando crescer?' Escolarização e gênero entre crianças de camadas populares urbanas. **Revista Brasileira de Estudos Pedagógicos RBEP-INEP**, 97, 2016. 179-194.

Silva, T.; Mendes, F.; Faria, C. A. Aplicação de um modelo de escolha discreta para análise da divisão modal em cidades de porte médio, Lisboa, 2010.

Simielli, L. E. R. Equidade e oportunidades educacionais: O acesso a professores no Brasil. **Education Policy Analysis Archives**, 25, 2017. 1–30.

Soares, J. F.; Alves, M. T. G.; Fonseca, J. A. Trajetórias educacionais como evidência da qualidade da educação básica brasileira. **Revista Brasileira de Estudos de População**, 38, 2021. 1-21.

Telles, E.; Paschel, T. Who is Black, White and Mixed Race? How Skin color, Status and Nation Shape Racial Classification in Latin America. **American Journal of Sociology**, 120, n. 3, 2014. 864-907.

Thonstad, T. **Analysing and projecting school enrolment in developing countries: a manual of methodology**. Unesco. Paris. 1980.

Todos Pela Educação. **Anuário Brasileiro da Educação Básica**. [S.l.]. 2021.

Unesco. **Sustainable development begins with education: how education can contribute to the proposed Post-2015 Goals**. Paris. 2014.

Unesco. **Incheon Declaration and Education Framework for Action: Towards Inclusive and Equitable Quality Education and Lifelong Education for All**. Brasília. 2016.

Unesco. **Qualidade da infraestrutura das escolas públicas do ensino fundamental no Brasil**. Brasília. 2019.

Unesco. **Inclusão, equidade e desigualdades entre estudantes das escolas públicas de ensino fundamental no Brasil**. [S.l.]. 2022.

Valbuena, J.; Gil, M. Effects of grade retention policies: a literature review of empirical studies applying causal inference. **Journal of Economic Surveys**, 00, n. 0, 2020. 1–44.

Vinha, T. et al. **Da escola para a vida em sociedade: o valor da convivência democrática**. Americana: Adonis, 2017.

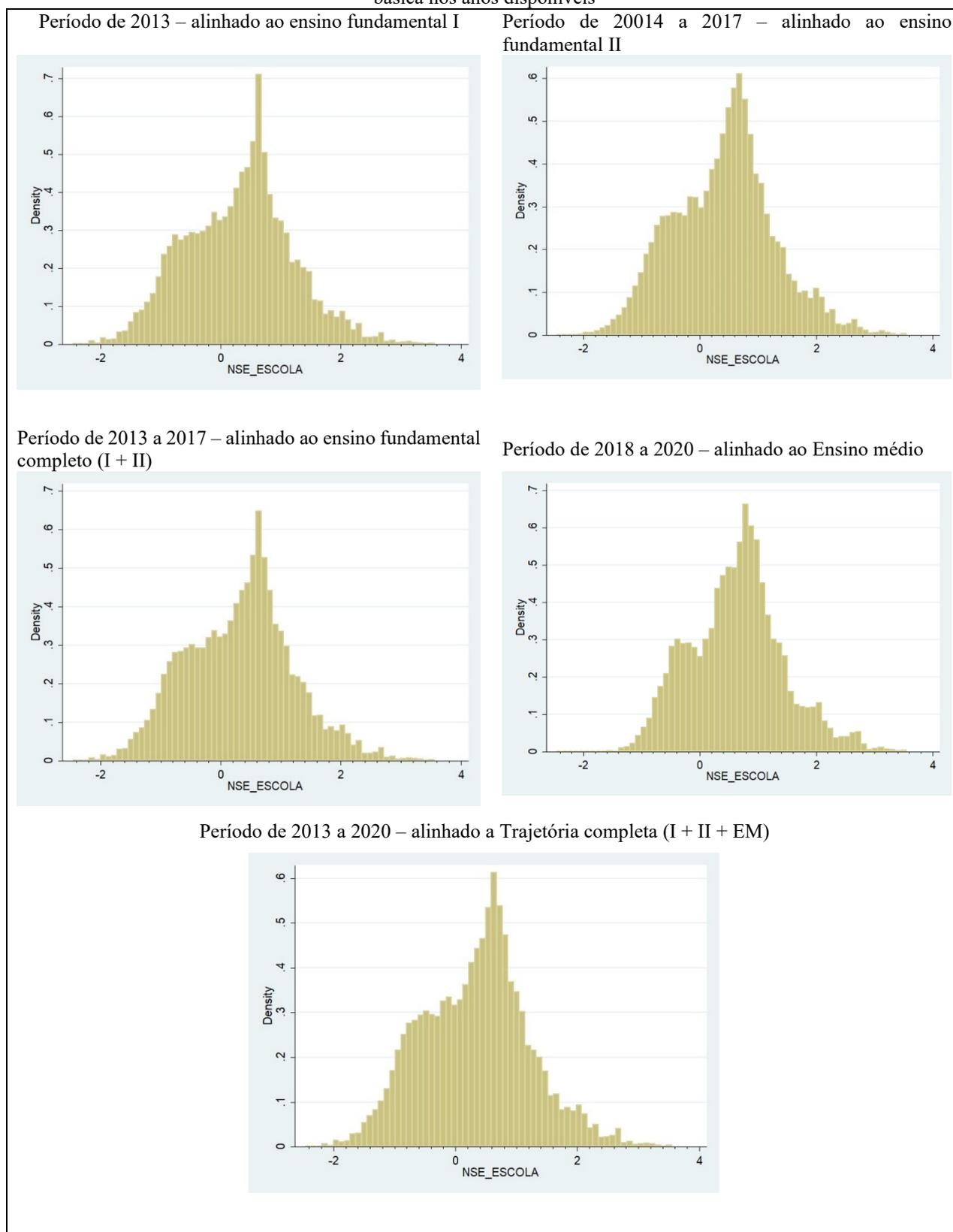
Willms, J. D. **Monitoring School Performance: A Guide for Educators**. Washington: The Falmer Press, 1992.

Xavier, F. P.; Neves, J. A. B. Estrutura social e transmissão intergeracional de status: uma análise hierárquica. **Revista Brasileira de Estudos de População (impresso)**, 29, 2012. 259-275.

Zago, N. Processos de escolarização nos meios populares - As contradições da obrigatoriedade escolar. In: NOGUEIRA, M. A. . R. G. e. Z. N. **Família e escola. Trajetórias de escolarização em camadas populares e médias**. 2. ed. Petrópolis: Vozes, 2003.

APÊNDICE A

Figura 5 - Distribuição do nível socioeconômico das escolas frequentadas pela coorte por etapas da educação básica nos anos disponíveis



Fonte: Elaboração própria com dados do Banco de fluxo/Inep

APÊNDICE B

Tabela 31 - Regressão logística multinomial do modelo de desigualdades de origem sobre as trajetórias educacionais completas (2009-2020)

Iteration 0: log likelihood = -5030643.5

Iteration 1: log likelihood = -4859525.6

Iteration 2: log likelihood = -4854874.4

Iteration 3: log likelihood = -4854845.7

Iteration 4: log likelihood = -4854845.7

Multinomial logistic regression

Number of obs = 3206518

LR chi2(52) = 367519.97

Log likelihood = -4846883.5

Prob > chi2 = 0.0000

Pseudo R2 = 0.0365

indicador	Coef	Std. Err.	z	P> z	[95% Conf, Interval]	
1	(base outcome)					
2						
masculino	.3599595	.0037371	96.32	0.000	.3526349	.3672842
no_declarada3	.2255604	.0052384	43.06	0.000	.2152934	.2358274
preta_indgena3	.5358831	.0100544	53.30	0.000	.5161768	.5555894
parda3	.2209268	.0048829	45.24	0.000	.2113564	.2304972
norte	.2518102	.0072366	34.80	0.000	.2376267	.2659937
nordeste	.4187933	.0062004	67.54	0.000	.4066407	.4309459
sul	.03443	.0060127	5.73	0.000	.0226453	.0462147
centroeste	-.1155247	.0072987	-15.83	0.000	-.1298298	-.1012195
est_pem	.3521887	.0066956	52.60	0.000	.3390655	.3653119
mun_pem	.4372174	.0069677	62.75	0.000	.4235609	.4508738
t1	.1152484	.0083771	13.76	0.000	.0988296	.1316671
t2	.1612934	.0058761	27.45	0.000	.1497763	.1728104
tsi	-.0070913	.0694356	-0.10	0.919	-.1431826	.1290001
_cons	-1.447.015	.0062259	-232.42	0.000	-1.459.217	-1.434.812
3						
masculino	.7904362	.003712	212.94	0.000	.7831609	.7977116
no_declarada3	.4101263	.0053109	77.22	0.000	.3997172	.4205355
preta_indgena3	.962931	.0093659	102.81	0.000	.9445742	.9812879
parda3	.3866056	.0049451	78.18	0.000	.3769133	.3962979
norte	.4945546	.0070554	70.10	0.000	.4807263	.5083828
nordeste	.6334002	.0062182	101.86	0.000	.6212127	.6455877
sul	.4107083	.0060242	68.18	0.000	.398901	.4225155
centroeste	.0708374	.0073292	9.67	0.000	.0564725	.0852024

