



UFRJ

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
PROGRAMA DE PÓS-GRADUAÇÃO EM EDUCAÇÃO



MONIQUE DAS GRAÇAS LIMA PINHEIRO

ANÁLISE DA TRAJETÓRIA ESCOLAR NOS ANOS FINAIS DE ALUNOS DA REDE PÚBLICA
MUNICIPAL DO RIO DE JANEIRO

Linha de Pesquisa: Políticas e Instituições Educacionais
Orientador: Prof. Dr. Tiago Lisboa Bartholo

**RIO DE JANEIRO
2018**

MONIQUE DAS GRAÇAS LIMA PINHEIRO

**ANÁLISE DA TRAJETÓRIA ESCOLAR NOS ANOS FINAIS DE ALUNOS DA REDE PÚBLICA
MUNICIPAL DO RIO DE JANEIRO**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Educação, na Linha de Políticas e Instituições Educacionais, da Universidade Federal do Rio de Janeiro, como requisito parcial à obtenção do título de Mestre em Educação.

Orientador: Prof. Dr. Tiago Lisboa Bartholo

Rio de Janeiro
2018

CIP - Catalogação na Publicação

P654a Pinheiro, Monique das Graças Lima
Análise da trajetória escolar nos anos finais de
alunos da rede pública municipal do Rio de Janeiro
/ Monique das Graças Lima Pinheiro. -- Rio de
Janeiro, 2018.
199 f.

Orientador: Tiago Lisboa Bartholo.
Dissertação (mestrado) - Universidade Federal do
Rio de Janeiro, Faculdade de Educação, Programa de
Pós-Graduação em Educação, 2018.

1. Trajetória Escolar. 2. Análise de
Sobrevivência. 3. Repetência. I. Bartholo, Tiago
Lisboa, orient. II. Título.



Universidade Federal do Rio de Janeiro

Centro de Filosofia e Ciências Humanas
Faculdade de Educação
Programa de Pós-Graduação em Educação

A Dissertação "**Análise da trajetória escolar nos anos finais de alunos da rede pública municipal do Rio de Janeiro**"

Mestrando(a): **Monique das Graças Lima Pinheiro**

Orientado(a) pelo(a): **Prof(a). Dr(a). Tiago Lisboa Bartholo**

E aprovada por todos os membros da Banca Examinadora, foi aceita pela Faculdade de Educação da Universidade Federal do Rio de Janeiro e homologada pelo Conselho de Ensino para Graduados e Pesquisa, como requisito parcial à obtenção do título de

MESTRE EM EDUCAÇÃO

Rio de Janeiro, 27 de julho de 2018.

Banca Examinadora:

Prof(a). Dr(a). Tiago Lisboa Bartholo- Presidente

Prof(a). Dr(a). Mariane Campelo Koslinski

Prof(a). Dr(a). Eduardo Ribeiro da Silva

Dedico este trabalho a todas as crianças e jovens desse país, que merecem uma educação de qualidade, para que se tornem cidadãos e, acima de tudo, pessoas do Bem. Todos nós merecemos um país melhor!

AGRADECIMENTOS

Prometo tentar se sucinta, mas este momento é um bastante especial para mim, onde de coração aberto agradeço a muitas pessoas que me ajudaram de alguma forma na conclusão desse trabalho, seja com a mão na massa, seja no incentivo e na torcida ou até em, simplesmente compreender, sem cobranças, os muitos momentos em que estive ausente.

Agradeço primeiramente ao meu marido, Edimilson Pinheiro, que foi meu parceiro e incentivador sempre, que soube me dar colo quando tudo parecia muito difícil, que soube me ajudar a focar e a centrar, quando dava vontade de jogar tudo para o alto, que me apoiou em todas as decisões, que cuidou de mim e fez tudo que podia para me ajudar e, principalmente, que aguentou firme a “pauleira” que é conviver com uma pessoa surtada, cansada, perfeccionista e atrasada...

Aos meus pais pela boa estrutura familiar e afetiva que me proporcionam, se concluí mais essa etapa, foi graças ao “belo trabalho” que fizeram e ainda fazem por mim.

À minha Tia Denise que também colocou a mão na massa, me ajudando na revisão, mas, mais que isso, esteve sempre ao meu lado e, não duvido nada, se pudesse, até faria esse trabalho por mim.

À minha família de sangue e à minha família de coração (meus amigos) que souberam entender a minha ausência, mas que continuaram sendo meus parceiros e meus incentivadores.

Aos meus colegas de trabalho, em especial ao meu chefe Natalino, que me apoiaram e seguraram as pontas para que cursasse o Mestrado e, mais ainda, sem questionar entendeu que precisei me ausentar durante um período, mesmo durante o “caos”, para que este trabalho realmente acontecesse. Carlinha, apesar de não estar mais na nossa equipe, você faz parte desse grupo, #SomosSempreNIEE.

Edu, você também merece um obrigado mais que especial, sem você, tenho certeza de que esse trabalho não teria terminado desse jeito, devo muito a você, obrigada por me ajudar com o método, obrigada por me ajudar com os dados e mais que tudo, obrigada pela disponibilidade para tirar as minhas dúvidas todas as vezes que precisei, a qualquer hora e dia.

Ao meu orientador Tiago, que foi a peça motriz dessa trajetória, que me deu todo o apoio para entrar no curso, que acreditou em mim e não me deixou desistir quando tudo parecia muito difícil e que persistiu comigo até o fim, lado a lado.

Aos meus amigos Rosa e Daniel, que sem nunca nos vermos antes, nos encontramos nos primeiros dias de aula e, sem muito esforço, viramos parceiros de curso e, melhor ainda, amigos para sempre, mesmo quando cada um continuou seus próprios caminhos.

À Érica Jares por estar me ajudado a focar, a entender minhas questões e, sobretudo, a vencê-las, obrigada por ter tornado este trabalho um fato real e não somente mais um projeto apenas idealizado.

Às muitas pessoas queridas que trabalham ou já trabalharam na SME-Rio, obrigada por tudo que aprendi com vocês nessa jornada onde aprendemos a arte de olhar “a garrafa pelo outro lado do rótulo”.

Aos muitos professores que passaram por toda a minha vida, alguns foram exemplos, outros se tornaram amigos, mas todos foram essenciais para minha formação.

E por fim, aos meus alunos, que se não fossem eles hoje não seria quem sou!

[...]

“Manhê! Tirei um dez na prova

Me dei bem, tirei um cem e eu quero ver quem me reprova

Decorei toda lição

Não errei nenhuma questão

Não aprendi nada de bom

Mas tirei dez (boa filhão!)

Quase tudo que aprendi, amanhã eu já esqueci

Decorei, copiei, memorizei, mas não entendi

Decoreba: esse é o método de ensino

Eles me tratam como ameba e assim eu não raciocino

Não aprendo as causas e consequências

Só decoro os fatos

Desse jeito até história fica chato

[...]

Trecho da Música: Estudo Errado, Gabriel O Pensador

RESUMO

PINHEIRO, Monique das Graças Lima. **Análise da trajetória escolar nos anos finais de alunos da rede pública municipal do Rio de Janeiro**. Rio de Janeiro, 2018. Dissertação (Mestrado em Educação) – Programa de Pós-Graduação em Educação, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2018.

A dissertação analisa a trajetória escolar de uma coorte específica de alunos matriculados no 6º ano do ensino fundamental em 2012 na rede municipal da cidade do Rio de Janeiro, ao longo de quatro anos (2012-2015), e tem como objetivo principal identificar fatores que possam estar associados à repetência escolar dos alunos das escolas de rede pública municipal da cidade do Rio de Janeiro. O método utilizado é a análise de sobrevivência e todos os alunos da coorte são observados até que saiam da sua trajetória, seja por repetência ou outro movimento que interrompa o fluxo regular no segundo segmento do ensino fundamental. O estudo foi dividido em duas partes, sendo a primeira uma análise descritiva do perfil da coorte no momento inicial do estudo e da trajetória escolar percorrida por esses alunos, bimestre a bimestre, e a segunda uma análise de sobrevivência, onde foi possível comparar as funções de sobrevivência de diferentes grupos utilizando-se o Estimador de Kaplan-Meier e o Modelo de Regressão de Cox. O modelo final utiliza como variáveis independentes características do aluno e sua família, das escolas e das principais políticas educacionais na rede pública. As análises preliminares sugerem que: a) apenas 45,97% dos alunos permaneceram dentro do fluxo escolar esperado, sendo promovidos ano a ano até a conclusão do ensino fundamental no final do estudo; b) 29,18% dos alunos tiveram pelo menos uma repetência ao longo desses anos; c) 1,09% foram encaminhados para a Educação de Jovens e Adultos; d) 3,53% foram encaminhados para turmas de projeto de correção de fluxo; e) 20,24% saíram da rede, seja por transferências externas, abandonos ou evasões; f) a taxa de incidência média para a repetência nos anos finais é de 10,5%, sendo maior na transição do 6º para o 7º ano do ensino fundamental; g) o comparativo das funções de sobrevivência ratificou que os alunos repetentes diferem, de forma significativa, dos alunos promovidos em vários aspectos isoladamente; h) o Modelo de Regressão de Cox revelou o efeito das covariáveis sob a razão de risco de repetência da coorte.

Palavras-chave: trajetória escolar; análise de sobrevivência; repetência;

ABSTRACT

PINHEIRO, Monique das Graças Lima. . **Análise da trajetória escolar nos anos finais de alunos da rede pública municipal do Rio de Janeiro**. Rio de Janeiro, 2018. Dissertação (Mestrado em Educação) – Programa de Pós-Graduação em Educação, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2018.

The dissertation analyzes the school trajectory of a specific cohort of students enrolled in the 6th year of elementary education in 2012 in Rio de Janeiro public municipal schools over four years (2012-2015), and the main aim is to identify factors that may be associated with school repetition in the pupils of the municipal public schools of the city of Rio de Janeiro. The method used is the survival analysis and all the students of the cohort are observed until they leave their trajectory, either by repetition or another movement that interrupts the regular flow in the second segment of elementary school. The study has been divided in two parts: the first a descriptive analysis of the cohort profile at the initial moment of the study and of the school trajectory of these students, bimester to bimester, and the second a survival analysis, where it was possible to compare the survival functions of different groups using the Kaplan-Meier Estimator and the Cox Regression Model. The final model uses as independent variables characteristics of the students and their family, of the schools and of the main educational policies in the public municipal schools. Preliminary analyzes suggest that: a) only 45.97% of the students remained within the expected school flow, being promoted year by year until completion of elementary education at the end of the study; b) 29.18% of the students had at least one repetition over those years; c) 1.09% were relocated to Youth and Adult Education; d) 3.53% were referred to classes of flow correction project; e) 20.24% left the public municipal schools, either by external transfers, dropouts or evasions; f) the average incidence rate for repetition in the final years is 10.5%, being higher in the transition from the 6th to the 7th year of elementary education; g) the comparative of the survival functions have confirmed that retain pupils differ significantly from students promoted in various aspects singly; h) the Cox Regression Model revealed the effect of covariables under the hazard ratio of cohort's repetition.

Keywords: grade repetition; school trajectory; survival analysis;

LISTA DE QUADROS

Quadro 1: Indicador 2B - Percentual de pessoas de 16 anos com pelo menos o EF concluído – Brasil, regiões e Estados. Fonte: Brasil, 2018, p. 46-47	34
Quadro 2: Programas, Projetos e Ações instituídos para o Reforço Escolar. Fonte: Santos (2014, p.116-117).....	59
Quadro 3: Conceitos básicos da análise de sobrevivência no estudo	64
Quadro 4: Variáveis dependentes	70
Quadro 5: Covariáveis – Perfil socioeconômico dos alunos	71
Quadro 6: Covariáveis – Trajetória e desempenho escolar dos alunos	72
Quadro 7: Covariáveis– Política Pública da Escola.....	73
Quadro 8: Critérios de Enturmação - 6º Ano - 2012-2013. Fonte: SME-Rio ³	92
Quadro 9: Resumo da trajetória da coorte (2012-2015)	95

LISTA DE FIGURAS

Figura 1: Taxas de fluxo 2014/2015 x Taxas de Rendimento 2014 - Brasil. Fonte: DEED/INEP, 2017.....	40
Figura 2: Estrutura Básica. Fonte: RIO DE JANEIRO, 2016, p. 8.....	49
Figura 3: Estrutura do Ginásio Carioca. Fonte: Rioeduca.....	55
Figura 4: Esquema de acompanhamento da Coorte.....	65

LISTA DE GRÁFICOS

Gráfico 1: Taxa de Aprovação no Ensino Fundamental por Capital (2016). Fonte: MEC/INEP. Elaboração própria da autora.	23
Gráfico 2: Indicador 2A – Percentual de pessoas de 6 a 14 anos que frequentam ou já que concluíram o EF (Meta 2 - PNE 2014-2024). Fonte: Brasil, 2018, p. 41.....	30
Gráfico 3: Indicador 2A – Percentual de pessoas de 6 a 14 anos que frequentam ou já que concluíram o EF (Meta 2 PNE 2014-2024) - Brasil, Sudeste, Estado do Rio de Janeiro. Fonte: Brasil, 2018, p.42. Elaboração própria	31
Gráfico 4: Indicador 2B - Percentual de pessoas de 16 anos com pelo menos o EF concluído (Meta 2 - PNE 2014-2024). Fonte: Brasil, 2018, p. 45.....	32
Gráfico 5: Indicador 2B - Percentual de pessoas de 16 anos com pelo menos o EF concluído (Meta 2 - PNE 2014-2024) - Brasil, Sudeste, Estado do Rio de Janeiro. Fonte: Observatório do PNE (IBGE/Pnad), 2018. Elaboração própria.....	33
Gráfico 6: Taxas de Fluxo - Brasil. Fonte: INEP, 2017. Elaboração própria.....	41
Gráfico 7: Taxas de Fluxo – Região Sudeste. Fonte: INEP, 2017. Elaboração própria.....	42
Gráfico 8: Taxas de Fluxo – Estado do Rio de Janeiro. Fonte: INEP, 2017. Elaboração própria.	42
Gráfico 9: Taxas de Fluxo – Município do Rio de Janeiro. Fonte: INEP, 2017. Elaboração própria.....	43
Gráfico 10: Taxas de Fluxo – Município do Rio de Janeiro (Escolas Públicas). Fonte: INEP, 2017. Elaboração própria.....	43
Gráfico 11: Taxa de Abandono - Anos Finais - Rede x Escolas do Amanhã. Fonte: SME-Rio ..	52
Gráfico 12: Total de Escolas do 6º Ano Experimental. Fonte: Dados da SME-Rio. Elaboração própria.....	54
Gráfico 13: Sexo dos alunos	76
Gráfico 14: Idade da Coorte em 2012.....	77
Gráfico 15: Anos de Defasagem (Apenas para os alunos com Defasagem)	78
Gráfico 16: Sexo vs Raça dos alunos	79
Gráfico 17: Percentual de alunos do 6º Ano Experimental 2012, por ano.....	87
Gráfico 18: Resumo das Políticas Públicas.....	89
Gráfico 19: Resumo das trajetórias.....	95

Gráfico 20: Função de Risco EKM	98
Gráfico 21: Função de Sobrevivência, EMK	99
Gráfico 22: Função de Sobrevivência - Sexo	103
Gráfico 23: Função de Sobrevivência - Não Branco	103
Gráfico 24: Função de Sobrevivência - Sexo e Não Branco	103
Gráfico 25: Função de Sobrevivência - Defasado	103
Gráfico 26: Função de Sobrevivência - Máxima Escolaridade dos Pais.....	105
Gráfico 27: Função de Sobrevivência - Bolsa Família	105
Gráfico 28: Função de Sobrevivência - Ginásio Carioca	108
Gráfico 29: Função de Sobrevivência - 6º Ano Experimental x Ginásio Carioca	108

LISTA DE TABELAS

Tabela 1: Idade em 2012.....	76
Tabela 2: Análise Descritiva de Anos de Defasagem em 2012	77
Tabela 3: Raça dos alunos	78
Tabela 4: Tipo de deficiência - Alunos incluídos	79
Tabela 5: Família biparental.....	80
Tabela 6: Máxima Escolaridade de Pais	81
Tabela 7: Bolsa Família.....	82
Tabela 8: Cartão Família Carioca.....	82
Tabela 9: Médias das Provas Bimestrais	83
Tabela 10: Quantidade de Conceito I no COC 1 e 2, por ano.	84
Tabela 11: Frequência Anual dos alunos, por ano.....	84
Tabela 12: Movimentações por ano	85
Tabela 13: Distribuição da coorte no Ginásio Carioca	88
Tabela 14: Distribuição da coorte no Programa Escolas do Amanhã	88
Tabela 15: Distribuição da coorte para o Turno Único	89
Tabela 16: Trajetória Detalhada da Coorte em 2012	90
Tabela 17: Trajetória Detalhada da Coorte em 2012-2013	92
Tabela 18: Trajetória Detalhada da Coorte em 2013-2014	93
Tabela 19: Trajetória Detalhada da Coorte em 2014-2015	94
Tabela 20: Resultado final de 2015.....	94
Tabela 21: Descritiva da Análise de Sobrevivência	97
Tabela 22: Resumo dos dados de sobrevivência por ano.....	97
Tabela 23: Função de Sobrevivência de EKM	98
Tabela 24: Testes de comparação entre grupos.....	100
Tabela 25: Função de Sobrevivência - Sexo.	101

Tabela 26: Função de Sobrevivência - Não Brancos.....	101
Tabela 27: Função de Sobrevivência - Sexo e Não Branco.....	102
Tabela 28: Função de Sobrevivência – Defasado	103
Tabela 29: Função de Sobrevivência - Máxima Escolaridade Pais	104
Tabela 30: Função de Sobrevivência - Bolsa Família	105
Tabela 31: Função de Sobrevivência - Ginásio Carioca	106
Tabela 32: Função de Sobrevivência - 6º Ano experimental x Ginásio Carioca	107
Tabela 33: Regressão de Cox - Modelo 1	109
Tabela 34: Regressão de Cox - Modelo 2	111
Tabela 35: Regressão de Cox - Modelo 3	112
Tabela 36: Regressão de Cox - Modelo 4	113
Tabela 37: Regressão de Cox - Modelo 5	116
Tabela 38: Modelos de Regressão de Cox.....	119

LISTA DE ABREVIATURAS E SIGLAS

CFB	Constituição Federal Brasileira
AF	Anos Finais
AI	Anos Iniciais
AIEE	Assessoria Técnica de Informações Educacionais Estratégicas
EA	Escolas do Amanhã
EJA	Educação de Jovens e Adultos
FMI	Fundo Monetário Internacional
GC	Ginásio Carioca
GEO	Ginásio Experimental Olímpico
IBGE	Instituto Brasileiro de Geografia e Estatística
INEP	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
LDB	Lei de Diretrizes e Bases da Educação
MEC	Ministério da Educação
PISA	<i>Programme for International Student Assessment</i>
Pnad	Pesquisa Nacional por Amostra de Domicílios
Pnad-c	Pesquisa Nacional por Amostra de Domicílios Contínua
PNE	Plano Nacional de Educação
OCDE	Organização para Cooperação e Desenvolvimento Econômico
OIT	Organização Internacional do Trabalho
OPNE	Observatório do Plano Nacional de Educação
SCA	Sistema de Controle Acadêmico
SGA	Sistema de Gestão Acadêmica
SME-Rio	Secretaria Municipal de Educação da Cidade do Rio de Janeiro
TU	Turno Único

SUMÁRIO

RESUMO.....	8
ABSTRACT	9
LISTA DE QUADROS	10
LISTA DE FIGURAS.....	11
LISTA DE GRÁFICOS.....	12
LISTA DE TABELAS.....	14
LISTA DE ABREVIATURAS E SIGLAS	16
SUMÁRIO	17
1 INTRODUÇÃO.....	19
1.1 JUSTIFICATIVA.....	23
1.2 OBJETIVOS.....	26
1.2.1 Objetivo Geral.....	26
1.2.2 Objetivos Específicos	26
2 O CONTEXTO EDUCACIONAL DO BRASIL	27
3 FLUXO ESCOLAR	37
3.1 FATORES ASSOCIADOS À REPETÊNCIA ESCOLAR.....	44
4 A REDE MUNICIPAL DA CIDADE DO RIO DE JANEIRO	47
5 MÉTODOS	61
5.1 O DESENHO.....	61
5.1.1 Análise de Sobrevivência	61
5.2 BANCOS DE DADOS.....	69
5.2.1 As variáveis	70
5.2.2 O banco de dados para a Análise de Sobrevivência	73
6 RESULTADOS.....	75
6.1 ANÁLISE DESCRITIVA DAS VARIÁVEIS	75
6.1.1 Análise Descritiva do Perfil dos Alunos	75
6.1.2 Análise Descritiva do Desempenho Acadêmico dos Alunos.....	82
6.1.3 Análise Descritiva das Políticas Públicas das Escolas.....	85
6.2 ANÁLISE DESCRITIVA DA TRAJETÓRIA ESCOLAR DOS ALUNOS	89
6.3 ANÁLISE DE SOBREVIVÊNCIA	96
7 CONSIDERAÇÕES FINAIS	120
REFERÊNCIAS	125
ANEXOS.....	137
ANEXO A – ANÁLISE DE SOBREVIVÊNCIA – SINTAXE DO STATA.....	139

ANEXO B – GRÁFICOS COMPARATIVOS DAS FUNÇÕES DE SOBREVIVÊNCIA E SINTAXE DO STATA	141
ANEXO C – MODELAGEM DA REGRESSÃO DE COX.....	177

1 INTRODUÇÃO

Um sistema educacional ideal seria aquele que todos tivessem acesso à escola e ao aprendizado adequado, evitando o desperdício de recursos com repetências e abandonos. Portanto, dois dos maiores desafios do sistema educacional brasileiro é garantir o aprendizado e o fluxo escolar adequado dos alunos matriculados na rede pública e privado (FERNANDES, 2007).

Tavares Jr. *et al.* (2012) esclarece o binômio “acesso x fluxo”, destacando que o mesmo “avança-se numa dimensão e retroage-se concomitantemente em outra, gerando resultados por vezes mais perversos do que o cenário anterior”, pois a medida em que a escola passa a admitir novos alunos – principalmente aqueles oriundos de classes mais populares – essa mesma escola os reprovam, anulando os avanços obtidos em relação ao acesso. O mesmo é observado ao se iniciar a tentativa de melhoria desse fluxo¹ alguns anos após, onde se observou a queda na qualidade da educação, fazendo com que:

“A democratização da educação brasileira parece ser caracterizada por uma dualidade, um perfil estranho e dialético de realização social. Se por um lado avança na direção das principais demandas sociais, por outro, num mesmo movimento, também tende a anular esses avanços, gerando efeitos perversos em outra ponta. Tal traço acaba por reiterar a desigualdade, como elemento característico de nossa formação social. [...] É preciso compreender como inaceitáveis tais indicadores de rendimento (fluxo) e desempenho (aprendizagem). A aceitação passiva de tais resultados fortaleceria a força motriz reprodutivista que ainda hoje opera da escola brasileira, através de sua ‘Pedagogia da Repetência’” (TAVARES JR *et al.*, 2012).

A reprovação escolar é considerada por alguns pesquisadores como o maior problema do sistema educacional brasileiro (TAVARES JR. *et al.*, 2012), sendo um desafio enfrentado por gestores e educadores de diversos estados e municípios. Os dados do Censo Escolar 2016, divulgados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), mostram que dos mais de 27 milhões de alunos matriculados no ensino fundamental, 8,3% (2.299.937 alunos) foram reprovados no ano de 2016. Analisando os dois segmentos do ensino fundamental de forma isolada, encontramos taxas de reprovação de 5,9% (905.063 alunos) para os anos iniciais (AI) e 11,4% (1.394.874 alunos) para os anos finais (AF). Restringindo os cálculos apenas às escolas públicas, os dados ficam ainda mais

¹ Através de políticas de correção de fluxo e a implementação de ciclos com progressões automáticas.

alarmantes, sendo 6,8% (860.299 alunos) nos AI e 12,7% (1.321.718 alunos) nos AF, totalizando 2.182.017 alunos reprovados no ensino fundamental da rede pública do Brasil somente no ano de 2016. Os dados reforçam, no entanto, que apesar de maiores no sistema público de ensino, as taxas de reprovação são relevantes também no ensino privado, demonstrando assim ser um fenômeno que afeta a educação brasileira como um todo.

Fajardo (2018) cita que somente em 2016, o Brasil gastou R\$16 bilhões com a reprovação de alunos na educação básica, sendo R\$12 bilhões exclusivamente no ensino fundamental. O gasto adicional ocorre porque ao demandar que o aluno refaça o ano escolar, o número de anos que um aluno potencialmente ficará no sistema escolar aumenta. Após uma única reprovação é esperado que o aluno permanecesse dez anos na escola ao longo do ensino fundamental e não mais nove anos (tempo esperado em um fluxo ideal). Isso impacta o orçamento dos sistemas de ensino. Bruns *et al.* (2012) destacam que o Brasil tem a maior taxa de reprovação do mundo, excluindo apenas alguns poucos países de baixa renda da África, e que esse é um dado relevante que sugere a ineficácia das políticas educacionais, o desperdício de recursos e o tempo dos alunos. A reprovação escolar está associada ainda a outro fenômeno educacional amplamente estudado, a saber, o abandono escolar. Há evidências que sugerem que a reprovação impacta as chances do aluno abandonar a escola já que “aproximadamente 30% dos estudantes brasileiros desistem antes de concluírem o ensino fundamental, depois de terem permanecido na escola por 11 anos em média” (BRUNS *et al.*, 2012, p.47, tradução nossa). Esse seria um dos efeitos perversos não previstos da adoção de uma prática que visa fundamentalmente aumentar a oportunidade dos alunos aprenderem determinados conteúdos que não foram assimilados.

Apesar de ser uma prática comum no Brasil, a reprovação escolar não é exclusiva do nosso sistema educacional, podendo ser considerada como um fenômeno educacional universal. Por exemplo, na União Europeia, segundo os dados do *Programme for International Student Assessment (PISA) 2011*, 10,4% dos estudantes já repetiram pelo menos uma vez na escola secundária obrigatória², sendo a Espanha o país que apresenta as maiores taxas - 31,9% (KLAPPROTH, SCHALTZ, 2015). A UNESCO (2012) informa ainda que, em 2010, 70% dos alunos repetentes (mais de 22,5 milhões de alunos) nos AF no mundo residiam em três regiões: América Latina e Caribe, Sul e Oeste da Ásia e África Subsaariana.

² Esse nível de escolaridade seria similar aos anos finais do ensino fundamental no sistema educacional brasileiro.

Por ser um fenômeno de magnitude mundial, há muito tempo a reprovação escolar é recorrente no cenário das pesquisas nacionais e internacionais, inclusive com diferentes abordagens. Em 1994, Gomes-Neto e Hanushek (1994) estudaram as causas e consequências da reprovação nos estudantes rurais do nordeste brasileiro; Klein e Ribeiro (1995) retratam a persistência da “pedagogia da repetência” presente no Brasil desde os anos 40; Oliveira e Soares (2012) buscaram os determinantes da repetência escolar; Correa *et al.* (2014) pesquisaram os efeitos sobre o desempenho acadêmico em testes cognitivos dos alunos repetentes em comparação com os promovidos; Ribeiro *et al.* (2018) analisaram a crença dos professores em relação à reprovação; Bacchetto (2016) calculou o custo da reprovação para o Brasil; Chan *et al.* (2015) observaram os efeitos da reprovação na motivação dos estudantes no Texas, EUA; Klapproth *et al.* (2016) averiguaram os efeitos da reprovação, a curto e médio prazo, nos resultados acadêmicos e psicológicos de alunos dos AF de Luxemburgo.

Apesar do grande número de pesquisas sobre o tema, poucos estudos sugerem que a reprovação escolar pode ser uma prática eficiente no processo de recuperação do aprendizado dos alunos, indicando que a reprovação pode ser vista como algo positivo, pois pode proporcionar uma oportunidade adicional para os alunos recuperarem pelo menos parte dos conteúdos que não foram assimilados. Além disso, essa ideia é defendida por vários alunos repetentes que consideram a reprovação como algo que potencializa o aprendizado, em especial dos alunos com maior dificuldade (PATTO, 1999). Contudo, a maioria dos estudos conclui que a reprovação é algo negativo para o aluno, pois além de não garantir um maior aprendizado, é prejudicial em termos comportamentais e emocionais (JIMERSON, 2001).

Koch *et al.* (2007) detectaram em sua pesquisa que a maioria dos professores dizem que os alunos reprovam porque não querem aprender, sendo em sua maioria indisciplinados, desinteressados ou pertencentes a famílias consideradas ‘defasadas culturalmente’. Os argumentos apresentados pelos professores reforçam a ideia de que a reprovação não está diretamente relacionada ao trabalho pedagógico desenvolvido em sala de aula, mas sim com aspectos individuais dos alunos e/ou família.

Desse modo, quando o educando se depara com uma prática pedagógica inadequada na sala de aula, os conteúdos apresentados se tornam de difícil assimilação, são confundidos

“com outros conhecimentos já adquiridos e não se integram aos novos, gerando grande ‘confusão’ e tornando a elaboração do conhecimento mais demorada e difícil” (WEISS, 2012, p. 24). Sendo possível perceber que a reprovação é um fenômeno que afeta a vida acadêmica e profissional do discente, fato é que esse processo acaba interferindo na formação intelectual, na possibilidade de abandono do curso ou até na vida do indivíduo (BISINOTO *et al.*, 2010).

Ainda que os argumentos dos efeitos pedagógicos da repetência, para consolidar conhecimentos fundamentais da série anteriormente cursada, também já tenham sido amplamente refutados pelas pesquisas educacionais, a repetência ainda afeta a trajetória de um número elevado de alunos, causando a distorção idade-série e o represamento do fluxo escolar (CRAHAY, 2007), acarretando a baixa efetividade dos sistemas de ensino, em face do congestionamento do fluxo escolar, afeta o autoconceito, o relacionamento, a trajetória escolar e a expectativa de renda futura dos alunos (SOARES, J., 2007; OLIVEIRA, SOARES, 2012). Além disso, a retenção dos alunos pode ser um fator importante para a evasão da escola (RIANI *et al.*, 2012).

Os argumentos dos efeitos pedagógicos da repetência para consolidar conhecimentos fundamentais da série anteriormente cursada já foram amplamente refutados por pesquisas educacionais. Entretanto, a repetência ainda afeta a trajetória de um número elevado de alunos, causando a distorção idade série e o represamento do fluxo escolar (CRAHAY, 2007). Isto acarreta em uma baixa efetividade dos sistemas de ensino em face do congestionamento do fluxo escolar, além de afetar negativamente o autoconceito, relacionamento, trajetória escolar e expectativa de renda futura dos alunos (SOARES, J., 2007; OLIVEIRA, SOARES, 2012). Além disso, a retenção dos alunos pode ser um fator importante para a evasão da escola (RIANI *et al.*, 2012).

Portanto, ganha cada vez mais importância a formulação e a implementação de políticas públicas que focalizem os problemas retratados nos indicadores gerados pelos instrumentos avaliativos da educação capazes de diagnosticar a realidade e informar as demandas equalizadoras para o governo, seja na esfera nacional, estadual e municipal.

Busca-se, portanto, nesta pesquisa identificar fatores associados à repetência escolar ao longo dos AF do ensino fundamental na rede pública municipal da cidade do Rio de Janeiro. Nesse sentido, a questão do estudo é: “Quais são os fatores associados

(características dos alunos, das famílias e das escolas) à repetência escolar nos anos finais do ensino fundamental na rede pública municipal da cidade do Rio de Janeiro?”.

A dissertação foi organizada em sete capítulos, começando com a Introdução. No Capítulo 2 é apresentado um breve contexto educacional do Brasil. O fluxo escolar, assim como os fatores associados à reprovação escolar, é discutido no Capítulo 3. O Capítulo 4 é dedicado à Rede Municipal da Cidade do Rio de Janeiro, contextualizando-o presente estudo. O desenho do estudo é encontrado no Capítulo 5, com informações sobre os dados e suas variáveis, a amostra e os métodos utilizados. No Capítulo 6 os resultados encontrados são apresentados em três momentos: a análise descritiva da coorte estudada, análise da trajetória escolar desses alunos na rede e os resultados da análise de sobrevivência. Ao final, no capítulo 7, são apresentadas as considerações finais do estudo.

1.1 JUSTIFICATIVA

Diante dos fatos, acredita-se ser importante fazer uma investigação que utilize um desenho de pesquisa adequado para identificar fatores associados à reprovação escolar, em especial na realidade brasileira.

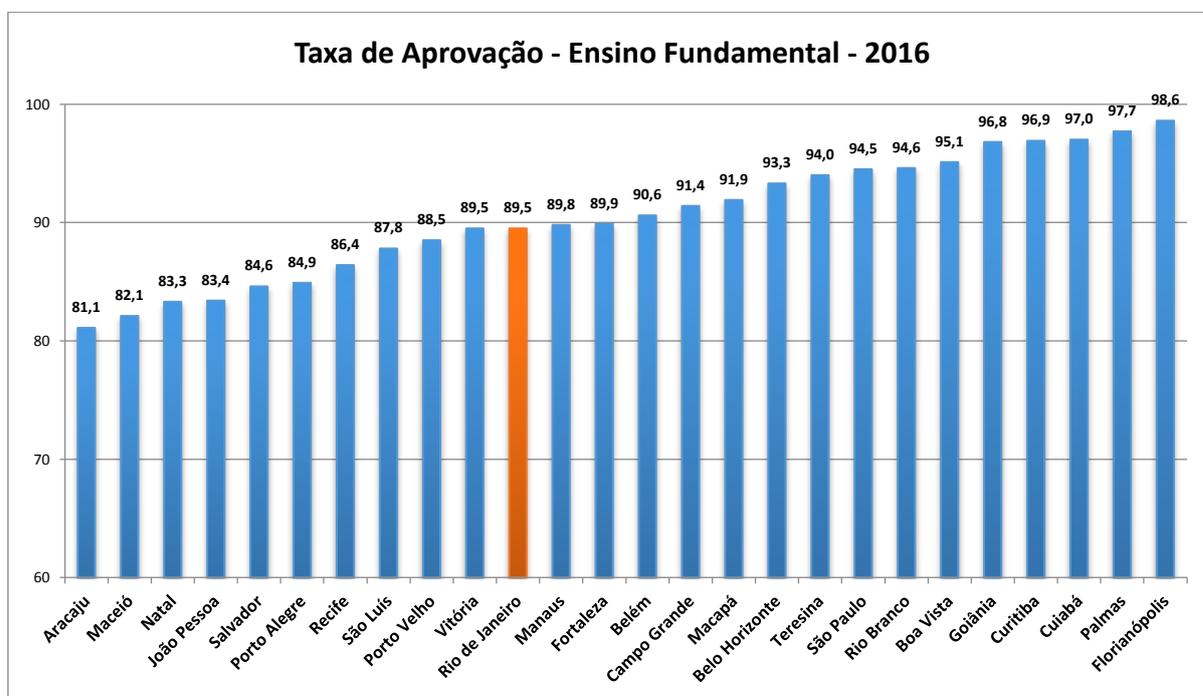


Gráfico 1: Taxa de Aprovação no Ensino Fundamental por Capital (2016). Fonte: MEC/INEP. Elaboração própria da autora.

A Rede Pública Municipal do Rio de Janeiro (SME-Rio) foi selecionada por ser considerada expressiva dentro do sistema educacional brasileiro, já que é a maior rede pública brasileira e possui uma das menores taxas de aprovação no Ensino Fundamental dentre as capitais brasileiras em 2016, conforme demonstrado no Gráfico 1.

Há ainda outro fato relevante que justifica a escolha da SME-Rio como campo empírico do estudo. O Laboratório de Pesquisa em Oportunidades Educacionais (LaPOpE) da UFRJ possui um acordo com a secretaria de educação da cidade do Rio de Janeiro para acesso e uso de dados administrativos com o objetivo de realizar pesquisas sobre desigualdades educacionais e impacto das principais políticas educacionais implementadas na rede. O acordo prevê devolutivas sistemáticas dos principais resultados para os gestores. Os pesquisadores do Laboratório têm acesso aos dados administrativos do Sistema de Gestão Acadêmica (SGA) da rede. Esses dados são mais ricos e apresentam qualidade superior do que o observado nos dados do Censo Escolar (Inep) ou outras bases disponíveis. Informações sobre o perfil dos alunos, seu desempenho escolar em avaliações próprias da SME-Rio, assim como as características da escola, serão incorporadas ao modelo para tentar explicar o fenômeno da repetência escolar. No entanto, o aspecto mais notável da base de dados é a possibilidade de identificar individualmente alunos. Isso é feito com uma chave única (código de matrícula) que é estável no tempo e bastante confiável. Essa característica da base de dados permite o uso de desenhos de pesquisa mais robustos.

Alguns autores relatam escassez de estudos com dados com características longitudinais. Isso é verdade no Brasil e em estudos internacionais no campo da educação e interfere diretamente na qualidade da evidência produzida (SHIRASU, ARRAES, 2016). A base de dados longitudinal do SGA fornece maior robustez para o desenho porque permite acompanhar individualmente cada aluno ao longo dos anos, observando mudança no tempo dos mesmos indivíduos. A qualidade superior dos dados do SGA pode ser descrita a partir de três pontos fundamentais: a) menor perda dos alunos ao longo do tempo, uma vez que o número de casos duplicados no cadastro (alunos com dois ou mais códigos de identificação) dos alunos é inferior a outras bases de dados nacionais; b) menor proporção de dados faltantes (*missing*); c) mais informações disponíveis sobre as características das famílias e dos resultados educacionais dos alunos. Essas características das bases de dados permitem analisar o fenômeno da repetência escolar com maior rigor e minimizar as ameaças aos resultados reportados. A dissertação irá acompanhar uma coorte de alunos ao longo do

tempo em um desenho longitudinal, diferente da maioria dos estudos sobre reprovação escolar realizados no Brasil que são de pequena escala ou com dados transversais.

O método de análise para observar a mudança no tempo será a análise de sobrevivência. Trata-se de método ainda pouco utilizado na pesquisa educacional e que permite analisar os preditores da repetência escolar, levando em consideração o tempo de ocorrência do fenômeno como uma variável dependente. A ideia é identificar os principais fatores associados ao fenômeno da repetência escolar e gerar informações úteis para que os gestores públicos possam formular planos de ação com objetivo de mitigar o problema.

De forma resumida há dois aspectos que reforçam a relevância da dissertação. O primeiro é a possibilidade de acompanhar uma coorte de alunos ao longo dos anos finais do ensino fundamental e associar a trajetória escolar (estamos aqui especialmente interessados na repetência escolar) com características individuais e do contexto. O segundo elemento é a modelagem utilizada para analisar o fenômeno. A análise de sobrevivência é muito utilizada em outras áreas de intervenção, como por exemplo, pesquisas no campo da saúde pública, mas ainda muito raras na pesquisa educacional. Acreditamos que essa é a modelagem mais adequada para estudar o fenômeno da repetência escolar. Os detalhes sobre o método serão apresentados no capítulo 5.

O estudo irá analisar uma coorte específica de alunos matriculados na rede pública municipal do Rio de Janeiro. Selecionamos alunos matriculados e no 6º ano da rede municipal do Rio de Janeiro em 2012 e avaliados no 1º COC deste ano letivo, que serão acompanhados ao longo de quatro anos, bimestre a bimestre – até a conclusão do ensino fundamental, ao final do ano letivo de 2015.

Assim, este estudo situa-se neste campo de interesse, tendo como foco da análise o estudo das trajetórias dos alunos matriculados no 6º ano do ensino fundamental em 2012. Na análise serão cotejadas as transições escolares relacionadas aos possíveis desfechos escolares ao longo dos quatro anos de acompanhamento, comparando-se semelhanças e diferenças entre os perfis dos alunos, de suas famílias e de políticas públicas implantadas na rede municipal do Rio de Janeiro.

1.2 OBJETIVOS

1.2.1 Objetivo Geral

O objetivo principal da pesquisa é identificar características dos alunos, das famílias e de políticas públicas associadas ao fenômeno de repetência escolar.

1.2.2 Objetivos Específicos

- Descrever a trajetória escolar da coorte estudada;
- Identificar as características dos alunos e suas famílias associadas à repetência escolar;
- Identificar fatores relacionados à trajetória escolar e desempenho acadêmico associado à repetência escolar;
- Identificar políticas públicas (6º Ano Experimental, Ginásio Carioca, Escolas do Amanhã) associadas à repetência escolar;
- Descrever o poder preditivo da Prova Bimestral e o de conceitos do primeiro semestre para estimar o risco de repetência escolar;

2 O CONTEXTO EDUCACIONAL DO BRASIL

Nas últimas décadas, várias mudanças de cunho econômico, político e social ocorreram na sociedade brasileira. A partir dos anos 80 assistiu-se a transição de um regime ditatorial para um regime democrático. Esta transição representou novas conquistas para os cidadãos, tais como: uma nova Constituição fundamentada nos princípios de igualdade e liberdade, o restabelecimento de eleições diretas para presidência da República e a garantia de direitos trabalhistas e sociais (BRASIL, 1988).

Neste contexto de conquistas, a educação formal brasileira também obteve as suas. A Constituição de 1988, segundo Vieira e Farias (2011, p.191), assegurou alguns avanços no âmbito educacional, como:

A consagração da educação como direito público subjetivo (Art. 208, 1º); o princípio da gestão democrática do ensino público (Art. 206, VI); o dever do estado em prover creche e pré-escola às crianças de 0 à 6 anos de idade (Art. 208, IV); a oferta de ensino noturno regular (art. 208, VI); o ensino fundamental obrigatório e gratuito inclusive para os que a ele não tiveram acesso em idade própria (Art. 208, I); o atendimento educacional especializado aos portadores de deficiências (Art. 208, III) (VIEIRA; FARIAS, 2011, p.191).

A Lei de Diretrizes e Bases da Educação nº 9394 (LDB 9394/96) – Lei Darcy Ribeiro – sancionada em dezembro de 1996, foi um marco no âmbito educacional, ratificando as leis constitucionais acima elencadas, bem como, dentre outros aspectos, instituiu uma nova organização para o sistema de ensino brasileiro, estruturando-o em dois grandes níveis de ensino: a educação básica (educação infantil, ensino fundamental, ensino médio e educação de jovens e adultos) e o ensino superior; estabeleceu o regime de colaboração entre a União, os Estados, o Distrito Federal e os Municípios para os respectivos sistemas de ensino; e determinou a elaboração de um Plano Nacional de Educação (PNE) que, em linhas gerais, refere-se a diretrizes e metas para a área educacional a cada dez anos (BRASIL, 1996).

Apesar de a determinação ter sido feita em 1996, o primeiro PNE só foi aprovado quase oito anos após, através da Lei nº 13.005, de 25 de junho de 2014, com vigência entre 2014 e 2024.

O Plano aglutinou um conjunto de aspirações e finalidades para a educação brasileira, que são expressas em suas 10 diretrizes, 20 metas e em um conjunto de 254 estratégias – dimensões específicas do documento que devem ser vistas de forma absolutamente integrada. Enquanto as diretrizes congregam macro-objetivos voltados à melhoria da educação brasileira, as metas delimitam

patamares educacionais concretos que devem ser alcançados no País. Por fim, as estratégias constituem os meios para o planejamento das políticas públicas, que visam, a princípio, o alcance definitivo das metas e, como resultado final do Plano, a consolidação das diretrizes (BRASIL, 2016).

Dentre as metas do PNE (2014-2014), Brasil (2014) destaca que nove delas foram pensadas para “garantia do direito à educação básica com qualidade, que dizem respeito ao acesso, à universalização da alfabetização e à ampliação da escolaridade e das oportunidades educacionais”, são elas:

- **Meta 1:** universalizar, até 2016, a educação infantil na pré-escola para as crianças de 4 (quatro) a 5 (cinco) anos de idade e ampliar a oferta de educação infantil em creches, de forma a atender, no mínimo, 50% (cinquenta por cento) das crianças de até 3 (três) anos até o final da vigência deste PNE;
- **Meta 2:** universalizar o ensino fundamental de 9 (nove) anos para toda a população de 6 (seis) a 14 (quatorze) anos e garantir que pelo menos 95% (noventa e cinco por cento) dos alunos concluam essa etapa na idade recomendada, até o último ano de vigência deste PNE;
- **Meta 3:** universalizar, até 2016, o atendimento escolar para toda a população de 15 (quinze) a 17 (dezessete) anos e elevar, até o final do período de vigência deste PNE, a taxa líquida de matrículas no ensino médio para 85% (oitenta e cinco por cento);
- **Meta 5:** alfabetizar todas as crianças, no máximo, até o final do 3º (terceiro) ano do ensino fundamental;
- **Meta 6:** oferecer educação em tempo integral em, no mínimo, 50% (cinquenta por cento) das escolas públicas, de forma a atender, pelo menos, 25% (vinte e cinco por cento) dos(as) alunos(as) da educação básica;
- **Meta 7:** fomentar a qualidade da educação básica em todas as etapas e modalidades, com melhoria do fluxo escolar e da aprendizagem, de modo a atingir as seguintes médias nacionais para o Ideb: 6,0 nos anos iniciais do ensino fundamental; 5,5 nos anos finais do ensino fundamental; 5,2 no ensino médio;
- **Meta 9:** elevar a taxa de alfabetização da população com 15 (quinze) anos ou mais para 93,5% (noventa e três inteiros e cinco décimos por cento) até 2015 e, até o final da vigência deste PNE, erradicar o analfabetismo absoluto e reduzir em 50% (cinquenta por cento) a taxa de analfabetismo funcional;

- **Meta 10:** oferecer, no mínimo, 25% (vinte e cinco por cento) das matrículas de educação de jovens e adultos, nos ensinos fundamental e médio, na forma integrada à educação profissional;
- **Meta 11:** triplicar as matrículas da educação profissional técnica de nível médio, assegurando a qualidade da oferta e pelo menos 50% (cinquenta por cento) da expansão no segmento público.

Dentre as metas supracitadas, a meta 2 é a que está mais relacionada a este estudo e por isso merece um detalhamento, pois explicita a importância não apenas do acesso ao ensino fundamental, mas a conclusão desta etapa de ensino na idade recomendada.

Para monitorar esta meta, dois indicadores são acompanhados: Indicador 2A – Percentual de pessoas de 6 a 14 anos que frequentam ou já que concluíram o ensino fundamental³ e Indicador 2B – Percentual de pessoas de 16 anos com pelo menos o ensino fundamental concluído, e ambos se utilizam de dados da Pnad/IBGE⁴, já que os dados do Censo Escolar não traz informações de pessoas que não estejam matriculadas em alguma escola (BRASIL, 2016).

No ‘Relatório do 2º ciclo de monitoramento das metas do PNE – 2018, o INEP/MEC traz a série histórica de tais indicadores no período de 2004 a 2015 calculados a partir da Pnad/IBGE e acrescenta uma nova série histórica no período de 2012 a 2017, onde tais indicadores foram calculados a partir da Pnad-c⁵. A justificativa da instituição para a troca da origem dos dados é que essa nova pesquisa amplia a representatividade da região Norte e das capitais brasileiras, entretanto, ela faz um alerta que, por serem duas pesquisas distintas, com desenhos amostrais específicos, as comparações dos indicadores de um ano para o outro devem ser feitas com cautela (BRASIL, 2018, p. 39).

No Gráfico 2, é possível acompanhar as duas séries históricas do Indicador 2A (Percentual de pessoas de 6 a 14 anos que frequentam ou já que concluíram o EF), a partir

³ Para melhorar a precisão do Indicador 2A, optou-se em utilizar a taxa de escolarização líquida ajustada, pois é mais específica para o ensino fundamental e leva em consideração também, as pessoas de 6 a 14 anos que já concluíram esta etapa de ensino. (BRASIL, 2016).

⁴ A Pesquisa Nacional por Amostra de Domicílios (Pnad) é realizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE) e obtém informações anuais sobre características demográficas e socioeconômicas da população

⁵ “A Pesquisa Nacional por Amostra de Domicílios Contínua (Pnad-c) é uma pesquisa por amostra probabilística de domicílios de abrangência nacional. [...] A PNAD Contínua segue um esquema de rotação de domicílios. Isso significa que cada domicílio selecionado será entrevistado cinco vezes, uma vez a cada trimestre, durante cinco trimestres consecutivos.” (IBGE, 2018). Para conhecer melhor acesse:

<https://ww2.ibge.gov.br/home/estatistica/indicadores/trabalhoerendimento/pnad_continua/default.shtm>

de 2004, conforme explicado no parágrafo anterior, onde nota-se que “há duas tendências: a primeira, de crescimento mais rápido, ocorreu entre 2004 e 2012; a segunda, de crescimento menos vigoroso, ocorreu a partir de 2012”, mostrando que o ensino fundamental está praticamente universalizado para as crianças de 6 a 14 anos. É pertinente ainda o esclarecimento de que entre os anos de 2004 e 2006, não havia ainda a obrigatoriedade de que as pessoas com seis anos estivessem matriculadas no ensino fundamental⁶, podendo essa a explicação para o crescimento mais acentuado do primeiro momento (BRASIL, 2018, p.42).

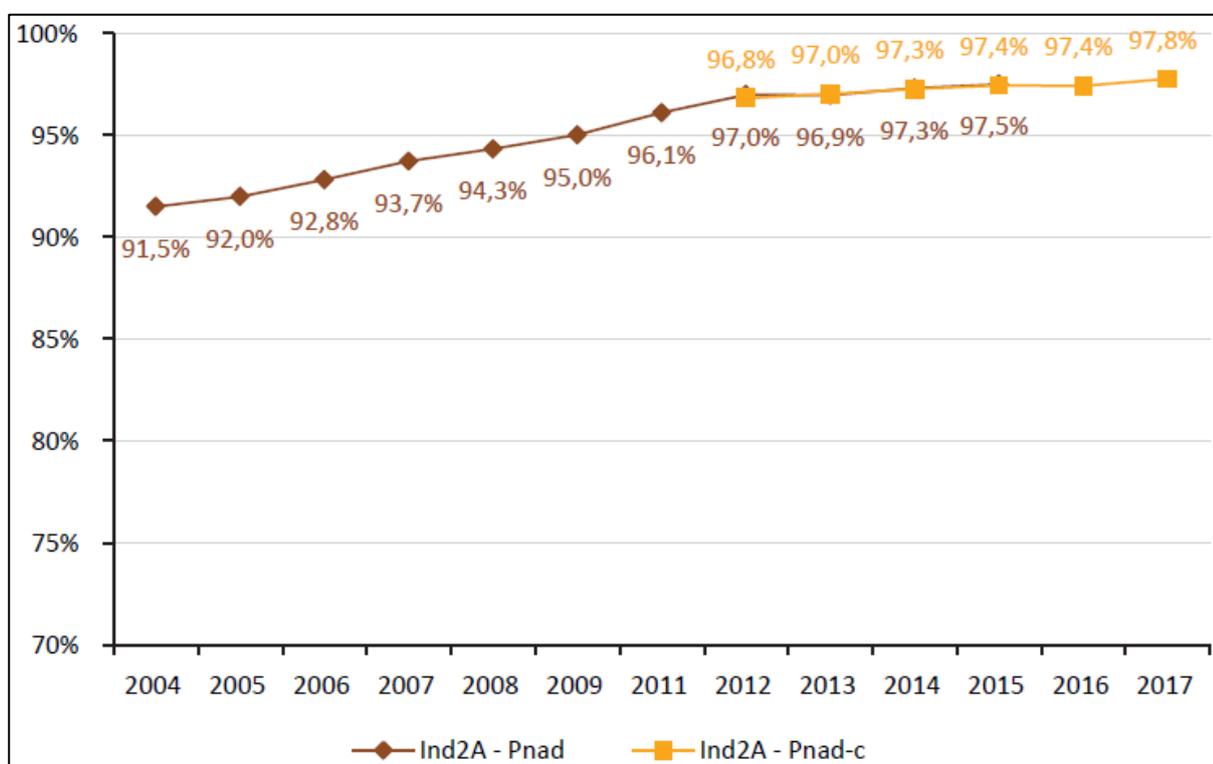


Gráfico 2: Indicador 2A – Percentual de pessoas de 6 a 14 anos que frequentam ou já que concluíram o EF (Meta 2 - PNE 2014-2024). Fonte: Brasil, 2018, p. 41.

Diante da diversidade brasileira e a fim de situar-se mais próximo à realidade do presente estudo, o Gráfico 3 ilustra os valores do Indicador 2A (Percentual de pessoas de 6 a 14 anos que frequentam ou já que concluíram o EF) para o Brasil, para a região Sudeste e para o estado do Rio de Janeiro⁷, no período de 2012 a 2017, calculados a partir da Pnad-c.

⁶ Tal obrigatoriedade se deu através da Lei nº 11.274, de 6 de fevereiro de 2006, começando sua implementação a partir de 2007.

⁷ Não há dados dos indicadores específicos para o município do Rio de Janeiro (apenas para a região metropolitana do Rio de Janeiro), por isso, optamos em trazer apenas a nível estadual.

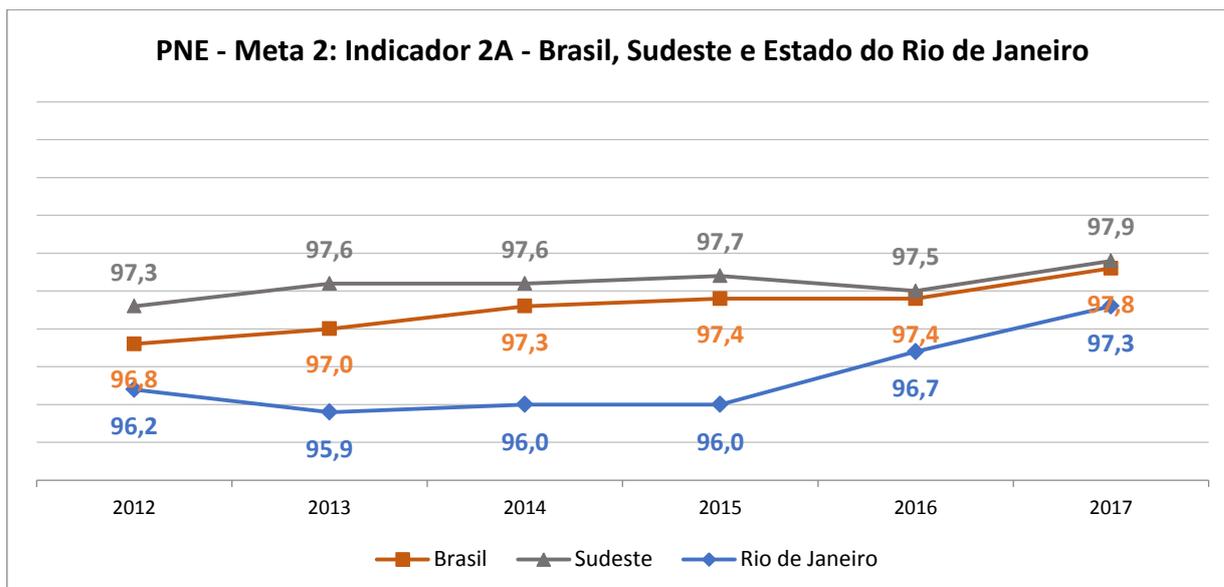


Gráfico 3: Indicador 2A – Percentual de pessoas de 6 a 14 anos que frequentam ou já que concluíram o EF (Meta 2 PNE 2014-2024) - Brasil, Sudeste, Estado do Rio de Janeiro. Fonte: Brasil, 2018, p.42. Elaboração própria

No Gráfico 3, se olharmos apenas os três últimos anos, observamos a curva mais íngreme para o estado do Rio de Janeiro⁸, quando comparado aos outros, isto é, a variação de 1,3 pontos percentuais para o estado do Rio de Janeiro é maior do que a variação de 0,2 ponto percentual para o Sudeste e de 0,4 ponto percentual para o Brasil. Se ampliarmos o olhar para toda a série histórica, ainda temos o estado do Rio de Janeiro com maior variação (1,1 pontos percentuais) do que o Sudeste (0,6 ponto percentual) e o Brasil (1,0 ponto percentual).

Apesar de o indicador demonstrar uma “quase” universalidade do ensino fundamental, é possível notar uma estabilidade nos valores do indicador, sugerindo uma dificuldade em atingir a meta de 100% em 2024 e fazendo-nos refletir sobre quem são essas cerca de 600 mil crianças de 6 a 14 anos que ainda se encontram fora do ensino fundamental, Cruz (2017) afirma que esse fato é preocupante, porque justamente as crianças que mais precisam são as que estão fora da escola, sendo geralmente as deficientes, as mais pobres e as que moram em lugares distantes.

Se o atingimento da meta do Indicador 2A (Percentual de pessoas de 6 a 14 anos que frequentam ou já que concluíram o EF) já é preocupante, a do Indicador 2B (Percentual de pessoas de 16 anos com pelo menos o EF concluído) é muito mais crítica. No Gráfico 4,

⁸ Vale lembrar aqui ainda, que a fonte de dados do indicador é a Pnad-c, que é uma pesquisa amostral e, portanto, não retrata a realidade exata dos dados.

podemos observar a série histórica do Indicador 2B (Percentual de pessoas de 16 anos com pelo menos o EF concluído) de 2004 a 2017, e, assim como a do indicador anterior, está dividido em duas linhas, por terem dados de diferentes pesquisas.

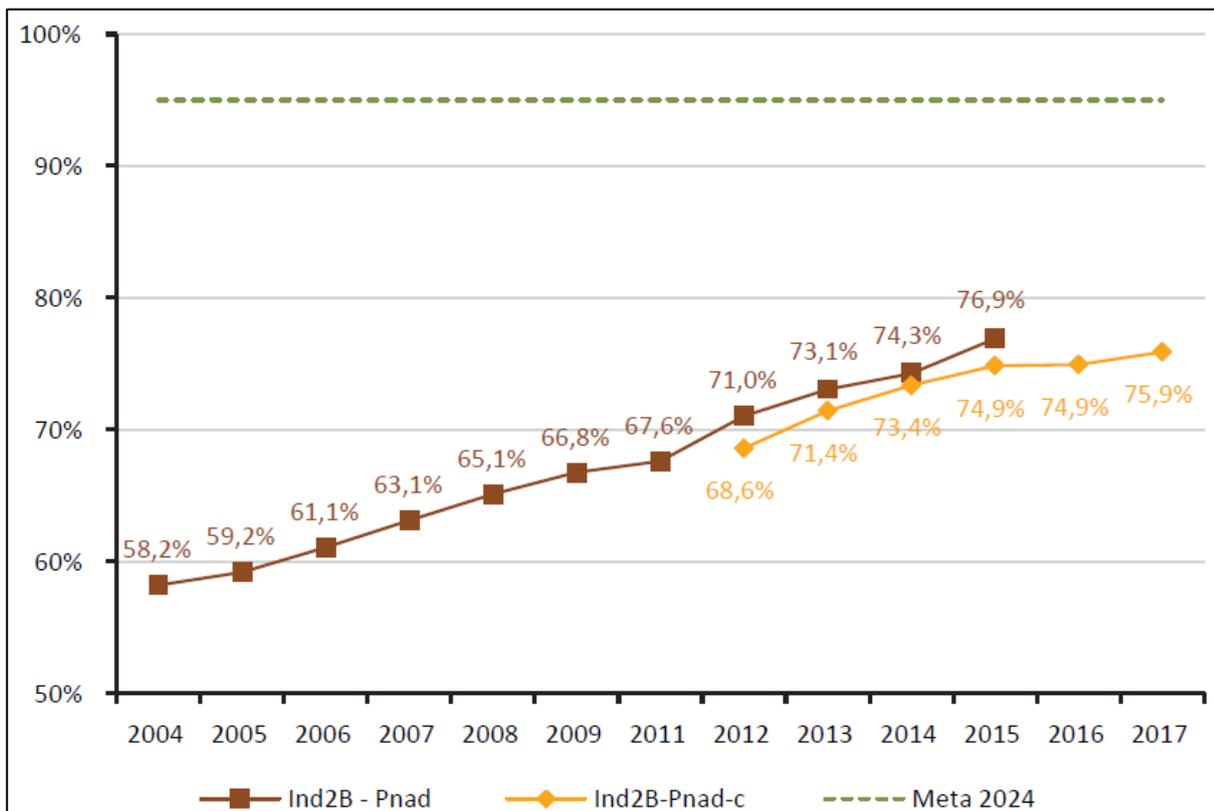


Gráfico 4: Indicador 2B - Percentual de pessoas de 16 anos com pelo menos o EF concluído (Meta 2 - PNE 2014-2024). Fonte: Brasil, 2018, p. 45

Observando o Gráfico 4, segundo a Pnad, no período de 2004 a 2015, o indicador 2B (Percentual de pessoas de 16 anos com pelo menos o EF concluído) aumentou 18,7 pontos percentuais, já segundo a Pnad-c, no período de 2012 a 2017, o indicador aumentou 7,3 pontos percentuais, segundo Brasil (2018, p. 44), a variação média observada no período da Pnad-c (de 1,5 pontos percentuais ao ano), é insuficiente para que a meta de 95% seja alcançada até 2024. Seguindo esse ritmo, a meta só seria atingida em 2030.

Seguindo o mesmo raciocínio de retratar melhor o contexto deste estudo, o Gráfico 5 demonstra os valores do Indicador 2B (Percentual de pessoas de 16 anos com pelo menos o EF concluído) para o Brasil, a região Sudeste e o Estado do Rio de Janeiro. Como é possível observar, neste indicador o estado do Rio de Janeiro também está abaixo do Brasil e da região Sudeste, sugerindo que no estado do Rio de Janeiro menos alunos concluem o ensino fundamental na idade recomendada (16 anos), quando comparado ao Brasil e ao Sudeste.

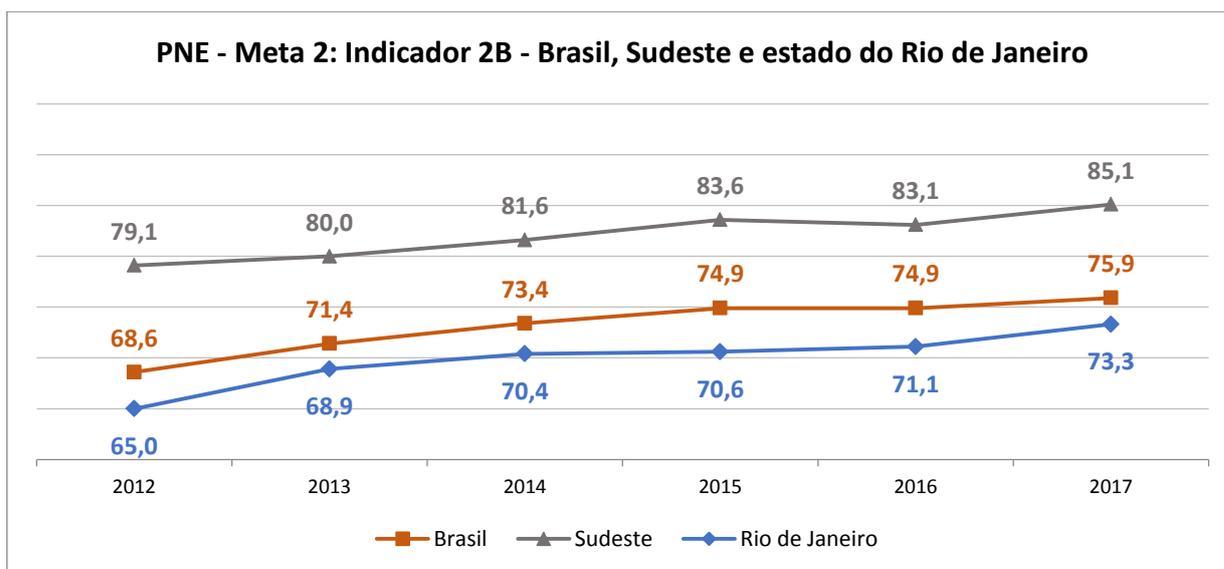


Gráfico 5: Indicador 2B - Percentual de pessoas de 16 anos com pelo menos o EF concluído (Meta 2 - PNE 2014-2024) - Brasil, Sudeste, Estado do Rio de Janeiro. Fonte: Observatório do PNE (IBGE/Pnad), 2018. Elaboração própria.

Brasil (2016) esclarece que este indicador resume várias outras situações no que diz respeito ao acesso, permanência e conclusão do ensino fundamental, “ou seja, os resultados deste indicador decorrem do acesso a essa etapa de ensino na idade recomendada, além das taxas de rendimento e de fluxo escolar”. Portanto, ao analisar o Indicador 2B, sugere-se cautela em sua inferência, pois um problema em um ou mais desses fatores, causa a não evolução do indicador, sendo necessário um olhar mais minucioso de cada um deles individualmente⁹.

Ainda assim, é válido apresentar os resultados deste indicador para todas as regiões e estados do Brasil, como podemos ver no Quadro 1, e através dele, deduzir que a conclusão do ensino fundamental na idade adequada é problema muito sério no país e merece atenção dos gestores com políticas públicas eficazes em todas as facetas que influenciam esse indicador, se não, no futuro, teremos uma escola onde apenas “depositamos” os alunos e não chegamos a lugar nenhum, deixando para trás o propósito principal da escola: ensinar. E, segundo Tavares Jr. *et al.* (2012), pode-se até gerar resultados ainda mais perversos do que a conjuntura anterior; ficando assim cada vez mais distante de um sistema educacional ideal, com garantia de aprendizado e fluxo escolar adequado a todos os alunos matriculados, seja na rede pública ou privado, conforme define Fernandes (2007).

⁹ Trataremos as taxas de fluxo escolar de forma mais detalhada no capítulo seguinte desta dissertação.

Quadro 1: Indicador 2B - Percentual de pessoas de 16 anos com pelo menos o EF concluído – Brasil, regiões e Estados. Fonte: Brasil, 2018, p. 46-47

Brasil/Região/UF	2012	2013	2014	2015	2016	2017	Varição 2012 - 2017 (p.p.)
Brasil	68,6%	71,4%	73,4%	74,9%	74,9%	75,9%	7,3
Norte	54,0%	59,4%	62,0%	66,8%	64,9%	68,4%	14,4
Rondônia	61,2%	70,5%	64,4%	73,4%	67,4%	72,5%	11,3
Acre	70,1%	71,6%	71,3%	77,0%	70,4%	75,6%	5,5
Amazonas	54,6%	58,7%	64,8%	66,5%	68,2%	77,4%	22,8
Roraima	79,4%	75,2%	75,5%	78,1%	75,5%	78,3%	-1,1
Pará	46,4%	53,1%	56,1%	61,0%	58,3%	62,9%	16,5
Amapá	67,4%	64,5%	68,8%	74,2%	75,9%	53,9%	-13,5
Tocantins	61,8%	70,2%	71,2%	78,8%	74,3%	74,6%	12,8
Nordeste	56,1%	60,5%	62,7%	63,8%	65,2%	66,2%	10,1
Maranhão	59,0%	64,5%	67,3%	68,3%	70,0%	66,8%	7,8
Piauí	66,4%	61,9%	59,4%	60,3%	60,9%	63,9%	-2,6
Ceará	69,0%	69,4%	71,7%	72,5%	77,1%	76,2%	7,2
Rio Grande do Norte	53,9%	61,0%	63,0%	63,4%	67,8%	68,7%	14,8
Paraíba	56,0%	60,0%	59,1%	57,4%	62,3%	64,4%	8,4
Pernambuco	59,7%	62,3%	67,7%	67,7%	70,6%	66,4%	6,8
Alagoas	47,7%	50,6%	54,9%	58,2%	60,0%	67,3%	19,6
Sergipe	43,2%	57,8%	52,2%	55,8%	58,3%	55,0%	11,8
Bahia	46,6%	54,6%	56,3%	59,3%	56,7%	61,7%	15,1
Sudeste	79,1%	80,0%	81,6%	83,6%	83,1%	85,1%	6,0
Minas Gerais	76,2%	75,3%	79,9%	78,3%	78,6%	83,6%	7,5
Espírito Santo	74,0%	71,7%	73,0%	72,8%	71,5%	68,7%	-5,3
Rio de Janeiro	65,0%	68,9%	70,4%	70,6%	71,1%	73,3%	8,3
São Paulo	86,7%	87,8%	88,2%	91,9%	90,3%	91,2%	4,5
Sul	74,4%	79,6%	79,8%	80,4%	79,2%	77,2%	2,7
Paraná	73,8%	78,5%	79,0%	81,5%	78,4%	78,3%	4,5
Santa Catarina	80,8%	87,0%	89,7%	89,3%	86,2%	82,7%	2,0
Rio Grande do Sul	71,4%	75,8%	74,6%	74,5%	75,5%	72,9%	1,5
Centro-Oeste	75,4%	76,4%	77,4%	76,1%	78,2%	77,1%	1,7
Mato Grosso do Sul	64,6%	67,6%	65,8%	66,4%	62,5%	71,3%	6,7
Mato Grosso	79,7%	79,6%	89,3%	84,9%	92,3%	88,6%	8,9
Goiás	76,9%	78,8%	76,9%	78,5%	78,1%	76,3%	-0,6
Distrito Federal	76,3%	75,1%	75,3%	69,5%	77,6%	72,3%	-4,0

E para finalizar essa parte, a crítica ao PNE apresentada por Schwartzman (2016) é trazida como alerta para os gestores educacionais, demonstrando que para alcançar o propósito principal do PNE¹⁰ não basta fazer uma “enorme lista de papai noel”, transcrever os desejos em formato de lei e simplesmente aumentar os gastos, porque isso por si só, agoura o PNE se tornar uma grande frustração, pois:

“A melhoria da educação depende de diagnósticos claros, baseados em evidências, das causas dos problemas, e políticas consistentes e de longo prazo para, entre outras coisas, melhorar o gerenciamento das redes e escolas públicas, assegurar a qualidade e a formação dos professores (que é muito diferente da simples titulação), introduzir currículos modernos e adequados, garantir a alfabetização das crianças no primeiro ano da educação básica e criar alternativas de formação no ensino médio; e nem sempre os interesses da educação do país e das corporações que habitam o setor educacional coincidem” (SCHWARTZMAN, 2016).

¹⁰ Que é consolidar um sistema educacional capaz de concretizar o direito à educação em sua integralidade, rompendo as barreiras de acesso e permanência, mitigando as desigualdades e garantindo a formação para o trabalho e exercício da cidadania (BRASIL, 2015, p. 9)

3 FLUXO ESCOLAR

A palavra fluxo pode ser definida, segundo o dicionário online de Michaelis (2018)¹¹, como “o ato ou efeito de fluir” ou ainda “escoamento ou movimento contínuo que segue um curso, uma corrente”.

Trazer tal conceito para a área da educação parece ser algo bastante intuitivo e de fácil compreensão e, apesar de muito utilizado na área, houve uma grande dificuldade para encontrar a definição formal do termo ‘fluxo escolar’, tendo sido encontrado apenas num glossário geral que trouxe a seguinte definição: “trata-se da análise do comportamento da progressão dos alunos pertencentes a uma coorte, em determinado nível de ensino seriado, em relação à sua condição de promovido, repetente ou evadido” (FULGENCIO, 2007, p. 233).

Portanto, ao falarmos de fluxo escolar, estamos tratando de transições (ou não) dentro do sistema de ensino – que é progressivo série a série – podendo a trajetória ser considerada regular (quando há promoção ano a ano, sem repetências ou evasões) ou não.

O interesse sobre fluxo escolar é uma preocupação antiga em nossa história, tendo sido estudado por vários pesquisadores, como exemplo: Anísio Teixeira de Freitas¹², Zaia Brandão, Philip Fletcher, Sérgio Costa Ribeiro, Ruben Klein e Fernando Tavares Junior.

Teixeira de Freitas (1940a, 1940b), um dos pioneiros no assunto, a partir dos dados do Censo Escolar da década de 1930, trouxe como alerta os altos índices de repetência e a evasão para o ensino fundamental da época, que era composto por três séries apenas e o autor (1947) encontrou incompatibilidade entre as estatísticas educacionais de fluxo e os dados demográficos.

Em 1975, Schiefelbein encontrou problemas nas estatísticas educacionais oficiais em praticamente todos os países da América Latina.

Nos anos 80, o estudo de Fletcher (1985) indicou evidências de que a estatística escolar subestimava as taxas de repetência e sobrestimava as taxas de evasão no ensino do 1º grau, propondo um novo modelo matemático para cálculo do fluxo escolar.