est_pem	.8029387	.0079332	101.21	0.000	.7873899	.8184875
mun_pem	1.329.416	.0079945	166.29	0.000	1.313.747	1.345.085
t1	.5016065	.0086683	57.87	0.000	.4846169	.518596
t2	.4584142	.0065058	70.46	0.000	.4456632	.4711653
tsi	.9176701	.0566207	16.21	0.000	.8066956	1.028.645
_cons	-2.787.337	.0079493	-350.64	0.000	-2.802.918	-2.771.757
4						
masculino	.2138015	.0030831	69.35	0.000	.2077586	.2198443
no_declarada3	.0821595	.0042238	19.45	0.000	.073881	.0904381
preta_indgena3	.3039076	.0087012	34.93	0.000	.2868535	.3209616
parda3	.0267325	.003976	6.72	0.000	.0189398	.0345252
norte	-.0396562	.0062159	-6.38	0.000	-.0518392	-.0274733
nordeste	.2854954	.0051209	55.75	0.000	.2754586	.2955322
sul	-.1428638	.0048621	-29.38	0.000	-.1523934	-.1333342
centroeste	-.3597223	.0060744	-59.22	0.000	-.3716279	-.3478168
est_pem	-.0144398	.0051973	-2.78	0.005	-.0246263	-.0042532
mun_pem	.0716651	.0054628	13.12	0.000	.0609582	.082372
t1	.1411321	.006896	20.47	0.000	.1276162	.154648
t2	.1562989	.0046699	33.47	0.000	.1471462	.1654517
tsi	.1291515	.0550989	2.34	0.019	.0211596	.2371433
_cons	-.1438525	.0045741	-31.45	0.000	-.1528176	-.1348873
masculino	.3862133	.0037611	102.69	0.000	.3788416	.393585
no_declarada3	.4366712	.0053999	80.87	0.000	.4260875	.4472548
preta_indgena3	.8439494	.009755	86.51	0.000	.82483	.8630688
parda3	.3128378	.0051355	60.92	0.000	.3027725	.3229032
norte	.3881314	.0072318	53.67	0.000	.3739573	.4023055
nordeste	.3780045	.0063289	59.73	0.000	.3656001	.390409
sul	.2277178	.0066167	34.42	0.000	.2147493	.2406863
centroeste	.0502445	.0075857	6.62	0.000	.0353768	.0651122
est_pem	-.4900364	.0073329	-66.83	0.000	-.5044087	-.4756641
mun_pem	.4910271	.007142	68.75	0.000	.477029	.5050252
t1	1.076.583	.0092986	115.78	0.000	1.058.358	1.094.808
t2	.7790204	.0070038	111.23	0.000	.7652931	.7927477
tsi	2.927.331	.0450309	65.01	0.000	2.839.072	301.559
_cons	-1.984.053	.0070408	-281.80	0.000	-1.997.852	-1.970.253

Tabela 32 - Modelo referente ao Gráfico 24 - Valores preditos marginais da trajetória regular (95% de intervalo de confiança)

Iteration 0: log likelihood = -5030643.5

Iteration 1: log likelihood = -4864462.1

Iteration 2: log likelihood = -4860057

Iteration 3: log likelihood = -4860030.3

Iteration 4: log likelihood = -4860030.3

Multinomial logistic regression

Number of obs = 3206518

LR chi2(36) = 341226.52

Prob > chi2 = 0.0000

Pseudo R2 = 0.0339

Log likelihood = -4860030.3

indicador	Coef	Std. Err.	z	P> z	[95% Conf, Interval]	
1	(base outcome)					
2						
1.sexo2	-.3593998	.0037361	-96.20	0.000	-.3667224	-.3520772
raca2						
1	-.5501869	.0100322	-54.84	0.000	-.5698497	-.5305241
2	-.3140524	.0096979	-32.38	0.000	-.3330599	-.2950449
regiao						
2	.1762514	.0068296	25.81	0.000	.1628657	.1896371
3	-.2493643	.006612	-37.71	0.000	-.2623236	-.236405
4	-.2343927	.0080024	-29.29	0.000	-.2500771	-.2187083
5	-.3579996	.0087308	-41.00	0.000	-.3751117	-.3408875
pem						
2	.4393472	.0055933	78.55	0.000	.4283845	.4503098
3	.5158046	.0056535	91.24	0.000	.504724	.5268852
_cons	-.2511056	.012148	-20.67	0.000	-.2749152	-.227296
3						
1.sexo2	-.7894299	.0037086	-212.87	0.000	-.7966986	-.7821612
raca2						

	-					
1	1.005.118	.009339	-107.63	0.000	-1.023.422	-.9868142
2	-.5744277	.0088712	-64.75	0.000	-.5918148	-.5570405
regiao						
2	.1671862	.0063344	26.39	0.000	.1547709	.1796015
3	-.5664544	.0063301	-89.49	0.000	-.5788612	-.5540477
4	-.2170788	.0076214	-28.48	0.000	-.2320165	-.2021412
5	-.4697047	.008411	-55.84	0.000	-.4861899	-.4532195
pem						
2	1.061.528	.0069623	152.47	0.000	1.047.882	1.075.173
3	1.595.576	.0068236	233.83	0.000	1.582.202	160.895
_cons	-.3143491	.0119855	-26.23	0.000	-.3378403	-.2908578
4						
1.sexo2	-.2132635	.003082	-69.20	0.000	-.2193041	-.2072228
raca2						
1	-.3217267	.0086821	-37.06	0.000	-.3387433	-.3047102
2	-.2581487	.0084851	-30.42	0.000	-.2747793	-.2415181
regiao						
2	.3370907	.0059776	56.39	0.000	.3253748	.3488066
3	.029883	.0057178	5.23	0.000	.0186764	.0410897
4	-.1288253	.0068225	-18.88	0.000	-.1421972	-.1154535
5	-.3144131	.0075147	-41.84	0.000	-.3291415	-.2996846
pem						
2	.0689942	.0042466	16.25	0.000	.060671	.0773173
3	.153195	.0043525	35.20	0.000	.1446643	.1617256
_cons	.3958001	.0104185	37.99	0.000	.3753801	.41622
5						
1.sexo2	-.3850692	.0037471	-102.77	0.000	-.3924133	-.3777251
raca2						
1	-.947752	.0096739	-97.97	0.000	-.9667124	-.9287915
2	-.5251233	.0091638	-57.30	0.000	-.543084	-.5071626

regiao						
2	.0394194	.0063152	6.24	0.000	.0270418	.0517969
3	-.6553255	.0063932	-102.50	0.000	-.667856	-.6427951
4	-.5225782	.0080512	-64.91	0.000	-.5383583	-.5067982
5	-.5346803	.0085814	-62.31	0.000	-.5514997	-.517861
pem						
2	-.0354991	.0060889	-5.83	0.000	-.0474332	-.0235651
3	.9898356	.0056413	175.46	0.000	.9787789	1.000.892
_cons	.1901812	.0116136	16.38	0.000	.167419	.2129433
Delta-method						
	Margin	Std. Err.	z	P>	[95% Conf. Interval]	
pem#raca2#sexo2						
1 0 0	.2118009	.0013146	161.11	0.000	.2092243	.2143776
1 0 1	.2758243	.0015737	175.27	0.000	.2727398	.2789087
1 1 0	.3214788	.0008199	392.10	0.000	.3198718	.3230858
1 1 1	.3925815	.0008883	441.96	0.000	.3908405	.3943225
1 2 0	.2793918	.0007289	383.33	0.000	.2779632	.2808203
1 2 1	.3504765	.0008107	432.31	0.000	.3488876	.3520655
2 0 0	.1596631	.0009917	160.99	0.000	.1577193	.1616069
2 0 1	.2244165	.0012768	175.77	0.000	.221914	.2269189
2 1 0	.2631181	.0006125	429.56	0.000	.2619176	.2643186
2 1 1	.3398283	.0006891	493.18	0.000	.3384778	.3411788
2 2 0	.2203603	.0004755	463.41	0.000	.2194283	.2212923
2 2 1	.2953051	.0005512	535.74	0.000	.2942247	.2963854
3 0 0	.1100978	.0007308	150.65	0.000	.1086654	.1115302
3 0 1	.1627748	.0010041	162.11	0.000	.1608067	.1647428
3 1 0	.2022015	.0005346	378.21	0.000	.2011537	.2032494
3 1 1	.2743794	.0006406	428.29	0.000	.2731238	.275635
3 2 0	.1607837	.000378	425.31	0.000	.1600427	.1615246
3 2 1	.2270059	.0004699	483.11	0.000	.226085	.2279269

Tabela 33 - Regressão linear e correlações referentes ao modelo básico - Ensino fundamental I

Source	SS	df	MS	Number of obs = 3197237
Model	601657715	11	54696155.9	F (11,3197225) =
Residual	1.0535e+093197225		329.500.312	Prob > F = 0.0000
Total	1.6551e+093197236	517.679.755		R-squared = 0.3635

Adj R-squared = 0.3635

Root MSE = 18152

traj_a	Coef.	Std. Err.	t	P> t	Beta
meninos_a	-.1385653	.000961	-144.19	0.000	-.0643578
pretos_a	-.3893437	.0011585	-336.08	0.000	-.1522728
norte	-1.710.936	.0391569	-436.94	0.000	-.236794
nordeste	-1.284.578	.0334038	-384.56	0.000	-.2615311
sul	-6.232.403	.0335667	-185.67	0.000	-.0907853
centroeste	-6.535.404	.0409819	-159.47	0.000	-.075273
est_pem	-1.051.834	.0394394	-26.67	0.000	-.0176451
mun_pem	-8.300.551	.0350722	-236.67	0.000	-.1729551
baixo_nse	-1.738.527	.0458168	-379.45	0.000	-.3560835
medio_nse	-784.564	.031898	-245.96	0.000	-.1722839
SI_nse	-3.560.598	.1682647	-211.61	0.000	-.0960536
_cons	9.951.737	.0575031	1730.64	0.000	.

predict y_pred

(option xb assumed; fitted values)

predict residuals, residuals

Pwcorr residuals afd_2013 ncg2013ird_2013 sig

residu~s	afd_2013	ncg2013	ird_2013	
residuals	1			
afd_2013	0.1510 0.0000	1		
ncg2013	0.0224 0.0000	0.0243 0.0000	1	
ird_2013	0.0542 0.0000	0.0098 0.0000	0.0853 0.0000	1

Oneway residuals afd_3, bonferroni tabulate

Summary of Residuals

afd_3	Mean	Std. Dev.	Freq.
1	.93275425	15.458.145	1097444
2	24.992.379	14.593.813	1218790
3	-37.760.473	22.042.649	712899
Total	.45483358	17.105.354	3029133

Analysis of Variance

Source	SS	df	MS	F	Prob > F
Between groups	18105850	2	9052925.01	31585.55	0.0000
Within groups		8,68197E+15	286.616.071		
Total	8,86303E+15	292.593.122			

Bartlett's test for equal variances: chi2(2) = 1.9e+05 Prob>chi2 = 0.000

Comparison of Residuals by afd_3 (Bonferroni)

Row	Mean-	Col
Mean	1	2
2	156.648	
	0.000	
3	-47.088	-627.529
	0.000	0.000

Oneway residuals ird_3, Bonferroni tabulate

Summary of Residuals

ird_3	Mean	Std. Dev.	Freq.
1	-12.339.211	19.406.702	850071
2	.95429307	16.163.894	1271049
3	.6609084	17.666.024	962477
Total	.25948326	17.601.009	3083597

Analysis of Variance

Source	SS	df	MS	F	Prob>F
Between groups	2664584.65	2	1332292.32	4312.58	0.0000
Within groups	9,5262E+15	308.931.594			
Total	9,55284E+15	309.795.509			

Bartlett's test for equal variances: $\chi^2(2) = 3.5e+04$ Prob> $\chi^2 = 0.000$

Comparison of Residuals by ird_3 (Bonferroni)

Row	Mean-	Col
Mean	1	2
2	218.821	
	0.000	
3	189.483	-.293385
	0.000	0.000

oneway residuals icg_3, bonferroni tabulate

Summary of Residuals

NCG2013	Mean	Std. Dev.	Freq.
0	-81.027.321	30.676.295	43272
1	-.75720356	20.219.288	971990
2	.81753751	16.951.398	1297332
3	.02938827	16.414.179	884643
Total	-1,81E-06	18.152.112	3197237

Analysis of Variance

Source	SS	df	MS	F	Prob > F
Between groups	4266147.67	3	1422049.22	4333.33	0.0000
Within groups	1.0492e+093197233	328.165.162			
Total	1.0535e+093197236	329.499.178			

Bartlett's test for equal variances: $\chi^2(3) = 8.8e+04$ Prob> $\chi^2 = 0.000$

Comparison of Residuals by NCG2013 (Bonferroni)

Row	Mean	Col 0	Col 1	Col 2
1		734.553		
		0.000		
2		892.027	157.474	
		0.000	0.000	
3		813.212	.786592	-.788149
		0.000	0.000	0.000

Tabela 34 - Regressão linear e correlações referentes ao modelo básico - Ensino fundamental I I

regress traj_a meninos_a pretos_a norte nordeste sul centroeste est_pem mun_pem baixo_nse medio_nse SI_nse, beta					
Source	SS	df	MS	Number of obs = 2841938	
Model	268345605	11	24395055	F(11,2841926) =99514.98	
Residual	6966683622841926		245.139515	Prob > F = 0.0000	
Total	9650139672841937	339.562055		R-squared = 0.2781	
				Adj R-squared = 0.2781	
				Root MSE = 15.657	
traj_a	Coef.	Std. Err.	t	P> t	Beta
meninos_a	-.1452665	.0010565	-137.50	0.000	-.069604
pretos_a	-.4139733	.001278	-323.93	0.000	-.1659026
norte	-7.046.634	.0366445	-192.30	0.000	-.1148975
nordeste	-8.341.882	.0309804	-269.26	0.000	-.2059383
sul	-6.951.379	.0300248	-231.52	0.000	-.1275573
centroeste	-3.086.213	.0366931	-84.11	0.000	-.0447388
est_pem	-7.819.858	.0354108	-220.83	0.000	-.2113196
mun_pem	-1.274.157	.0368065	-346.18	0.000	-.3373896
baixo_nse	-1.052.839	.0444794	-236.70	0.000	-.2658905
medio_nse	-640.692	.032399	-197.75	0.000	-.1738249
SI_nse	-1.905.687	.4103356	-46.44	0.000	-.0234822
_cons	9.595.497	.0597116	1606.97	0.000	.
predict y_pred					
(option xb assumed; fitted values)					
predict residuals, residuals					
Pwcorr residuals afd_2013 ncg2013ird_2013 sig					
residu~s	afd_2013	ncg2013	ird_2013		
residuals	1				
afd_	-0.0114	1		-0.0114	
	0.0000			0.0000	
ncg	-0.1047	0.0566	1	-0.1047	
	0.0000	0.0000		0.0000	
ird	-0.0460	-0.0035	0.1313	1	-0.0460

	0.0000	0.0000	0.0000	0.0000		
Oneway residuals afd_3, bonferroni tabulate						
Summary of Residuals						
afd_3	Mean	Std. Dev.		Freq.		
1	.06206206	14.518.876		993314		
2	11.224.904	19.841.663		520171		
3	-.01636011	13.202.409		1295470		
Total	.22226772	15.100.363		2808955		
Analysis of Variance						
Source	SS	df	MS	F	Prob > F	
Between groups	520.809.558	2	260.404.779	1142.95	0.0000	520.809.558
Within groups	6,3998E+15	227.835.726				
Total	6,405E+15	228.020.974				

Bartlett's test for equal variances: $\chi^2(2) = 1.9e+05$ Prob> $\chi^2 = 0.000$

Comparison of Residuals by afd_3 (Bonferroni)

Row Mean-|Col

Mean

2	106.043		106.043
	0.000		0.000
3			
	-.078422	-113.885	-.078422

Oneway residuals ird_3, Bonferroni tabulate

Summary of Residuals

ird_3	Mean	Std. Dev.	Freq.
1	.89745085	16.256.779	747813
2	-12.630.708	15.489.138	827535
3	.31599839	15.095.394	1179519
Total	-.0005023	15.559.977	2754867

Analysis of Variance

Source	SS	df	MS	F	Prob>F
Between groups	2040288.4	2	1020144.2	4226.43	0.0000
Within groups	6,64948E+15	241.372.436			
Total	6669885222754866	242.112873			

Bartlett's test for equal variances: $\chi^2(2) = 3.5e+04$ Prob> $\chi^2 = 0.000$

Comparison of Residuals by ird_3 (Bonferroni)

Row Mean- Col

Mean

2	-2.16052		
	0.000		
3			
	-.581452	1.57907	
	0.000	0.000	

oneway residuals icg_3, bonferroni tabulate					
Summary of Residuals					
NCG	Mean	Std. Dev.		Freq.	
0	19.590.297	19.241.057		13959	
1	-29.874.198	27.722.486		73977	
2	16.408.211	15.224.185		1635342	
3	-2.225.564	14.797.782		1118660	
Total	-9,37E-07	15.656.902		2841938	
Analysis of Variance					
Source	SS	df	MS	F	Prob > F
Between groups	10657488.2	3	3552496.06	14716.91	
Within groups	6,86011E+15	241.388.742			
Total	6,96668E+15	245.138.566			
Bartlett's test for equal variances: $\chi^2(3) = 8.8e+04$ Prob> $\chi^2 = 0.000$					
Comparison of Residuals by NCG2013 (Bonferroni)					
Row	Mean	Col			
	0	1	2		
1	-494.645				
	0.000				
2					
	-318209	462.824			
3	0.096	0.000			

Tabela 35 - Regressão linear e correlações referentes ao modelo básico - Ensino Médio

regress traj_a meninos_a pretos_a norte nordeste sul centroeste est_pem mun_pem baixo_nse medio_nse SI_nse, beta					
Source	SS	df	MS	Number of obs = 1885951	
Model	48990770.2	11	4453706.38	F(11,1885939) = 12533.38	
Residual	6701640831885939		355.347699	Prob > F = 0.0000	
Total	7191548541885950	381.322333		R-squared = 0.0681	
				Adj R-squared = 0.0681	
				Root MSE = 18.851	
traj_a	Coef.	Std. Err.	t	P> t	Beta
meninos_a	-.0840137	.0017649	-47.60	0.000	-.0344058
pretos_a	-.3995029	.0024326	-164.23	0.000	-.1179735
norte	4.165.249	.0561087	74.24	0.000	.0594224
nordeste	-7.457.528	.0451222	-165.27	0.000	-.1664785
sul	4.885.243	.0434759	112.37	0.000	.0849536
centroeste	5.981.428	.0524849	113.96	0.000	.0843967
est_pem	.810951	.0481212	16.85	0.000	.0162278
mun_pem	1.059.446	.1977079	53.59	0.000	.0384271

baixo_nse	2.120.605	.0640232	33.12	0.000	.0492341
medio_nse	-2.266.554	.048141	-47.08	0.000	-.0579531
SI_nse	2.057.029	.4135546	49.74	0.000	.0350646
_cons	5.385.682	.0935042	575.98	0.000	.

predict y_pred

(option xb assumed; fitted values)

predict residuals, residuals

Pwcorr residuals afd_2013 neg2013ird_2013 sig

residu~s	afd_2013	icg	ird_2013	
residuals	1			
afd	0.1492 0.0000	1		
icg	-0.1036 0.0000	-0.0002 0.8171	1	
ird	0.0243 0.0000	-0.0700 0.0000	0.0236 0.0000	1

Oneway residuals afd_3, bonferroni tabulate

Summary of Residuals

afd_3	Mean	Std. Dev.	Freq.
1	13.199.008	1.680.322	740923
2	1.297.114	17.474.315	830333
3	-65.254.208	24.357.781	309099
Total	13.199.008	1.680.322	740923

Analysis of Variance

Source	SS	df	MS	F	Prob > F
Between groups	15848842.8	2	7924421.39	23061.48	0.0000
Within groups	6,4613E+15	343.621.607			
Total					

Bartlett's test for equal variances: $\chi^2(2) = 1.9e+05$ Prob> $\chi^2 = 0.000$

Comparison of Residuals by afd_3 (Bonferroni)

Row Mean-|Col

Mean

2	-.022787 1.000	
3	-784.532 0.000	-782.253 0.000

Oneway residuals ird_3, Bonferroni tabulate

Summary of Residuals

ird_3	Mean	Std. Dev.	Freq.
1	10.130.994	20.082.024	511593
2	-.26138455	16.979.316	532388

3	-.84030003	18.669.871	784054
Total	-.15300803	1.862.561	1828035

Analysis of Variance

Source	SS	df	MS	F	Prob>F
Between groups	1072284.46	2	536.142.232	1548.08	0.0000
Within groups	6,33097E+15	346.327.159			

Total

Bartlett's test for equal variances: $\chi^2(2) = 3.5e+04$ Prob> $\chi^2 = 0.000$

Comparison of Residuals by ird_3 (Bonferroni)

Row Mean- Col

Mean

2	-127.448	
	0.000	
3		
	-18.534	-.578915
	0.000	0.000

oneway residuals icg_3, bonferroni tabulate

Summary of Residuals

ICG	Mean	Std. Dev.	Freq.
0	83.820.669	33.249.564	3948
1	59.501.506	1.852.844	122620
2	.47076307	18.492.062	1070006
3	-18.370.488	19.067.182	689377
Total	-6,03E-08	18.850.613	1885951

Analysis of Variance

Source	SS	df	MS	F	Prob > F
Between groups	7182263.35	3	2394087.78	6810.33	0.0000
Within groups	6,62982E+15	351.537.885			
Total	6,70164E+15	355.345.626			

Bartlett's test for equal variances: $\chi^2(3) = 8.8e+04$ Prob> $\chi^2 = 0.000$