O Profluxo, um modelo matemático foi proposto como um procedimento alternativo para analisar o fluxo dos alunos num sistema educacional seriado, onde as matrículas em um

¹¹ Disponível em: <<http://michaelis.uol.com.br/busca?id=oOPY>>. Acesso em 13 Jun. 2018

¹² Leia mais sobre o autor em <<https://memoria.ibge.gov.br/sinteses-historicas/pioneiros-do-ibge/mario-augusto-teixeira-de-freitas.html>>. Acesso em 13 Jun. 2018.

determinado ano são baseadas em matrículas na série do ano anterior e em coeficientes que descrevem os fluxos dos alunos entre as séries de um ano para o ano seguinte (FLETCHER, RIBEIRO, 1989, p.4).

Foi a partir daí é que se teve certeza de que a evasão não era o maior problema educacional do Brasil, mas sim a repetência, surgindo assim vários estudos importantes sobre o tema, tais como: a “Pedagogia da Repetência” de Sergio Costa Ribeiro (1991), “O Censo educacional e o modelo de fluxo: o problema da repetência” (KLEIN, RIBEIRO, 1991) e “A Pedagogia da Repetência ao longo das décadas” (KLEIN, RIBEIRO, 1995). Fazendo assim com que o tema ficasse em destaque nos debates educacionais e recebendo políticas públicas para tentativa de solução desse problema.

Em 2002, através da publicação “Geografia da Educação Brasileira 2001”, o INEP/MEC define fluxo de uma forma completamente distinta e esclarece que o número de alunos promovidos, repetentes e evadidos é estimado de acordo com o modelo de fluxo escolar “Análise e Projeções de Matrícula nos Países em Desenvolvimento” (Tore Thonstad, Paris, 1986) e corrigido para o sistema educacional brasileiro (Klein, Rio de Janeiro, 1995).

Com o passar do tempo, o tema não esmaeceu e continuou recebendo novas análises e novas propostas de metodologias de cálculos (KLEIN, 2003; GOLGHER, 2004; GOLGHER, RIOS-NETO, 2005; RIOS-NETO *et al.*, 2005; SOARES, 2002).

No “Dicionário de Indicadores Educacionais: Fórmulas de cálculo” divulgado também pelo INEP/MEC (2004, p.20), as taxas de fluxo são definidas como sendo o percentual de alunos promovidos, alunos repetentes e alunos evadidos, e que seriam calculadas anualmente, tendo como fórmulas para cálculo:

$$TP_{sm} = \frac{Prom_{s+1,m+1} \times 100}{M_{sm}}$$

$$TR_{sm} = \frac{Rept_{s,m+1} \times 100}{M_{sm}}$$

$$TEV_{sm} = 100 - (TP_{sm} + TR_{sm})$$

Onde,

TP_{sm} = taxa de promoção na série s no ano m ;

TR_{sm} = taxa de repetência na série s no ano m ;

TEV_{sm} = taxa de evasão na série s no ano m ;

$Prom_{s+1,m+1}$ = número de promovidos, no fluxo, para a série $s+1$ no ano $m+1$;

$Rept_{s,m+1}$ = repetentes na série s no ano $m+1$;

M_{sm} = matrícula inicial na série s no ano m ;

m = ano;

s = ensino fundamental (série, 1ª a 4ª, 5ª a 8ª, total), ensino médio (série, total).

Apesar de o indicador ter periodicidade anual, o MEC/INEP ficou um longo período sem divulgá-lo, retornando apenas em 2017, quando divulgou todas as taxas para o período 2007/2008, 2008/2009, 2010/2011, 2012/2013, 2013/2014 e 2014/2015, para o Brasil, as regiões, as Unidades da Federação e os Municípios. A novidade para esta publicação está na distinção entre os alunos evadidos e aqueles que migraram para a Educação de Jovens e Adultos¹³ (EJA). O INEP/MEC ainda explica que esse trabalho longitudinal de acompanhamento do aluno só foi possível devido às mudanças da metodologia do Censo da Educação Básica em 2007, onde as informações passaram a ser individuais.

Antes de prosseguir, é importante esclarecer a diferença entre as taxas de rendimento (aprovação, reprovação e abandono) e as taxas de fluxo, pois comumente são confundidas ou utilizadas como sinônimos. As taxas de rendimento se referem ao resultado final dos alunos ao final de cada ano letivo (situação final do aluno declarada na segunda etapa da coleta do Censo Escolar), isto é, se foram aprovados ou reprovados ou abandonaram durante o ano letivo. As taxas de fluxo mensuram a transição do aluno entre dois anos consecutivos.

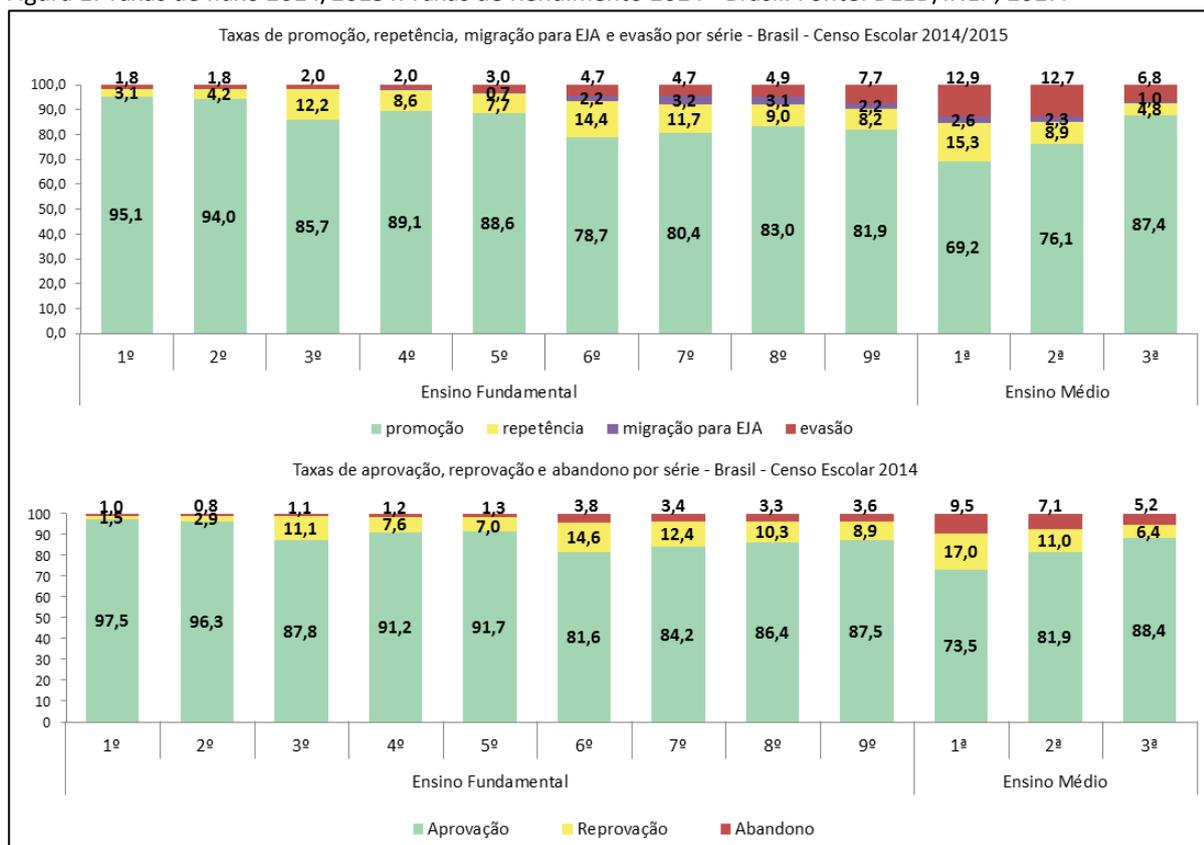
É claro que se espera que haja uma congruência entre elas, por exemplo, que os alunos aprovados sejam promovidos e os reprovados sejam repetentes, entretanto não é incomum encontrar alunos aprovados-repetentes, reprovados-promovidos, abandono-

¹³ Evitando assim a superestimação das taxas de evasões, já que quando o aluno migra para o EJA, ele não deixa o sistema educacional, ele apenas muda de modalidade de ensino.

promovido, etc. Por essas peculiaridades, cada taxa é única, tem seus próprios objetivos e mede fenômenos diferentes, não podendo ser usadas como sinônimos¹⁴.

Para ilustrar a diferença entre as taxas de fluxo e de rendimento, na Figura 1 é possível observar as taxas de fluxo para a transição 2014/2015 (acima) e as taxas de rendimento de 2014 do Brasil (embaixo).

Figura 1: Taxas de fluxo 2014/2015 x Taxas de Rendimento 2014 - Brasil. Fonte: DEED/INEP, 2017.



A seguir, apresentaremos as taxas de fluxo do ensino fundamental das três últimas transições (2012/2013, 2013/2014 e 2014/2015) divulgadas pelo INEP/MEC, para o Brasil, o Sudeste, o estado do Rio de Janeiro e o município do Rio de Janeiro. Estas taxas são compostas por dados de todas as escolas, sejam privadas ou públicas de qualquer natureza (federal, estadual e municipal). Então, para se chegar o mais próximo possível do contexto deste estudo, optamos também por trazer as taxas de fluxo para as escolas públicas localizadas no município do Rio de Janeiro, tais dados não são exclusivos SME-Rio, mas podemos afirmar que as taxas da SME-Rio são muito próximas a estas, pois a rede é

¹⁴ Por isso, aqui neste estudo nosso foco é a repetência, pois estamos interessados em trajetórias/transição e não em rendimento.

responsável pela grande maioria dos alunos do ensino fundamental da rede pública do município¹⁵.

No Gráfico 6 e Gráfico 7, as taxas de fluxo do Brasil e do Sudeste respectivamente, podemos observar pequenas variações nas taxas entre os anos, mostrando um mesmo comportamento de evolução (a taxa de promoção aumenta entre a primeira e a segunda transição e diminui na última, mas ainda assim estando acima da primeira transição). Um comportamento parecido é encontrado no Gráfico 8 (Estado do Rio de Janeiro), entretanto a queda para a última transição é maior, fazendo que seja mais baixa do que a da primeira transição apresentada. Já para o município do Rio de Janeiro, nos Gráfico 9 e Gráfico 10, as taxas de promoção estão em declínio ano após ano e as taxas de repetência estão aumentando para o mesmo período. Um fato que não pode passar despercebido é a estabilidade das taxas de migração para o EJA (não variando mais que 0,2 pontos percentuais) e das taxas de evasão, onde se repetem em praticamente todas as regiões entre os dois primeiros anos e sobem no último ano, de forma mais intensa nos dois resultados do Rio de Janeiro.

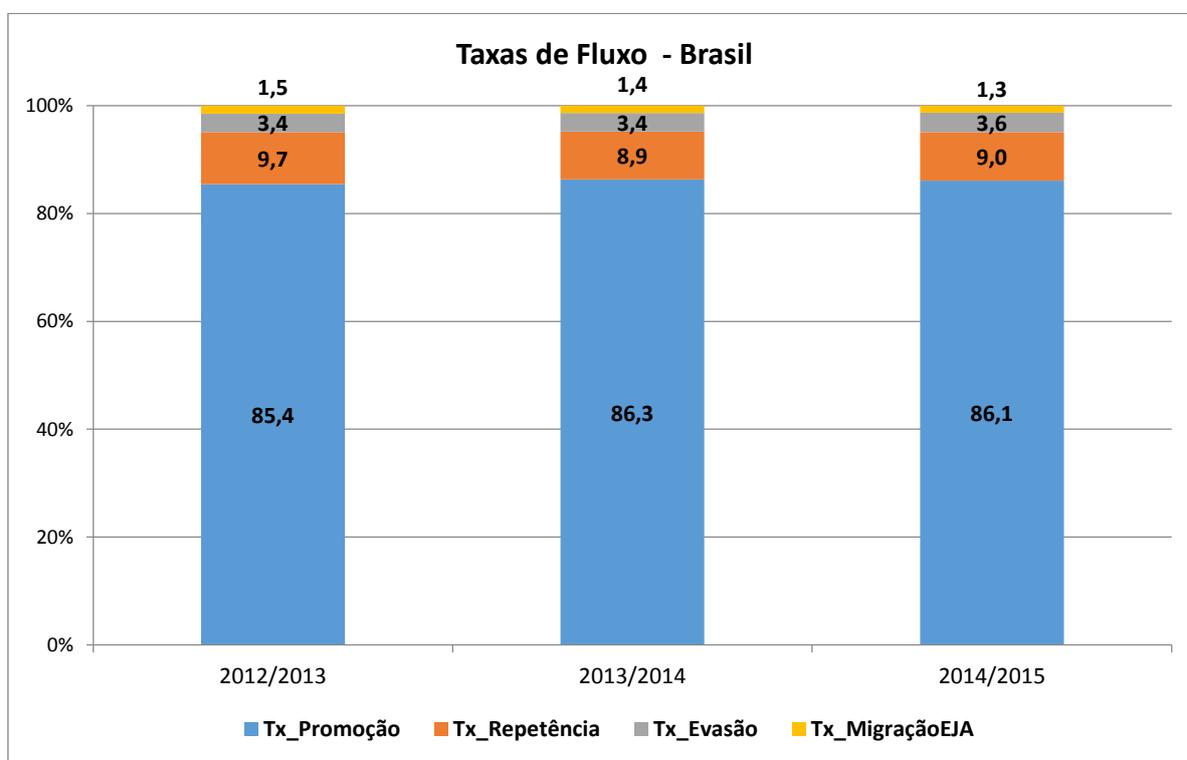


Gráfico 6: Taxas de Fluxo - Brasil. Fonte: INEP, 2017. Elaboração própria.

¹⁵ Segundo os dados da Sinopse Estatística da Educação Básica 2015, a rede municipal do Rio de Janeiro possui 96,95% dos alunos da rede pública do município do Rio de Janeiro (Rede Pública = 496.564 alunos, sendo 10.235 (2,06%) alunos da rede federal, 4.919 (0,99%) alunos da rede estadual e 481.410 (96,95%) alunos da rede municipal) (INEP, 2016).

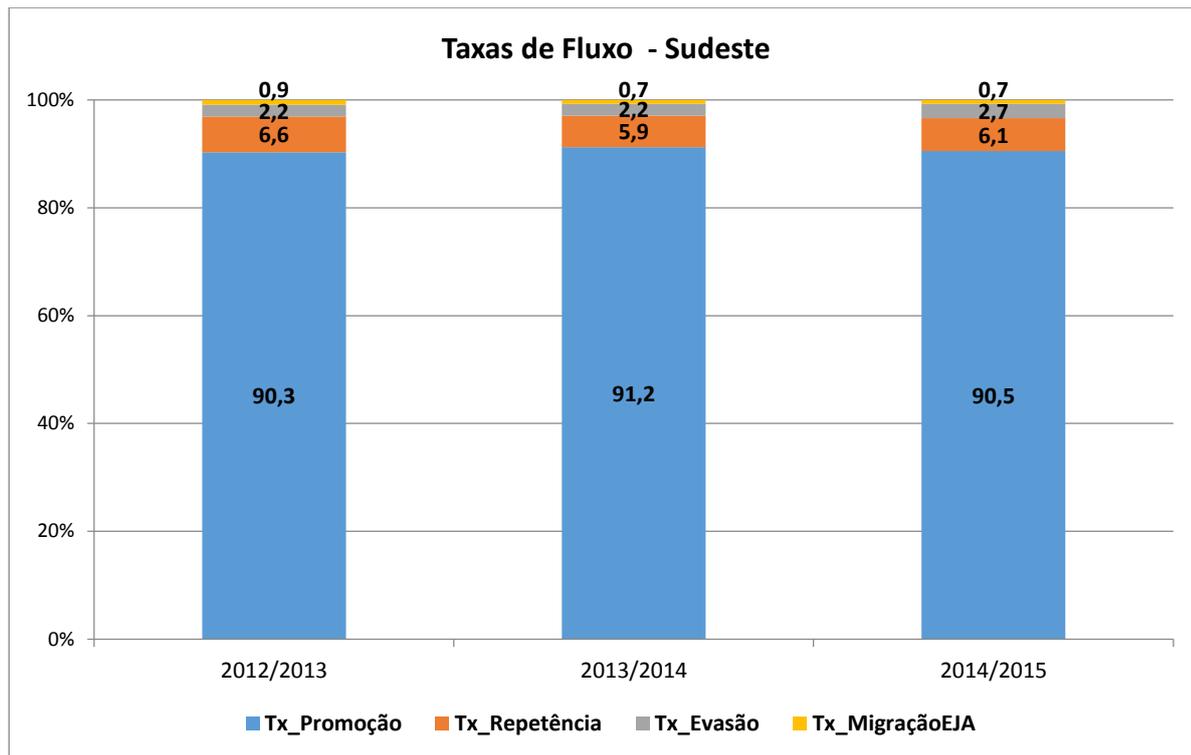


Gráfico 7: Taxas de Fluxo – Região Sudeste. Fonte: INEP, 2017. Elaboração própria.

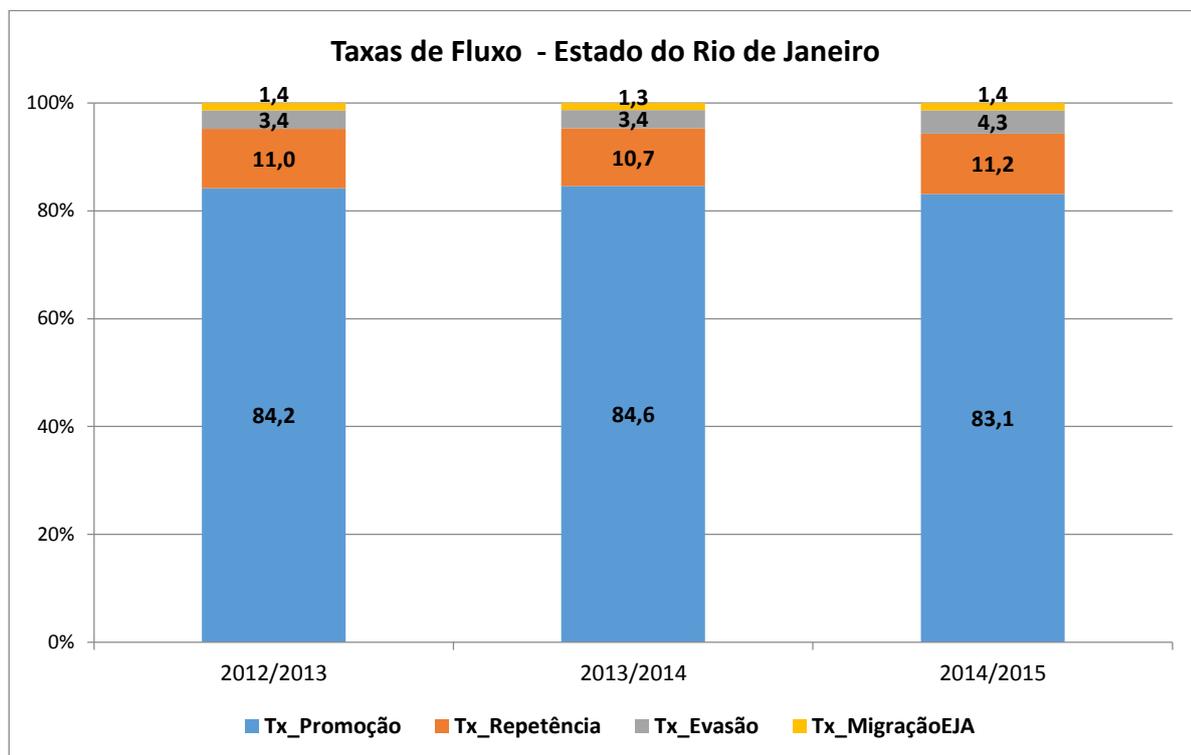


Gráfico 8: Taxas de Fluxo – Estado do Rio de Janeiro. Fonte: INEP, 2017. Elaboração própria.

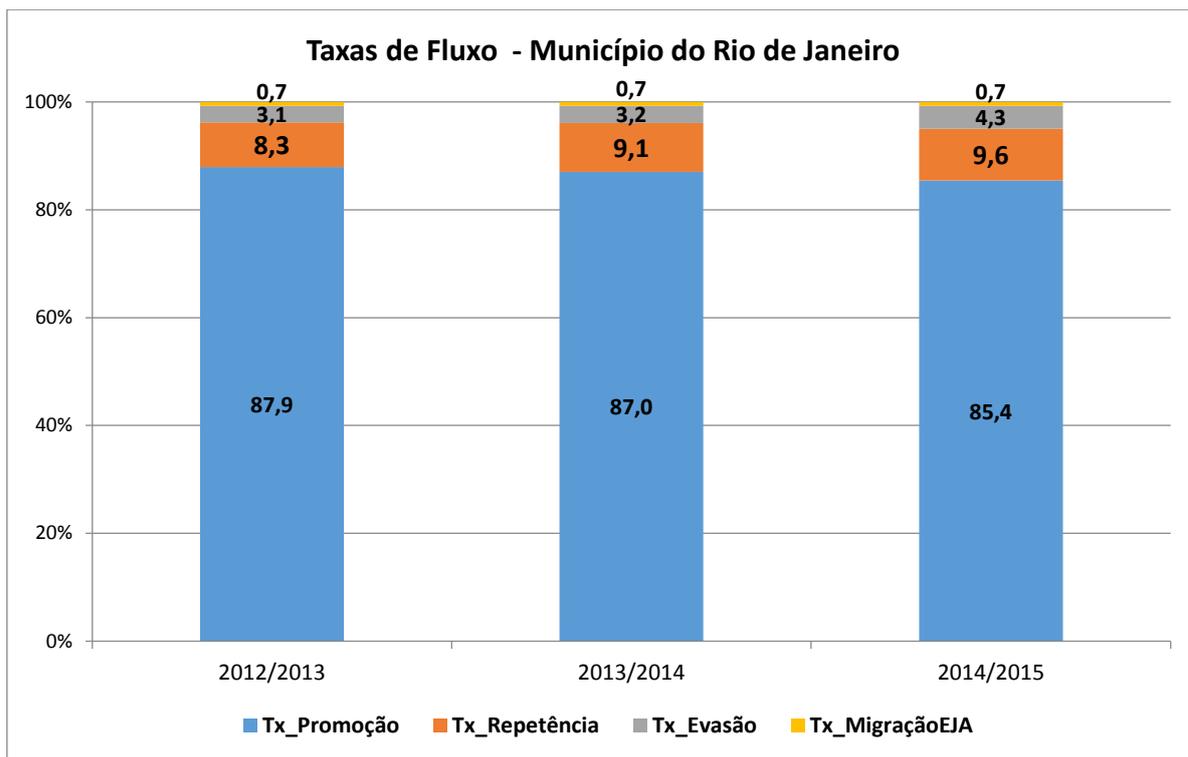


Gráfico 9: Taxas de Fluxo – Município do Rio de Janeiro. Fonte: INEP, 2017. Elaboração própria

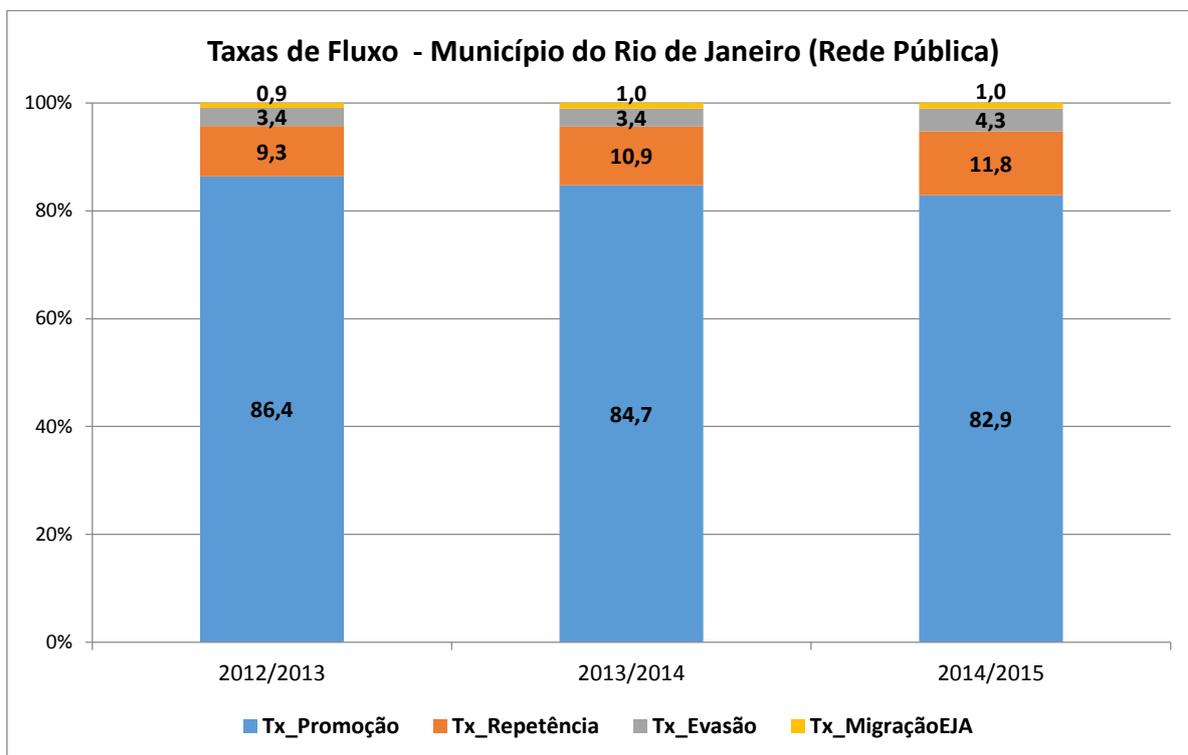


Gráfico 10: Taxas de Fluxo – Município do Rio de Janeiro (Escolas Públicas). Fonte: INEP, 2017. Elaboração própria.

3.1 FATORES ASSOCIADOS À REPETÊNCIA ESCOLAR

Já vimos que reprovação e repetência escolar possuem conceitos distintos, entretanto, neste momento, não vamos nos ater aos conceitos técnicos e utilizaremos o termo repetência para tratar de ambas as situações, pois o objetivo principal de agora é fazer uma apanhado sobre o que a literatura nos informa sobre tais fenômenos.

A repetência é um problema antigo e de magnitude mundial, por isso é recorrente nas pesquisas nacionais e internacionais há algum tempo. Por ser um fenômeno complexo e multifacetado, vem sendo abordado por diferentes prismas, onde os pesquisadores buscam evidências de melhora no desempenho acadêmico ou ainda se causa algum tipo de alteração psicológica dos alunos repetentes, entre vários outros.

Historicamente, na grande maioria dos estudos, os resultados apontam para mais desvantagens (prejuízos) do que vantagens (benefícios) quando falamos de repetência.

Jackson (1975) sugeriu em seu estudo que a repetência poderia até trazer benefícios para os alunos, mas que a promoção aparentava gerar muito mais benefícios quando comparado à repetência.

Holmes e Matthews (1984) reuniram 44 estudos sobre o tema que comparavam grupos de alunos retidos com promovidos, e através da meta-análise¹⁶ chegaram à mesma conclusão de autor supracitado, ressaltando ainda que potenciais efeitos negativos consistentemente superam os resultados positivos.

Jimerson (2001), que também realizou uma meta-análise de 20 estudos publicados entre 1990 e 1999, chegou ao mesmo entendimento, afirmando ainda que tais achados são compatíveis com as pesquisas dos anos 1970 e 1980 e ainda suplementa que, diante dos resultados regulares ao longo do tempo, as pesquisas educacionais poderiam se ater a eficácia de estratégias alternativas como remediação para a repetência do que ficar presos aos estudos do fenômeno em si.

Calafate (2017, p. 27), no entanto, alerta que a comparação de alunos retidos e promovidos pode ser uma análise enviesada, já que uma das dificuldades dos estudos é justamente as medidas pós-teste, pois os grupos de alunos foram expostos a experiências diferenciadas, já que estão em séries diferentes, ou seja, aprendendo conteúdos diferentes,

¹⁶ “meta-análise consiste em colocar diferentes estudos juntos em um mesmo banco de dados e utilizar metodologias analíticas e estatísticas para explicar a variância dos resultados utilizando fatores comuns aos estudos” (ROSCOE; JENKINS, 2005).

assim, o autor ainda esclarece que o fenômeno da repetência tem particularidades que a torna “inimiga” dos desenhos experimentais, dificultando assim as investigações das relações causais e de impactos reais.

No cenário das pesquisas educacionais brasileiras, em 1994, Gomes-Neto e Hanushek, estudando as causas e consequências da repetência para os estudantes das escolas rurais do nordeste brasileiro, observaram dois fatores determinantes para o fenômeno: o baixo desempenho acadêmico dos alunos e a escassez de séries mais avançadas nas escolas.

Oliveira e Soares (2012) buscaram os determinantes da repetência escolar através dos dados do Censo Escolar, através da metodologia de Logit, estimaram as chances de repetência de acordo com as características dos alunos e das escolas, duas das conclusões foram que alunos do sexo masculino estão mais susceptíveis à repetência do que as meninas e que estudantes das escolas municipais também estão mais expostos à repetência.

Correa *et al.* (2014) realizaram um estudo observacional, comparando dois grupos pareados de alunos repetentes e os promovidos, através de testes cognitivos de Língua Portuguesa e Matemática por três anos seguidos, os resultados por série indicam favorecimento para os repetentes, entretanto, com o passar do tempo, os repetentes passam a demonstrar que aprendem menos do que os promovidos, ou seja, as vantagens vão se dissolvendo ao longo do tempo.

Bachetto (2016) calculou o custo da reprovação para o Brasil, utilizando como referência o valor por aluno do Fundeb 2012, chegando ao valor estimado de R\$ 8,8 bilhões, que significa 9,1% do total do Fundeb 2012, entretanto, o autor afirma que este valor não é o final para ao Brasil, “pois a esses valores serão adicionados os complementos dos Estados e municípios que são praticamente impossíveis de serem computados, devido à grande quantidade de orçamentos municipais”. O autor ainda conclui que os estados de São Paulo, Bahia e Rio de Janeiro foram os que apresentaram os maiores custos com a repetência.

Ainda sobre o custo da reprovação, com dados um pouco mais recentes, em reportagem ao jornal G1, Fajardo (2018) esclarece que somente em 2016, o Brasil gastou R\$ 16 bilhões com a repetência na educação básica, sendo R\$12 bilhões exclusivamente no ensino fundamental.

A fim de tentar reduzir as taxas de repetências no Brasil, alguns sistemas de ensino adotaram políticas de ciclo ou de progressão continuada, entretanto tais políticas não foram bem aceitas pela sociedade, alegando que afetaria diretamente a qualidade do ensino.

Entretanto, Sergei Soares (2007) provou que as políticas de progressão continuada não exercem nenhum impacto negativo sobre o desempenho acadêmico dos alunos. Koch *et al.* (2007) também observaram a política de progressão continuada nas escolas públicas de Florianópolis – Santa Catarina e concluíram que a implantação desta política não alterou a qualidade do ensino fundamental. O estudo de Ferrão *et al.* (2002) também obteve os resultados na mesma direção, à nível escola, a promoção automática indicou que contribui para o correção da defasagem idade-série sem sacrificar a qualidade da educação e ainda indica que o desempenho escolar do aluno em risco tende a melhorar se ele for promovido, outro achado importante foi em relação aos “alunos desfavorecidos socioeconomicamente” de escolas com promoção automática, que não tiveram evidências de que seus resultados foram inferiores se comparados aos seus colegas.

Como pudemos observar, a repetência carece de atenção por partes dos gestores educacionais, pois a grande maioria dos estudos, mesmo quando se utilizam de diferentes abordagens, concluem que a reprovação não traz benefícios para os alunos e muito menos para os sistemas de ensino, causando inclusive desperdício de dinheiro, pois inflam os sistemas e demandam mais infraestruturas e profissionais, mesmo quando o principal argumento para a repetência está no âmbito pedagógico, onde os alunos que não aprenderam os conhecimentos necessários para avançar para a etapa seguinte devem repetir esta etapa mais uma vez, e assim consolidar os conhecimentos fundamentais.

Koch *et al.* (2007) acrescentaram ainda que, em sua pesquisa, a maioria dos professores dizem que os alunos reprovam porque não querem aprender, tais argumentos apresentados pelos professores reforçam a ideia de que a reprovação não está diretamente relacionada ao trabalho pedagógico desenvolvido em sala de aula, mas sim com aspectos individuais dos alunos e/ou família e assim tirando das escolas a sua parcela de culpa pelas altas taxas de repetências. Sergei Soares (2007) ainda acrescenta uma reflexão pertinente sobre a repetência, indicando que a mesma é utilizada nas escolas como incentivo para os alunos estudarem e afirma que “é lamentável que uma escola não consiga motivar seus alunos de outro modo que não com a ameaça da repetência”.

4 A REDE MUNICIPAL DA CIDADE DO RIO DE JANEIRO

A Sinopse Estatística da Educação Básica 2017 informa que, em relação à quantidade de unidades escolares, a rede municipal de ensino da cidade do Rio de Janeiro é o maior sistema educacional público do Brasil¹⁷, com 1.540 unidades escolares e atendendo a 649.708 alunos (INEP, 2018). Entretanto, na página eletrônica da SME-Rio¹⁸, a informação diferencia um pouco, indicando 1.537 unidades escolares, com 654.949 alunos. Essa pequena divergência pode ser devido a momentos diferentes de coleta da informação, já que no site da SME-Rio há uma indicação de que a última atualização dos dados foi feita em julho de 2017.

Como o presente estudo foi realizado no período de 2012 a 2015 e faz parte de uma gestão anterior da SME-Rio, é válido contextualizar, ainda que brevemente, as principais políticas e ações implantadas por tal gestão, pois elas podem ter sido um diferencial na trajetória dos alunos da coorte estudada.

Em 2009, quando o então prefeito Eduardo Paes assumiu o cargo, a professora Claudia Costin foi selecionada para assumir a SME-Rio e uma nova política de educação foi adotada para a cidade, dentre as principais medidas adotadas estão: as orientações curriculares, os cadernos pedagógicos e o sistema de avaliação (MORAES, SPINDOLA, 2017).

Já a própria Prefeitura, em seu portfólio sobre a Educação¹⁹, indica que as três medidas que foram fundamentais “para garantir uma educação pública de qualidade” seriam: currículo básico para todas as escolas, realização de provas bimestrais e participação em avaliações externas e adoção de um programa de reforço escolar (RIO DE JANEIRO, 2016, p.6).

As orientações curriculares são compostas apenas de objetivos, conteúdos e habilidades que deveriam ser trabalhadas em cada ano escolar a cada bimestre do ano letivo. Os cadernos pedagógicos são “apostilas” desenvolvidas pela própria SME-Rio, com base nas orientações curriculares e entregue a todos os alunos da rede. E o sistema de avaliação é composto por avaliações bimestrais e pela Prova Rio (MORAES, SPINDOLA, 2017). As provas bimestrais são únicas para toda a rede e permitem que se “conheçam, com

¹⁷ A rede municipal da cidade de São Paulo possui 1.506 e a estadual possui 1.177 unidades escolares.

¹⁸ Disponível em: <<http://www.rio.rj.gov.br/web/sme/educacao-em-numeros>>. Acesso em: 02 jul.2018.

¹⁹ Cadernos de Políticas Públicas: Educação. Disponível em:

<http://www.rio.rj.gov.br/dlstatic/10112/4379008/4130264/EducacaoPORT1211.pdf>

precisão, os pontos de atenção e a evolução dos estudantes” nas disciplinas de Português, Matemática, Ciências e Redação (RIO DE JANEIRO, 2016, p.9).

Santos (2014), que teve como objetivo na sua tese, estudar a política de gestão do sistema educacional público da cidade do Rio de Janeiro no período de 2009 a 2012, também encontrou um conjunto de ações que foi sendo implantado na rede destacando as seguintes: o fim da aprovação automática, a criação do IDERIO e a adoção da política de responsabilização, a criação de uma nova matriz curricular e a implantação de programas específicos como o Programa Escolas do Amanhã, o Ginásio Carioca e Programa Reforço Escolar. O autor ainda explica que toda estruturação dessa nova política educacional tinha como foco a melhoria da qualidade da educação ofertada à sociedade, levando em consideração os maiores desafios encontrados na avaliação diagnóstica da área de educação, feita para a confecção do Plano Estratégico 2009-2012.

A prefeitura ainda destaca que, paralelo a todas essas ações, também deu início a uma reforma “pioneira e ambiciosa” da rede com a implantação de mais um programa, o Programa Escolas de Tempo Integral, onde é previsto que as escolas tenham sua carga horária ampliada para sete horas diárias e que sejam organizadas por segmento dentro de um mesmo território²⁰, sendo o Espaço de Desenvolvimento Infantil (EDI) para a Educação Infantil, o Primário Carioca para as turmas de 1º ao 6º ano e o Ginásio Carioca para as turmas de 7º ao 9º ano (RIO DE JANEIRO, 2016, p.6-7).

Vale ressaltar ainda que, inicialmente, não foi pensada dessa forma a reorganização da rede, pois no Plano Estratégico da Prefeitura do Rio de Janeiro 2013-2016, a iniciativa estratégica chamada de “Escola Carioca em Tempo Integral”, previa três segmentos para o ensino fundamental, subdividindo-o em três ciclos de três anos escolares cada, chamados de Casa de Alfabetização (1º ano 3º ano), Primário Carioca (4º ao 6º ano) e Ginásio Carioca (7º ao 9º ano) (RIO DE JANEIRO, 2013). Não foi encontrado nenhum documento, notícia ou entrevista que justificasse o “abandono” das Casas de Alfabetização, agregando-as com os Primários Cariocas.

Um resumo da estrutura proposta pela SME-Rio está na **Erro! Fonte de referência não encontrada.** a seguir.

²⁰ A SME-Rio chama cada território deste de microrregiões e esclarece que a cidade foi dividida em 233 microrregiões a partir de um estudo da rede em 2013 (RIO DE JANEIRO, 2016, p. 7)

Figura 2: Estrutura Básica. Fonte: RIO DE JANEIRO, 2016, p. 8



Como é possível perceber, muitas ações foram implantadas ao longo de toda essa gestão, algumas delas se perpetuam até o momento presente, outras foram sendo descontinuadas ou até mesmo esquecidas ainda que durante a gestão anterior. Ainda assim, os alunos que compõe a coorte estudada foram também atores destas políticas, experimentaram as ações e podem, também, ter tido suas trajetórias escolares influenciadas por elas. Em função disso, algumas dessas ações (aquelas que julgamos serem mais pertinentes aos anos finais) foram incorporadas neste estudo e merecem um pouco mais de detalhamento que veremos a seguir.

a) Programa Escolas do Amanhã:

O Programa Escolas do Amanhã²¹ (EA) foi implantado em 155 unidades escolares localizadas em áreas mais vulneráveis, contava com a cooperação técnica da Unesco e possuía cinco pilares básicos: “educação em tempo integral, formas dinâmicas e instigantes de aprendizagem, atenção à saúde, participação da comunidade e implantação de uma cultura de resultados” (RIO DE JANEIRO: SECRETARIA MUNICIPAL DE EDUCAÇÃO, 2016, p. 12), se estruturando em seis ações: Mais Educação, Reforço Escolar, Projeto Uerê-Mello, Bairro Educador, Programa Saúde na Escola e Cientistas do Amanhã (GAWRYSZEWSKI, 2015) e tendo como resultados esperados, de acordo com os Planos Estratégicos da Prefeitura do Rio de Janeiro de 2009-2012 e 2013-2016, respectivamente: “melhoria da qualidade do ensino, melhoria do desempenho dos estudantes e redução da evasão escolar” e “melhoria na qualidade do ensino público, com a redução da evasão escolar” (RIO DE JANEIRO, 2009; 2013).

Christovão (2018) esclarece que a ideia de vulnerabilidade por parte da SME-Rio está associada ao contexto de violência do entorno, sendo este prejudicial aos processos de ensino e aprendizagem dos alunos, por isso, a ampliação da permanência dos alunos nas escolas poderia mitigar a exposição aos riscos iminentes, dando indícios da crença que SME-Rio tem no potencial que a educação tem em transformar a realidade social diante às desigualdades sociais.

O estudo de Gawryszewski (2015) indica que tal modelo se baseia no “aluno em tempo integral” e não na organização das unidades escolares a fim de oferecer condições que permitam um trabalho em tempo integral, demonstrando assim uma grande fragilidade do programa, a saber:

“A consequência que tem se verificado é que não existe uma coesão entre o que é ensinado nos dois turnos, ocasionando não mais do que um enxerto de horário, em que o turno estaria (pelo menos, supostamente) amparado pelas diretrizes contidas no Projeto Político-Pedagógico da instituição e o contraturno, tempo das atividades extraescolares, seria realizado sem qualquer referência aos processos pedagógicos realizados pelos professores”.

Souza (2015) também encontrou problemas no programa em questão no seu estudo, indicando uma “falta de identidade” do mesmo, pois atividades eram ofertadas em quantidade e qualidade diferenciada, dependendo da unidade escolar, sugerindo uma

²¹ Resolução nº 1038, de 24 de agosto de 2009.

possível adaptação a cada realidade ou ainda uma improvisação e até mesmo uma precariedade do programa.

Christovão (2018) corrobora com tal opinião e sugere que as fragilidades do programa podem ter sido causadas pela falta de um desenho bem elaborado.

Em relação aos resultados esperados, a tese de Christovão (2018) revela que, para os indicadores dos anos iniciais (IDEB e IDERio), entre 2009 e 2015 “as variações diversas quanto aos resultados [...] não podem ser explicados pela implementação do programa nas escolas selecionadas pela rede de ensino”, não sendo possível afirmar que o programa em questão acarreta efeitos de melhoramentos no desempenho dos alunos.

Na página do Rioeduca²² da SME-Rio sobre o programa, também é possível encontrar resultados comparativos²³ da Rede com as escolas do Programa Escolas do Amanhã, mostrando gráficos paralelos, ao longo de 2009 a 2015, na maior parte dos indicadores (IDEB, IDERio e AlfabetizaRio). A taxa de abandono dos Anos Finais foi o indicador que mais chamou a atenção, conforme Gráfico 11, o gráfico das Escolas do Amanhã apresenta uma trajetória descendente ao longo dos três primeiros anos do programa, que apesar de também ser a trajetória da Rede, se mostrou mais acentuado, sugerindo até uma possível confluência em longo prazo, entretanto, a partir de 2013, as curvas de ambos os grupos ascendem, mantendo um paralelismo.

A estabilização de tal resultado pode ser explicada pelos achados de Souza (2015) que indicaram que com o passar do tempo o Programa Escolas do Amanhã foi se dissipando e restringindo-se apenas ao Programa Mais Educação do Governo Federal. Christovão (2016) também compartilha da mesma impressão, indicando que “nos últimos anos (2016, 2017), [o programa] apresenta um desenrolar que aponta para seu término”, Bouzada (2016) complementa observando uma “descontinuidade das propostas, a falta de avaliação e o rompimento com o planejamento antes mesmo que se obtenham resultados palpáveis”. Por fim, Souza (2015) conclui que “será preciso um grande aperfeiçoamento nas bases

²² <http://www.rioeduca.net/programasAcoes.php?id=19>. Acesso em: 05 jul. 2018.

²³ Como os resultados da SME-Rio são apenas descritivos entre os grupos, não havendo nenhum tipo de estatística que garanta a significância dessas diferenças ou até mesmo o controle dos grupos por métodos estatísticos, sugere-se que os resultados do estudo de Christovão (2018) parecem ser mais adequados pela sua metodologia robusta, garantindo sua validade interna, isto é, são mais confiáveis para conclusões sobre o programa.

estruturais do sistema educacional, envolvendo pessoal, edificações e currículo, a fim de que os possíveis avanços gerados por programas como o *Escolas do Amanhã* sejam efetivos”.



Gráfico 11: Taxa de Abandono - Anos Finais - Rede x Escolas do Amanhã. Fonte: SME-Rio

Na própria página eletrônica do programa²⁴, embora mantida a sua identidade (logomarca, nome e data de criação - 2009), o programa possui as mesmas características do Programa Escolas de Tempo Integral apresentado no portfólio sobre Educação da prefeitura (RIO DE JANEIRO, 2016, p.6).

b) Projeto 6º Ano Experimental:

Essa ação, diferente das outras aqui tratadas não é considerada um programa por parte da SME-Rio e também não foi encontrada nenhuma regulamentação oficial (Resolução, Decreto, etc.). É um projeto que se iniciou em 2011 com caráter de “piloto” em 53 turmas (20 escolas²⁵) cujo objetivo era incorporar o 6º ano aos anos iniciais, garantindo aos alunos “uma transição sem sobressaltos para o último segmento do ensino fundamental” (RIO DE JANEIRO, 2016, p.7), onde os alunos permanecem apenas com um único docente para todas as disciplinas (exceto Inglês, Educação Física e Artes).

²⁴ <http://www.escolasdoamanha.com.br/>

²⁵ O critério de seleção das primeiras escolas se deu pela proximidade aos Ginásios Cariocas, já que a intenção inicial era “de oferecer um suporte para aqueles alunos que fossem entrar n Ginásio Carioca” (SANTOS *et al.*, 2017).

Santos (2016, p.17) esclarece que dentre as transições escolares que o aluno precisa realizar durante a educação básica, a do 6º ano do ensino fundamental é diferente por ser também a transição de inúmeros fatores, como por exemplo, o início da adolescência e acrescenta ainda questões pedagógicas relativas a essa transição:

“[...]os problemas como o maior número de professores, maior número de matérias, o convívio com alunos maiores, maior distanciamento dos professores e dos responsáveis, possível troca de escola e, principalmente, mudanças bruscas nas metodologias adotadas, são alguns fatores que podem contribuir para tornar a passagem do último ano do EF I [anos iniciais] para o primeiro do EF II [anos finais] um momento desafiador para os educandos. Vale ressaltar que esses educandos serão exigidos a se ajustarem o mais rápido possível para que não sofram prejuízos em relação ao nível de desenvolvimento esperado para o 6º ano.”

Em entrevista ao Jornal O Extra, em 25/11/2014, a então secretária de Educação da SME-Rio, Profª. Helena Bomeny, afirmou que “o garoto não tem maturidade e não consegue corresponder bem em todas as disciplinas”, disse ainda que durante os quatro anos de projeto (2011 a 2014) a média das notas do 6º ano experimental é 50% acima da média quando comparada ao restante de rede²⁶ e que, por isso, o projeto seria expandido (BOMENY, 2014).

A pergunta que fica é se com esse projeto a SME-Rio consegue realmente atingir seus objetivos ou se apenas retarda o problema da transição em mais um ano. Essa questão só poderá ser respondida de forma confiável se for realizado um estudo robusto e com um bom desenho.

Falando em estudo científico, na literatura, só foi encontrado um único estudo de Santos *et al.* (2017) que visava avaliar o impacto em curto prazo do projeto das escolas que começaram o projeto no ano de 2014, utilizando a diferença entre as notas dos alunos nas Provas Bimestrais no final do 2º bimestre de 2013 (5º ano) e no final do 2º bimestre de 2014 (6º ano), portanto, a metodologia adotada foi a “diferenças-em-diferenças”, comparando um conjunto de escolas semelhantes²⁷, com um grupo de controle (sem o projeto) e um grupo experimental (que recebeu o projeto em 2014), tendo sido encontrado impacto positivo, de grande magnitude, sobre o nível de proficiência dos alunos, sendo maior em

²⁶ A SME-Rio não esclarece qual metodologia é utilizada para a comparação, se é apenas uma comparação entre notas e índices dos grupos, sem qualquer tratamento estatístico ou se há um rigor científico necessário para que se possa garantir a validade do estudo.

²⁷ Foi utilizado o método do pareamento por escore de propensão, com as seguintes dimensões: proporção de meninos, proporção de brancos, proporção de professores com nível superior, variação da nota média de português e matemática entre o 5º e 6º anos antes da adoção do projeto, número de alunos e razão professor/aluno.

Matemática do que em Português. Além desses achados, também foi encontrado “melhoria do clima em sala de aula, maior autonomia pedagógica dos docentes e melhor capacitação dos professores do 6º Ano Experimental²⁸” (SANTOS *et al.*, 2017).

Mesmo sem garantias científicas da eficácia do projeto, baseando-se apenas em indícios positivos, a SME-Rio vem expandindo o projeto na rede ano após ano (conforme apresentado no Gráfico 12), inclusive já fazendo parte da proposta de reforma oficial da SME-Rio.

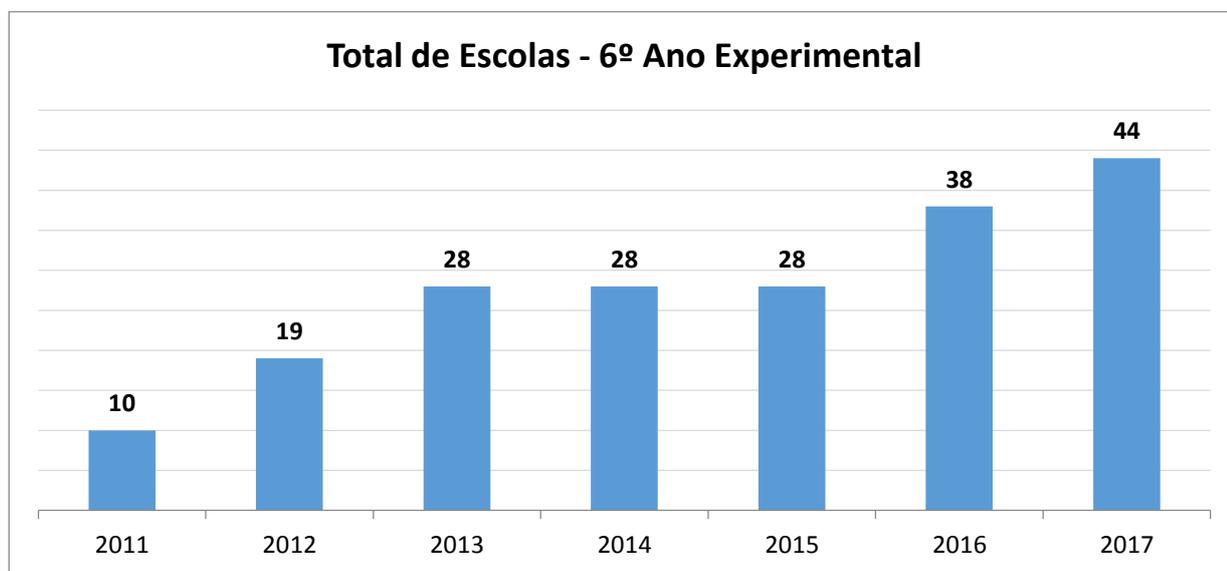


Gráfico 12: Total de Escolas do 6º Ano Experimental. Fonte: Dados da SME-Rio. Elaboração própria.

c) Programa Ginásio Carioca:

O Programa Ginásio Carioca²⁹ (GC) foi implantado na rede em 2011, inicialmente em 10 escolas de 7º a 9º ano – como tinha característica de experimento ficou conhecido como GEC³⁰ – e que tinha como arcabouço:

“a introdução de novos métodos e práticas no cenário educacional que viabilizassem a existência de uma escola voltada para a excelência acadêmica e para a formação de jovens competentes, autônomos, solidários e corresponsáveis por sua própria suficiência, bem como pela transformação da comunidade e da sociedade em que vivem” (SANTOS, 2014).

Em uma notícia na página eletrônica da SME³¹, datada de 16/12/2016, o idealizador do projeto e presidente do Instituto de Corresponsabilidade pela Educação (ICE), Marcos

²⁸ O projeto prevê formação semanal dos docentes do projeto e formação para a direção das escolas.

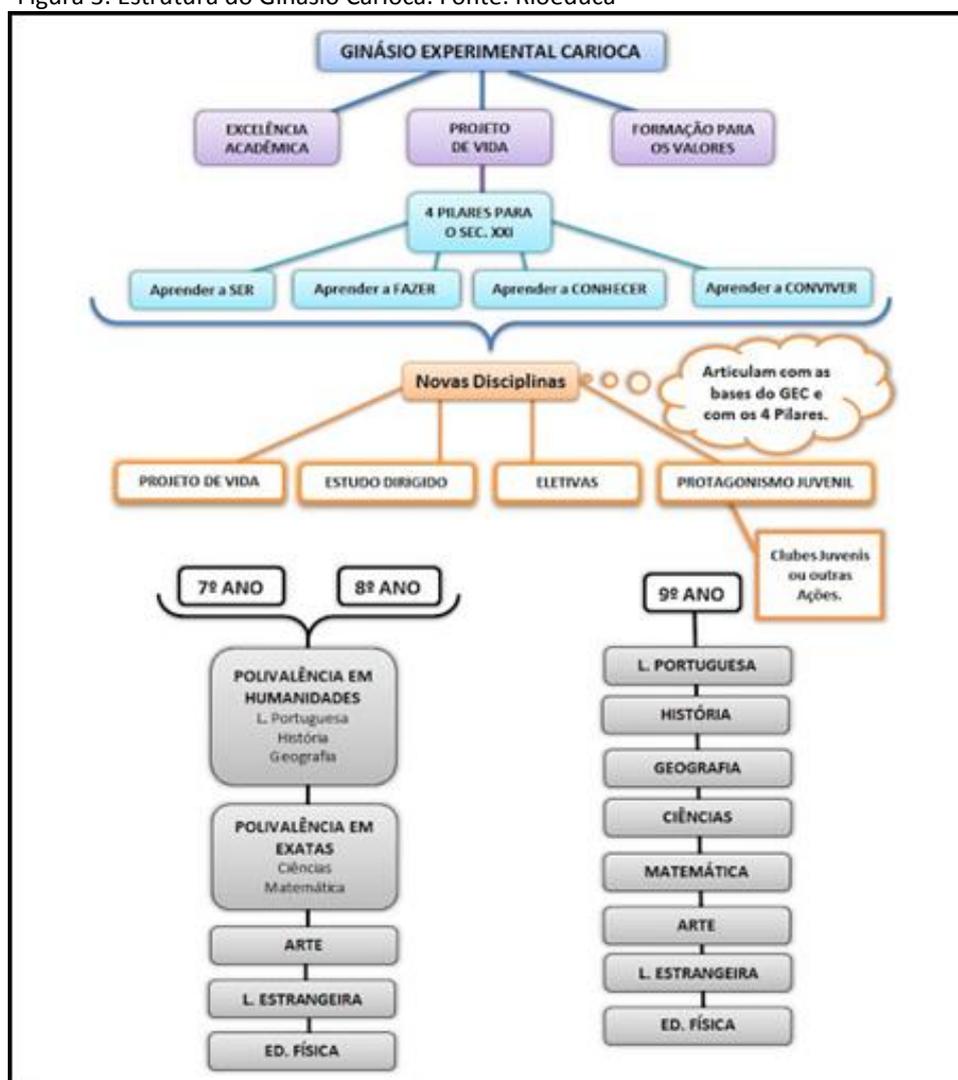
²⁹ Decreto nº 32.672, de 18 de agosto de 2010

³⁰ Ginásio Experimental Carioca

Magalhães, explica que a iniciativa surgiu a partir de um projeto já implantado em Pernambuco (Ginásio Pernambucano – Ensino Médio).

O programa tem como objetivo inovar a na matriz curricular, na metodologia e nos modelos pedagógico e de gestão, incorporando técnicas de planejamento, execução, monitoramento e avaliação de seus processos, tendo como premissa a educação integral, com sete horas diárias e atividades eletivas e docentes de 40h semanais (SECRETARIA MUNICIPAL DE EDUCAÇÃO, 2016).

Figura 3: Estrutura do Ginásio Carioca. Fonte: Rioeduca³²



Na Em uma notícia na página eletrônica da SME, datada de 16/12/2016, o idealizador do projeto e presidente do Instituto de Corresponsabilidade pela Educação (ICE), Marcos

³¹ <http://www.rio.rj.gov.br/web/sme/exibeconteudo?id=6612676>

³² Disponível em: <<http://www.rioeduca.net/blogViews.php?bid=14&id=3293>>. Acesso em 05 jul. 2018.

Magalhães, explica que a iniciativa surgiu a partir de um projeto já implantado em Pernambuco (Ginásio Pernambucano – Ensino Médio).

O programa tem como objetivo inovar a na matriz curricular, na metodologia e nos modelos pedagógico e de gestão, incorporando técnicas de planejamento, execução, monitoramento e avaliação de seus processos, tendo como premissa a educação integral, com sete horas diárias e atividades eletivas e docentes de 40h semanais (SECRETARIA MUNICIPAL DE EDUCAÇÃO, 2016).

Figura 3, encontrada na página eletrônica do Rioeduca, é possível entender melhor a estrutura prevista para os Ginásios Cariocas.

A expansão do programa foi gradativa sendo mais 9 unidades em 2012 e outras 9 em 2013, depois, somente em 2016 voltou a ganhar mais 10 unidades e, em 2017, outras 6 novas unidades.

Em 2013, seis unidades receberam uma estrutura vocacionada, um para as Artes (Ginásio das Artes – GEA), um para a Música (Ginásio do Samba – GES), um para as Novas Tecnologias (Ginásio Experimental de Novas Tecnologias – GENTE) e três unidades para o Esporte (Ginásio Experimental Olímpico – GEO)³³.

O GEA tem integração com o Museu de Arte do Rio (MAR) e é para os alunos com aptidão e interesse pelo universo artístico, com especialização em ateliês de desenho, pintura, cor, figura, tecelagem, 3D, artes gráficas, mídias, dança, teatro e artes cênicas; os alunos recebem sete tempos de aulas semanais dedicados a atividades artísticas (RIO DE JANEIRO, 2016, p.32).

Já o GES funciona em Ramos e seis tempos semanais de aula são dedicados a atividades específicas de música, como por exemplo: percussão, instrumentos de corda, teclas e sopro, aulas de canto e oficinas para desenvolvimento da criatividade musical (RIO DE JANEIRO, 2016, p.32).

O GEO é o único deles que recebe alunos do 6º ano e visa formar aluno-atleta-cidadão, baseado nos valores olímpicos e no desenvolvimento de habilidades esportivas, os alunos possuem 2 horas diárias de práticas esportivas. Atualmente, a rede possui quatro GEOs em funcionamento (em Santa Tereza, Caju, Pedra de Guaratiba e Ilha do Governador)

³³ Apesar de todo esse diferencial, no nosso estudo todas essas escolas foram consideradas como GCs apenas.

com atividades esportivas de atletismo, tênis de mesa, vôlei, handebol, xadrez, natação e luta olímpica (RIO DE JANEIRO, 2016, p.34).

O GENTE se localiza na Rocinha e tem como diferencial, além do uso de tablets e smartphones – pois a internet é parte central do aprendizado dos alunos, não existem turmas em salas de aula tradicionais (com quadro negro e carteiras individuais enfileiradas) – os alunos são agrupados por times, que misturam alunos do 7º ao 9º ano, sob a monitoria de professores, onde cada aluno deve seguir seu plano de estudo personalizado (TERRA, 2016).

Ainda sobre o GENTE, Ouchana (2014) escreve um artigo para a Revista Educação intitulado “Qual o futuro do Projeto GENTE?” que trata sobre o futuro desse modelo de escola, explicando que experiências pontuais desse tipo são comuns no Brasil, entretanto, o grande problema está em como transformar esses pilotos em políticas públicas de verdade, ou seja, como replicar o projeto de modo que se possa realmente resolver os problemas educacionais de grandes redes, não se tornando apenas uma “vitrine” que, no final das contas, acaba aprofundando as desigualdades já que apenas um pequeno grupo de alunos tem acesso a esses modelos.

Tal reflexão acima pode ser ampliada para todo o programa de Ginásio Carioca e uma prova disso pode estar sendo vista na expansão, onde ao longo de seis anos do programa (2011 a 2017) o número de escolas é apenas um pouco mais que quatro vezes o número inicial³⁴, muito diferente, por exemplo, do projeto 6º Ano Experimental, conforme apresentado no item anterior.

Um fato que chamou a atenção é que este programa é tido como um dos “carros-chefes” da SME-Rio e, assim como o sexto ano experimental, faz parte da reforma educacional proposta pela própria secretária, mas não há sido encontrado nenhum estudo do programa, sobre sua eficiência, ou até mesmo sobre impacto, e mais uma vez a tomada de decisão dos gestores se baseia em indícios e não em evidências comprovadas cientificamente, não havendo garantias da manutenção dos resultados quanto a sua expansão.

Sob esse ponto, Gertler *et al.* (2016, p. 3, tradução nossa) esclarecem que comprovar se os objetivos pretendidos foram alcançados ou não, quando se desenvolve algum

³⁴ Conforme já informado, o programa iniciou com 10 escolas em 2010 e, até 2017, tinha 44 escolas.

programa ou alguma política, “é uma questão crucial de política pública, mas que não é frequentemente examinada” e ainda acrescentam:

“O monitoramento e a avaliação estão no centro da formulação de políticas baseadas em evidências. Eles fornecem um conjunto básico de ferramentas que as partes interessadas podem usar para verificar e melhorar a qualidade, a eficiência e a eficácia de políticas e programas em vários estágios de implementação - ou, em outras palavras, para se concentrar nos resultados. No nível de gerenciamento do programa, há uma necessidade de entender quais opções de projeto de programa são mais eficazes em termos de custo ou argumentar para os tomadores de decisão que os programas estão atingindo os resultados pretendidos para obter alocações orçamentárias para continuar ou expandi-los.” (GERTLER *et al.*, 2016, p. 4, tradução nossa)

d) Programa Reforço Escolar:

O Programa Reforço Escolar foi uma das iniciativas estratégicas presente no plano estratégico da prefeitura do Rio de Janeiro 2009-2012 e foi pensado com base numa avaliação diagnóstica realizada na rede em 2009 em que foram encontrados 28 mil alunos analfabetos funcionais e defasagem de aprendizagem em Língua Portuguesa e Matemática em praticamente todo o ensino fundamental, sendo uma das estratégias propostas a criação de turmas dedicadas ao reforço escolar (RIO DE JANEIRO, 2009, p.52).

Calafate (2007, p. 48) que trata em sua dissertação sobre a política de correção de fluxo, esclarece que o Programa Reforço Escolar estava dividido em duas áreas de atuação, sendo uma para “correção de fluxo escolar” e outra para “manutenção da aprendizagem adequada”, a primeira área era voltada para os alunos identificados como analfabetos funcionais e para aqueles que tinham defasagem idade-série ficariam em turmas específicas conhecidas como “projetos de correção de fluxo” e a segunda área para os alunos com defasagem de aprendizagem, onde atividades de reforço eram feitas na própria aula ou no contraturno do aluno.

O mesmo autor ainda elucida que tal iniciativa não é inédita na rede, pois algum tipo de projeto de aceleração já existe na rede há vários anos, como por exemplo, as “classes de progressão” que aparecem nas normativas de 2001.

Santos (2014, p. 115) esclarece ainda que esse programa ocorreu através de parcerias da SME-Rio com várias organizações não governamentais – como o Instituto Ayrton Senna, Fundação Roberto Marinho, Fundação Victor Civita, Fundação Lemann, Instituto Alfa e Beto, Instituto Sangari, entre outros – através da compra “de materiais ou pacotes didático-pedagógicos completos (material, treinamento, consultoria...)”

Este autor fez um quadro (replicado no Quadro 2 a seguir) com o resumo das ações que compõem o Programa Reforço Escolar no período de 2009 a 2012³⁵, seu público alvo e os parceiros que eram responsáveis e ainda enfatiza que mais tarde, a SME-Rio, altera o nome de alguns projetos e assume o nome “Aceleração” para os projetos que visam à correção da distorção idade-série e “Realfabetização” para os projetos que visam à consolidação da alfabetização junto com a correção da distorção idade-série.

Quadro 2: Programas, Projetos e Ações instituídos para o Reforço Escolar. Fonte: Santos (2014, p.116-117)

Projeto	Especificação	Parceiro
“Se Liga”	Realfabetização de alunos do 3º, 4º e 5º anos com defasagem idade/ano escolar.	Instituto Ayrton Senna
“Acelera Brasil”	Aceleração de aprendizagem para alunos do 3º e 4º anos com defasagem idade/ano escolar.	Instituto Ayrton Senna
“Fórmula da Vitória”	Realfabetização de alunos do 6º ano, com defasagem idade/ano escolar, com nível básico de alfabetização.	Instituto Ayrton Senna
“Ponto de Partida”	Realfabetização de alunos do 6º ano com defasagem idade/ano escolar ou analfabetos funcionais.	Instituto Ayrton Senna
“Autonomia Carioca - Complementar de Matemática”	Reforço escolar em Matemática, no contra turno, para alunos do 9º ano com baixo desempenho.	Fundação Roberto Marinho
“Autonomia Carioca – Aceleração do 6º ano”	Aceleração de aprendizagem para alunos do 6º ano com defasagem idade/ano escolar.	Fundação Roberto Marinho
“Autonomia Carioca - Aceleração do 7º e 8º anos”	Aceleração de aprendizagem para alunos do 7º e 8º anos com defasagem idade/ano escolar.	Fundação Roberto Marinho
Tecendo o Saber	Aceleração de aprendizagem para alunos do 5º ano com defasagem idade/ano escolar.	Fundação Roberto Marinho
Nenhuma criança a menos	Alunos oriundos de turmas de projetos que ainda estavam consolidando o processo de alfabetização.	XXX
Nenhum jovem a menos	Alunos, com defasagem idade/série que precisam, ainda, de um apoio mais cuidadoso após o processo de consolidação da leitura e da escrita.	XXX

Em sua dissertação, Calafate (2017, p. 47) ainda nos traz um alerta em relação a quantidade de alunos que estão nas turmas de projeto de correção de fluxo, pois, embora possa aparecer uma fatia pequena em relação a quantidade de alunos do ensino

³⁵ Período de estudo da sua tese.

fundamental da rede municipal do Rio de Janeiro, em 2013, cerca de 30 mil alunos estavam matriculados em alguma turma deste projeto.

Sendo assim, podemos verificar que mais uma vez temos uma política pública com uma grande abrangência, ou seja, que atende a um número considerável de alunos e por isso deve ser feita com responsabilidade, tendo garantias de que esse tipo de intervenção produz melhoras no desempenho escolar dos alunos, minimizando a reprodução de clivagens de certas categorias socioeconômicas, que por si só já são desfavoráveis ao desempenho educacional.

Neste estudo, o Programa de Reforço Escolar, em particular as turmas de projeto de correção de fluxo, não está sendo considerado como uma variável que incorpora aos modelos de Políticas Públicas, mas como uma forma de retirada do aluno do fluxo regular do ensino fundamental, se tornando assim imprescindível a sua apropriação – ainda que de forma resumida – para o entendimento da trajetória escolar da coorte aqui pesquisada, já que Calafate (2017, p.155) observa em suas considerações finais que se pode “garantir que os programas do segundo segmento funcionam como *‘trackings’*³⁶ curtos para facilitar o fim do Ensino Fundamental”.

³⁶ “Processos em grande escala que tendem a ocupar integralmente uma rede de ensino, atuando no sentido de agrupar os alunos em diferentes *modalidades de ensino*” (CALAFATE, 2017, p. 38).

5 MÉTODOS

5.1 O DESENHO

O banco de dados utilizado na dissertação tem característica longitudinal e acompanha uma coorte específica de alunos entre os anos de 2012 e 2015. Podemos dividir as análises que serão apresentadas em três partes: a) análise descritiva das variáveis; b) análise descritiva da trajetória da coorte; c) análise de sobrevivência.

A coorte estudada é composta por todos os alunos matriculados no 6º ano na rede da SME-Rio, avaliados no 1º Bimestre (isto é, que receberam algum conceito global no COC 1) no ano letivo de 2012, somando ao todo 59.018 alunos.

Para este estudo, o aluno é observado, anualmente, em todos os quatro anos de acompanhamento, até que ele saia do ensino fundamental, seja por repetência (evento pesquisado) ou por saídas do fluxo regular dos anos finais do ensino fundamental da SME-Rio, como abandono, evasão, transferência para outras modalidades de ensino dentro da própria rede ou transferência para outras redes de ensino.

Nas duas primeiras partes do estudo serão apresentadas análises descritivas da coorte, sendo inicialmente focada nas características dos alunos, de suas famílias e de políticas públicas e, em seguida, na trajetória percorrida pelos alunos ao longo do período estudado, utilizando o *software* estatístico SPSS.

A terceira parte do estudo será a análise de sobrevivência³⁷ propriamente dita, onde além de observar os resultados gerais encontrados, vamos comparar as funções de sobrevivência de diferentes grupos e estimar modelos através da Regressão de Cox. Para esta parte, o *software* estatístico escolhido foi o Stata.

5.1.1 Análise de Sobrevivência

A análise de sobrevivência teve sua origem no final da década de 1950 nas áreas das ciências biomédicas e da engenharia. Ambas são áreas de intervenção, assim como a educação, e precisam analisar dados agregados de sobrevivência, sendo a primeira em ensaios clínicos e a segunda na decomposição de máquinas e componentes eletrônicos

³⁷ A análise de sobrevivência será abordada com mais detalhes a seguir.

(OUD, 2014). Segundo Colosimo e Giolo (2006, p. 1-3), essa é uma das áreas da estatística que mais cresceu nos vinte anos finais do século passado, sendo utilizada em diversas áreas, como por exemplo: criminalistas, que estudam o tempo entre a libertação de presos a ocorrência de novos crimes; demógrafos, que analisam nascimentos, mortes, casamento, divórcios e migrações; ciências sociais, que se concentram em mudanças de empregos, desempregos, promoções e aposentadorias.

Oud (2014) esclarece que esse uso em diversas áreas fez com que tais procedimentos de análise recebessem diferentes nomes, tais como: análise da história de eventos, modelagem de duração do evento, modelagem do tempo de falha, modelagem de riscos, modelagem de tempo de resposta e modelagem ao longo da vida. Nas ciências sociais, o termo mais utilizado é a análise de história de eventos, tendo sido difundido a partir da publicação de dois livros em 1984: um livro introdutório de Allison e outro mais avançado de Tuma e Hannan (BERNARDI, 2006, p. 7). Aqui neste estudo, vamos adotar o termo análise de sobrevivência, por ser o mais utilizado no Brasil.

Na literatura brasileira é raro encontrar estudos no campo das ciências sociais ou humanas que se utilizam essa metodologia. Lins *et al.* (2017, p. 281-282) pesquisaram nas revistas DADOS, *Brazilian Political Science Review* (BPSR), *Revista de Sociologia e Política* (RSP) e *Revista Brasileira de Ciências Sociais* (RBCS) e só encontraram dois trabalhos, concluindo que a análise de sobrevivência “ainda não foi incorporada ao arsenal metodológico das Ciências Sociais”.

O que difere na análise de sobrevivência de outras técnicas, é que sua variável resposta é o tempo até a ocorrência do evento de interesse, conhecido como tempo de falha, e tem, por sua natureza, características longitudinais (COLOSIMO, GIOLO, 2006, p. 1-2). Além de possuir uma maior capacidade para estudar processos de mudança de estado do que outras técnicas de análise dinâmica, pois estas últimas se atêm à associação entre uma variável dependente e uma variável independente, enquanto a análise de sobrevivência permite observar as mudanças na variável independente e avaliar como essas mudanças influenciam a probabilidade de que uma mudança na variável dependente ocorra mais tarde (BERNARDI, 2006, p. 25, tradução nossa), sendo considerada por alguns autores como uma “nova abordagem para investigação empírica das relações causais” (BERNARDI, 2006, p. 25 *apud* BLOSSFELD, ROHWER, 2001).

Em termos gerais, a análise de sobrevivência permite investigar processos que se caracterizam por: a) a unidade de análise pode alterar de um estado de origem para um estado de destino; b) a mudança de estado não está restrita a um momento determinado no tempo; e c) esta mudança pode ser influenciada por um conjunto de variáveis independentes que podem ser constante no tempo ou não (BERNARDI, 2006, p. 13 *apud* COLEMAN, 1981).

De acordo com Lins *et al.* (2017) *apud* Kleinbaum e Klein (2005), a análise de sobrevivência possui três objetivos: a) estimar e interpretar funções de sobrevivência; b) comparar funções de sobrevivência; e c) avaliar a relação das variáveis explicativas e o tempo de sobrevivência.

A análise de sobrevivência é caracterizada por dados específicos que recebem nomes próprios, para facilitar a compreensão do método, faz-se necessário a definição desses, tais como: estado, evento, duração e censura.

Estado é uma característica qualitativa da unidade de análise, para exemplificar, podemos citar: vida ou morte, ter ou não ter uma moléstia, empregado ou desempregado, solteiro ou casado ou divorciado ou viúvo (OUD, 2014).

Evento é uma mudança qualitativa da unidade de análise do estado de origem para um estado de destino de interesse, em algum momento específico do tempo (BERNARDI, 2006, p. 13), também são conhecidos como falha, por serem eventos indesejáveis na maioria dos casos (COLOSIMO, GIOLO, 2006, p.7).