Comparison of Residuals by ICG (Bonferroni)

Row Mean- Col

Mean	0	1	2
1	-243.192		
	0.000		
2			
	-79.113	-547.939	
3	0.000	0.000	

Tabela 36 - Regressão linear e correlações referentes ao modelo para características da escola - Ensino fundamental I

Source	SS	df	MS	Number of obs = 3197237	
Model	643659057	20	32182952.8	F(20,3197216) = .	
Residual	1.0115e+093197216		316.364391	Prob > F = 0.0000	
Total	1.6551e+093197236	517.679755		R-squared = 0.3889	
				Adj R-squared = 0.3948	
				Root MSE = 17.787	
traj_a	Coef.	Std. Err.	t	P> t	Beta
meninos_a	-.1361566	.0009418	-144.57	0.000	-.0632391
pretos_a	-.3591136	.0011413	-314.65	0.000	-.1404498
norte	-165.511	.03848	-430.12	0.000	-.2290676
nordeste	-117.057	.0329957	-354.76	0.000	-.2383198
sul	-6.375.562	.0330544	-192.88	0.000	-.0928707
centroeste	-7.360.176	.0405567	-181.48	0.000	-.0847725
est_pem	-3.112.836	.0408332	-76.23	0.000	-.0522195
mun_pem	-1.000.194	.0363472	-275.18	0.000	-.2084064
baixo_nse	-1.394.227	.0459965	-303.12	0.000	-.2855644
medio_nse	-6.511.691	.0315671	-206.28	0.000	-.1429915
SI_nse	-2.817.452	.1672317	-168.48	0.000	-.0760058
Baixa					
adequação	-7.926.561	.0293871	-269.73	0.000	-.1450103
Média					
Adequação	-2.092.768	.0238066	-87.91	0.000	-.0436712
SI adequação	-1.220.584	.0550343	-221.79	0.000	-.119732
Baixa					
Regularidade	-1.112.495	.0452536	-24.58	0.000	-.0216019
Média					
Regularidade	.2626116	.0244271	10.75	0.000	.0095857
SI Regularidade	-6.278.689	.0781632	-80.33	0.000	-.0510926
Baixa					
complexidade	.2844955	.0266598	10.67	0.000	.0057516
Média					
complexidade	1.402.946	.0260077	53.94	0.000	.030278
SI					
complexidade	6.515.082	.1211862	53.76	0.000	.0330862
_cons	1.012.506	.0715836	1414.44	0.000	

Tabela 37 - Regressão linear e correlações referentes ao modelo para características da escola - Ensino fundamental I I					
Source	SS	df	MS	Number of obs = 2841938	
Model	303983555	20	15199177.7	F(22,2841915) =64594.75	
Residual	6,6103E+15		232.600.183	Prob > F = 0.0000	
Total	9650139672841937	339.562055		R-squared = 0.315	
				Adj R-squared = 0.315	
				Root MSE = 15.251	
traj_a	Coef.	Std. Err.	t	P> t	Beta
meninos_a	-.1393724	.0010328	-134.94	0.000	-.0667799
pretos_a	-.3971325	.001251	-317.45	0.000	-.1591535
norte	-6.600.568	.0361732	-182.47	0.000	-.1076243
nordeste	-7.793.933	.0315883	-246.73	0.000	-.1924109
sul	-7.391.727	.0294575	-250.93	0.000	-.1356376
centroeste	-315.777	.0369231	-85.52	0.000	-.0457761
est_pem	-6.861.572	.0364616	-188.19	0.000	-.1854234
mun_pem	-1.210.463	.0368675	-328.33	0.000	-.3205238
baixo_nse	-1.000.225	.0447281	-223.62	0.000	-.252603
medio_nse	-6.224.178	.0317471	-196.06	0.000	-.168867
SI_nse	-1.811.233	.4009899	-45.17	0.000	-.0223184
Baixa adequação	.7799843	.0317156	24.59	0.000	.0163679
Média					
Adequação	.1667828	.0220506	7.56	0.000	.0043156
SI adequação	-3.464.679	.1140131	-303.88	0.000	-.2013754
Baixa					
Regularidade	155.311	.0260825	59.55	0.000	.037113
Média					
Regularidade	1.468.743	.0227552	64.55	0.000	.039273
SI Regularidade	.0644971	.0592998	1.09	0.277	.0006032
Baixa					
complexidade	2.421.078	.0595456	40.66	0.000	.0209201
Média					
complexidade	4.244.387	.0200974	211.19	0.000	.1138482
SI Complexidade	4.011.855	.1812137	221.39	0.000	.1522076
Constante					
	9.098.419	.0616864	1474.95	0.000	.

Tabela 38 - Regressão linear e correlações referentes ao modelo para características da escola - Ensino Médio

Source	SS	df	MS	Number of obs = 1885951	
Model	82306585	20	4115329.25	F(20,1885930) =12186.93	
Residual	6,36848E+15	337.683.938	6,36848E+15	Prob > F = 0.0000	
Total	7,19155E+15	381.322.333		R-squared = 0.1144	

Adj R-squared = 0.1144

Root MSE = 18.376

traj_a	Coef.	Std. Err.	t	P> t	Beta
meninos_a	-.0767218	.0017282	-44.39	0.000	-.0314196
pretos_a	-.3291235	.0024011	-137.07	0.000	-.0971904
norte	3.879.911	.0551844	70.31	0.000	.0553517
nordeste	-5.856.509	.0460053	-127.30	0.000	-.1307381
sul	5.042.467	.042699	118.09	0.000	.0876877
centroeste	1.019.007	.0538856	189.11	0.000	.1437798
est_pem	2.130.241	.0482588	44.14	0.000	.0426278
mun_pem	1.172.207	.1930524	60.72	0.000	.042517
baixo_nse	4.892	.0639008	76.56	0.000	.1135776
medio_nse	-1.182.527	.0476716	-24.81	0.000	-.0302358
SI_nse	1.109.796	.5121544	21.67	0.000	.0189178
Baixa adequação	-1.026.877	.0445749	-230.37	0.000	-.1946644
Média					
Adequação	-.4031819	.0299329	-13.47	0.000	-.0102495
SI adequação	-4.645.596	.4537765	-102.38	0.000	-.1293969
Baixa					
Regularidade	2.313.988	.0555034	41.69	0.000	.0533382
Média					
Regularidade	-2.374.608	.0344757	-68.88	0.000	-.0929195
SI Regularidade	2.167.644	.0844755	25.66	0.000	.0191515
Baixa					
complexidade	7.679.722	.0586351	130.97	0.000	.0969654
Média					
complexidade	2.562.127	.0299951	85.42	0.000	.0650052
SI complexidade	5.651.056	.5901265	95.76	0.000	.1322671
_cons	5.162.029	.0958986	538.28	0.000	.

ANEXO A

agrupando etapas

```
brasil2007_v3 <- brasil2007_1x3r %>%
```

```
  mutate( etapaEnsino = case_when( FK_COD_ETAPA_ENSINO == 1 ~ 1,
```

```
    FK_COD_ETAPA_ENSINO == 2 ~ 2,
```

FK_COD_ETAPA_ENSINO == 14 ~ 10,
FK_COD_ETAPA_ENSINO == 4 ~ 20,
FK_COD_ETAPA_ENSINO == 15 ~ 20,
FK_COD_ETAPA_ENSINO == 5 ~ 30,
FK_COD_ETAPA_ENSINO == 16 ~ 30,
FK_COD_ETAPA_ENSINO == 6 ~ 40,
FK_COD_ETAPA_ENSINO == 17 ~ 40,
FK_COD_ETAPA_ENSINO == 7 ~ 50,
FK_COD_ETAPA_ENSINO == 18 ~ 50,
FK_COD_ETAPA_ENSINO == 8 ~ 60,
FK_COD_ETAPA_ENSINO == 19 ~ 60,
FK_COD_ETAPA_ENSINO == 9 ~ 70,
FK_COD_ETAPA_ENSINO == 20 ~ 70,
FK_COD_ETAPA_ENSINO == 10 ~ 80,
FK_COD_ETAPA_ENSINO == 21 ~ 80,
FK_COD_ETAPA_ENSINO == 11 ~ 90,
FK_COD_ETAPA_ENSINO == 41 ~ 90,
FK_COD_ETAPA_ENSINO == 25 ~ 100,
FK_COD_ETAPA_ENSINO == 26 ~ 200,
FK_COD_ETAPA_ENSINO == 27 ~ 300,
FK_COD_ETAPA_ENSINO == 28 ~ 300,
FK_COD_ETAPA_ENSINO == 29 ~ 1000,
FK_COD_ETAPA_ENSINO == 30 ~ 100,
FK_COD_ETAPA_ENSINO == 31 ~ 200,
FK_COD_ETAPA_ENSINO == 32 ~ 300,
FK_COD_ETAPA_ENSINO == 33 ~ 300,
FK_COD_ETAPA_ENSINO == 34 ~ 1000,
FK_COD_ETAPA_ENSINO == 35 ~ 100,
FK_COD_ETAPA_ENSINO == 36 ~ 200,
FK_COD_ETAPA_ENSINO == 37 ~ 300,
FK_COD_ETAPA_ENSINO == 38 ~ 300,
FK_COD_ETAPA_ENSINO == 39 ~ 9999,
FK_COD_ETAPA_ENSINO == 40 ~ 9999,
FK_COD_ETAPA_ENSINO == 43 ~ 1000,
FK_COD_ETAPA_ENSINO == 44 ~ 1000,
FK_COD_ETAPA_ENSINO == 45 ~ 1000,
FK_COD_ETAPA_ENSINO == 46 ~ 1000,
FK_COD_ETAPA_ENSINO == 47 ~ 1000,
FK_COD_ETAPA_ENSINO == 48 ~ 1000,
FK_COD_ETAPA_ENSINO == 49 ~ 1000,

```

FK_COD_ETAPA_ENSINO == 50 ~ 1000,
FK_COD_ETAPA_ENSINO == 51 ~ 1000,
FK_COD_ETAPA_ENSINO == 52 ~ 1000,
FK_COD_ETAPA_ENSINO == 53 ~ 1000,
FK_COD_ETAPA_ENSINO == 54 ~ 1000,
FK_COD_ETAPA_ENSINO == 55 ~ 1000,
FK_COD_ETAPA_ENSINO == 57 ~ 1000,
FK_COD_ETAPA_ENSINO == 59 ~ 1000,
FK_COD_ETAPA_ENSINO == 60 ~ 1000,
FK_COD_ETAPA_ENSINO == 61 ~ 1000,
FK_COD_ETAPA_ENSINO == 62 ~ 1000,
FK_COD_ETAPA_ENSINO == 65 ~ 1000,
FK_COD_ETAPA_ENSINO == 67 ~ 1000,
FK_COD_ETAPA_ENSINO == 68 ~ 1000,
FK_COD_ETAPA_ENSINO == 69 ~ 1000,
FK_COD_ETAPA_ENSINO == 70 ~ 1000,
FK_COD_ETAPA_ENSINO == 71 ~ 1000,
FK_COD_ETAPA_ENSINO == 73 ~ 1000,
FK_COD_ETAPA_ENSINO == 74 ~ 1000,
TRUE ~ NaN ))

```

```

brasil2008_v3 <- brasil2008_1x3r %>%

```

```

  mutate( etapaEnsino = case_when( FK_COD_ETAPA_ENSINO == 1 ~ 1,
    FK_COD_ETAPA_ENSINO == 2 ~ 2,
    FK_COD_ETAPA_ENSINO == 14 ~ 10,
    FK_COD_ETAPA_ENSINO == 4 ~ 20,
    FK_COD_ETAPA_ENSINO == 15 ~ 20,
    FK_COD_ETAPA_ENSINO == 5 ~ 30,
    FK_COD_ETAPA_ENSINO == 16 ~ 30,
    FK_COD_ETAPA_ENSINO == 6 ~ 40,
    FK_COD_ETAPA_ENSINO == 17 ~ 40,
    FK_COD_ETAPA_ENSINO == 7 ~ 50,
    FK_COD_ETAPA_ENSINO == 18 ~ 50,
    FK_COD_ETAPA_ENSINO == 8 ~ 60,
    FK_COD_ETAPA_ENSINO == 19 ~ 60,
    FK_COD_ETAPA_ENSINO == 9 ~ 70,
    FK_COD_ETAPA_ENSINO == 20 ~ 70,
    FK_COD_ETAPA_ENSINO == 10 ~ 80,
    FK_COD_ETAPA_ENSINO == 21 ~ 80,
    FK_COD_ETAPA_ENSINO == 11 ~ 90,

```

FK_COD_ETAPA_ENSINO == 41 ~ 90,
FK_COD_ETAPA_ENSINO == 25 ~ 100,
FK_COD_ETAPA_ENSINO == 26 ~ 200,
FK_COD_ETAPA_ENSINO == 27 ~ 300,
FK_COD_ETAPA_ENSINO == 28 ~ 300,
FK_COD_ETAPA_ENSINO == 29 ~ 1000,
FK_COD_ETAPA_ENSINO == 30 ~ 100,
FK_COD_ETAPA_ENSINO == 31 ~ 200,
FK_COD_ETAPA_ENSINO == 32 ~ 300,
FK_COD_ETAPA_ENSINO == 33 ~ 300,
FK_COD_ETAPA_ENSINO == 34 ~ 1000,
FK_COD_ETAPA_ENSINO == 35 ~ 100,
FK_COD_ETAPA_ENSINO == 36 ~ 200,
FK_COD_ETAPA_ENSINO == 37 ~ 300,
FK_COD_ETAPA_ENSINO == 38 ~ 300,
FK_COD_ETAPA_ENSINO == 39 ~ 9999,
FK_COD_ETAPA_ENSINO == 40 ~ 9999,
FK_COD_ETAPA_ENSINO == 43 ~ 1000,
FK_COD_ETAPA_ENSINO == 44 ~ 1000,
FK_COD_ETAPA_ENSINO == 45 ~ 1000,
FK_COD_ETAPA_ENSINO == 46 ~ 1000,
FK_COD_ETAPA_ENSINO == 47 ~ 1000,
FK_COD_ETAPA_ENSINO == 48 ~ 1000,
FK_COD_ETAPA_ENSINO == 49 ~ 1000,
FK_COD_ETAPA_ENSINO == 50 ~ 1000,
FK_COD_ETAPA_ENSINO == 51 ~ 1000,
FK_COD_ETAPA_ENSINO == 52 ~ 1000,
FK_COD_ETAPA_ENSINO == 53 ~ 1000,
FK_COD_ETAPA_ENSINO == 54 ~ 1000,
FK_COD_ETAPA_ENSINO == 55 ~ 1000,
FK_COD_ETAPA_ENSINO == 57 ~ 1000,
FK_COD_ETAPA_ENSINO == 59 ~ 1000,
FK_COD_ETAPA_ENSINO == 60 ~ 1000,
FK_COD_ETAPA_ENSINO == 61 ~ 1000,
FK_COD_ETAPA_ENSINO == 62 ~ 1000,
FK_COD_ETAPA_ENSINO == 65 ~ 1000,
FK_COD_ETAPA_ENSINO == 67 ~ 1000,
FK_COD_ETAPA_ENSINO == 68 ~ 1000,
FK_COD_ETAPA_ENSINO == 69 ~ 1000,
FK_COD_ETAPA_ENSINO == 70 ~ 1000,

```

FK_COD_ETAPA_ENSINO == 71 ~ 1000,
FK_COD_ETAPA_ENSINO == 73 ~ 1000,
FK_COD_ETAPA_ENSINO == 74 ~ 1000,
TRUE ~ NaN ))

```

```

brasil2009_v3 <- brasil2009_1x3r %>%

```

```

  mutate( etapaEnsino = case_when( FK_COD_ETAPA_ENSINO == 1 ~ 1,
    FK_COD_ETAPA_ENSINO == 2 ~ 2,
    FK_COD_ETAPA_ENSINO == 14 ~ 10,
    FK_COD_ETAPA_ENSINO == 4 ~ 20,
    FK_COD_ETAPA_ENSINO == 15 ~ 20,
    FK_COD_ETAPA_ENSINO == 5 ~ 30,
    FK_COD_ETAPA_ENSINO == 16 ~ 30,
    FK_COD_ETAPA_ENSINO == 6 ~ 40,
    FK_COD_ETAPA_ENSINO == 17 ~ 40,
    FK_COD_ETAPA_ENSINO == 7 ~ 50,
    FK_COD_ETAPA_ENSINO == 18 ~ 50,
    FK_COD_ETAPA_ENSINO == 8 ~ 60,
    FK_COD_ETAPA_ENSINO == 19 ~ 60,
    FK_COD_ETAPA_ENSINO == 9 ~ 70,
    FK_COD_ETAPA_ENSINO == 20 ~ 70,
    FK_COD_ETAPA_ENSINO == 10 ~ 80,
    FK_COD_ETAPA_ENSINO == 21 ~ 80,
    FK_COD_ETAPA_ENSINO == 11 ~ 90,
    FK_COD_ETAPA_ENSINO == 41 ~ 90,
    FK_COD_ETAPA_ENSINO == 25 ~ 100,
    FK_COD_ETAPA_ENSINO == 26 ~ 200,
    FK_COD_ETAPA_ENSINO == 27 ~ 300,
    FK_COD_ETAPA_ENSINO == 28 ~ 300,
    FK_COD_ETAPA_ENSINO == 29 ~ 1000,
    FK_COD_ETAPA_ENSINO == 30 ~ 100,
    FK_COD_ETAPA_ENSINO == 31 ~ 200,
    FK_COD_ETAPA_ENSINO == 32 ~ 300,
    FK_COD_ETAPA_ENSINO == 33 ~ 300,
    FK_COD_ETAPA_ENSINO == 34 ~ 1000,
    FK_COD_ETAPA_ENSINO == 35 ~ 100,
    FK_COD_ETAPA_ENSINO == 36 ~ 200,
    FK_COD_ETAPA_ENSINO == 37 ~ 300,
    FK_COD_ETAPA_ENSINO == 38 ~ 300,
    FK_COD_ETAPA_ENSINO == 39 ~ 9999,

```

```

FK_COD_ETAPA_ENSINO == 40 ~ 9999,
FK_COD_ETAPA_ENSINO == 43 ~ 1000,
FK_COD_ETAPA_ENSINO == 44 ~ 1000,
FK_COD_ETAPA_ENSINO == 45 ~ 1000,
FK_COD_ETAPA_ENSINO == 46 ~ 1000,
FK_COD_ETAPA_ENSINO == 47 ~ 1000,
FK_COD_ETAPA_ENSINO == 48 ~ 1000,
FK_COD_ETAPA_ENSINO == 49 ~ 1000,
FK_COD_ETAPA_ENSINO == 50 ~ 1000,
FK_COD_ETAPA_ENSINO == 51 ~ 1000,
FK_COD_ETAPA_ENSINO == 52 ~ 1000,
FK_COD_ETAPA_ENSINO == 53 ~ 1000,
FK_COD_ETAPA_ENSINO == 54 ~ 1000,
FK_COD_ETAPA_ENSINO == 55 ~ 1000,
FK_COD_ETAPA_ENSINO == 57 ~ 1000,
FK_COD_ETAPA_ENSINO == 59 ~ 1000,
FK_COD_ETAPA_ENSINO == 60 ~ 1000,
FK_COD_ETAPA_ENSINO == 61 ~ 1000,
FK_COD_ETAPA_ENSINO == 62 ~ 1000,
FK_COD_ETAPA_ENSINO == 65 ~ 1000,
FK_COD_ETAPA_ENSINO == 67 ~ 1000,
FK_COD_ETAPA_ENSINO == 68 ~ 1000,
FK_COD_ETAPA_ENSINO == 69 ~ 1000,
FK_COD_ETAPA_ENSINO == 70 ~ 1000,
FK_COD_ETAPA_ENSINO == 71 ~ 1000,
FK_COD_ETAPA_ENSINO == 73 ~ 1000,
FK_COD_ETAPA_ENSINO == 74 ~ 1000,
TRUE ~ NaN ))

```

```

brasil2010_v3 <- brasil2010_1x3r %>%

```

```

  mutate( etapaEnsino = case_when( FK_COD_ETAPA_ENSINO == 1 ~ 1,
    FK_COD_ETAPA_ENSINO == 2 ~ 2,
    FK_COD_ETAPA_ENSINO == 14 ~ 10,
    FK_COD_ETAPA_ENSINO == 4 ~ 20,
    FK_COD_ETAPA_ENSINO == 15 ~ 20,
    FK_COD_ETAPA_ENSINO == 5 ~ 30,
    FK_COD_ETAPA_ENSINO == 16 ~ 30,
    FK_COD_ETAPA_ENSINO == 6 ~ 40,
    FK_COD_ETAPA_ENSINO == 17 ~ 40,
    FK_COD_ETAPA_ENSINO == 7 ~ 50,