Duração é o intervalo de tempo que a unidade de análise permanece no estado de origem até o momento da ocorrência do evento, isto é, quando a unidade de análise passa para o estado de destino (BERNARDI, 2006, p. 14) ou até a ocorrência de uma censura.

Censura é uma observação parcial da resposta devido à interrupção, por algum motivo, da observação completa do evento, ou seja, quando se possui alguma informação sobre o tempo de sobrevivência da unidade de análise, mas não o tempo exato de ocorrência do evento e se o mesmo ocorreu ou não (LINS *et al.*, 2017).

Segundo Colosimo e Giolo (2006, p. 2), a presença de censura é a principal característica dos dados da análise de sobrevivência, pois, este método incorpora na análise estatística a informação contida nos dados censurados, portanto, sem a presença de

censuras, outras técnicas estatísticas poderiam ser utilizadas de forma satisfatória para análise dos mesmos.

Os mesmos autores ainda destacam:

“Ressalta-se o fato de que, mesmo censurados, todos os resultados provenientes de um estudo de sobrevivência devem ser usados na análise estatística. Duas razões justificam tal procedimento: (i) mesmo sendo incompletas, as observações censuradas fornecem informações sobre o tempo de vida de pacientes; (ii) a omissão das censuras no cálculo das estatísticas de interesse pode acarretar conclusões viciadas.” (COLOSIMO, GIOLO, 2006, p. 8)

Devido à importância das censuras acima destacada, faz-se fundamental um aprofundamento no conhecimento das mesmas. As censuras podem ser classificadas em dois tipos, de acordo com Bernardi (2006, p.21-22):

- Censuras à direita ocorrem quando o período de observação termina e o evento de interesse ainda não aconteceu ou quando não há mais a possibilidade de se continuar coletando informações da unidade de análise, mesmo durante o período de observação.
- Censuras à esquerda acontecem quando ao se iniciar o período de observação o evento já tenha ocorrido para a unidade de análise.

A análise de sobrevivência permite lidar sem problemas com os casos de censuras à direita, mas nunca com a censura à esquerda (BERNARDI, 2006, p.22-23; LINS *et al.*, 2017).

No Quadro 3, os conceitos básicos da análise de sobrevivência para este estudo estão resumidos, facilitando assim a compreensão dos termos que serão utilizados a partir de agora.

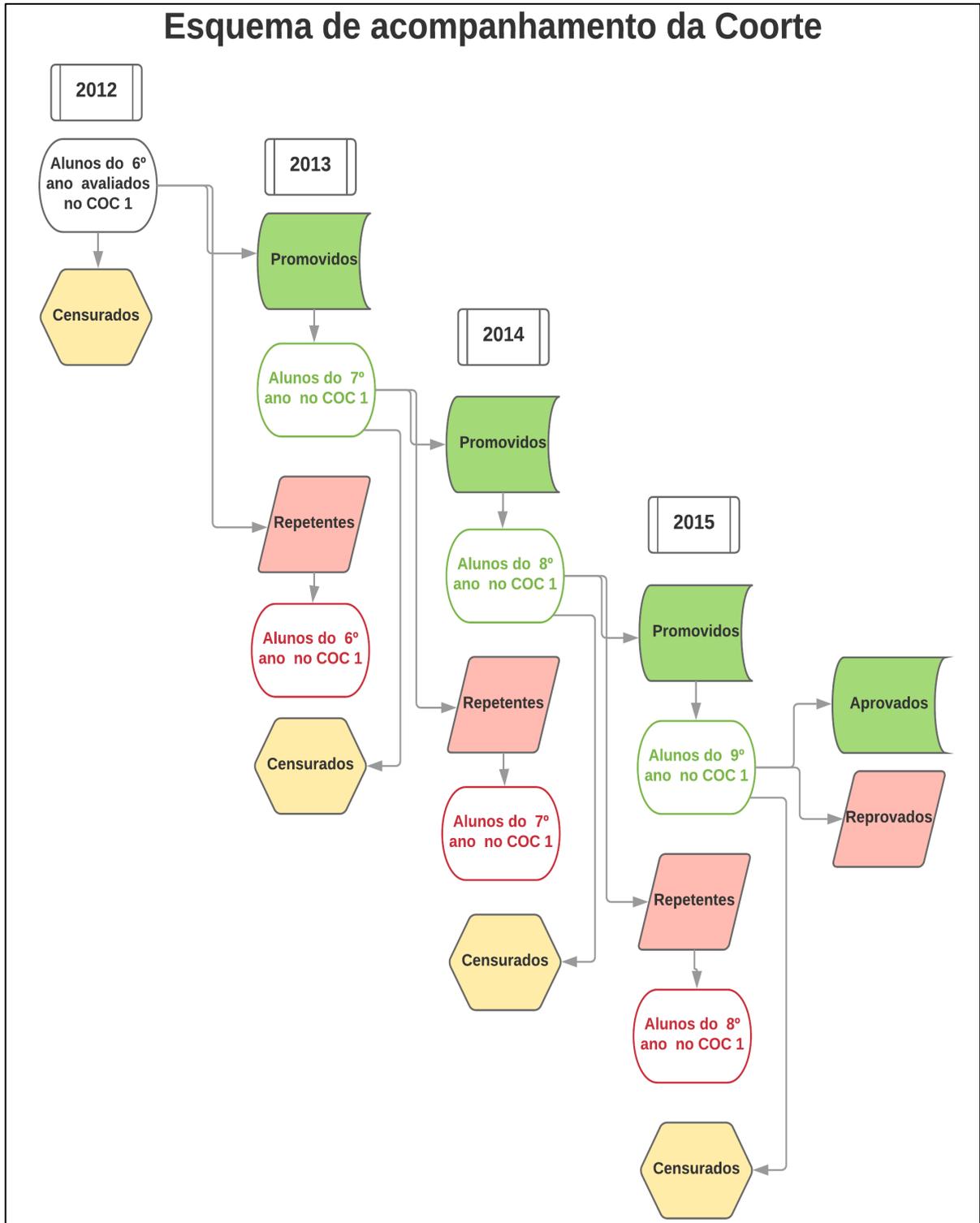
Quadro 3: Conceitos básicos da análise de sobrevivência no estudo

Conceitos básicos da Análise de Sobrevivência	
Conceito	Descrição
Unidade de análise	Aluno
Escala de medida	Anos
Período de observação	2012 a 2015
Evento	Primeira repetência dentro do período observado
Tempo de ocorrência	4 tempos (T=1 a T=4)
Censuras	Todas as possibilidades de perda do acompanhamento do aluno ou de saídas do fluxo escolar do ensino fundamental, como: <ul style="list-style-type: none"> ▪ Transferências para outras redes (particular ou pública); ▪ Transferências para outras modalidades de ensino que não seja o Ensino Fundamental (EJA, Classe Especial ou Classes de projeto de correção de fluxo) dentro da SME-Rio; ▪ Abandono ou evasão; ▪ Falecimento

Portanto, utilizando os termos específicos desta metodologia, podemos observar de forma ilustrativa, na

Figura 4 o esquema de acompanhamento da coorte que foi utilizado para este estudo.

Figura 4: Esquema de acompanhamento da Coorte



Bernardi (2006, p. 30) define a Função de sobrevivência e a função de densidade como sendo um dos fundamentos estatístico da análise de sobrevivência, sendo a função de sobrevivência³⁸ definida como a probabilidade que a unidade de análise tenha sobrevivido ao tempo e a função de densidade³⁹ como a probabilidade de uma unidade de análise não sobreviver ao tempo; esta última é também conhecida como função de distribuição acumulada (COLOSIMO; GIOLO, 2006, p.21).

O Estimador de Kaplan-Meier (EKM) é uma técnica não paramétrica, proposta pela dupla que o nomeia em 1958, e também conhecida como estimador limite-produto onde leva em consideração que, para uma unidade de análise sobreviver a n intervalos de tempo, antes ela teria que ter sobrevivido a cada um dos intervalos de tempo anterior a n , por exemplo, um aluno que tenha sobrevivido até o tempo 3, antes, obrigatoriamente, ele teve que sobreviver ao tempo 1 e 2, por isso, o estimador é calculado a partir do produto de todas as probabilidades de sobreviver a cada momento (LIMA JR *et al.*, 2012), tendo como fórmula:

$$\hat{S}_{KM}(t) = \prod_{t_i=0}^t \left(1 - \frac{d_i}{n_i}\right)$$

Onde,

t_i é um intervalo de tempo i ;

d é o número de eventos ocorridos em t_i ;

n_i é o número total de unidades de análise que estavam em risco no tempo t_i .

Colosimo e Giolo (2006, p. 40) citam as principais propriedades do EKM: “i) não-viciado para amostras grandes, ii) é francamente consistente, iii) converge assintoticamente para um processo gaussiano e iv) é estimador de máxima verossimilhança de $S(t)$.” e confirmam a sua superioridade em relação a outros estimadores existentes, indicando assim que o uso de tal estimador é o mais indicado para estudos de análise de sobrevivência.

Lima Junior *et al.* (2012) inteiram que a utilização do EKM, como uma abordagem não-paramétrica à análise de sobrevivência, permite comparar categorias, observando as

³⁸ $S(t) = Pr(T > t) = 1 - F(t)$

³⁹ $F(t) = Pr(T \leq t)$

curvas que cada grupo assume e, através de testes de significância estatística, poder afirmar se há ou não diferença entre os grupos.

O teste *Logrank* é o mais utilizado nos estudos sobre análise de sobrevivência e consiste na diferença entre o número observado de falhas em cada grupo e uma quantidade que poderia ser o número esperado de falhas sob a hipótese nula (COLOSIMO, GIOLO, 2006, p. 55), entretanto, ele é mais sensível nas diferenças entre as funções de sobrevivência ao final do eixo temporal, por isso, para complementar deve-se utilizar o teste de Wilcoxon⁴⁰ conjuntamente, pois este é mais sensível às diferenças no início do tempo (BERNARDI, 2006, p. 60).

Em ambos os testes, a hipótese nula baseia-se no fato de que a diferença observada entre as funções que estão sendo comparadas se dá pela aleatoriedade da amostra. Como tais testes se distribuem na estatística χ^2 com grau de liberdade igual ao número de grupos menos um, portanto, quando o resultado dos testes for estatisticamente significativo, rejeita-se a hipótese nula e assume que as diferenças encontradas não são aleatórias (BERNARDI, 2006, p.60).

Todavia, estas análises estratificadas – apesar da simplicidade de aplicação e de interpretação – possuem limitações, principalmente no que tange as covariáveis, pois alguns grupos podem ter poucas observações, fazendo com que as comparações sejam inviabilizadas (COLOSIMO, GIOLO, 2006, p. 116).

Uma solução para tal fato está na utilização de um modelo de regressão apropriado para dados censurados, onde é possível acondicionar os efeitos das covariáveis⁴¹. Tais modelos, em análise de sobrevivência, podem ser classificados como paramétricos ou semiparamétricos (COLOSIMO, GIOLO, 2006, p. 116).

Nos modelos paramétricos, deve-se especificar a distribuição do tempo dos eventos, como exemplo, pode-se citar o modelo exponencial, o modelo de Weibull, Gompertz e Logística. Já nos modelos semiparamétricos, temos o modelo de riscos proporcionais de Cox (ou simplesmente modelo de Regressão de Cox), onde atende à classificação de paramétrico por especificar um modelo de regressão com uma forma funcional específica e de não-

⁴⁰ “O teste de Wilcoxon, em particular, pondera pelo número de indivíduos sob risco que dependa da experiência de sobrevivência, assim como da de censura.” (COLOSIMO e GIOLO, 2006, p.63)

⁴¹ Também são chamadas de variáveis explicativas ou variáveis independentes, mas a partir daqui utilizaremos apenas o termo covariáveis .

paramétrico por não especificar de forma exata a distribuição dos tempos dos eventos (ALLISON, 2014, p.5).

A Regressão de Cox tem sido bastante utilizada, devido a sua flexibilidade e também por admitir a incorporação de covariáveis que se modificam com o tempo de forma fácil (COLOSIMO, GIOLO, 2006, p. 116).

As covariáveis que se modificam com o tempo⁴² são as variáveis que podem assumir valores diferentes em tempos diferentes do estudo e, na análise de sobrevivência, permite que se investigue como uma mudança na variável x no tempo t influencia a propensão a ocorrer o evento num tempo seguinte ao tempo t . Tais covariáveis podem ser referentes às características individuais das unidades de análise ou referentes a fatores contextuais que operem a nível macro (BERNARDI, 2006, p.24).

Retornando à Regressão de Cox, como ela pertence à família dos modelos proporcionais, pressupõe que os efeitos das covariáveis produzem somente o deslocamento da distribuição dos tempos dos eventos para cima ou para baixo no gráfico, mas que não modificam a sua forma, ou seja, os efeitos se mantêm constantes ao longo do tempo observado (BERNARDI, 2006, p. 87). Por isso, cada uma das covariáveis utilizada na modelagem, deve ter sua proporcionalidade testada, pois “sua violação pode acarretar sérios vícios na estimação dos coeficientes do modelo” (COLOSIMO, GIOLO, 2006, p. 166 *apud* STRUTHERS, KALBFLEISCH, 1986).

Allison (2014, p. 43) sugere que o teste de resíduos de Schoenfeld é a técnica mais simples para avaliar se as covariáveis atendem ou não este pressuposto. O autor ainda esclarece que se os riscos são realmente proporcionais, o teste deverá indicar a não correlação das covariáveis com o tempo ou com qualquer uma função de tempo.

De forma geral, de acordo com Lima *et al.* (2012), a Regressão de Cox pode ser definida pela seguinte expressão:

$$h(t, x, B) = h_0(t).r(x, B), \quad \text{com } r(x, B) = e^{B.x}$$

Onde,

$h_0(t)$ é a taxa de base de risco

x é o vetor das covariáveis

B é o vetor dos parâmetros ajustáveis

⁴² Pode ser conhecida também como covariáveis dependentes do tempo, conforme Colosimo e Giolo (2006, p. 116).

Sendo assim, o efeito de cada covariável é de acelerar (quando B é positivo) ou de desacelerar (quando B é negativo) a função de risco e sua melhor interpretação é feita através da razão de risco no tempo ($RR = e^{Bx}$), isto é, quando o a razão de risco (RR) é maior do que um (B positivo), significa que as covariáveis aumentam o risco ao evento e quando RR é menor do que um (B negativo), as covariáveis diminuem o risco ao evento. Por exemplo, um $RR = 1,29$, significa que o risco é aumentado em 29% ($1,29 - 1,00$) quando comparado à referência e já o $RR = 0,81$, significa que o risco é diminuído em 19% ($1,00 - 0,81 = 0,19$).

O último ponto que merece ser esclarecido acerca da análise de sobrevivência é que as predições sobre quando o evento acontecerá no futuro são muito limitadas e não confiáveis, especialmente quando estão fora do intervalo de dados observados, pois não se sabe como a função de sobrevivência se comporta após o período do estudo (ALLISON, 2014, p. 50-51). Portanto, a Regressão de Cox é uma função de sobrevivência predita, cabendo apenas para os intervalos de tempos observados.

5.2 BANCOS DE DADOS

A dissertação utiliza dados secundários, provenientes do Sistema de Gestão Acadêmica (SGA), do Sistema de Desempenho Escolar (Desesc) e da Assessoria Técnica de Informações Educacionais Estratégicas (AIEE) da Secretaria Municipal de Educação da Cidade do Rio de Janeiro (SME-Rio) e indicadores disponibilizados pelo do MEC/INEP.

O banco de dados do SGA traz informações pessoais dos alunos, assim como toda a sua trajetória escolar a cada bimestre, incluindo seu desempenho acadêmico, sua frequência escolar dentro da rede de ensino.

Já no banco de dados do Desesc – que teve seu início no 2º bimestre de 2012 – é possível encontrar os resultados dos alunos nas avaliações de rede, conhecidas como “Prova Bimestral”. Essas avaliações são compostas por questões de múltipla escolha, elaboradas pela própria SME-Rio, e aplicadas a todos os alunos do ensino fundamental. Tem como objetivo principal mensurar os descritores trabalhados ao longo de cada bimestre. No entanto, como não se trata de uma avaliação padronizada, o estudo da evolução (comparações no tempo) das mesmas deve ser evitado. Ainda assim, é uma ferramenta que não deveria ser descartada por pesquisadores do campo da educação, uma vez que já está

consolidada na Rede, e seus resultados podem ser bons preditores para repetência escolar. Nesse sentido, os resultados obtidos no primeiro ou segundo bimestres poderiam indicar os alunos em risco, ou seja, com maior probabilidade de reprovação ao final do ano letivo corrente. E é justamente por sua regularidade, pois são realizadas com todas as turmas dos Anos Finais do ensino fundamental em todos os bimestres, que se optou em utilizar as provas bimestrais como uma das variáveis de desempenho acadêmico dos alunos.

A AIEE forneceu as informações sobre as características das escolas e seus históricos, como por exemplo, se a escola pertenceu a algum programa/projeto desenvolvido na SME-Rio e durante qual(is) ano(s) letivo(s).

Quanto ao MEC/INEP, optou por utilizar o indicador de complexidade de gestão⁴³ (ICG), que é indicador que “Assume-se que a complexidade da gestão escolar se concretiza em quatro características: (1) porte da escola; (2) número de turnos de funcionamento; (3) complexidade das etapas ofertadas pela escola e (4) número de etapas/modalidades oferecidas”.

5.2.1 As variáveis

Para este estudo, utilizam-se variáveis que caracterizam o perfil socioeconômico dos alunos, sua trajetória escolar, seu desempenho escolar, políticas públicas das escolas e variáveis específicas para a modelagem da análise de sobrevivência.

Para facilitar a compreensão, foram elencadas todas as variáveis nos quadros abaixo:

(i) Variáveis dependentes (resposta):

Quadro 4: Variáveis dependentes

Variável	Tipo da Variável	Descrição da Variável
Tempo	numérica	Momento no Tempo, sendo: 1 = Ano 2012 2 = Ano 2013 3 = Ano 2014 4 = Ano 2015
Evento	categórica bivariada (dummy)	Se aluno repetiu Não = 0 Sim = 1

⁴³ Nota técnica disponível em:

<http://download.inep.gov.br/informacoes_estatisticas/indicadores_educacionais/2014/escola_complexidade_gestao/nota_tecnica_indicador_escola_complexidade_gestao.pdf>. Acesso em: 05 jun. 2018.

(ii) Variáveis independentes (covariáveis):

a) Perfil Socioeconômico dos alunos:

Quadro 5: Covariáveis – Perfil socioeconômico dos alunos

Variável	Tipo da Variável	Descrição da Variável
Idade	Numérica	Idade do aluno em 31/12/2012
Sexo	categórica bivariada (dummy)	Masculino = 0 Feminino = 1
Raça/Cor	categórica	Branca = 1 Parda = 2 Preta = 3 Amarela = 4 Indígena = 5 Sem Informação = 999
Não-branco	categórica bivariada (dummy)	Branco = 0 Não-branco = 1
NIS	categórica bivariada (dummy)	Possui NIS Não = 0 Sim = 1
Bolsa_Família	categórica bivariada (dummy)	Beneficiário do programa social Bolsa Família Não = 0 Sim = 1
Cartão Família Carioca	categórica bivariada (dummy)	Beneficiário do programa social Cartão Família Carioca Não = 0 Sim = 1
Deficiente	categórica bivariada (dummy)	Possui deficiência Não = 0 Sim = 1
Max_Esc_Pais	categórica bivariada (dummy)	Máxima Escolaridade entre os pais com no mínimo o Ensino Médio completo Não = 0 Sim = 1
Família_Biparental	categórica bivariada (dummy)	Aluno que mora com pai e mãe Não = 0 Sim = 1

b) Trajetória e desempenho escolar dos alunos

Quadro 6: Covariáveis – Trajetória e desempenho escolar dos alunos

Variável	Tipo da Variável	Descrição da Variável
PB_LPL_1	numérica	Nota na Prova Bimestral de Língua Portuguesa Leitura (LPL) do COC 1, podendo variar entre 0,0 e 10,0. Sem informação ou com falta na prova = 999.
PB_MAT_1	numérica	Nota na Prova Bimestral de Matemática (MAT) do COC 1, podendo variar entre 0,0 e 10,0. Sem informação ou com falta na prova = 999.
PB_LPL_1e2	numérica	Média das notas nas Provas Bimestrais de Língua Portuguesa Leitura (LPL) do COC 1 e COC 2, podendo variar entre 0,0 e 10,0. Para os alunos como apenas um dos resultados, utilizou-se este resultado. Sem informação ou com falta na prova = 999.
PB_MAT_1e2	numérica	Média das notas nas Provas Bimestrais de Matemática (MAT) do COC 1 e COC 2, podendo variar entre 0,0 e 10,0. Para os alunos como apenas um dos resultados, utilizou-se este resultado. Sem informação ou com falta na prova = 999.
Conc_I_1e2	numérica	Total de Conceitos I obtidos pelo aluno no COC 1 e COC 2, podendo variar entre 0 e 2. Sem informação = 999.
Freq_Anual	numérica	Percentual de frequência no ano letivo, podendo variar entre 0 e 100. Sem informação = 999.
MovTurma	numérica	Total de transferências de turma do aluno na mesma escola por ano, comparando os bimestres.
MovEscola	numérica	Total de transferências de escola do aluno por ano, comparando os bimestres.

c) Política Pública da Escola

Quadro 7: Covariáveis– Política Pública da Escola

Variável	Tipo da Variável	Descrição da Variável
6ºAExp_2012	categórica bivariada (dummy)	A escola do aluno em 2012 era de 6º Ano Experimental Não = 0 Sim = 1
GC	categórica bivariada (dummy)	A escola do aluno era de Ginásio Carioca Não = 0 Sim = 1
TU	categórica bivariada (dummy)	O turno da turma do aluno era de Turno Único Não = 0 Sim = 1
EA	categórica bivariada (dummy)	A escola do aluno era do programa Escolas do Amanhã Não = 0 Sim = 1
ICG	categórica	Índice de Complexidade de Gestão, com níveis que variam de 1 a 6.

5.2.2 O banco de dados para a Análise de Sobrevida

O banco de dados para aplicação da análise de sobrevivência deve ser construído em formato vertical, onde existe uma linha com as informações de cada aluno para cada ano em que ele sobrevive. Portanto, todos os alunos acompanhados ganharam uma linha com o Tempo 1 e suas variáveis fixas e as que modificavam com o tempo com as informações do ano letivo de 2012. Para todos aqueles que sobreviveram ao Tempo 1 – isto é, que não repetiram entre 2012 e 2013, ou que não censuraram no Tempo 1 – ganharam a Linha 2 no banco de dados, com as variáveis fixas repetidas e as variáveis que se modificam no tempo atualizadas com informações referentes ao ano de 2013. O mesmo procedimento foi repetido para os anos de 2014 e 2015. Ao final da organização do banco de dados, temos um banco cujo total de linhas é igual ao total de alunos sobreviventes, multiplicado pelo tempo que eles permaneceram no fluxo regular. Allison (2014, p. 11), trata essas unidades de medida como sendo pessoas-ano.

6 RESULTADOS

6.1 ANÁLISE DESCRITIVA DAS VARIÁVEIS

Nesta parte da dissertação, serão apresentados os resultados descritivos das variáveis estudadas. Como é grande o número de variáveis e também para facilitar a compreensão e auxiliar na organização do estudo, as variáveis foram subdivididas em três blocos temáticos: Perfil dos alunos, Desempenho acadêmico dos alunos e Políticas Públicas das escolas.

6.1.1 Análise Descritiva do Perfil dos Alunos

A descrição do perfil dos alunos tem como objetivo principal conhecer melhor a coorte estudada, em especial as características dos alunos e suas famílias que, segundo a literatura educacional, estão associadas com desempenho e trajetória escolar (GORARD; SEE, 2013)

Para a caracterização dos alunos, utilizaremos como variáveis: sexo, idade, raça e deficiência; para caracterizar as famílias, utilizaremos as variáveis referentes à máxima escolaridade dos pais, se o aluno mora com ambos os responsáveis (pai e mãe) e se a família é beneficiária de algum programa social de transferência de renda.

Vale ressaltar que não há como afirmar qual período do tempo exato algumas dessas informações foram coletadas, podendo ter sido no momento da matrícula do aluno na rede, no momento de alguma transferência interna ou até em um momento qualquer que o responsável do aluno desejasse atualizar as informações de seu cadastro. Isso provavelmente aumenta a imprecisão do registro da informação (erro), mas nem por isso, podemos renegar a importância de tais variáveis na caracterização das famílias. Outras bases de dados disponíveis, por exemplo, o Censo Escolar, não possuem muitas algumas das variáveis descritas no estudo⁴⁴ e, em muitos casos, a qualidade da informação é inferior ao observado no SGA (BARTHOLO, 2014).

Vamos começar a análise pela variável sexo que, conforme representado no Gráfico 13, 29.985 alunos são do sexo masculino, enquanto 28.034 são do sexo feminino, demonstrando assim que a nossa amostra está equilibrada em relação a esta variável.

⁴⁴ Um bom exemplo é a informação sobre a escolaridade do pai e da mãe do aluno (BARTHOLO, 2014).

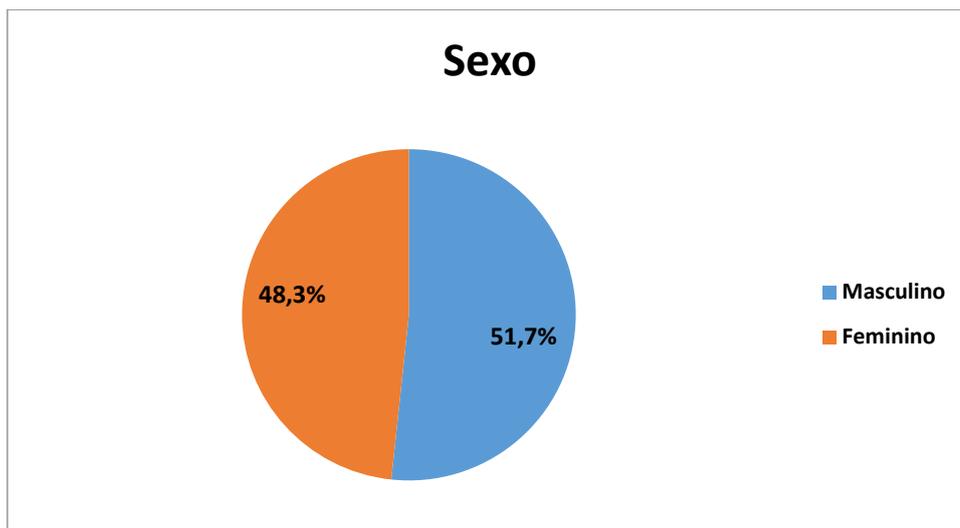


Gráfico 13: Sexo dos alunos

Para a variável idade, utilizamos como base o ano inicial do estudo, seguindo a Portaria de enturmação da SME-Rio para o ano de 2012⁴⁵, que tem como parâmetro a data de referência de 31 de março de 2012. A coorte apresentou idade entre 9 e 31 anos, tendo como média 11,58 anos \pm 0,966, conforme apresentado na Tabela 1.

Tabela 1: Idade em 2012

Idade em 2012								
	N	Amplitude	Mínimo	Máximo	Média		Desvio Padrão	Variância
	Estat.	Estat.	Estat.	Estat.	Estat.	Erro Padrão	Estat.	Estat.
Idade em 2012	58.019	22	9	31	11,58	0,004	0,966	0,934

A idade ideal para o 6º ano é 11 anos, levando em consideração que o aluno deveria ter seis anos completos em 2007 para iniciar o ensino fundamental. O Gráfico 14 apresenta a distribuição por idade da coorte e podemos observar que 59,3% dos alunos está dentro da idade ideal (até 11 anos), 27,5% tinham um ano a mais do que a idade ideal e 13,2% dos alunos tinham 2 anos ou mais da idade ideal.

⁴⁵ Portaria E/SUBE/CED nº 09, de 18 de novembro de 2011. Disponível em: http://doweb.rio.rj.gov.br/visualizar_pdf.php?edi_id=420&page=14

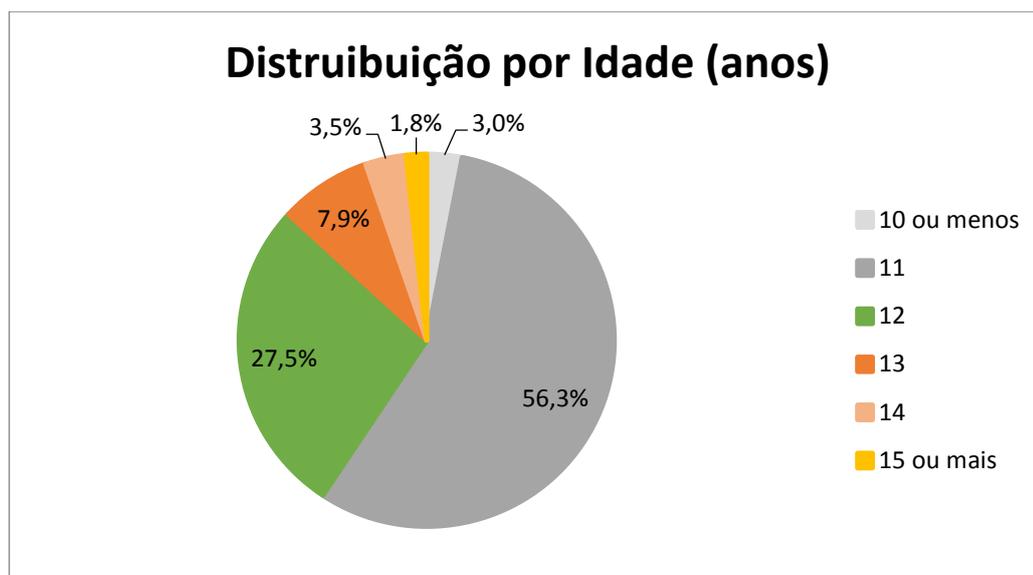


Gráfico 14: Idade da Coorte em 2012

A definição do MEC indica que o aluno está em situação de defasagem idade-série quando apresenta 2 anos ou mais de idade em relação à idade ideal, portanto, 7.651 alunos (13,2%) dos alunos da coorte são considerados defasados. Com o objetivo de conhecer a intensidade do fenômeno para a coorte analisada, calculamos os “anos de defasagem”, considerando a diferença entre a idade do aluno em 2012 e a idade ideal para o mesmo ano. Os resultados são apresentados na Tabela 2. Os dados indicam média de $2,57 \pm 0,842$ anos de defasagem, variando entre 2 e 20 anos.

Tabela 2: Análise Descritiva de Anos de Defasagem em 2012

Anos de Defasagem em 2012								
	N	Amplitude	Mínimo	Máximo	Média		Desvio Padrão	Variância
	Estat.	Estat.	Estat.	Estat.	Estat.	Erro Padrão	Estat.	Estat.
Anos de Defasagem	7.651	18	2	20	2,57	0,01	0,842	0,71

No Gráfico 15, nota-se que 60% dos alunos estão com o tempo mínimo de defasagem (2 anos), outros 27% tem 3 anos de defasagem, 10% tem 4 anos de defasagem e os 3% restante tem 5 anos ou mais de defasagem. Desses alunos, 62,7% eram do sexo masculino.

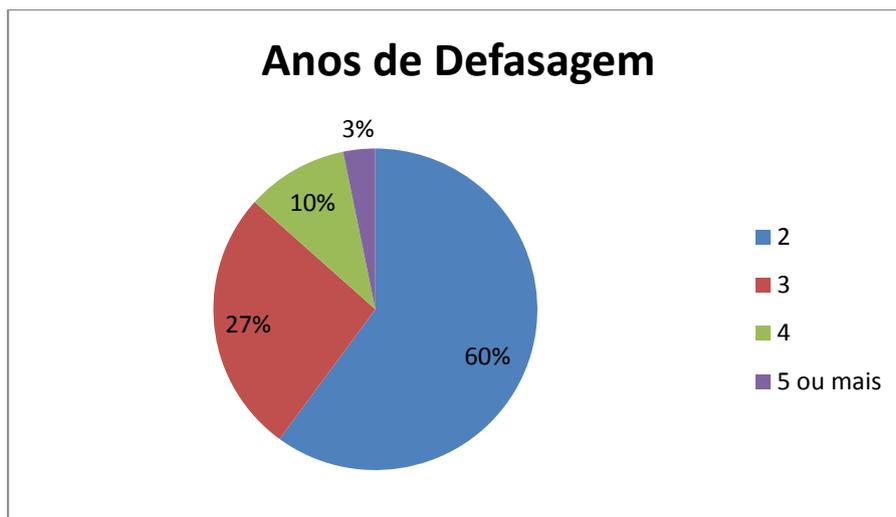


Gráfico 15: Anos de Defasagem (Apenas para os alunos com Defasagem)

Os dados para a variável cor/raça indicam que a maioria dos alunos é declarada como da Raça “Parda” (49,3%) ou “Branca” (34,6%); 11,5% dos alunos são negros e 0,3% são amarelos ou indígenas. Outros 4,3% dos alunos não têm sua raça em seu cadastro, tendo sido considerado como *missing* para este estudo. Na Tabela 3, encontramos esta estatística de forma mais detalhada.

Tabela 3: Raça dos alunos

		Raça			
		Frequência	Percentual	Percentual Válido	Percentual Acumulado
Válido	Branca	20.060	34,6	36,1	36,1
	Parda	28.585	49,3	51,5	87,6
	Preta	6.681	11,5	12,0	99,7
	Amarela	105	0,2	0,2	99,8
	Indígena	84	0,1	0,2	100,0
	Total	55.515	95,7	100,0	
Missing	999	2.504	4,3		
Total		58.019	100,0		

Por questões metodológicas, optamos também por utilizar uma variável dicotômica (*dummy*) que expressasse apenas os “não brancos”. Alunos brancos foram codificados como 0 (zero) e todas as outras recebem a codificação 1 (um).

O Gráfico 16 apresenta a distribuição das raças pelo sexo, demonstrando um equilíbrio entre o sexo, mesmo quando controlado pela raça dos alunos.

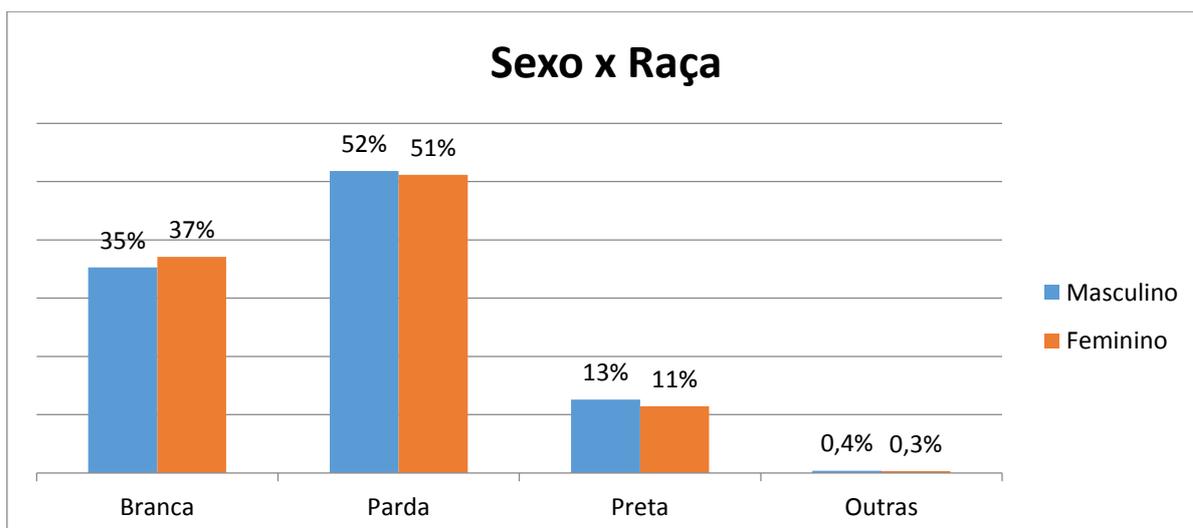


Gráfico 16: Sexo vs Raça dos alunos

Em relação à inclusão, foram encontrados apenas 452 alunos incluídos no 6º ano em 2012, representando 0,8% da amostra, os tipos de deficiências encontrados estão listados na Tabela 4, onde observamos que a deficiência intelectual é a mais comum entre as deficiências. Pelos dados solicitados, não temos como aprofundar sobre a vida escolar desses alunos, como se, por exemplo, recebiam Atendimento Educacional Especializado, conforme previsto no Estatuto da Pessoa com Deficiência⁴⁶. Devido ao pequeno número de casos, optamos apenas por utilizar uma variável dicotômica (dummy) indicando se o aluno era incluído ou não, não considerando o tipo de deficiência do mesmo, embora suponhamos que o tipo de deficiência ter influência na trajetória escolar dos alunos.