```

FK_COD_ETAPA_ENSINO == 18 ~ 50,
FK_COD_ETAPA_ENSINO == 8 ~ 60,
FK_COD_ETAPA_ENSINO == 19 ~ 60,
FK_COD_ETAPA_ENSINO == 9 ~ 70,
FK_COD_ETAPA_ENSINO == 20 ~ 70,
FK_COD_ETAPA_ENSINO == 10 ~ 80,
FK_COD_ETAPA_ENSINO == 21 ~ 80,
FK_COD_ETAPA_ENSINO == 11 ~ 90,
FK_COD_ETAPA_ENSINO == 41 ~ 90,
FK_COD_ETAPA_ENSINO == 25 ~ 100,
FK_COD_ETAPA_ENSINO == 26 ~ 200,
FK_COD_ETAPA_ENSINO == 27 ~ 300,
FK_COD_ETAPA_ENSINO == 28 ~ 300,
FK_COD_ETAPA_ENSINO == 29 ~ 1000,
FK_COD_ETAPA_ENSINO == 30 ~ 100,
FK_COD_ETAPA_ENSINO == 31 ~ 200,
FK_COD_ETAPA_ENSINO == 32 ~ 300,
FK_COD_ETAPA_ENSINO == 33 ~ 300,
FK_COD_ETAPA_ENSINO == 34 ~ 1000,
FK_COD_ETAPA_ENSINO == 35 ~ 100,
FK_COD_ETAPA_ENSINO == 36 ~ 200,
FK_COD_ETAPA_ENSINO == 37 ~ 300,
FK_COD_ETAPA_ENSINO == 38 ~ 300,
FK_COD_ETAPA_ENSINO == 39 ~ 9999,
FK_COD_ETAPA_ENSINO == 40 ~ 9999,
FK_COD_ETAPA_ENSINO == 43 ~ 1000,
FK_COD_ETAPA_ENSINO == 44 ~ 1000,
FK_COD_ETAPA_ENSINO == 45 ~ 1000,
FK_COD_ETAPA_ENSINO == 46 ~ 1000,
FK_COD_ETAPA_ENSINO == 47 ~ 1000,
FK_COD_ETAPA_ENSINO == 48 ~ 1000,
FK_COD_ETAPA_ENSINO == 49 ~ 1000,
FK_COD_ETAPA_ENSINO == 50 ~ 1000,
FK_COD_ETAPA_ENSINO == 51 ~ 1000,
FK_COD_ETAPA_ENSINO == 52 ~ 1000,
FK_COD_ETAPA_ENSINO == 53 ~ 1000,
FK_COD_ETAPA_ENSINO == 54 ~ 1000,
FK_COD_ETAPA_ENSINO == 55 ~ 1000,
FK_COD_ETAPA_ENSINO == 57 ~ 1000,
FK_COD_ETAPA_ENSINO == 59 ~ 1000,

```
FK_COD_ETAPA_ENSINO == 60 ~ 1000,  
FK_COD_ETAPA_ENSINO == 61 ~ 1000,  
FK_COD_ETAPA_ENSINO == 62 ~ 1000,  
FK_COD_ETAPA_ENSINO == 65 ~ 1000,  
FK_COD_ETAPA_ENSINO == 67 ~ 1000,  
FK_COD_ETAPA_ENSINO == 68 ~ 1000,  
FK_COD_ETAPA_ENSINO == 69 ~ 1000,  
FK_COD_ETAPA_ENSINO == 70 ~ 1000,  
FK_COD_ETAPA_ENSINO == 71 ~ 1000,  
FK_COD_ETAPA_ENSINO == 73 ~ 1000,  
FK_COD_ETAPA_ENSINO == 74 ~ 1000,  
TRUE ~ NaN ))
```

```
brasil2011_v3 <- brasil2011_1x3r %>%
```

```
  mutate( etapaEnsino = case_when( FK_COD_ETAPA_ENSINO == 1 ~ 1,  
    FK_COD_ETAPA_ENSINO == 2 ~ 2,  
    FK_COD_ETAPA_ENSINO == 14 ~ 10,  
    FK_COD_ETAPA_ENSINO == 4 ~ 20,  
    FK_COD_ETAPA_ENSINO == 15 ~ 20,  
    FK_COD_ETAPA_ENSINO == 5 ~ 30,  
    FK_COD_ETAPA_ENSINO == 16 ~ 30,  
    FK_COD_ETAPA_ENSINO == 6 ~ 40,  
    FK_COD_ETAPA_ENSINO == 17 ~ 40,  
    FK_COD_ETAPA_ENSINO == 7 ~ 50,  
    FK_COD_ETAPA_ENSINO == 18 ~ 50,  
    FK_COD_ETAPA_ENSINO == 8 ~ 60,  
    FK_COD_ETAPA_ENSINO == 19 ~ 60,  
    FK_COD_ETAPA_ENSINO == 9 ~ 70,  
    FK_COD_ETAPA_ENSINO == 20 ~ 70,  
    FK_COD_ETAPA_ENSINO == 10 ~ 80,  
    FK_COD_ETAPA_ENSINO == 21 ~ 80,  
    FK_COD_ETAPA_ENSINO == 11 ~ 90,  
    FK_COD_ETAPA_ENSINO == 41 ~ 90,  
    FK_COD_ETAPA_ENSINO == 25 ~ 100,  
    FK_COD_ETAPA_ENSINO == 26 ~ 200,  
    FK_COD_ETAPA_ENSINO == 27 ~ 300,  
    FK_COD_ETAPA_ENSINO == 28 ~ 300,  
    FK_COD_ETAPA_ENSINO == 29 ~ 1000,  
    FK_COD_ETAPA_ENSINO == 30 ~ 100,  
    FK_COD_ETAPA_ENSINO == 31 ~ 200,
```


FK_COD_ETAPA_ENSINO == 14 ~ 10,
FK_COD_ETAPA_ENSINO == 4 ~ 20,
FK_COD_ETAPA_ENSINO == 15 ~ 20,
FK_COD_ETAPA_ENSINO == 5 ~ 30,
FK_COD_ETAPA_ENSINO == 16 ~ 30,
FK_COD_ETAPA_ENSINO == 6 ~ 40,
FK_COD_ETAPA_ENSINO == 17 ~ 40,
FK_COD_ETAPA_ENSINO == 7 ~ 50,
FK_COD_ETAPA_ENSINO == 18 ~ 50,
FK_COD_ETAPA_ENSINO == 8 ~ 60,
FK_COD_ETAPA_ENSINO == 19 ~ 60,
FK_COD_ETAPA_ENSINO == 9 ~ 70,
FK_COD_ETAPA_ENSINO == 20 ~ 70,
FK_COD_ETAPA_ENSINO == 10 ~ 80,
FK_COD_ETAPA_ENSINO == 21 ~ 80,
FK_COD_ETAPA_ENSINO == 11 ~ 90,
FK_COD_ETAPA_ENSINO == 41 ~ 90,
FK_COD_ETAPA_ENSINO == 25 ~ 100,
FK_COD_ETAPA_ENSINO == 26 ~ 200,
FK_COD_ETAPA_ENSINO == 27 ~ 300,
FK_COD_ETAPA_ENSINO == 28 ~ 300,
FK_COD_ETAPA_ENSINO == 29 ~ 1000,
FK_COD_ETAPA_ENSINO == 30 ~ 100,
FK_COD_ETAPA_ENSINO == 31 ~ 200,
FK_COD_ETAPA_ENSINO == 32 ~ 300,
FK_COD_ETAPA_ENSINO == 33 ~ 300,
FK_COD_ETAPA_ENSINO == 34 ~ 1000,
FK_COD_ETAPA_ENSINO == 35 ~ 100,
FK_COD_ETAPA_ENSINO == 36 ~ 200,
FK_COD_ETAPA_ENSINO == 37 ~ 300,
FK_COD_ETAPA_ENSINO == 38 ~ 300,
FK_COD_ETAPA_ENSINO == 39 ~ 9999,
FK_COD_ETAPA_ENSINO == 40 ~ 9999,
FK_COD_ETAPA_ENSINO == 43 ~ 1000,
FK_COD_ETAPA_ENSINO == 44 ~ 1000,
FK_COD_ETAPA_ENSINO == 45 ~ 1000,
FK_COD_ETAPA_ENSINO == 46 ~ 1000,
FK_COD_ETAPA_ENSINO == 47 ~ 1000,
FK_COD_ETAPA_ENSINO == 48 ~ 1000,
FK_COD_ETAPA_ENSINO == 49 ~ 1000,

```

FK_COD_ETAPA_ENSINO == 50 ~ 1000,
FK_COD_ETAPA_ENSINO == 51 ~ 1000,
FK_COD_ETAPA_ENSINO == 52 ~ 1000,
FK_COD_ETAPA_ENSINO == 53 ~ 1000,
FK_COD_ETAPA_ENSINO == 54 ~ 1000,
FK_COD_ETAPA_ENSINO == 55 ~ 1000,
FK_COD_ETAPA_ENSINO == 57 ~ 1000,
FK_COD_ETAPA_ENSINO == 59 ~ 1000,
FK_COD_ETAPA_ENSINO == 60 ~ 1000,
FK_COD_ETAPA_ENSINO == 61 ~ 1000,
FK_COD_ETAPA_ENSINO == 62 ~ 1000,
FK_COD_ETAPA_ENSINO == 65 ~ 1000,
FK_COD_ETAPA_ENSINO == 67 ~ 1000,
FK_COD_ETAPA_ENSINO == 68 ~ 1000,
FK_COD_ETAPA_ENSINO == 69 ~ 1000,
FK_COD_ETAPA_ENSINO == 70 ~ 1000,
FK_COD_ETAPA_ENSINO == 71 ~ 1000,
FK_COD_ETAPA_ENSINO == 73 ~ 1000,
FK_COD_ETAPA_ENSINO == 74 ~ 1000,
TRUE ~ NaN ))