Tabela 4: Tipo de deficiência - Alunos incluídos

		Tipo de Deficiência			
		Frequência	Percentual	Percentual Válido	Percentual Acumulado
Válido	Sem deficiência	57.567	99,2	99,2	99,2
	Deficiência Física	82	0,1	0,1	99,4
	Deficiência Intelectual	176	0,3	0,3	99,7
	Deficiência Visual	37	0,1	0,1	99,7
	Deficiência Auditiva	64	0,1	0,1	99,8
	Transtorno Global de Desenvolvimento	75	0,1	0,1	100,0
	Altas Habilidades	6	0,0	0,0	100,0
	Deficiência Múltipla	12	0,0	0,0	100,0
	Total	58.056	100,0	100,0	

⁴⁶ Lei nº 13.146, de 6 de julho de 2015. Disponível em: < http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/lei/l13146.htm>. Acesso em: 05 jun. 2018.

Para caracterizar a família do aluno, construímos duas variáveis dicotômicas a partir de informações presentes no cadastro dos alunos: família biparental e máxima escolaridade dos pais. A primeira foi construída a partir da informação se o aluno morava ou não com ambos os responsáveis – os alunos que moravam com apenas um dos responsáveis ou com nenhum deles foram codificados como 0 (zero). Vale esclarecer que essa informação é um *checkbox* no SGA podendo ou não ser assinalado, portanto, é possível que o valor esteja superestimado, pois não podemos afirmar categoricamente que se por não estar assinalada é porque realmente o aluno não mora com seus pais ou porque seria um *missing* neste campo.

A outra variável foi construída a partir da informação de escolaridade de cada um dos pais. Para a construção, os seguintes procedimentos foram feitos: primeiro, a escolaridade de cada responsável foi codificada; em seguida, foi computada a maior escolaridade entre as duas, nos casos de *missing* de um dos responsáveis, fica valendo a do outro responsável exclusivamente; o último passo foi codificar 1 para aqueles que tiveram sua máxima escolaridade o “Ensino Médio Completo” pelo menos⁴⁷ e 0 para outros. Os casos de *missing* para a escolaridade de ambos os pais foram mantidos como *missing*.

A Tabela 5 apresenta a frequência da variável biparental, demonstrando que 44,1% dos alunos moram com ambos os responsáveis. Já na Tabela 6 encontramos que 31,2% dos alunos têm com maior escolaridade de seus pais o ensino médio incompleto e 55,7% pelo menos um de seus pais tem, ao menos, o ensino médio completo. Não podemos deixar de registrar que temos 13,1% de *missing* nessa variável.

Tabela 5: Família biparental

Família Biparental					
		Frequência	Percentual	Percentual Válido	Percentual Acumulado
Válido	Não	32.407	55,9	55,9	55,9
	Sim	25.612	44,1	44,1	100,0
	Total	58.019	100,0	100,0	

⁴⁷ Inclusive os de ensino superior completo e incompleto

Tabela 6: Máxima Escolaridade de Pais

Máxima Escolaridade de Pais com Ensino Médio Completo					
		Frequência	Percentual	Percentual Válido	Percentual Acumulado
Válido	Não	32.333	55,7	64,1	64,1
	Sim	18.089	31,2	35,9	100,0
	Total	50.422	86,9	100,0	
Missing	9	7.597	13,1		
Total		58.019	100,0		

Pensando na caracterização do perfil socioeconômico das famílias, apresentamos outras três variáveis dicotômicas na dissertação. A primeira variável informa se o aluno possui ou não o Número de Identificação Social (NIS)⁴⁸ em seu cadastro. Com essa informação, podemos afirmar apenas que esse aluno já foi – ou ainda é – beneficiário de algum programa social, mas não sabemos qual, quando, ou por quanto tempo esse benefício é recebido. É importante que estudos futuros analisem a pertinência de medir o NIS, Bolsa Família ou Cartão Família Carioca não apenas como uma variável dicotômica (já foi exposto ou não), mas sim como uma variável categórica que indica quantos anos o aluno foi exposto à situação de desvantagem. Estudos realizados na Inglaterra com os dados sobre *Free School Meal* sugerem um incremento importante no poder explicativo da variável utilizando a mesma como categórica (GORARD, 2012; TAYLOR, 2017)

As outras variáveis, também *dummies*, são para os beneficiários dos programas Bolsa Família e Cartão Família Carioca. A mesma limitação citada anteriormente cabe para estas variáveis, como essas informações são assinaladas apenas por *checkbox*, não temos o histórico desses benefícios, sabemos apenas que no momento da consulta dos dados, em 2015, essas opções estavam assinaladas.

Na Tabela 7, verificamos que 39,9% da amostra são beneficiários do Programa Bolsa Família, e na Tabela 8, que 24,5% são beneficiários do Programa Cartão Família Carioca.

⁴⁸ O NIS é fornecido pela Caixa Econômica Federal às pessoas beneficiadas de algum projeto social.

Tabela 7: Bolsa Família

Bolsa Família					
		Frequência	Percentual	Percentual Válido	Percentual Acumulado
Válido	Não	34.861	60,1	60,1	60,1
	Sim	23.158	39,9	39,9	100,0
	Total	58.019	100,0	100,0	

Tabela 8: Cartão Família Carioca

Cartão Família Carioca					
		Frequência	Percentual	Percentual Válido	Percentual Acumulado
Válido	Não	43.803	75,5	75,5	75,5
	Sim	14.216	24,5	24,5	100,0
	Total	58.019	100,0	100,0	

6.1.2 Análise Descritiva do Desempenho Acadêmico dos Alunos

Nesta seção, vamos descrever as variáveis referentes ao desempenho acadêmico da coorte, com resultados de avaliações internas (Provas Bimestrais), quantidade de conceitos 'I' (Insuficiente), frequência anual dos alunos e quantidade de movimentações entre turmas e entre escolas. Todas essas variáveis foram caracterizadas como variáveis que se modificam no tempo, isto é, os valores das mesmas não são fixos, podendo variar de ano para ano.

O objetivo é tentar captar de alguma forma o desempenho acadêmico dos alunos da nossa coorte ao longo dos anos finais do ensino fundamental. Optamos por não utilizar as Provas Bimestrais e os conceitos "I" de todos os bimestres do ano, justamente por entender que poderia haver uma colinearidade entre tais variáveis e o resultado final do aluno para cada ano. Sendo assim, utilizamos os resultados apenas do primeiro semestre para cada ano de cada variável (primeiro e segundo bimestres).

Para a Prova Bimestral, trabalhamos com quatro variáveis, sendo duas delas para as provas bimestrais de Língua Portuguesa Leitura (PB_LPL_1) e Matemática (PB_MAT_1) do COC 1 e, as outras duas, para a média entre as provas bimestrais do COC 1 e 2, também para as mesmas disciplinas. Para essas variáveis, quando o aluno não tinha uma das duas notas, foi utilizado apenas a única nota que o aluno possuía. Para os alunos em que as notas não foram encontradas na base de dados ou para aqueles que faltaram à prova, deixamos no banco de dados como *missing*.

Para cada ano analisado, os resultados são referentes àquela série acompanhada naquele ano, portanto, os resultados de 2012 são das provas bimestrais do 6º ano, os resultados de 2013 são das provas do 7º ano, os de 2014 são das provas do 8º ano e os de 2015 são das provas de 9º ano.

Na Tabela 9 é possível ver as médias, os desvios padrões e a quantidade de notas que compõem essas médias para cada uma das notas. Como já dito anteriormente, as provas bimestrais não são padronizadas e por isso não devem ser comparadas em momentos diferentes do tempo, mas, ainda assim, o fato das notas de Matemática serem sempre menores que as de Língua Portuguesa Leitura chama a atenção.

Tabela 9: Médias das Provas Bimestrais

PROVAS BIMESTRAIS					
Tempo		PB LP-Leitura 1º COC	PB Matemática 1º COC	PB LP-Leitura 1º e 2º COC	PB Matemática 1º e 2º COC
2012 (6º ano)	Média	7,437	6,900	7,126	6,490
	N	53.539	54.069	56.519	56.582
	Desvio Padrão	2,2638	1,9455	2,0040	1,8149
2013 (7º ano)	Média	6,248	4,907	5,967	4,708
	N	39.953	39.923	41.226	41.219
	Desvio Padrão	1,8918	1,6651	1,6791	1,5361
2014 (8º ano)	Média	5,438	3,498	4,866	3,687
	N	31.927	31.199	33.249	33.050
	Desvio Padrão	1,7978	1,7317	1,5776	1,4585
2015 (9º ano)	Média	6,184	4,094	5,898	4,225
	N	27.056	27.052	27.991	27.941
	Desvio Padrão	1,9347	1,9464	1,6795	1,8486

Ao final de cada bimestre, o aluno recebe um conceito global⁴⁹, este conceito é dado em conjunto por todos os professores do aluno, podendo ser: Muito Bom (MB), Bom (B), Regular (R) e Insuficiente (I), ficando retido o aluno que tirar o conceito I ao término do ano letivo.

Para a variável sobre conceito, utilizamos a soma de conceitos 'I' que os alunos tiveram no COC 1 e 2 de cada ano, portanto, se o aluno não teve conceito I nestes bimestres,

⁴⁹ Resolução SME n.º 1123, de 24 de janeiro de 2011, estabelece diretrizes para a avaliação escolar na Rede Pública do Sistema Municipal de Ensino da Cidade do Rio de Janeiro e dá outras providências. Disponível em: <http://doweb.rio.rj.gov.br/visualizar_pdf.php?edi_id=486&page=6>. Acesso em: 05 jun. 2018.

o valor da variável é 0, se ele teve conceito I em apenas um dos bimestres, o valor recebido foi 1 e o valor 2 foi para o caso dele ter recebido conceito I em ambos os bimestres. Analisando os resultados da Tabela 10, observamos que não há grandes variações entre as médias para cada ano, mesmo quando há uma variação na quantidade de alunos presentes em cada tempo.

Tabela 10: Quantidade de Conceito I no COC 1 e 2, por ano.

Nº Conceito I – 1º e 2º COC			
Tempo	Média	N	Desvio Padrão
2012	0,46	58.019	0,738
2013	0,45	41.440	0,722
2014	0,39	33.707	0,657
2015	0,38	28.134	0,666

A frequência escolar é outro ponto importante do desempenho acadêmico do aluno, pois, conforme estabelecido a LDB 9394/96, a aprovação está atrelada à frequência mínima de 75% do total de horas letivas entre outros fatores (BRASIL, 1996).

As médias das frequências anuais da presente coorte parecem ser bem estáveis ao longo dos quatro anos acompanhados, conforme demonstrado na Tabela 11.

Tabela 11: Frequência Anual dos alunos, por ano.

Frequência Anual dos alunos			
Tempo	Mean	N	Std. Deviation
2012	93,28	53.996	6,542
2013	93,84	41.440	7,044
2014	93,49	33.683	5,793
2015	93,17	28.134	5,661

As transferências dos alunos ao longo de sua trajetória escolar, em especial, quando acontece no decorrer do ano letivo, podem influenciar o emocional, o social e até o desempenho escolar dos alunos (SILVA, 2015), por esta razão, duas variáveis acerca deste assunto foi calculada para este estudo: uma para retratar a movimentação dos alunos entre turmas da mesma escola e a segunda para contabilizar a movimentação dos alunos entre escolas da rede.

Para construir estas variáveis utilizou-se a seguinte lógica: comparava a escola do aluno no COC 1 com a escola do mesmo aluno no COC 2, se as escolas fossem diferentes, contava 1 movimentação para a variável MovEscola; se caso as escolas fossem iguais, mas

quem diferenciava era a turma, contava 1 movimentação para a variável MovTurma; e, por fim, caso não houvessem diferenças nem entre escolas, nem entre turmas, não somava nenhuma movimentação nas duas variáveis. O mesmo procedimento foi repetido entre os COCs subsequentes (COC 2 x COC 3 e COC 3 x 4), portanto, para cada ano observado, cada variável podia variar de 0 a 3 movimentações por ano. Uma limitação para nosso estudo que merece ser esclarecida é que, por questões de tempo, optamos em não trabalhar com o banco de movimentações do SGA, as movimentações foram computadas apenas entre bimestres, olhando a turma e a escola dos alunos em cada bimestre do ano letivo, por isso, é provável que esta variável esteja subestimada, já que entre os COCs podem ter havido movimentações que não foram captadas, quando se analisava apenas a escola que o aluno estava matriculado ao final de cada bimestre.

Os resultados apresentados na Tabela 12, apresentam resultados estáveis ao longo dos períodos, tendo sido maiores as médias entre as movimentações no ano letivo de 2014.

Tabela 12: Movimentações por ano

Movimentações			
Tempo		Movimentações Bimestrais Entre Escolas	Movimentações Bimestrais Entre Turmas
2012	Média	0,018	0,032
	N	58.019	58.019
	Desvio Padrão	0,1365	0,1797
2013	Média	0,018	0,030
	N	41.440	41.440
	Desvio Padrão	0,1374	0,1740
2014	Média	0,022	0,047
	N	33.707	33.707
	Desvio Padrão	0,1533	0,2172
2015	Média	0,013	0,033
	N	28.134	28.134
	Desvio Padrão	0,1152	0,1860

6.1.3 Análise Descritiva das Políticas Públicas das Escolas

Nesta seção, a análise será sobre as variáveis referentes às Políticas Públicas implementadas pela da SME-Rio durante o período observado neste estudo, selecionamos quatro programas/projetos: 6º Ano Experimental, Ginásio Carioca, Escolas do Amanhã e Turno Único. Todos foram apresentados no Capítulo 4 desta dissertação.

Todas as variáveis são dicotômicas e foram criadas considerando a escola que o aluno estava matriculado no COC4 de cada ano ou, nos casos de censuras, na última escola que o aluno esteve matriculado antes de sair do ensino fundamental regular ou da SME-Rio. Optamos definir a participação do aluno nestas Políticas Públicas no último período do ano, por entender que o nosso evento de interesse (repetência) – mesmo que ainda seja o desfecho final de um ano letivo inteiro – é determinado pela escola que o mesmo estava matriculado no final do ano letivo.

O 6º Ano Experimental é um projeto exclusivo, como o próprio nome diz, para o 6º ano do ensino regular e por isso, os alunos que terminaram o ano letivo de 2012 em uma escola pertencente a este projeto, ganhou o código 1 na sua variável e os outros alunos o código 0, e esta mesma codificação seguiu como os alunos ao longo dos anos de estudo, sendo considerada como uma variável que não se modifica com o tempo (variável fixa).

As variáveis dos outros programas tiveram codificações próprias ano a ano, tendo sido consideradas para o modelo, como variáveis que se modificam no tempo. A seguir um exemplo para tornar mais claro o procedimento: se o aluno estava numa escola de Ginásio Carioca em 2013, ele recebeu 1 para este período, mas se em 2014 ele transferiu para uma escola que não era Ginásio Carioca, ele recebe 0 para o período seguinte; então, o mesmo aluno pode ter 0 ou 1 em tempos de análise diferentes.

As variáveis sobre 6º Ano Experimental, Ginásio Carioca e Escolas do Amanhã foram construídas a partir de uma lista de escolas por programa, fornecida pelo AIEE. A variável sobre Turno Único foi criada a partir do turno que o aluno estava matriculado⁵⁰, seguindo a mesma lógica utilizada para as outras variáveis, conforme já explicitado anteriormente.

É pertinente esclarecer que esses programas não são excludentes, isto é, o aluno que está em um desses programas pode perfeitamente fazer parte de outro programa sem problema, inclusive todos os alunos do Ginásio Carioca também são do Turno Único, pois o tempo integral é uma das premissas do Ginásio Carioca.

Em 2012, o 6º Ano Experimental ainda era um programa incipiente, estando no seu segundo ano de implantação e abrangendo apenas 61 escolas. Na coorte estudada, apenas 3.693 alunos eram de escolas de 6º Ano Experimental, o que representa apenas 6,4% da amostra inicial. No Gráfico 17, apresentamos a distribuição dos alunos que tinham estudado, pelo menos o último COC de 2012, em uma escola com o projeto, ao longo dos quatro anos

⁵⁰ Esta informação estava presente no bando de dados do SGA fornecido pela SME-Rio

(tempo), lembrando que esta variável é fixa, não havendo alterações de ano para ano, portanto, o aluno considerado do 6º Ano Experimental em 2012 é o mesmo em 2013, 2014 e 2015.

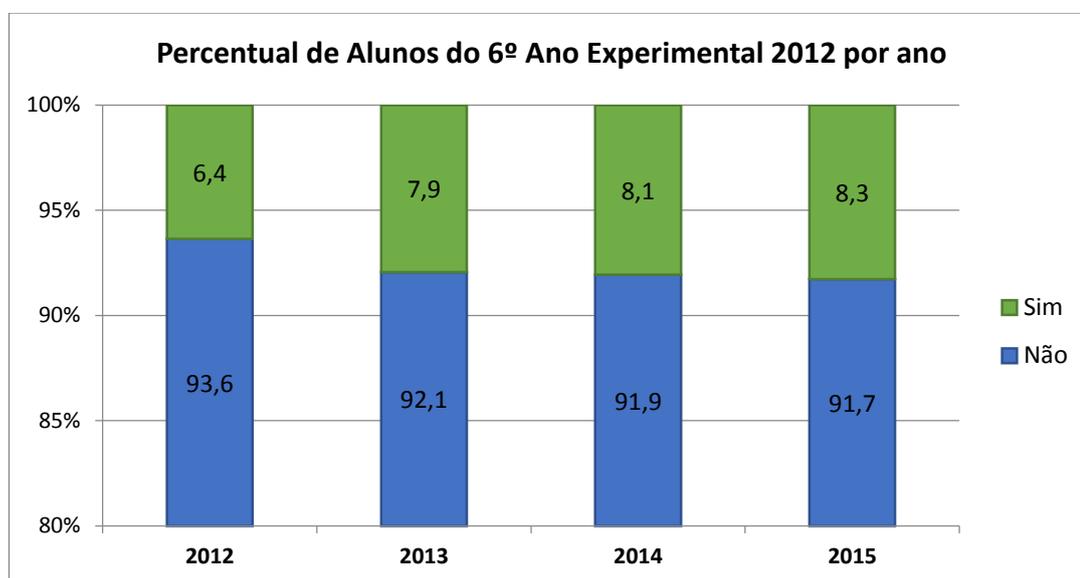


Gráfico 17: Percentual de alunos do 6º Ano Experimental 2012, por ano.

O percentual do 6º Ano Experimental 2012 vai aumentando de forma gradativa ao longo dos quatro anos (tempo), isto é, os alunos que eram do 6º Ano Experimental em 2012 representam cada vez mais em relação a todos os alunos que prosseguiram no fluxo regular da rede. Mas, apesar desses dados, neste momento da análise, ainda é precoce assumir qualquer conclusão em relação ao projeto.

O Ginásio Carioca também estava em seu segundo ano de implantação em 2012, tendo apenas 29 escolas. Como o projeto é voltado para alunos do 7º ao 9º ano, exceto nos Ginásios Experimentais Olímpicos (GEO) que também possuem turmas de 6º ano, em 2012 temos poucos alunos nele. Nos anos seguintes, o percentual de alunos da coorte fica em torno de 6%, com aumentos gradativos. É válido deixar claro que o número de escolas do projeto também foi aumentando ano a ano, fato este que pode estar influenciando nesta distribuição.

Vale ainda também reforçar o fato de estarmos acompanhando neste estudo apenas os alunos que estavam no 6º ano em 2012 na rede, que não saíram por um bimestre inteiro e que continuaram seu fluxo no ensino regular como esperado ano a ano. Portanto, o que

temos na Tabela 13, não é um panorama geral do projeto, mas sim, um panorama do acompanhamento da nossa coorte em relação ao projeto somente.

Tabela 13: Distribuição da coorte no Ginásio Carioca

Ginásio Carioca					
		Ano			
		2012	2013	2014	2015
Ginásio Carioca	Não	58.000 (99,97%)	38.992 (94,1%)	31.580 (93,7%)	26.268 (93,4%)
	Sim	19 (0,03%)	2.448 (5,9%)	2.127 (6,3%)	1.866 (6,6%)
Total		58.019	41.440	33.707	28.134

Em relação ao programa Escolas do Amanhã, a distribuição dos alunos ao longo dos anos se apresenta em ligeiro declínio, conforme apresentado na Tabela 14, entretanto, diferente do Ginásio Carioca, o número de escolas não variou ao longo dos anos, tendo sido as mesmas 155 escolas em todo o programa.

Tabela 14: Distribuição da coorte no Programa Escolas do Amanhã

Escolas do Amanhã					
		Ano			
		2012	2013	2014	2015
Escolas do Amanhã	Não	49.315 (85,0%)	36.283 (87,6%)	29.583 (87,8%)	24.763 (88,0%)
	Sim	8.704 (15,0%)	5.157 (12,4%)	4.124 (12,2%)	3.371 (12,0%)
Total		58.019	41.440	33.707	28.134

A

Tabela 15 apresenta a distribuição dos alunos pelas turmas de Turno Único, isto é, por aquelas turmas que no banco de dados vieram com o código "I" (integral). A ampliação da política de tempo integral na rede também foi adotada e ampliada ao longo do período analisado na dissertação.

Tabela 15: Distribuição da coorte para o Turno Único

		Turno Único			
		Ano			
		2012	2013	2014	2015
Turno Único	Não	56.870 (98,0%)	38.250 (92,3%)	30.832 (91,5%)	25.668 (91,2%)
	Sim	1.149 (2,0%)	3.190 (7,7%)	2.875 (8,5%)	2.466 (8,8%)
Total		58.019	41.440	33.707	28.134

E por fim, o Gráfico 18 apresenta as quatro variáveis que compõem essa seção, de forma resumida, onde podemos observar a distribuição da coorte por cada um dos programas/projetos contemplados neste estudo, ao longo do período pesquisado.

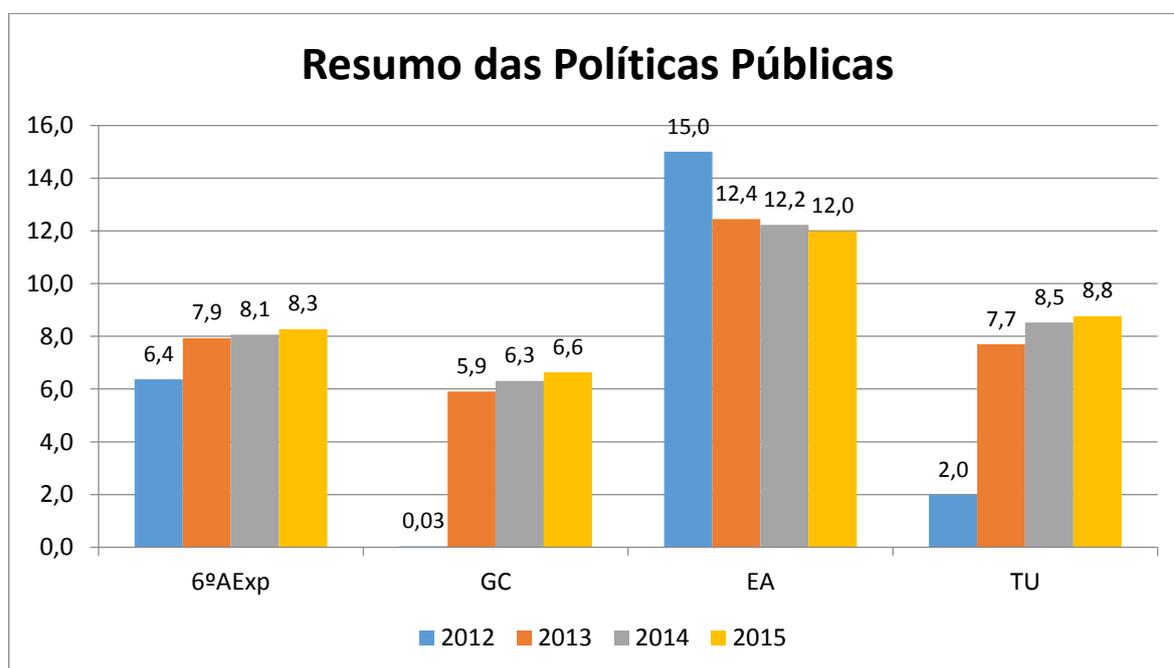


Gráfico 18: Resumo das Políticas Públicas

6.2 ANÁLISE DESCRITIVA DA TRAJETÓRIA ESCOLAR DOS ALUNOS

A seção apresenta dados descritivos para a trajetória da coorte estudada. Cada aluno foi acompanhado bimestre a bimestre, durante os anos de 2012, 2013, 2014 e 2015, tendo sido verificado qual série o aluno estava matriculado. Vale lembrar que os dados não são

acumulativos, portanto, um aluno censurado ou repetente é retirado da amostra, não sendo mais acompanhado a partir deste momento. Segue um exemplo, para facilitar a compreensão: um aluno que sai da SME-Rio no COC 2 de 2012 não será mais acompanhado neste estudo, mesmo que retorne para a Rede. O mesmo ocorre com os alunos da coorte que sejam reprovados em qualquer um dos anos investigados, reclassificados para outra série ou ainda encaminhados para qualquer um dos programas de correção de fluxo existentes na Rede. A seguir, a trajetória da coorte é apresentada ano a ano.

Na Tabela 16, podemos observar de forma mais detalhada as censuras ocorridas em cada bimestre do ano 2012. Observa-se que 3 alunos foram reclassificados para o 7º ano, 9 alunos foram encaminhados para Projetos de Correção de Fluxo (Realfabetização 1, Realfabetização 2B, Aceleração 1 e Aceleração 2A), 31 alunos mudaram de modalidade para a Educação de Jovens e Adultos da rede (PEJA), 1 aluno foi encaminhado para a Classe Especial (CE) e 3.728 alunos saíram da rede, podendo ter sido por abandonos ou transferências externas. Como não foi utilizado o Banco de Movimentação da SME-Rio não podemos afirmar o tipo de movimentação de cada aluno. Essa é uma limitação desse estudo, no entanto, análises sobre a movimentação dos alunos da SME-Rio realizadas em anos anteriores sugerem que a maior parte das saídas está relacionada à transferência dos alunos para outras redes de ensino (rede privada, outras redes públicas municipais ou rede federal de ensino) (BARTHOLO 2014a; BARTHOLO, 2014b).

Tabela 16: Trajetória Detalhada da Coorte em 2012

TRAJETÓRIA DA COORTE EM 2012									
Período	6º ano	Censurados							TOTAL
		7º ano	Realfa1	Realfa2B	Acelera2A	PEJA	CE	Saídas	
COC1	58.019	0	0	0	0	0	0	0	0
COC2	56.397	1	0	2	2	13	0	1.604	1.622
COC3	55.057	1	1	1	0	18	1	1.318	1.340
COC4	54.247	1	0	2	1	0	0	806	810
FINAL	54.247	3	1	5	3	31	1	3.728	3.772

Vale ainda ressaltar que a perda do aluno na base de dados também pode se dar por problemas no momento da extração ou ainda por uma duplicação no código único de identificação dos alunos. Esse erro impede que consigamos acompanhar a trajetória dos alunos no sistema, uma vez que dois ou mais códigos são atribuídos ao mesmo aluno e gera a impressão que o aluno abandonou a rede ou desapareceu da base de dados, para

minimizar tal situação também foram confrontados os dados do nome completo, data de nascimento e nome da mãe dos alunos, mas ainda assim, não podemos garantir que todos esses casos foram identificados.

No momento da extração dos dados para a dissertação, a SME-Rio estava fazendo migração de sua base de dados entre sistemas – saindo do SCA e passando para o SGA. Isso pode ter maximizado o problema da duplicação ou “desaparecimento” dos alunos no sistema. Para ter real proporção de tal problema, foram encontrados na base de dados 391 alunos que não tem dados para o COC2 de 2012, mas que retornam ao 6º ano da rede no COC3 e outros 99 no COC3 que retornam ao 6º ano no COC4, como não podemos afirmar se são perdas de dados na base de dados ou se tais alunos efetivamente saíram da rede e retornaram (o que seria realmente um caso de censura para este estudo), optou-se em considerar todos censurados.

Resumindo, observa-se que, 2,8% da amostra foi censurada no COC 2, 2,3% no COC 3 e 1,4% no COC 4, portanto, 93,5% da amostra cursou o 6º ano durante todo o ano letivo de 2012 na rede municipal. Os dados são consistentes com o encontrado por Bartholo (2014a; 2014b) que sugere que mais ou menos 6% da coorte migram para outras redes de ensino ao longo de um ano. Esses números foram obtidos analisando a movimentação de alunos entre 2006 e 2011 na SME-Rio.

Em 2013, bimestre a bimestre, repetiu-se o procedimento com todos os 54.247 alunos que terminaram o ano letivo de 2012. Antes de prosseguir no acompanhamento dos alunos no ano letivo de 2013, é necessário fazer um adendo sobre a enturmação dos alunos em 2013. O Quadro 8 abaixo, transcreve os critérios estabelecidos na Circular nº 166 de 28 de novembro de 2012⁵¹ para o 6º ano. Nele é possível observar que os alunos poderiam ter encaminhamentos diferentes dependendo principalmente do seu conceito final no 6º ano e da sua idade (considerados com defasagem idade-série, os nascidos até 31/03/1999).

A importância do esclarecimento destes critérios se dá, pois havia uma suspeita de que o encaminhamento dos alunos para as turmas de projetos poderia ser uma forma de repetência que não estaria sendo captada neste estudo, podendo assim, enviesar os dados aqui analisados. Confrontando então as informações do Quadro 8 com as da Tabela 17, optou-se em considerar que os 1.461 alunos encaminhados para o projeto de

⁵¹ Disponível em: <<http://portais.rioeduca.rio.gov.br/Portais/4/Arquivos/CIRCULAR%20CED%20166.doc>>. Acesso em: 05 jun. 2018.

Realfabetização 2 também como repetentes, já que conceitualmente eles haviam sido reprovados no ano letivo anterior; por esta razão, a coluna desta série foi destacada da mesma forma que a do 6º ano na Tabela 17. Para os alunos encaminhados para outros projetos de correção de fluxo considerou-se apenas censurados por sair do fluxo do ensino fundamental regular.

Quadro 8: Critérios de Enturmação - 6º Ano - 2012-2013. Fonte: SME-Rio³

CRITÉRIOS DE ENTURMAÇÃO PARA 2013 ³			
2012	CONCEITO FINAL	CRITÉRIO(S) DATA-BASE EM 2013: 31/03/2013	ENTURMAÇÃO 2013
6º ANO	MB, B ou R	não defasado, com frequência igual ou superior a 75%.	7º ANO
	I	defasado ou não, independentemente da frequência, alfabetizado	6º ANO
	I	defasado ou não, independentemente da frequência, não alfabetizado	REALFABETIZAÇÃO 2
	MB ou B	defasado (nascido até 31/03/1998), com frequência igual ou superior a 75%.	ACELERAÇÃO 3 (se houver vaga; caso negativo, 7º Ano)
	R	defasado (nascido até 31/03/1999), com frequência igual ou superior a 75%.	7º ANO ou PROJETO 7º ANO – NJM
	MB, B ou R	defasado ou não, com frequência inferior a 75%.	6º ANO (com possibilidade de reclassificação)

Portanto, foram considerados promovidos aqueles que estavam matriculados no 7º ano, repetentes aqueles que estavam no 6º ano novamente e os matriculados nas turmas de Realfabetização2 e censurados todos os outros tipos de encaminhamento ou saída da rede da SME-Rio.

Tabela 17: Trajetória Detalhada da Coorte em 2012-2013

TRAJETÓRIA DA COORTE EM 2012-2013												
Período	Repetentes		Promov.	Censurados								
	6º ano	Realf a2	7º ano	8º ano	7º NJM	Realfa 1	Acele ra 2A	Acele ra 3	PEJA	CE	Saídas	TOTAL
COC1	5.747	1.461	42.604	1	1.353	4	4	218	166	3	2.686	4.435
COC2	0	0	42.170	0	3	0	0	0	9	0	422	434
COC3	0	0	42.815	0	2	0	0	0	3	0	350	355
COC4	0	0	41.440	0	0	0	0	0	0	0	375	375
FINAL	5.747	1.461	41.440	1	1.358	0	4	218	178	3	3.833	5.599

Na Tabela 17, nota-se que 7.208 alunos foram considerados repetentes, o que representa 16,92% dos alunos que terminaram o ano letivo de 2012 ou 12,42% da coorte

inicial. O total de alunos censurados também merece atenção, pois representa 10,32% dos alunos que haviam terminado o ano letivo anterior, sendo a saída da Rede no COC 01 a razão da grande maioria dos censurados (8,02%).

Continuando a análise no ano letivo seguinte, o ano de 2014, para os 41.440 alunos que terminaram o ano letivo de 2013, isto é, que permaneciam dentro do fluxo escolar ideal (7º ano) até o final do ano, a Tabela 18 apresenta a trajetória dos mesmos em 2014.

Tabela 18: Trajetória Detalhada da Coorte em 2013-2014

TRAJETÓRIA DA COORTE EM 2013-2014									
Período	Repetentes	Promovidos	Censurados						
	7º ano	8º ano	7º NJM	Realfa2	Acelera 2	Acelera 3	PEJA	Saídas	TOTAL
COC1	4.446	34.861	15	1	3	428	141	1.545	2.133
COC2	0	34.490	0	0	0	3	25	343	371
COC3	0	34.085	0	0	0	1	20	384	405
COC4	0	33.707	0	0	0	0	0	378	378
FINAL	4.446	33.707	15	1	3	432	186	2.650	3.287

Como os critérios de enturmação para este ano letivo não previam encaminhamento dos alunos reprovados para turmas de projeto de correção de fluxo⁵², não houve a necessidade de fazer a adequação feita anteriormente para este ano letivo. Com isso, temos 4.446 alunos repetentes no 7º ano, o que representa 10,73% dos alunos que terminaram o ano letivo de 2013 ou 7,66% da coorte inicial e 3.287 alunos censurados, que representa 7,93% da coorte que terminou o ano letivo anterior ou 5,67% da coorte inicial.

Prosseguindo o acompanhamento no ano de 2015, com os 33.707 alunos terminaram o ano letivo anterior. Uma situação similar àquela ocorrida no ano letivo de 2013 ocorre novamente na transição deste ano. Os alunos encaminhados para as turmas de Aceleração 3 do projeto de correção de fluxo são os alunos defasados e reprovados em 2014⁵³, portanto, seguindo a mesma lógica de transição do ano 2012-2013, esses alunos também foram considerados como repetentes para este estudo.

Dos 33.707 alunos que terminaram o ano letivo de 2014 matriculados no 8º ano, apenas 28.143 (48,51% da coorte inicial) terminaram o ano letivo de 2015 no 9º ano, 3.816

⁵² Estava previsto apenas o encaminhamento dos aprovados e defasados para a turma de Aceleração 3

⁵³ PORTARIA E/SUBE/CED Nº 15, de 24 de outubro de 2014. Disponível em:

<http://doweb.rio.rj.gov.br/visualizar_pdf.php?edi_id=2579&page=20>. Acesso em: 05 jun. 2018.

alunos (6,58%) repetiram ou foram encaminhados para turmas de Aceleração 3 e outros 1.757 alunos (3,03%) foram censurados, conforme detalhamento na Tabela 19.

Tabela 19: Trajetória Detalhada da Coorte em 2014-2015

TRAJETÓRIA DA COORTE EM 2014-2015						
Período	Repetentes		Promovidos	Censurados		
	8º ano	Acelera 3	9º ano	PEJA	Saídas	TOTAL
COC1	2.890	926	28.985	141	765	906
COC2	0	0	28.687	61	237	298
COC3	0	0	28.407	25	255	280
COC4	0	0	28.143	0	273	273
FINAL	2.890	926	28.143	227	1.530	1.757

Como a previsão de fluxo desta coorte era que a mesma concluísse o ensino fundamental no final do ano de 2015 e a SME-Rio também não oferece o ensino médio para confirmar a promoção dos alunos aprovados no ano em questão, para finalizar o acompanhamento dos 28.143 alunos que chegaram ao COC 4 de 2015 matriculados no 9º ano, foi analisado apenas o resultado final (rendimento) destes alunos (aprovados, reprovados ou reprovados por falta). Na Tabela 20, o resultado final para o ano de 2015:

Tabela 20: Resultado final de 2015

RESULTADO FINAL DE 2015		
Resultado Final	Nº. de Alunos	Percentual
Aprovados	26.673	94,78%
Reprovados	1.461	5,19%
Não avaliados	9	0,03%
FINAL	28.143	100,00%

De forma geral, de todos os alunos acompanhados, apenas 45,97% dos alunos permaneceram dentro do fluxo escolar esperado, sendo promovidos ano a ano até a conclusão do ensino fundamental no final do estudo. 16.931 alunos tiveram pelo menos uma repetência ao longo desses anos⁵⁴ e outros 14.415 alunos foram censurados – ou por serem encaminhados a outros tipos de modalidade de ensino (PEJA, Classe Especial, alguns projetos ou readequação de matrícula) ou por saírem da Rede (mesmo que retornasse a seguir), conforme demonstrado no Gráfico 19.

⁵⁴ Neste estudo, não foi controlado se o aluno já havia ou não tido alguma repetência ao longo de sua trajetória escolar antes do período observado, assim como também não foi o propósito do estudo observar múltiplas repetências, portanto, no momento que acontecia a primeira repetência neste período, o aluno deixava de ser acompanhado.

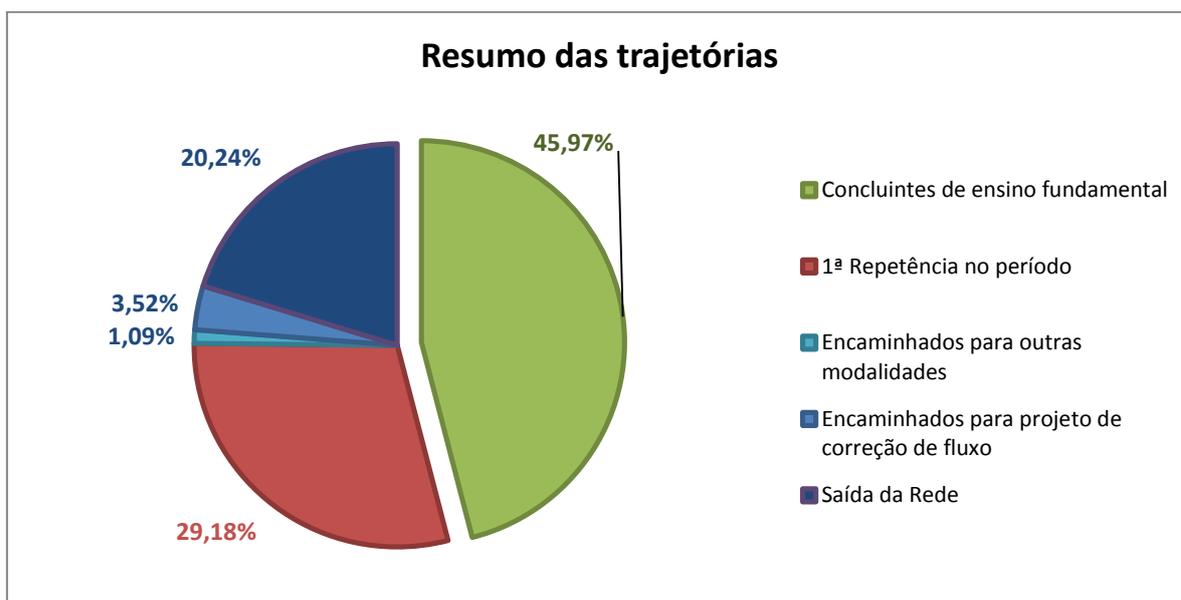


Gráfico 19: Resumo das trajetórias

Quadro 9: Resumo da trajetória da coorte (2012-2015)

Resumo da trajetória da coorte				
Situação/Período	2012-2013	2013-2014	2014-2015	2015
Promovidos	42.604 73,43% (7º Ano)	34.861 60,09% (8º Ano)	28.985 49,96% (9º Ano)	26.673 45,97% (Conclusão EF)
Repetentes	7.208 12,42% (6º Ano + Realfa 2)	4.446 7,66% (7º Ano)	3.816 6,58% (8º Ano + Acelera 3)	1.461 2,52% (9º Ano ⁵⁵)
Outras Modalidades	205 0,35%	153 0,26%	186 0,32%	86 0,15%
Projetos de Correção de Fluxo	1.588 2,74%	452 0,78%	04 0,01%	00 0,00%
Saída da Rede	6.414 11,05%	2.692 4,64%	1.870 3,22%	765 1,32%
TOTAL	58.019 100,00%	42.604 73,43%	34.861 60,09%	28.985 49,96%

Para finalizar esta parte do estudo, apresentamos o Quadro 9 com o resumo de todas as transições ocorridas na coorte estudada, utilizando com base para cada período o COC 1 dos anos de 2012 a 2015 e também o resultado dos alunos no final do ano letivo de 2015.

⁵⁵ Como já explicitado anteriormente, como não o ano letivo de 2016 não foi observado, conceitualmente não podemos classificar os alunos de repetentes, mas sim de reprovados, justamente por olhar apenas seu resultado final no ano letivo de 2015.

Nas linhas trouxemos as situações possíveis para este estudo: promovidos, repetentes e censurados, que podem ser encaminhamento para outras modalidades de ensino dentro da rede da SME-Rio (PEJA, Classe Especial ou outra série do ensino fundamental), encaminhamentos para projetos de correção de fluxo que não exigissem por Portarias de enturmação que fossem apenas para alunos reprovados no ano anterior, e as saídas da Rede, que podem ter sido por transferências externas, abandonos ou evasões.

6.3 ANÁLISE DE SOBREVIVÊNCIA

Nesta seção serão apresentados os resultados da análise de sobrevivência. Inicialmente apresentaremos as estatísticas descritivas e a função de sobrevivência geral do estudo, seu gráfico e sua interpretação. Em seguida, serão feitas as comparações entre as funções de sobrevivência de grupos distintos, utilizando o Estimador de Kaplan-Meier (EKM), com objetivo de identificar se possuir ou não determinada característica ou pertencer ou não a determinado grupo pode alterar de forma significativa o risco de repetência da coorte. E, por fim, serão expostos e interpretados os modelos de regressão de Cox que têm como finalidade analisar como algumas características podem influenciar positiva ou negativamente os riscos de repetência de tais alunos.

A Tabela 21 apresenta uma descrição resumida dos dados de sobrevivência. Na coluna Total há a indicação do total de alunos que compõem a coorte analisada (58.019 alunos), dos registros no banco de dados (161.300 registros para cada aluno para cada ano que ele sobreviveu) e 16.931 repetências registradas. Nas colunas indicadas por *per subject*, os resultados são referentes ao tempo (de 1 a 4), onde temos os valores das médias, dos mínimos, das medianas e dos máximos. A média do tempo encontrado foi de 2,78 anos. Como neste estudo toda a coorte começou ao mesmo tempo, a linha *first entry time* (tempo de entrada) encontram-se todos no mesmo tempo inicial $t=0$. Os tempos de saídas (*final exit time*) variam ao longo dos tempos estudados, ou seja, do tempo 1 ao 4, o mesmo ocorre o tempo em risco (*time at risk*), já que nos quatro tempos estudados haviam alunos passíveis à repetência. Na linha de falha (*failure*), ou evento, como tenho sido usado ao longo do estudo, a média, que também indica proporcionalidade por ser uma variável dicotômica dummy, nos traz que 29,18% dos alunos repetiram ao longo dos quatro anos.

Tabela 21: Descritiva da Análise de Sobrevivência

Category	total	per subject			
		mean	min	median	max
no. of subjects	58019				
no. of records	161300	2.780124	1	3	4
(first) entry time		0	0	0	0
(final) exit time		2.780124	1	3	4
subjects with gap	0				
time on gap if gap	0
time at risk	161300	2.780124	1	3	4
failures	16931	.2918182	0	0	1

A Tabela 22 apresenta os dados gerais que compuseram o banco de dados para a análise. Nela observamos quantos alunos estavam em risco (*time at risk*), a taxa de incidência do evento (*incidence rate*) e o total de sujeitos para cada ano.

Tabela 22: Resumo dos dados de sobrevivência por ano

Tempo	time at risk	incidence rate	no. of subjects	Survival time		
				25%	50%	75%
1	58019	.1242352	58019	.	.	.
2	41440	.1072876	41440	.	.	.
3	33707	.1132109	33707	.	.	.
4	28134	.05193	28134	.	.	.
total	161300	.1049659	58019	3	.	.

A taxa de incidência ou taxa de falha é calculada através da razão dos sujeitos que sofreram o evento pelo total de alunos que estavam em risco em cada tempo, portanto, a Tabela 22 nos revela que a maior incidência de reprovação está no tempo 1, isto é, na transição do 6º para o 7º ano do ensino fundamental, onde 12,4% dos alunos da coorte repetiram, a menor incidência está no tempo 4, onde 5,2% dos alunos foram reprovados no 9º ano. Analisando todo o período do segundo segmento do ensino fundamental (anos finais), a taxa de incidência para esta coorte é de 10,5% em geral.

Lins *et al.* (2017) recomenda que para “aumentar a robustez das análises”, é de suma importância apresentar a função de risco (taxa de falha) do estudo assim como a função de

sobrevivência que será apresentada a seguir, já a função de risco é possível encontra-la no Gráfico 20.

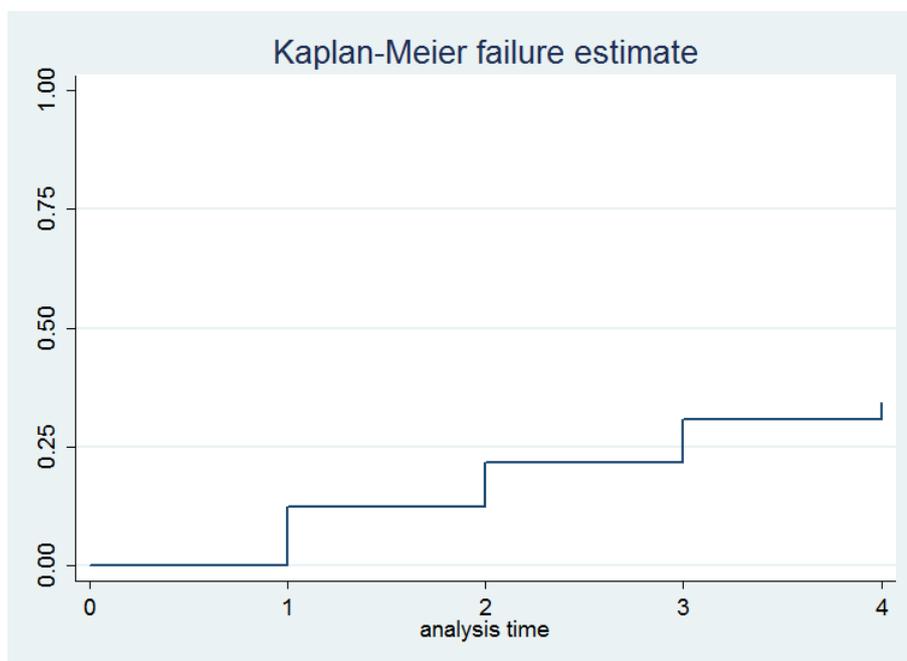


Gráfico 20: Função de Risco EKM

A função de sobrevivência é definida como a probabilidade de um aluno não ser reprovado até certo tempo. Para simplificar, em outras palavras, poderia se dizer que é a probabilidade de que, decorrido um tempo t , esse aluno não tenha repetido.

As estimativas da função de sobrevivência e seus respectivos intervalos de confiança, calculados a partir do Estimador de Kaplan-Meier (EKM), encontram-se na Tabela 23. Nela podemos encontrar na coluna 'N. Inicial' o total de alunos que estavam em risco (ou seja, que podiam sofrer a reprovação), o total de alunos que sofreram o evento (reprovados), o total de alunos censurados, a função de sobrevivência, seu erro padrão e o seu intervalo de Confiança a 95% em cada momento do tempo.

Tabela 23: Função de Sobrevivência de EKM

Tempo	N. Inicial	Reprovados	Censurados	Função de Sobrevivência	Erro padrão	I.C. 95%
1	58.019	7.208	9.371	0,8758	0,0014	0,8731 0,8784
2	41.440	4.446	3.287	0,7818	0,0018	0,7782 0,7853
3	33.707	3.816	1.757	0,6933	0,0021	0,6892 0,6974
4	28.134	1.461	27.000 ⁵⁶	0,6573	0,0022	0,6530 0,6616

⁵⁶ O sistema entende que é censurado todos os casos que não tem sofreram o evento (repetência) até o final do estudo, por este motivo, temos uma valor muito alto na coluna da censura no tempo 4.

Ainda na Tabela 23, observando a função de sobrevivência para cada tempo, no tempo 1, a probabilidade condicional do aluno não repetir é de 0,8758, com o passar do tempo, essa probabilidade de sobrevivência vai decrescendo gradativamente. Como este estimador é também conhecido por estimador limite-produto, por levar em consideração que para chegar no tempo $t + 1$ este aluno já teve que sobreviver anteriormente ao tempo t . Na prática, por exemplo, a função de sobrevivência de um aluno que sobreviveu ao tempo 3 é o produto das probabilidades de sobrevivências nos tempos 1, 2 e 3.

O Gráfico 21 apresenta a função de sobrevivência de EKM no qual cada degrau representa um dos tempos analisados e a “altura de cada degrau” seria a representação gráfica de quantos sofreram o evento em cada tempo, ou seja, de quantos não sobreviveram aquele tempo.

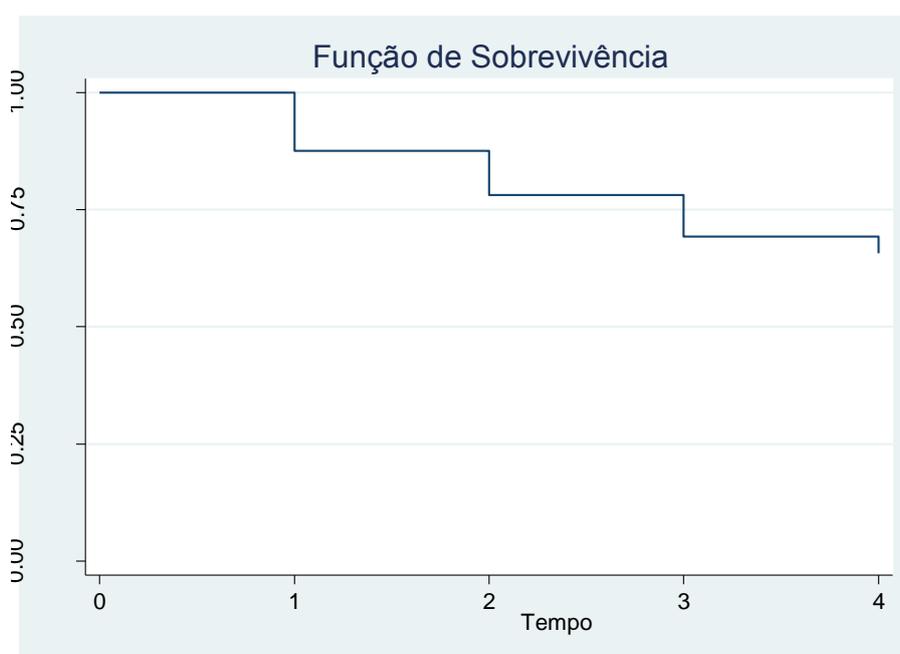


Gráfico 21: Função de Sobrevivência, EMK

A partir de agora, continuaremos com as funções de sobrevivência, entretanto, nosso objetivo é comparar tais funções por grupos de acordo com as variáveis que possuímos em nosso banco de dados, onde a pergunta que queremos responder é: existe diferença no risco de repetência para tais grupos?

Cabe reforçar que, para as comparações que serão feitas neste momento, não há controle de variáveis, portanto, são feitas individualmente variável por variável.

Para avaliar se as possíveis diferenças encontradas visualmente nos gráficos das funções de sobrevivências de cada grupo são estatisticamente significativas empregaremos o teste *Long-rank* e o teste de Wilcoxon, onde testaremos a hipótese nula (H_0) de que as diferenças observadas entre as funções são devidas a aleatoriedade da amostra, ou seja, de que não há diferença entre elas. Portanto, valores de p menores que 0,05 indicam que a existência de diferenças entre os grupos, podendo rejeitar a hipótese nula. Na Tabela 24, encontramos os resultados de tais testes de alguns dos grupos comparados e já se pode confirmar que há diferença entre os grupos.

Tabela 24: Testes de comparação entre grupos

Testes de comparação entre grupos					
Variáveis	Valores	Long-rank		Wilcoxon	
		$\chi^2(1)$	valor p	$\chi^2(1)$	valor p
Sexo	Masculino x Feminino	718,38	0,0000	774,47	0,0000
Não-Branco	Branco x Não Branco	308,37	0,0000	319,28	0,0000
Sexo	Meninos brancos x Meninos não brancos x				
Não-Branco	Meninas brancas x Meninas não brancas	1.019,25	0,0000	1.074,67	0,0000
Defasados	Não defasados x Defasados	1.072,01	0,0000	1.104,03	0,0000
	Ens. Médio incompleto x Ens. Médio				
Max_Esc_Pais	completo	733,07	0,0000	756,98	0,0000
PB_LPL_1e2	Prova Bimestral – LPL – COCs 1 e 2	6.528,31	0,0000	7.090,23	0,0000
GC	Escola regular x Ginásio Carioca	276,09	0,0000	289,92	0,0000
6ºAExp + GC	6º Ano Experimental x Ginásio Carioca	376,46	0,0000	492,94	0,0000

A primeira comparação entre meninos e meninas, já foi estudada por diversos pesquisadores (SILVA, HASENBALG, 2002; SOUZA *et al.*, 2011; OLIVEIRA, SOARES, 2012; KLAPPROTH, SCHALTZ, 2015; TAVARES JR *et al.*, 2015). Analisando a Tabela 25 e Gráfico 22, com as funções de sobrevivências no mesmo tempo entre os grupos, observamos que as funções das meninas têm sempre valores maiores do que as dos meninos, indicando que as meninas repetem menos do que os meninos. É possível observar que essa diferença aumenta com o passar dos tempos estudados, pois no tempo 1 a diferença era 6,3 pontos percentuais e no tempo 4 chegou a 10,6 pontos percentuais, confirmando os resultados de pesquisas anteriores.

Tabela 25: Função de Sobrevivência - Sexo.

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
Masculino							
1	29985	4640	5095	0.8453	0.0021	0.8411	0.8493
2	20250	2544	1617	0.7391	0.0027	0.7338	0.7443
3	16089	2071	849	0.6439	0.0030	0.6379	0.6499
4	13169	791	1.2e+04	0.6053	0.0032	0.5990	0.6114
Feminino							
1	28034	2568	4276	0.9084	0.0017	0.9050	0.9117
2	21190	1902	1670	0.8269	0.0024	0.8221	0.8315
3	17618	1745	908	0.7450	0.0028	0.7394	0.7505
4	14965	670	1.4e+04	0.7116	0.0030	0.7057	0.7174

Estudos sobre desigualdades educacionais utilizam de forma recorrente a variável cor/raça para analisar a trajetória escolar dos alunos. Nessa dissertação optou-se pela codificação dos alunos, em dois grupos: “Não branco”, todos os alunos com raças diferentes de brancos e os alunos codificados como brancos. Na Tabela 26, os resultados apresentados na coluna da Função de Sobrevivência (*Survival Function*) apontam que os alunos brancos possuem funções de sobrevivências mais longas quando comparados aos não brancos, assim como ocorreu entre as meninas, corroborando com outros achados (SILVA, HASENBALG, 2002; TAVARES JR *et al.*, 2015), que revelam que alunos não brancos apresentam maior risco de reprovação. O Gráfico 23 sugere que há um aumento da diferença dos “degraus” no gráfico.

Tabela 26: Função de Sobrevivência - Não Brancos

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
Branco							
1	20060	1953	2969	0.9026	0.0021	0.8985	0.9067
2	15138	1388	1163	0.8199	0.0028	0.8142	0.8254
3	12587	1229	641	0.7398	0.0034	0.7332	0.7463
4	10717	479	1.0e+04	0.7068	0.0035	0.6998	0.7136
Não Branco							
1	35455	4885	5960	0.8622	0.0018	0.8586	0.8658
2	24610	2881	1968	0.7613	0.0024	0.7566	0.7659
3	19761	2411	1033	0.6684	0.0027	0.6630	0.6738
4	16317	921	1.5e+04	0.6307	0.0029	0.6250	0.6363

Com base nos resultados descritivos dessas duas variáveis, optou-se em agregá-las no mesmo gráfico para verificar as funções de sobrevivência dos quatro grupos formados: meninos brancos, meninos não brancos, meninas brancas e meninas não brancas. No Gráfico 24, observamos que a curva mais curta é dos meninos não brancos, seguida pela

curva dos meninos brancos, as curvas das meninas são mais longas do que as dos meninos, sendo a curva das meninas brancas a mais longa no gráfico. Tais resultados sugerem que o sexo influencia mais do que a raça quando tais grupos são comparados entre si.

Tabela 27: Função de Sobrevivência - Sexo e Não Branco

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
Masculino Branco							
1	10108	1236	1547	0.8777	0.0033	0.8712	0.8840
2	7325	761	546	0.7865	0.0043	0.7780	0.7948
3	6018	684	304	0.6971	0.0050	0.6873	0.7068
4	5030	263	4767	0.6607	0.0052	0.6504	0.6708
Masculino Não Branco							
1	18572	3165	3296	0.8296	0.0028	0.8241	0.8349
2	12111	1692	996	0.7137	0.0035	0.7067	0.7205
3	9423	1300	504	0.6152	0.0040	0.6074	0.6229
4	7619	499	7120	0.5749	0.0041	0.5669	0.5829
Feminino Branco							
1	9952	717	1422	0.9280	0.0026	0.9227	0.9329
2	7813	627	617	0.8535	0.0037	0.8460	0.8606
3	6569	545	337	0.7827	0.0045	0.7737	0.7913
4	5687	216	5471	0.7529	0.0047	0.7435	0.7621
Feminino Não Branco							
1	16883	1720	2664	0.8981	0.0023	0.8935	0.9026
2	12499	1189	972	0.8127	0.0032	0.8064	0.8188
3	10338	1111	529	0.7253	0.0038	0.7179	0.7326
4	8698	422	8276	0.6902	0.0039	0.6824	0.6978

Outra variável de interesse é se o aluno é defasado ou não. Sabemos que há uma grande chance dos alunos defasados serem alunos que já repetiram alguma série em sua trajetória escolar pregressa ou que abandonaram/evadiram seus estudos. Entretanto, como já explicado anteriormente, não podemos garantir aqui neste estudo tal informação. Estudos indicam que quando um aluno repete, as chances dele repetir novamente aumentam, por isso, é importante analisar a função de sobrevivência pra esses alunos. Lembrando que os alunos foram considerados defasados quando possuíam pelo menos dois anos além da idade ideal em 2012. Os resultados dessa comparação encontram-se na Tabela 28 e no Gráfico 25 e, conforme esperado, os alunos defasados possuem uma função de sobrevivência menos longa do que aqueles que não estão defasados, no tempo 4, a diferença chega a ser de 19,5 pontos percentuais, sendo que neste mesmo tempo, os alunos defasados têm menos que a metade da probabilidade de não repetir. Tais resultados são compatíveis com os encontrados por Souza *et al.* (2011) e Oliveira e Soares (2012).

Tabela 28: Função de Sobrevivência – Defasado

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
Não							
1	50368	5432	5349	0.8922	0.0014	0.8894	0.8948
2	39587	4091	2496	0.8000	0.0018	0.7963	0.8035
3	33000	3708	1627	0.7101	0.0021	0.7058	0.7143
4	27665	1419	2.6e+04	0.6736	0.0022	0.6692	0.6780
Sim							
1	7651	1776	4022	0.7679	0.0048	0.7583	0.7772
2	1853	355	791	0.6208	0.0080	0.6048	0.6363
3	707	108	130	0.5259	0.0108	0.5045	0.5469
4	469	42	427	0.4788	0.0120	0.4550	0.5022

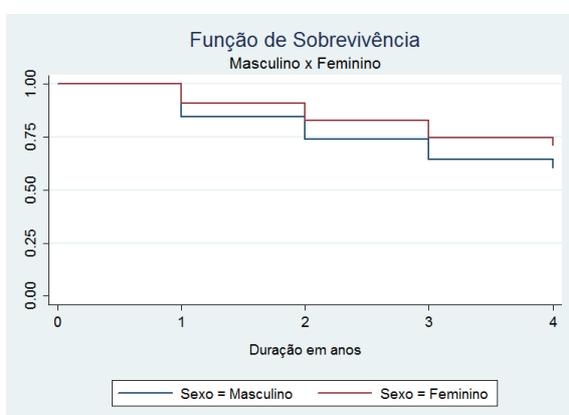


Gráfico 22: Função de Sobrevivência - Sexo

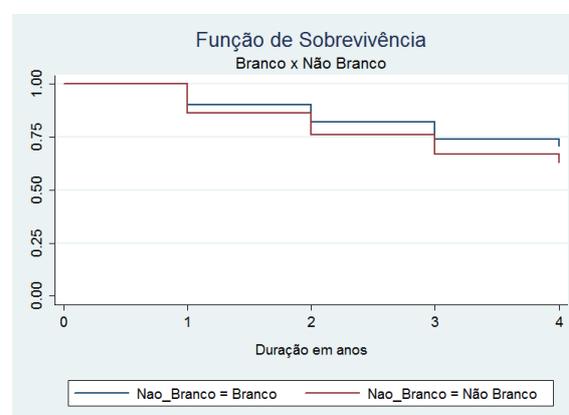


Gráfico 23: Função de Sobrevivência - Não Branco

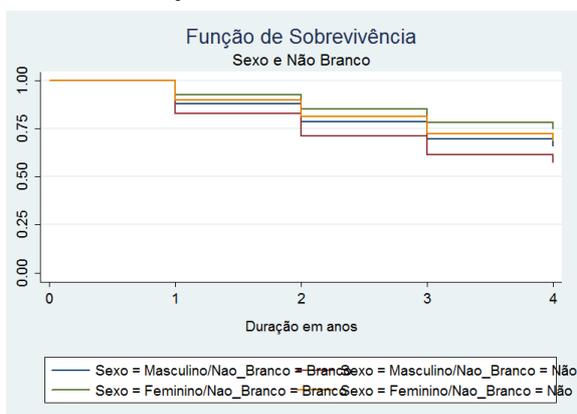


Gráfico 24: Função de Sobrevivência - Sexo e Não Branco

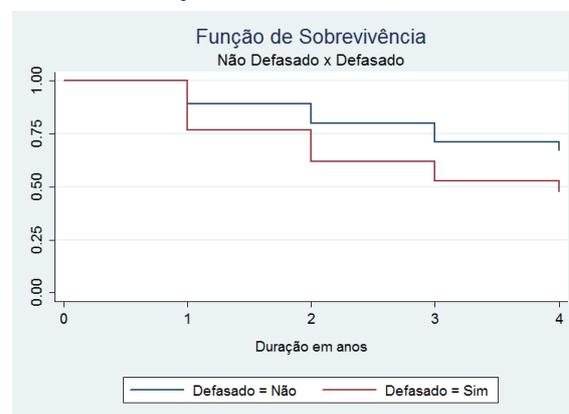


Gráfico 25: Função de Sobrevivência - Defasado

A escolaridade da mãe e do pai é uma das características familiares mais associadas com a trajetória escolar dos alunos. Nessa dissertação optou-se em analisar a escolaridade máxima dos pais, com o objetivo de minimizar os dados faltantes (missing data) no momento da análise (Bartholo, 2014). Esta variável é uma variável categórica com cinco valores (ensino fundamental incompleto, ensino fundamental completo, ensino médio, ensino superior incompleto e ensino superior completo) e que foi transformada em uma

variável dicotômica: o aluno recebe 1 (sim), quando a escolaridade máxima dos pais é igual ou maior que o Ensino Médio completo; o aluno recebe 0 (zero) quando a máxima escolaridade dos pais é inferior ao ensino médio completo. Na Tabela 29 e no Gráfico 26, encontramos os resultados desses grupos. É possível observar que os resultados corroboram achados em estudos anteriores no campo da educação que sugerem uma correlação positiva entre a escolaridade dos pais e as trajetórias escolares dos filhos. Alunos cujos pais têm escolaridade menor apresentam uma maior probabilidade de reprovarem ao longo do segundo segmento do ensino fundamental (SOUZA *et al.*, 2011; GORARD; SEE, 2013).

Tabela 29: Função de Sobrevivência - Máxima Escolaridade Pais

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
Não							
1	32333	4727	5327	0.8538	0.0020	0.8499	0.8576
2	22279	2705	1738	0.7501	0.0025	0.7451	0.7551
3	17836	2318	930	0.6526	0.0029	0.6469	0.6583
4	14588	849	1.4e+04	0.6147	0.0030	0.6087	0.6206
Sim							
1	18089	1492	2448	0.9175	0.0020	0.9134	0.9214
2	14149	1106	1051	0.8458	0.0028	0.8402	0.8512
3	11992	1062	584	0.7709	0.0034	0.7642	0.7774
4	10346	443	9903	0.7379	0.0036	0.7308	0.7448

Outra característica das famílias amplamente estudada é o impacto da renda no desempenho e trajetória escolar dos alunos. Não há informação sobre a renda familiar nas bases da SME, no entanto, há informação se a família é beneficiária do programa Bolsa Família. Tratamos essa variável como uma proxy de pobreza e outros estudos no Brasil e no exterior utilizam estratégia semelhante quando tem interesse em estimar o aprendizado ou a trajetória escolar dos alunos (Bartholo; Costa, 2016). Na Tabela 30 e no Gráfico 27, podemos observar que apesar de haver diferença entre tais grupos, essas diferenças são sutis, não passando de 5 pontos percentuais em nenhum momento do tempo. É preciso destacar que há possível erro de mensuração nessa variável, uma vez que a falta de informação (*missing data*) é codificada como “não beneficiário do programa Bolsa Família”. Há o risco de subnotificação nas bases da SME-RJ (Bartholo, 2014).

Tabela 30: Função de Sobrevivência - Bolsa Família

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
Não							
1	34861	3978	7056	0.8859	0.0017	0.8825	0.8892
2	23827	2458	2176	0.7945	0.0023	0.7899	0.7990
3	19193	2032	1068	0.7104	0.0027	0.7050	0.7157
4	16093	735	1.5e+04	0.6779	0.0028	0.6723	0.6835
Sim							
1	23158	3230	2315	0.8605	0.0023	0.8560	0.8649
2	17613	1988	1111	0.7634	0.0029	0.7577	0.7690
3	14514	1784	689	0.6696	0.0033	0.6631	0.6759
4	12041	726	1.1e+04	0.6292	0.0034	0.6225	0.6358

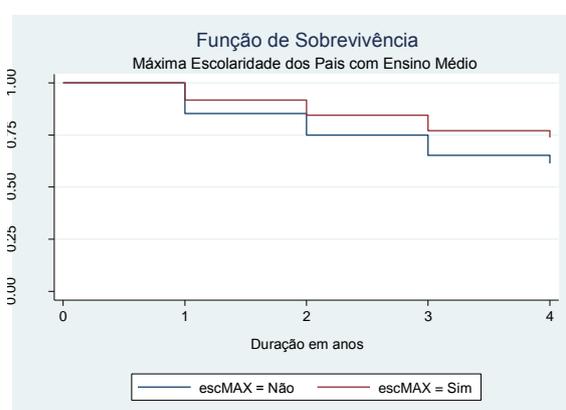


Gráfico 26: Função de Sobrevivência - Máxima Escolaridade dos Pais

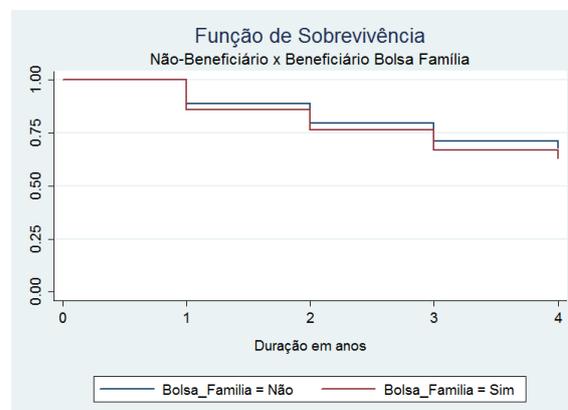


Gráfico 27: Função de Sobrevivência - Bolsa Família

Tratando agora de Políticas Públicas da SME-Rio, selecionou a variável dicotômica sobre o Ginásio Carioca, essa política foi pensada para atender a demanda inerente ao público da faixa etária dos anos finais, conforme apresentado no Capítulo 4.

Antes de prosseguir, cabe uma observação em relação a esta variável, já que a mesma é dependente do tempo, ou seja, ela pode se modificar ao longo do tempo, portanto, os alunos com “sim” no tempo 1 eram apenas os alunos que estavam nos Ginásios Cariocas em 2012, os alunos com “sim” no tempo 2, eram apenas os alunos que estavam nos Ginásios Cariocas em 2013 e assim sucessivamente. Esta peculiaridade faz com que tenhamos bastante cautela em utilizar o EKM, já que não foi encontrada na literatura como tal metodologia se comporta em relação a variáveis desse tipo, entretanto, ao observar que apenas 303 alunos tiveram modificações dos valores dessa variável ao longo do tempo, optou-se por prosseguir assim mesmo, mas não podemos deixar de reforçar a prudência necessária para as conclusões.

Na Tabela 31 e Gráfico 28, os resultados são apresentados demonstrando uma vantagem na probabilidade de não repetir para os alunos dos Ginásios Cariocas em todos os momentos, se compararmos as funções dos grupos no mesmo momento, temos uma diferença de 19,6%, 24,3% e 24,67, respectivamente nos tempos 2, 3 e 4⁵⁷. No tempo 1, a probabilidade foi 1,00 até porque só haviam 19 alunos na base de dados e todos foram promovidos, já que esta política é voltada para o 7º ao 9º ano, tendo 6º ano apenas os ginásios vocacionados para o esporte (GEO).

Tabela 31: Função de Sobrevivência - Ginásio Carioca

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
Não							
1	58000	7208	1.2e+04	0.8757	0.0014	0.8730	0.8784
2	38992	4382	3030	0.7773	0.0019	0.7736	0.7809
3	31580	3718	1594	0.6858	0.0022	0.6815	0.6900
4	26268	1395	2.5e+04	0.6494	0.0023	0.6449	0.6538
Sim							
1	19	0	-2429	1.0000	.	.	.
2	2448	64	257	0.9739	0.0032	0.9667	0.9795
3	2127	98	163	0.9290	0.0054	0.9176	0.9388
4	1866	66	1800	0.8961	0.0065	0.8825	0.9082

Para terminar a análise preliminar com a comparação das funções de sobrevivência optou-se em combinar duas variáveis sobre Políticas Públicas da SME-Rio, por entender que elas são complementares: 6º Ano Experimental 2012 e o Ginásio Carioca e, por isso, apresentamos quatro grupos para comparação: alunos que não participaram de nenhuma das duas políticas (Não/Não), alunos que frequentaram o 6º ano regular em 2012 e migraram para o Ginásio Carioca no 7º, 8º e 9º ano (Não/Sim), alunos que estudaram em escolas que ofereciam o 6º ano experimental em 2012 e migraram posteriormente para escolas regulares (Sim/Não) e alunos matriculados no 6º ano experimental 2012 e que foram remanejados para o Ginásio Carioca (Sim/Sim).

A Tabela 32 e Gráfico 29 apresentam as funções de sobrevivência para os quatro grupos supracitados. Os resultados corroboram os achados do estudo do Itaú Social (2017)⁵⁸ que sugere que a probabilidade de não repetir é muito maior para os alunos que

⁵⁷ Optou-se em não fazer a diferença no tempo 1, já que o grupo do GEC no tempo 1 é muito pequeno, não valendo a pena a comparação.

⁵⁸ Estudo ainda não publicado.

frequentaram o Ginásio Carioca, independente de qual 6º ano cursaram anteriormente. Os alunos do 6º ano experimental que vão para escolas regulares (não são Ginásios Cariocas) tem um largo “degrau” no tempo 2, sugerindo uma maior reprovação no 7º ano, praticamente se igualando ao grupo que não experimentou nenhuma das duas políticas momento 4 do tempo – a diferença entre as curvas no tempo 4 é de apenas 0,5 pontos percentuais.