```

```

brasil2013_v3 <- brasil2013_1x3r %>%

```

```

  mutate( etapaEnsino = case_when( FK_COD_ETAPA_ENSINO == 1 ~ 1,
    FK_COD_ETAPA_ENSINO == 2 ~ 2,
    FK_COD_ETAPA_ENSINO == 14 ~ 10,
    FK_COD_ETAPA_ENSINO == 4 ~ 20,
    FK_COD_ETAPA_ENSINO == 15 ~ 20,
    FK_COD_ETAPA_ENSINO == 5 ~ 30,
    FK_COD_ETAPA_ENSINO == 16 ~ 30,
    FK_COD_ETAPA_ENSINO == 6 ~ 40,
    FK_COD_ETAPA_ENSINO == 17 ~ 40,
    FK_COD_ETAPA_ENSINO == 7 ~ 50,
    FK_COD_ETAPA_ENSINO == 18 ~ 50,
    FK_COD_ETAPA_ENSINO == 8 ~ 60,
    FK_COD_ETAPA_ENSINO == 19 ~ 60,
    FK_COD_ETAPA_ENSINO == 9 ~ 70,
    FK_COD_ETAPA_ENSINO == 20 ~ 70,
    FK_COD_ETAPA_ENSINO == 10 ~ 80,
    FK_COD_ETAPA_ENSINO == 21 ~ 80,
    FK_COD_ETAPA_ENSINO == 11 ~ 90,

```

FK_COD_ETAPA_ENSINO == 41 ~ 90,
FK_COD_ETAPA_ENSINO == 25 ~ 100,
FK_COD_ETAPA_ENSINO == 26 ~ 200,
FK_COD_ETAPA_ENSINO == 27 ~ 300,
FK_COD_ETAPA_ENSINO == 28 ~ 300,
FK_COD_ETAPA_ENSINO == 29 ~ 1000,
FK_COD_ETAPA_ENSINO == 30 ~ 100,
FK_COD_ETAPA_ENSINO == 31 ~ 200,
FK_COD_ETAPA_ENSINO == 32 ~ 300,
FK_COD_ETAPA_ENSINO == 33 ~ 300,
FK_COD_ETAPA_ENSINO == 34 ~ 1000,
FK_COD_ETAPA_ENSINO == 35 ~ 100,
FK_COD_ETAPA_ENSINO == 36 ~ 200,
FK_COD_ETAPA_ENSINO == 37 ~ 300,
FK_COD_ETAPA_ENSINO == 38 ~ 300,
FK_COD_ETAPA_ENSINO == 39 ~ 9999,
FK_COD_ETAPA_ENSINO == 40 ~ 9999,
FK_COD_ETAPA_ENSINO == 43 ~ 1000,
FK_COD_ETAPA_ENSINO == 44 ~ 1000,
FK_COD_ETAPA_ENSINO == 45 ~ 1000,
FK_COD_ETAPA_ENSINO == 46 ~ 1000,
FK_COD_ETAPA_ENSINO == 47 ~ 1000,
FK_COD_ETAPA_ENSINO == 48 ~ 1000,
FK_COD_ETAPA_ENSINO == 49 ~ 1000,
FK_COD_ETAPA_ENSINO == 50 ~ 1000,
FK_COD_ETAPA_ENSINO == 51 ~ 1000,
FK_COD_ETAPA_ENSINO == 52 ~ 1000,
FK_COD_ETAPA_ENSINO == 53 ~ 1000,
FK_COD_ETAPA_ENSINO == 54 ~ 1000,
FK_COD_ETAPA_ENSINO == 55 ~ 1000,
FK_COD_ETAPA_ENSINO == 57 ~ 1000,
FK_COD_ETAPA_ENSINO == 59 ~ 1000,
FK_COD_ETAPA_ENSINO == 60 ~ 1000,
FK_COD_ETAPA_ENSINO == 61 ~ 1000,
FK_COD_ETAPA_ENSINO == 62 ~ 1000,
FK_COD_ETAPA_ENSINO == 65 ~ 1000,
FK_COD_ETAPA_ENSINO == 67 ~ 1000,
FK_COD_ETAPA_ENSINO == 68 ~ 1000,
FK_COD_ETAPA_ENSINO == 69 ~ 1000,
FK_COD_ETAPA_ENSINO == 70 ~ 1000,

```

FK_COD_ETAPA_ENSINO == 71 ~ 1000,
FK_COD_ETAPA_ENSINO == 73 ~ 1000,
FK_COD_ETAPA_ENSINO == 74 ~ 1000,
TRUE ~ NaN ))

```

```

brasil2014_v3 <- brasil2014_1x3r %>%

```

```

  mutate( etapaEnsino = case_when( FK_COD_ETAPA_ENSINO == 1 ~ 1,
    FK_COD_ETAPA_ENSINO == 2 ~ 2,
    FK_COD_ETAPA_ENSINO == 14 ~ 10,
    FK_COD_ETAPA_ENSINO == 4 ~ 20,
    FK_COD_ETAPA_ENSINO == 15 ~ 20,
    FK_COD_ETAPA_ENSINO == 5 ~ 30,
    FK_COD_ETAPA_ENSINO == 16 ~ 30,
    FK_COD_ETAPA_ENSINO == 6 ~ 40,
    FK_COD_ETAPA_ENSINO == 17 ~ 40,
    FK_COD_ETAPA_ENSINO == 7 ~ 50,
    FK_COD_ETAPA_ENSINO == 18 ~ 50,
    FK_COD_ETAPA_ENSINO == 8 ~ 60,
    FK_COD_ETAPA_ENSINO == 19 ~ 60,
    FK_COD_ETAPA_ENSINO == 9 ~ 70,
    FK_COD_ETAPA_ENSINO == 20 ~ 70,
    FK_COD_ETAPA_ENSINO == 10 ~ 80,
    FK_COD_ETAPA_ENSINO == 21 ~ 80,
    FK_COD_ETAPA_ENSINO == 11 ~ 90,
    FK_COD_ETAPA_ENSINO == 41 ~ 90,
    FK_COD_ETAPA_ENSINO == 25 ~ 100,
    FK_COD_ETAPA_ENSINO == 26 ~ 200,
    FK_COD_ETAPA_ENSINO == 27 ~ 300,
    FK_COD_ETAPA_ENSINO == 28 ~ 300,
    FK_COD_ETAPA_ENSINO == 29 ~ 1000,
    FK_COD_ETAPA_ENSINO == 30 ~ 100,
    FK_COD_ETAPA_ENSINO == 31 ~ 200,
    FK_COD_ETAPA_ENSINO == 32 ~ 300,
    FK_COD_ETAPA_ENSINO == 33 ~ 300,
    FK_COD_ETAPA_ENSINO == 34 ~ 1000,
    FK_COD_ETAPA_ENSINO == 35 ~ 100,
    FK_COD_ETAPA_ENSINO == 36 ~ 200,
    FK_COD_ETAPA_ENSINO == 37 ~ 300,
    FK_COD_ETAPA_ENSINO == 38 ~ 300,
    FK_COD_ETAPA_ENSINO == 39 ~ 9999,

```

```

FK_COD_ETAPA_ENSINO == 40 ~ 9999,
FK_COD_ETAPA_ENSINO == 43 ~ 1000,
FK_COD_ETAPA_ENSINO == 44 ~ 1000,
FK_COD_ETAPA_ENSINO == 45 ~ 1000,
FK_COD_ETAPA_ENSINO == 46 ~ 1000,
FK_COD_ETAPA_ENSINO == 47 ~ 1000,
FK_COD_ETAPA_ENSINO == 48 ~ 1000,
FK_COD_ETAPA_ENSINO == 49 ~ 1000,
FK_COD_ETAPA_ENSINO == 50 ~ 1000,
FK_COD_ETAPA_ENSINO == 51 ~ 1000,
FK_COD_ETAPA_ENSINO == 52 ~ 1000,
FK_COD_ETAPA_ENSINO == 53 ~ 1000,
FK_COD_ETAPA_ENSINO == 54 ~ 1000,
FK_COD_ETAPA_ENSINO == 55 ~ 1000,
FK_COD_ETAPA_ENSINO == 57 ~ 1000,
FK_COD_ETAPA_ENSINO == 59 ~ 1000,
FK_COD_ETAPA_ENSINO == 60 ~ 1000,
FK_COD_ETAPA_ENSINO == 61 ~ 1000,
FK_COD_ETAPA_ENSINO == 62 ~ 1000,
FK_COD_ETAPA_ENSINO == 65 ~ 1000,
FK_COD_ETAPA_ENSINO == 67 ~ 1000,
FK_COD_ETAPA_ENSINO == 68 ~ 1000,
FK_COD_ETAPA_ENSINO == 69 ~ 1000,
FK_COD_ETAPA_ENSINO == 70 ~ 1000,
FK_COD_ETAPA_ENSINO == 71 ~ 1000,
FK_COD_ETAPA_ENSINO == 73 ~ 1000,
FK_COD_ETAPA_ENSINO == 74 ~ 1000,
TRUE ~ NaN ))

```

```

brasil2015_v3 <- brasil2015_1x3r %>%

```

```

  mutate( etapaEnsino = case_when( TP_ETAPA_ENSINO == 1 ~ 1,
    TP_ETAPA_ENSINO == 2 ~ 2,
    TP_ETAPA_ENSINO == 14 ~ 10,
    TP_ETAPA_ENSINO == 4 ~ 20,
    TP_ETAPA_ENSINO == 15 ~ 20,
    TP_ETAPA_ENSINO == 5 ~ 30,
    TP_ETAPA_ENSINO == 16 ~ 30,
    TP_ETAPA_ENSINO == 6 ~ 40,
    TP_ETAPA_ENSINO == 17 ~ 40,
    TP_ETAPA_ENSINO == 7 ~ 50,

```

TP_ETAPA_ENSINO == 18 ~ 50,
TP_ETAPA_ENSINO == 8 ~ 60,
TP_ETAPA_ENSINO == 19 ~ 60,
TP_ETAPA_ENSINO == 9 ~ 70,
TP_ETAPA_ENSINO == 20 ~ 70,
TP_ETAPA_ENSINO == 10 ~ 80,
TP_ETAPA_ENSINO == 21 ~ 80,
TP_ETAPA_ENSINO == 11 ~ 90,
TP_ETAPA_ENSINO == 41 ~ 90,
TP_ETAPA_ENSINO == 25 ~ 100,
TP_ETAPA_ENSINO == 26 ~ 200,
TP_ETAPA_ENSINO == 27 ~ 300,
TP_ETAPA_ENSINO == 28 ~ 300,
TP_ETAPA_ENSINO == 29 ~ 1000,
TP_ETAPA_ENSINO == 30 ~ 100,
TP_ETAPA_ENSINO == 31 ~ 200,
TP_ETAPA_ENSINO == 32 ~ 300,
TP_ETAPA_ENSINO == 33 ~ 300,
TP_ETAPA_ENSINO == 34 ~ 1000,
TP_ETAPA_ENSINO == 35 ~ 100,
TP_ETAPA_ENSINO == 36 ~ 200,
TP_ETAPA_ENSINO == 37 ~ 300,
TP_ETAPA_ENSINO == 38 ~ 300,
TP_ETAPA_ENSINO == 39 ~ 9999,
TP_ETAPA_ENSINO == 40 ~ 9999,
TP_ETAPA_ENSINO == 43 ~ 1000,
TP_ETAPA_ENSINO == 44 ~ 1000,
TP_ETAPA_ENSINO == 45 ~ 1000,
TP_ETAPA_ENSINO == 46 ~ 1000,
TP_ETAPA_ENSINO == 47 ~ 1000,
TP_ETAPA_ENSINO == 48 ~ 1000,
TP_ETAPA_ENSINO == 49 ~ 1000,
TP_ETAPA_ENSINO == 50 ~ 1000,
TP_ETAPA_ENSINO == 51 ~ 1000,
TP_ETAPA_ENSINO == 52 ~ 1000,
TP_ETAPA_ENSINO == 53 ~ 1000,
TP_ETAPA_ENSINO == 54 ~ 1000,
TP_ETAPA_ENSINO == 55 ~ 1000,
TP_ETAPA_ENSINO == 57 ~ 1000,
TP_ETAPA_ENSINO == 59 ~ 1000,

```

TP_ETAPA_ENSINO == 60 ~ 1000,
TP_ETAPA_ENSINO == 61 ~ 1000,
TP_ETAPA_ENSINO == 62 ~ 1000,
TP_ETAPA_ENSINO == 65 ~ 1000,
TP_ETAPA_ENSINO == 67 ~ 1000,
TP_ETAPA_ENSINO == 68 ~ 1000,
TP_ETAPA_ENSINO == 69 ~ 1000,
TP_ETAPA_ENSINO == 70 ~ 1000,
TP_ETAPA_ENSINO == 71 ~ 1000,
TP_ETAPA_ENSINO == 73 ~ 1000,
TP_ETAPA_ENSINO == 74 ~ 1000,
TRUE ~ NaN ))

```

```

brasil2016_v3 <- brasil2016_1x3r %>%

```

```

  mutate( etapaEnsino = case_when( TP_ETAPA_ENSINO == 1 ~ 1,

```

```

    TP_ETAPA_ENSINO == 2 ~ 2,
    TP_ETAPA_ENSINO == 14 ~ 10,
    TP_ETAPA_ENSINO == 4 ~ 20,
    TP_ETAPA_ENSINO == 15 ~ 20,
    TP_ETAPA_ENSINO == 5 ~ 30,
    TP_ETAPA_ENSINO == 16 ~ 30,
    TP_ETAPA_ENSINO == 6 ~ 40,
    TP_ETAPA_ENSINO == 17 ~ 40,
    TP_ETAPA_ENSINO == 7 ~ 50,
    TP_ETAPA_ENSINO == 18 ~ 50,
    TP_ETAPA_ENSINO == 8 ~ 60,
    TP_ETAPA_ENSINO == 19 ~ 60,
    TP_ETAPA_ENSINO == 9 ~ 70,
    TP_ETAPA_ENSINO == 20 ~ 70,
    TP_ETAPA_ENSINO == 10 ~ 80,
    TP_ETAPA_ENSINO == 21 ~ 80,
    TP_ETAPA_ENSINO == 11 ~ 90,
    TP_ETAPA_ENSINO == 41 ~ 90,
    TP_ETAPA_ENSINO == 25 ~ 100,
    TP_ETAPA_ENSINO == 26 ~ 200,
    TP_ETAPA_ENSINO == 27 ~ 300,
    TP_ETAPA_ENSINO == 28 ~ 300,
    TP_ETAPA_ENSINO == 29 ~ 1000,
    TP_ETAPA_ENSINO == 30 ~ 100,
    TP_ETAPA_ENSINO == 31 ~ 200,

```

```
TP_ETAPA_ENSINO == 32 ~ 300,  
TP_ETAPA_ENSINO == 33 ~ 300,  
TP_ETAPA_ENSINO == 34 ~ 1000,  
TP_ETAPA_ENSINO == 35 ~ 100,  
TP_ETAPA_ENSINO == 36 ~ 200,  
TP_ETAPA_ENSINO == 37 ~ 300,  
TP_ETAPA_ENSINO == 38 ~ 300,  
TP_ETAPA_ENSINO == 39 ~ 9999,  
TP_ETAPA_ENSINO == 40 ~ 9999,  
TP_ETAPA_ENSINO == 43 ~ 1000,  
TP_ETAPA_ENSINO == 44 ~ 1000,  
TP_ETAPA_ENSINO == 45 ~ 1000,  
TP_ETAPA_ENSINO == 46 ~ 1000,  
TP_ETAPA_ENSINO == 47 ~ 1000,  
TP_ETAPA_ENSINO == 48 ~ 1000,  
TP_ETAPA_ENSINO == 49 ~ 1000,  
TP_ETAPA_ENSINO == 50 ~ 1000,  
TP_ETAPA_ENSINO == 51 ~ 1000,  
TP_ETAPA_ENSINO == 52 ~ 1000,  
TP_ETAPA_ENSINO == 53 ~ 1000,  
TP_ETAPA_ENSINO == 54 ~ 1000,  
TP_ETAPA_ENSINO == 55 ~ 1000,  
TP_ETAPA_ENSINO == 57 ~ 1000,  
TP_ETAPA_ENSINO == 59 ~ 1000,  
TP_ETAPA_ENSINO == 60 ~ 1000,  
TP_ETAPA_ENSINO == 61 ~ 1000,  
TP_ETAPA_ENSINO == 62 ~ 1000,  
TP_ETAPA_ENSINO == 65 ~ 1000,  
TP_ETAPA_ENSINO == 67 ~ 1000,  
TP_ETAPA_ENSINO == 68 ~ 1000,  
TP_ETAPA_ENSINO == 69 ~ 1000,  
TP_ETAPA_ENSINO == 70 ~ 1000,  
TP_ETAPA_ENSINO == 71 ~ 1000,  
TP_ETAPA_ENSINO == 73 ~ 1000,  
TP_ETAPA_ENSINO == 74 ~ 1000,  
TRUE ~ NaN ))
```

```
brasil2017_v3 <- brasil2017_1x3r %>%
```

```
  mutate( etapaEnsino = case_when( TP_ETAPA_ENSINO == 1 ~ 1,  
    TP_ETAPA_ENSINO == 2 ~ 2,
```

TP_ETAPA_ENSINO == 14 ~ 10,
TP_ETAPA_ENSINO == 4 ~ 20,
TP_ETAPA_ENSINO == 15 ~ 20,
TP_ETAPA_ENSINO == 5 ~ 30,
TP_ETAPA_ENSINO == 16 ~ 30,
TP_ETAPA_ENSINO == 6 ~ 40,
TP_ETAPA_ENSINO == 17 ~ 40,
TP_ETAPA_ENSINO == 7 ~ 50,
TP_ETAPA_ENSINO == 18 ~ 50,
TP_ETAPA_ENSINO == 8 ~ 60,
TP_ETAPA_ENSINO == 19 ~ 60,
TP_ETAPA_ENSINO == 9 ~ 70,
TP_ETAPA_ENSINO == 20 ~ 70,
TP_ETAPA_ENSINO == 10 ~ 80,
TP_ETAPA_ENSINO == 21 ~ 80,
TP_ETAPA_ENSINO == 11 ~ 90,
TP_ETAPA_ENSINO == 41 ~ 90,
TP_ETAPA_ENSINO == 25 ~ 100,
TP_ETAPA_ENSINO == 26 ~ 200,
TP_ETAPA_ENSINO == 27 ~ 300,
TP_ETAPA_ENSINO == 28 ~ 300,
TP_ETAPA_ENSINO == 29 ~ 1000,
TP_ETAPA_ENSINO == 30 ~ 100,
TP_ETAPA_ENSINO == 31 ~ 200,
TP_ETAPA_ENSINO == 32 ~ 300,
TP_ETAPA_ENSINO == 33 ~ 300,
TP_ETAPA_ENSINO == 34 ~ 1000,
TP_ETAPA_ENSINO == 35 ~ 100,
TP_ETAPA_ENSINO == 36 ~ 200,
TP_ETAPA_ENSINO == 37 ~ 300,
TP_ETAPA_ENSINO == 38 ~ 300,
TP_ETAPA_ENSINO == 39 ~ 9999,
TP_ETAPA_ENSINO == 40 ~ 9999,
TP_ETAPA_ENSINO == 43 ~ 1000,
TP_ETAPA_ENSINO == 44 ~ 1000,
TP_ETAPA_ENSINO == 45 ~ 1000,
TP_ETAPA_ENSINO == 46 ~ 1000,
TP_ETAPA_ENSINO == 47 ~ 1000,
TP_ETAPA_ENSINO == 48 ~ 1000,
TP_ETAPA_ENSINO == 49 ~ 1000,

TP_ETAPA_ENSINO == 50 ~ 1000,
TP_ETAPA_ENSINO == 51 ~ 1000,
TP_ETAPA_ENSINO == 52 ~ 1000,
TP_ETAPA_ENSINO == 53 ~ 1000,
TP_ETAPA_ENSINO == 54 ~ 1000,
TP_ETAPA_ENSINO == 55 ~ 1000,
TP_ETAPA_ENSINO == 57 ~ 1000,
TP_ETAPA_ENSINO == 59 ~ 1000,
TP_ETAPA_ENSINO == 60 ~ 1000,
TP_ETAPA_ENSINO == 61 ~ 1000,
TP_ETAPA_ENSINO == 62 ~ 1000,
TP_ETAPA_ENSINO == 65 ~ 1000,
TP_ETAPA_ENSINO == 67 ~ 1000,
TP_ETAPA_ENSINO == 68 ~ 1000,
TP_ETAPA_ENSINO == 69 ~ 1000,
TP_ETAPA_ENSINO == 70 ~ 1000,
TP_ETAPA_ENSINO == 71 ~ 1000,
TP_ETAPA_ENSINO == 73 ~ 1000,
TP_ETAPA_ENSINO == 74 ~ 1000,
TRUE ~ NaN))