Tabela 32: Função de Sobrevivência - 6º Ano experimental x Ginásio Carioca

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
Não Não							
1	54307	7174	9810	0.8679	0.0015	0.8650	0.8707
2	37323	4093	3003	0.7727	0.0019	0.7690	0.7764
3	30227	3524	1600	0.6826	0.0022	0.6783	0.6869
4	25103	1312	2.4e+04	0.6470	0.0023	0.6424	0.6514
Não Sim							
1	19	0	-810	1.0000	.	.	.
2	829	22	46	0.9735	0.0056	0.9600	0.9824
3	761	40	18	0.9223	0.0095	0.9014	0.9389
4	703	19	684	0.8974	0.0108	0.8740	0.9166
Sim Não							
1	3693	34	1990	0.9908	0.0016	0.9871	0.9934
2	1669	289	27	0.8192	0.0093	0.8002	0.8366
3	1353	194	-6	0.7018	0.0111	0.6793	0.7230
4	1165	83	1082	0.6518	0.0116	0.6285	0.6740
Sim Sim							
1	0	0	-1619	1.0000	.	.	.
2	1619	42	211	0.9741	0.0040	0.9651	0.9808
3	1366	58	145	0.9327	0.0065	0.9187	0.9444
4	1163	47	1116	0.8950	0.0083	0.8776	0.9101

Há aqui uma sugestão inicial de políticas que podem ter efeito protetor sobre as trajetórias escolares dos alunos, minimizando o risco de repetência. O maior efeito protetor é dos Ginásios Cariocas, no entanto, a combinação da exposição às duas políticas parece gerar os melhores resultados quando pensamos em trajetórias escolares regulares.

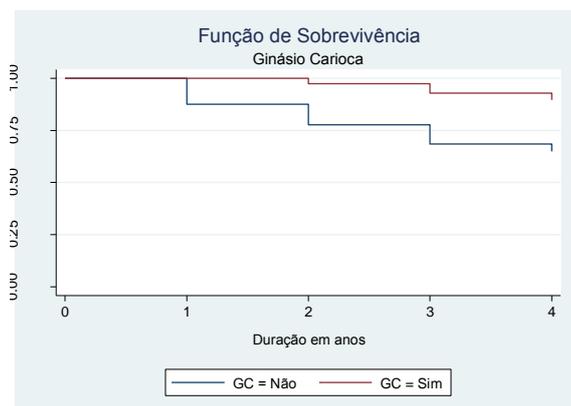


Gráfico 28: Função de Sobrevivência - Ginásio Carioca

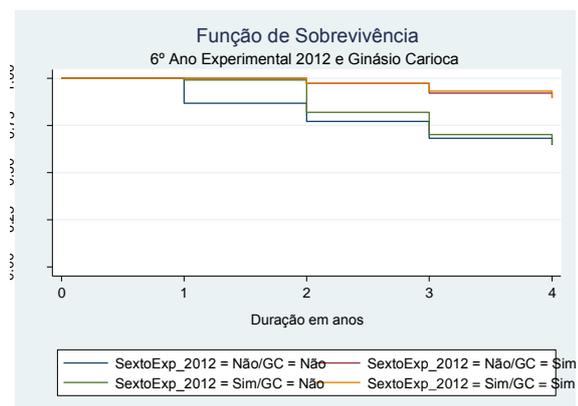


Gráfico 29: Função de Sobrevivência - 6º Ano Experimental x Ginásio Carioca

Os modelos subsequentes da dissertação irão estimar o potencial efeito das políticas controlando pelas características do aluno e família. O risco aqui é de um possível viés de seleção por parte das escolas que oferecem as políticas no momento da matrícula dos alunos. Por exemplo, apenas alunos sem defasagem idade série podem se inscrever em escolas que oferecem o 6º ano experimental. Essas regras formais ou informais⁵⁹ podem interferir no resultado dos programas com um viés de seleção dos alunos.

Resumindo, as funções de sobrevivência foram significativas mais longas para: a) os estudantes do sexo feminino, b) os alunos brancos, c) os alunos não defasados, d) alunos cujos pais têm escolaridade mais alta, e) os não beneficiários do programa social Bolsa Família, f) alunos dos Ginásios Cariocas. Portanto, podemos inferir preliminarmente que os alunos repetentes diferem dos alunos promovidos em vários aspectos.

As tabelas e gráficos de outras comparações estão disponíveis no Anexo B desta dissertação.

Passando para a fase final deste estudo, a Regressão de Cox foi calculada para cinco modelos⁶⁰, abaixo descritos e que irão ser abordados individualmente:

- **Modelo 1:** com variáveis sobre o Perfil Socioeconômico dos alunos;
- **Modelo 2:** com variáveis sobre o desempenho acadêmico dos alunos;
- **Modelo 3:** com variáveis sobre políticas públicas das escolas dos alunos;
- **Modelo 4:** com as variáveis de Perfil Socioeconômico e desempenho acadêmico dos alunos, unindo o Modelo 1 com o Modelo 2;

⁵⁹ Podemos também pensar em critérios não explícitos de seleção de alunos pelas escolas da rede pública. O chamado quase mercado escolar já foi estudado por outros pesquisadores do Lapope (BRUEL; BARTHOLO, 2012; COSTA; KOSLINSKI, 2011; COSTA; KOSLINSKI 2012).

⁶⁰ A construção dos modelos seguiu a orientação de Colosimo e Giolo (2006) e estão disponíveis no Anexo C desta dissertação.

- **Modelo 5:** com todas as variáveis disponíveis, partindo do Modelo 4 e agregando-o ao Modelo 3.

Cada modelo foi composto por variáveis que foram estatisticamente significativas e que atenderam à suposição de riscos proporcionais, exigida para o Modelo de Cox, conforme explicitado no capítulo 5.

Os resultados do Modelo 1 são descritos na Tabela 33 e sua interpretação logo a seguir.

Tabela 33: Regressão de Cox - Modelo 1

Modelo 1							
No. of subjects =	48.598				Number of obs =	136.705	
No. of failures =	14.128						
Time at risk =	136.705				LR chi2(9) =	2.709,49	
Log likelihood =	-147.339,34				Prob > chi2 =	0,00000	
Variável	Coefic. (b)	Razões de Risco (e ^b)	Erro Padrão	z	P>z	IC 95%	
Idade	0,24009	1,27136	0,00765	39,92	0,00000	1,25646	1,28643
Sexo	-0,52114	0,59385	0,02289	-13,52	0,00000	0,55064	0,64044
Sexo*Tempo	0,06942	1,07189	0,01834	4,06	0,00000	1,03653	1,10845
Não Branco	0,17767	1,19444	0,02180	9,74	0,00000	1,15247	1,23793
Bolsa Família	0,11044	1,11677	0,01912	6,45	0,00000	1,07993	1,15488
Biparental	-0,18751	0,82902	0,03143	-4,95	0,00000	0,76965	0,89298
Biparental*Tempo	-0,03733	0,96336	0,01631	-2,21	0,02700	0,93193	0,99586
Máx. Escolaridade dos Pais	-0,50178	0,60545	0,02592	-11,72	0,00000	0,55672	0,65844
Máx. Escolaridade dos Pais*Tempo	0,06207	1,06404	0,01973	3,35	0,00100	1,02607	1,10342

Interpretando os resultados das razões de risco para cada variável deste modelo, temos as seguintes conclusões:

- **Idade:** para cada ano de idade a mais que o aluno possui, o risco de repetência é aumentado em 27,1%, demonstrando assim que alunos mais velhos correm mais risco de repetir quando comparados a alunos com menos idade.
- **Sexo:** para este modelo, além da variável Sexo, tivemos que utilizar a variável Sexo com interação com o Tempo (Sexo*Tempo), pois as razões de risco não eram proporcionais ao longo do tempo, portanto, teremos duas interpretações para esta variável, a interpretação do resultado da variável Sexo que é a razão de risco no tempo inicial, sendo considerado o efeito principal, e depois a interpretação da variável Sexo*Tempo, que é o quanto esse efeito é mudado com o passar do tempo.

Como a codificação era 1 para as meninas, o resultado indica que as meninas têm 40,6%⁶¹ a menos de risco de repetirem quando comparadas aos meninos no momento inicial do estudo, mas com o passar do tempo, essa diferença vai diminuindo, passando para 36,3%⁶² no tempo 1; 31,8% no tempo 2; 26,9% no tempo 3; e 21,6% no tempo 4.

- **Não Branco:** os alunos não brancos possuem 19,4% a mais de risco de repetirem do que os alunos brancos.
- **Bolsa Família:** os alunos que são beneficiários do programa social Bolsa Família, possuem 11,7% a mais de risco de repetir do que aqueles que não são beneficiários deste programa.
- **Biparental:** os alunos que moram com seus dois pais têm, no início do estudo, 17,1% a menos de risco de repetir do que aqueles que moram com um dos seus pais ou com nenhum deles. Como também há interação dessa variável com o tempo, a cada ano adicional, essa disparidade continua aumentando, sendo de 20,1% no tempo 1; 23,1% no tempo 2; 25,9% no tempo 3; e 28,6% no tempo 4.
- **Máxima Escolaridade dos Pais:** para esta variável, o código 1 foi atribuído para os alunos em que a escolaridade máxima entre seus pais era pelo menos o ensino médio completo, portanto, os valores indicam que para o momento inicial do estudo, os alunos com pelo um dos pais com no mínimo o ensino médio completo possuem 39,5% a menos de risco de repetir do que os outros alunos. Com a interação, o efeito vai diminuindo, tendo como valores: 35,6% no tempo 1; 31,5% no tempo 2; 27,1% no tempo 3; e 22,4% no tempo 4.

Os resultados corroboram com estudos anteriores com desenho transversal que estimaram as probabilidades de reprovação de alunos no ensino básico. Esse modelo difere das funções de sobrevivência apresentadas anteriormente porque os coeficientes apresentados estão controlados pelas demais variáveis apresentadas no modelo. Ou seja, o efeito da escolaridade dos pais sobre o risco de repetência é estimado considerando todas as demais características da família estáveis/controladas.

⁶¹ Para coeficientes menores que 1, o cálculo deve ser feito 1 menos o exponencial de beta encontrado. Neste caso a $1 - 0,59385 = 0,4061548$ ou 40,6%.

⁶² A fórmula utilizada para o cálculo foi $RR_t = e^{b_p + (t \cdot b_i)}$, onde b_p = coeficiente da variável principal, b_i = coeficiente da variável com interação do tempo e t = tempo.

Os resultados do Modelo 2 encontram-se na Tabela 34 e sua interpretação logo a seguir. Ele foi feito com variáveis sobre o desempenho acadêmico dos alunos. Vale lembrar que são variáveis que se modificam com o tempo, não tendo seus valores fixos em cada momento do tempo estudado.

Tabela 34: Regressão de Cox - Modelo 2

Modelo 2							
No. of subjects =	56.270				Number of obs =	157.712	
No. of failures =	16.116						
Time at risk =	157.712				LR chi2(9) =	10.798,31	
Log likelihood =	-166.614,07				Prob > chi2 =	0,00000	
Variável	Coefic. (b)	Razões de Risco (e ^b)	Erro Padrão	z	P>z	IC 95%	
PB_LPL_1e2	-0,19254	0,82486	0,00771	-20,59	0,00000	0,80988	0,84012
PB_LPL_1e2*Tempo	-0,03949	0,96128	0,00444	-8,56	0,00000	0,95262	0,97001
PB_MAT_1e2	-0,27596	0,75885	0,00448	-46,72	0,00000	0,75011	0,76768
Mov_Turma	0,36295	1,43757	0,04505	11,58	0,00000	1,35193	1,52864

Abaixo a interpretação das razões de risco para as variáveis deste modelo.

- **PB_LPL_1e2:** essa variável é referente à média do aluno na Prova Bimestral de Língua Portuguesa Leitura no COC 1 e 2. Para cada ponto a mais na média, o risco de repetência diminui em 17,5% no momento inicial do estudo, sendo que a cada ano, a disparidade vai crescendo⁶³, sendo de 20,7% no tempo 1; 23,8% no tempo 2; 26,7% no tempo 3; e 29,6% no tempo 4.
- **PB_MAT_1e2:** essa variável é referente à média do aluno na Prova Bimestral de Matemática e se mostrou com risco proporcional ao tempo, não havendo necessidade em utilizar a sua interação com o tempo. A razão de risco indica que cada ponto acrescido à média neste prova, diminui o risco de repetência em 24,1%.
- **Mov_Turma:** esta variável foi construída a partir das somas das movimentações de turmas da mesma escola entre os quatro bimestres e, segundo o resultado, para cada movimentação acrescida, o risco dos alunos em repetir aumenta em 43,8%.

Todas as variáveis apresentadas neste modelo possuem altos valores de razões de risco, parecendo influenciar o fluxo escolar dos alunos de forma intensa e apontando para as direções esperadas. As médias de provas bimestrais no primeiro semestre podem ser

⁶³ A lógica do cálculo, segue o mesmo explicado na variável "Sexo" do Modelo 1, com a fórmula indicada na nota anterior a essa.

utilizadas como um indicador de previsão de repetência, apesar de tal afirmativa parecer óbvia no contexto escolar, já que o baixo rendimento escolar seria a premissa principal para justificar a reprovação escolar. Entretanto, esse dado cria um questionamento acerca da recuperação⁶⁴, pois parece que ou não tem sido feita ou se feita, não tem sido efetiva.

O modelo 3 encontra-se detalhado na Tabela 35 e ele foi pensado para políticas públicas das escolas empregadas na SME-Rio.

Tabela 35: Regressão de Cox - Modelo 3

Modelo 3							
No. of subjects =	58.019				Number of obs =	161.300	
No. of failures =	16.931						
Time at risk =	161.300				LR chi2(9) =	376,46	
Log likelihood =	-180.893,90				Prob > chi2 =	0,00000	
Variável	Coefic. (b)	Razões de Risco (e ^b)	Erro Padrão	z	P>z	IC 95%	
Ginásio Carioca Ginásio	-2,64522	0,07099	0,01958	-9,59	0,00000	0,04134	0,12190
Carioca*Tempo	0,57963	1,78539	0,15997	6,47	0,00000	1,49783	2,12814
ICG	-0,07966	0,92343	0,01664	-4,42	0,00000	0,89139	0,95663
ICG*Tempo	0,03358	1,03415	0,00870	3,99	0,00000	1,01724	1,05134

Das variáveis que tínhamos disponíveis no banco de dados, apenas a variável do Ginásio Carioca e o Índice de Complexidade de Gestão foram significativas para compor tal modelo. A seguir, as interpretações dos resultados para este modelo.

- **Ginásio Carioca:** os alunos dos Ginásios Cariocas têm menos risco de repetir quando comparados aos alunos das escolas que não são deste programa. No momento inicial do estudo este risco está em 92,9% a menos, mas esse risco vai aumentando com o passar do tempo, sendo de 87,3% no tempo 1; 77,4% no tempo 2; 59,6% no tempo 3; e 27,9% no tempo 4. Como o número de alunos que eram do Ginásio Carioca no 6º ano é muito pequeno em relação à coorte, conforme já descrito anteriormente, é prudente olhar os resultados nos tempos iniciais deste estudo com cautela.
- **ICG:** este indicador é composto com 6 níveis ordenados, portanto, o nível 1 é indicador de uma escola pouco complexa, enquanto o nível 6 indica escolas muito complexas. O comportamento das razões de chance ao longo do tempo chama

⁶⁴ Segundo a alínea “e” do inciso V, do art. 24 da Lei nº 9.394/96 “a 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”.

atenção, pois na primeira metade do estudo, os resultados indicam que o risco de repetência são menores e na outra metade os riscos são aumentados, vejamos como é cada razão de risco no tempo: No tempo 1, a cada nível do ICG acrescido para as escolas, o risco dos seus alunos em repetir é de 4,5% a menos quando comparados com os alunos de escolas do nível anterior; no tempo 2, o risco é de apenas 1,2% a menos para cada nível do ICG; já no tempo 3, as razões de risco são invertidas, sendo que para cada nível do ICG aumentado, há um aumento de 2,1% no risco de repetir; e no tempo 4, os riscos são de 5,6% a mais de repetir para cada nível somado no ICG. Tais resultados indicam que em séries mais avançadas, escolas mais complexas em relação à sua gestão aumentam o risco de repetência.

O Modelo 4 teve como objetivo unir as variáveis socioeconômicas com as de desempenho acadêmico dos alunos. Na Tabela 36, encontram-se os resultados e logo a seguir a interpretação dos mesmos.

Tabela 36: Regressão de Cox - Modelo 4

Modelo 4							
No. of subjects =	47.229				Number of obs =	133.773	
No. of failures =	13.459						
Time at risk =	133.773				LR chi2(9) =	9.961,14	
Log likelihood =	-136.388,87				Prob > chi2 =	0,00000	
Variável	Coefic. (b)	Razões de Risco (e^b)	Erro Padrão	z	P>z	IC 95%	
Idade	0,32678	1,38650	0,02476	18,30	0,00000	1,33880	1,43589
Anos de Defasagem	-0,27716	0,75793	0,01361	-15,43	0,00000	0,73171	0,78509
Sexo	-0,35021	0,70454	0,02817	-8,76	0,00000	0,65143	0,76197
Sexo*Tempo	0,03570	1,03634	0,01840	2,01	0,04400	1,00091	1,07304
Não Branco	0,04084	1,04168	0,01952	2,18	0,02900	1,00412	1,08065
Bolsa Família Cartão Família	0,05015	1,05143	0,02421	2,18	0,02900	1,00503	1,09997
Carioca	-0,06932	0,93303	0,02365	-2,73	0,00600	0,88781	0,98055
Biparental	-0,08412	0,91932	0,03561	-2,17	0,03000	0,85212	0,99183
Biparental*Tempo	-0,05343	0,94797	0,01648	-3,07	0,00200	0,91621	0,98083
Máx. Escolaridade dos Pais	-0,15772	0,85409	0,01685	-7,99	0,00000	0,82169	0,88776
PB_LPL_1e2	-0,16918	0,84436	0,00891	-16,03	0,00000	0,82707	0,86200
PB_LPL_1e2*Tempo	-0,03698	0,96369	0,00491	-7,26	0,00000	0,95412	0,97337
PB_MAT_1e2	-0,26960	0,76369	0,00501	-41,06	0,00000	0,75392	0,77358
Mov_Turma	0,32442	1,38323	0,04759	9,43	0,00000	1,29304	1,47972

Como neste momento estamos combinando dois modelos anteriores, é interessante além de interpretar as razões de risco, também observar como as variáveis se comportaram perante esta nova interação.

- **Idade:** o aumento da idade continua aumentando o risco dos alunos para repetir, sendo que neste modelo o risco de repetência aumentou para 38,6% para cada ano acrescido.
- **Anos de Defasagem⁶⁵:** essa variável foi calculada apenas para os alunos com defasagem idade-série (2 anos ou mais além da idade considerada ideal para a série, isto é, com pelo menos 13 anos) e é definida pela diferença da idade do ano em 2012 menos a idade considerada ideal pra o 6º ano (11 anos). Esta variável apontou para a direção diferente do esperado, até porque está contradizendo o resultado da variável Idade. O resultado indica que para cada ano a mais de defasagem, diminui o risco de repetência em 24,2%, este dado pode estar relacionado à maturação dos alunos.
- **Sexo:** o sexo apresenta o mesmo comportamento do modelo 1, inclusive a interação da variável com tempo também continua indicando que com o passar do tempo, o risco das meninas vai diminuindo em relação aos meninos, entretanto, neste modelo, a disparidade é menor do que ao modelo anterior. No tempo inicial, o risco das meninas é 29,5% menor do que a dos meninos e com o passar do tempo, temos os seguintes resultados: 27,0% no tempo 1; 24,3% no tempo 2; 21,6% no tempo 3; e 18,7% no tempo 4.
- **Não branco:** neste modelo, esta variável perdeu sua magnitude, trazendo sua razão de risco para 4,2% a mais em relação aos alunos brancos. Lembrando que no modelo 1, essa razão era de 19,4%.
- **Bolsa Família:** esta variável teve o mesmo comportamento da variável de raça, mantendo a mesma direção, mas perdendo sua magnitude. Aqui, os beneficiários do programa Bolsa Família têm 5,1% a mais de risco de repetirem do que os não beneficiários, contra os 11,7% do modelo 1.
- **Cartão Família Carioca⁶⁶:** o CFC é um programa social do município do Rio de Janeiro que complementa a renda dos beneficiários do Bolsa Família (Decreto 32.887, de 08 de

⁶⁵ A variável Anos_Defasagem foi acrescentada ao Modelo 4 para que a variável Idade_2012 atenda a suposição de proporcionalidade (para maiores detalhes, ver Anexo 3).

⁶⁶ Optou-se por acrescentar a variável CFC para tentar melhorar a significância do programa Bolsa Família, que se perdeu ao construir o Modelo 4 (para maiores detalhes, ver Anexo 3).

outubro de 2010). Os beneficiários desse programa têm 6,7% de risco a menos de repetir quando comparados aos não beneficiários. Uma informação importante sobre este programa é que é previsto um benefício adicional de R\$50,00 por bimestre conforme o desempenho escolar das crianças e adolescentes da família⁶⁷, este fato pode ajudar a explicar a direção da razão de risco ser diferente do programa Bolsa Família.

- **Biparental:** o fato dos alunos morarem com seus dois pais continua diminuindo o risco dos alunos de repetir uma série, entretanto nesse modelo, a razão de risco no momento inicial é bem menor, sendo de apenas 8,1% a menos, o risco de repetir do que os que não moram com seus dois pais (no modelo 1, esse resultado era de 17,1%), ao longo do tempo, os riscos vão diminuindo, sendo: 12,9% no tempo 1; 17,4% no tempo 2; 21,7% no tempo 3; e 25,8% no tempo 4. No modelo 1, no tempo 4 os alunos tinham 28,6% a menos de risco, portanto, apesar de diferir bastante no momento inicial, não é o mesmo que ocorre como passar do tempo.
- **Máxima Escolaridade dos Pais:** neste modelo esta variável apresentou proporcionalidade de risco ao longo do tempo, não sendo mais necessário trazer a sua interação com o tempo. Os alunos com pelo menos um dos pais com no mínimo o ensino médio completo tem 14,6% a menos de risco de repetirem quando comparados aos outros alunos. Perceba que essa razão é bem menor do que todas apresentadas para esta mesma variável no modelo 1.
- **PB_LPL_1e2:** neste modelo, essa variável teve praticamente o mesmo resultado do modelo 2, sendo que para cada ponto a mais na média, o risco de repetência é 15,6% menor no momento inicial e de 18,6% no tempo 1; 21,6% no tempo 2; 24,4% no tempo 3; e 27,2% no tempo 4, tendo em média 2 pontos percentuais a menos do que no modelo 2.
- **PB_MAT_1e2:** esta variável também ficou bem parecida como no modelo 2, indicando que a cada ponto na média do aluno, diminui o risco de repetência em 23,6% (no modelo 2 era de 24,1%).

⁶⁷ Disponível em: <<http://www.rio.rj.gov.br/web/cvl/exibeconteudo?article-id=2248211>> . Acesso em 25 Jun. 2018

- **Mov_Turma:** cada transferência de turma do aluno, aumenta o risco de repetência em 38,3%, esta variável continuou tendo uma razão de risco alta, sendo um pouco menor do que o apresentado no modelo 2 (43,8%).

Com os resultados do modelo 4, observamos que unindo as variáveis de perfil socioeconômico com as desempenho acadêmico, não houve grandes diferenças na interpretação das razões de risco, além do mais, para esse modelo, algumas variáveis do modelo 1 saíram e outras novas variáveis entraram⁶⁸.

Para o nosso último modelo – o modelo 5 – partimos do modelo 4 e agregamos as variáveis do modelo 3. Para esta modelagem não houve nenhuma necessidade de ajuste de variáveis. Os resultados do modelo encontram-se na Tabela 37.

Tabela 37: Regressão de Cox - Modelo 5

Modelo 5							
No. of subjects =	47.229				Number of obs =	133.773	
No. of failures =	134.59						
Time at risk =	133.773				LR chi2(9) =	10.060,51	
Log likelihood =	-136.339,18				Prob > chi2 =	0,000000	
Variável	Coefic. (b)	Razões de Risco (e ^b)	Erro Padrão	z	P>z	IC 95%	
Idade	0,32585	1,38521	0,02474	18,25	0,00000	1,33756	1,43455
Anos de Defasagem	-0,27647	0,75846	0,01362	-15,39	0,00000	0,73222	0,78564
Sexo	-0,34894	0,70544	0,02820	-8,73	0,00000	0,65227	0,76294
Sexo*Tempo	0,03509	1,03571	0,01838	1,98	0,04800	1,00030	1,07238
Não Branco	0,04291	1,04384	0,01956	2,29	0,02200	1,00620	1,08289
Bolsa Família	0,05100	1,05233	0,02423	2,21	0,02700	1,00588	1,10091
Cartão Família Carioca	-0,06887	0,93345	0,02366	-2,72	0,00700	0,88820	0,98100
Biparental	-0,08585	0,91773	0,03555	-2,22	0,02700	0,85063	0,99012
Biparental*Tempo	-0,05394	0,94749	0,01648	-3,10	0,00200	0,91574	0,98035
Máx. Escolaridade dos Pais	-0,15811	0,85376	0,01684	-8,01	0,00000	0,82138	0,88741
PB_LPL_1e2	-0,17077	0,84301	0,00892	-16,15	0,00000	0,82572	0,86067
PB_LPL_1e2*Tempo	-0,03597	0,96467	0,00493	-7,03	0,00000	0,95505	0,97439
PB_MAT_1e2	-0,26617	0,76631	0,00505	-40,40	0,00000	0,75648	0,77627
Mov_Turma	0,32975	1,39062	0,04785	9,58	0,00000	1,29992	1,48765
Ginásio Carioca	-2,33404	0,09690	0,03062	-7,39	0,00000	0,05216	0,18002
Ginásio Carioca*Tempo	0,65548	1,92608	0,19715	6,40	0,00000	1,57597	2,35396
ICG	-0,05351	0,94789	0,01892	-2,68	0,00700	0,91152	0,98572
ICG*Tempo	0,02948	1,02992	0,00965	3,15	0,00200	1,01118	1,04900

⁶⁸ Conforme já informado anteriormente, o processo de modelagem encontra-se no Anexo desta dissertação.

Neste nosso último modelo, faremos o mesmo procedimento do modelo 4, analisaremos as razões de chance de cada variável e comparando com as mesmas no modelo 4 e no modelo 3.

- **Idade:** a razão de risco dessa variável praticamente não sofreu nenhuma alteração, para cada ano a mais que o aluno possui, seu risco em repetir é acrescido de 38,5%. No modelo 4, este valor era de 38,6%.
- **Anos de Defasagem:** o resultado desta variável é praticamente igual (variando apenas a partir da terceira casa decimal), para cada ano adicional de defasagem, o risco diminui em 24,15%.
- **Sexo:** a variável sexo também teve os mesmos resultados, tanto no momento inicial, quanto na sua interação com o tempo, sendo de 29,5% no momento inicial e 26,9% no tempo 1; 24,3% no tempo 2; 21,6% no tempo 3; e 18,8% no tempo 4.
- **Não Branco:** o aluno não branco tem 4,4% a mais de risco do que o aluno branco. Este resultado também está próximo ao encontrado no Modelo 4 – 4,2%.
- **Bolsa Família:** os beneficiários possuem 5,2% mais risco do que os não beneficiários, aumentando apenas 0,1 ponto percentual em relação ao modelo 4.
- **Cartão Família Carioca:** os beneficiários deste programa têm 6,7% a menos de risco de repetir do que os não beneficiários, sendo exatamente o mesmo valor encontrado no modelo 4.
- **Biparental:** os resultados desta variável também foram bem próximos aos achados no modelo 4, sendo de 8,2% para o momento inicial; de 13,0% para o tempo 1; de 17,6% para o tempo 2; de 21,9% para o tempo 3; e de 26,0% para o tempo 4.
- **Máxima Escolaridade dos Pais:** também similar ao modelo 4, indicando que os alunos com pelo menos um dos pais com no mínimo o ensino médio completo tem 14,6% a menos de risco de repetirem quando comparados aos outros alunos.
- **PB_LPL_1e2:** cada ponto a mais na média, o risco de repetência é 15,7% menor no momento inicial e de 18,7% no tempo 1; 21,5% no tempo 2; 24,3% no tempo 3; e 27,0% no tempo 4
- **PB_MAT_1e2:** cada ponto a mais na média do aluno, diminui o risco de repetência em 23,4%.

- **Mov_Turma:** cada transferência de turma a mais, aumenta o risco de repetência em 39,1%, aumentando em 0,8 pontos percentuais quando comparado ao modelo 4.
- **Ginásio Carioca:** No momento inicial do estudo, os alunos possuem 90,3% a menos de risco em repetir do que os outros alunos que não são de escolas deste programa, já no tempo 1, o risco dos alunos aumenta, passando para 81,3% no tempo 1; 64,1% no tempo 2; 30,8% no tempo 3; e no tempo 4, curiosamente, os alunos dos Ginásios Cariocas apresentam um maior risco de repetência (33,4%) quando comparados aos alunos das outras escolas.
- **ICG:** esta variável apresentou o mesmo comportamento do modelo 3, sendo que no momento inicial, a cada nível de ICG que a escola aumentava, o risco de repetência de seus alunos era 5,2% menor quando comparado a alunos de escolas com um nível menor. No tempo 1, esse risco diminui para 2,4; a partir do tempo 2, as razões de risco se invertem passando aumentar a cada nível de ICG que a escola aumento, tendo 0,5% a mais no tempo 2; 3,6% no tempo 3; e 6,7% no tempo 4.

Portanto, conforme vimos, ao unir ao modelo 4 o modelo 3 (e fazer o modelo 5), praticamente não houve alterações nas variáveis de perfil socioeconômico dos alunos e seus desempenhos acadêmicos, todas continuaram apontando para a mesma direção e inclusive não alteram muito sua magnitude.

Na **Erro! Fonte de referência não encontrada.**, encontram-se os resumos dos cinco modelos, os coeficientes (b), as razões de risco (RR) e a estatística z^{69} das variáveis de cada um dos modelos e no final da tabela, o número de alunos observados (n), o resultado do qui-quadrado (χ^2) e o Log Verossimilhança Parcial de cada modelo.

⁶⁹ O valor Z é uma estatística de teste para testes Z que mede a diferença entre uma estatística observada e seu parâmetro de população em unidades de erro padrão.

Tabela 38: Modelos de Regressão de Cox

Variáveis Explicativas	Modelo 1			Modelo 2			Modelo 3			Modelo 4			Modelo 5		
	Razões de Risco (e ^b)	Coefic. (b)	z	Razões de Risco (e ^b)	Coefic. (b)	z	Razões de Risco (e ^b)	Coefic. (b)	z	Razões de Risco (e ^b)	Coefic. (b)	z	Razões de Risco (e ^b)	Coefic. (b)	Z
Idade_2012	1,27136	0,24009	39,92*							1,38650	0,32678	18,3*	1,38521	0,32585	18,25*
Anos_Defasagem										0,75793	-0,27716	-15,43*	0,75846	-0,27647	-15,39*
Sexo	0,59385	-0,52114	-13,52*							0,70454	-0,35021	-8,76*	0,70544	-0,34894	-8,73*
sexo_t	1,07189	0,06942	4,06*							1,03634	0,03570	2,01*	1,03571	0,03509	1,98*
Nao_Branco	1,19444	0,17767	9,74*							1,04168	0,04084	2,18*	1,04384	0,04291	2,29*
Bolsa_Familia	1,11677	0,11044	6,45*							1,05143	0,05015	2,18*	1,05233	0,05100	2,21*
CFC										0,93303	-0,06932	-2,73*	0,93345	-0,06887	-2,72*
Biparental	0,82902	-0,18751	-4,95*							0,91932	-0,08412	-2,17*	0,91773	-0,08585	-2,22*
Biparental_t	0,96336	-0,03733	-2,21*							0,94797	-0,05343	-3,07*	0,94749	-0,05394	-3,1*
Max_Esc_Pais	0,60545	-0,50178	-11,72*							0,85409	-0,15772	-7,99*	0,85376	-0,15811	-8,01*
Max_Esc_Pais_t	1,06404	0,06207	3,35*												
PB_LPL_1e2				0,82486	-0,19254	-20,59*				0,84436	-0,16918	-16,03*	0,84301	-0,17077	-16,15*
PB_MAT_1e2				0,96128	-0,03949	-8,56*				0,76369	-0,26960	-41,06*	0,76631	-0,26617	-40,4*
PB_LPL_1e2_t				0,75885	-0,27596	-46,72*				0,96369	-0,03698	-7,26*	0,96467	-0,03597	-7,03*
Mov_Turma				1,43757	0,36295	11,58*				1,38323	0,32442	9,43*	1,39062	0,32975	9,58*
GC							0,07099	-2,64522	-9,59*				0,09690	-2,33404	-7,39*
GC_t							1,78539	0,57963	6,47*				1,92608	0,65548	6,4*
ICG							0,92343	-0,07966	-4,42*				0,94789	-0,05351	-2,68*
ICG_t							1,03415	0,03358	3,90*				1,02992	0,02948	3,15*
n		48.598			56.270			58.019			47.229			47.229	
χ ²		2.709,49*			10.798,31*			376,46*			9.961,14*			10.060,51*	
Log Likelihood		-147.339,34			-166.614,07			-180.893,90*			-136.388,87			-136.339,18	

7 CONSIDERAÇÕES FINAIS

O fenômeno da repetência escolar é objeto de estudos em diferentes redes de ensino no Brasil e no exterior. Entretanto, essa dissertação tenta analisar a repetência utilizando uma nova metodologia e explora a possibilidade de trabalhar com dados longitudinais fornecidos pela SME-Rio. Ressaltamos, portanto, três aspectos inovadores da pesquisa: a) uso da base de dados com característica longitudinal para analisar fatores associados à repetência escolar⁷⁰; b) uso da análise de sobrevivência; c) uso da Prova Bimestral e conceitos bimestral como estimador no modelo.

Os resultados descritivos da trajetória dos alunos indicaram que menos da metade dos alunos da nossa coorte (45,97%) conseguiu ter uma trajetória regular ao longo das transições nos anos finais do ensino fundamental e que outros 16.931 alunos (29,18%) sofreram pelo menos uma repetência ao longo deste período. Tais achados reforçam o tamanho do problema que a rede pública municipal enfrenta, pois de nada adianta termos as crianças e jovens nas escolas se não conseguimos garantir o aprendizado dos alunos. O debate sobre a qualidade da educação deve necessariamente considerar os achados dessa dissertação e outros estudos que demonstram que garantir o acesso à escola é insuficiente para lidar com os problemas observados nas redes públicas no Brasil.

Já na análise de sobrevivência, os achados estão compatíveis com a literatura do campo. Observamos que algumas características por si só já influenciam a trajetória dos alunos, como sexo, raça, escolaridade dos pais, dentre outras. O modelo de regressão de Cox permite analisar como tais características aumentam ou diminuem o risco dos alunos em experimentar uma repetência, e, novamente, a grande maioria das variáveis se comportou como esperado. Por exemplo, as características associadas à desvantagem social influenciam negativamente, ou seja, possuir tais características aumenta o risco do aluno repetir ao longo das transições dos anos finais do ensino fundamental.

Pensando em políticas públicas de forma macro, encontramos indícios de que os resultados das Provas Bimestrais no primeiro semestre do ano letivo podem ser bons preditores de repetências no futuro, sugerindo então que a recuperação dos alunos ao longo

⁷⁰ Sabemos que outros estudos realizados na rede pública municipal do Rio de Janeiro utilizaram base de dados com característica longitudinal para estimar diferentes fenômenos, mas é forçoso ressaltar que a maioria dos estudos utiliza dados transversais nas suas análises (BARTHOLO, 2014a; MARINO, 2014; CASTRO, 2016).

do ano letivo pode não estar sendo feita de forma eficaz. Identificar os alunos em situação de risco antes do desfecho desfavorável, nesse caso a reprovação, é uma das condições necessárias para a mudança das altas taxas de repetência. Acreditamos que esse é um dos resultados mais relevantes dessa dissertação. No Brasil outros estudos já utilizaram dados das avaliações externas para estimar o risco de reprovação, no entanto, esses dados são pouco úteis para ajudar o gestor na tomada de decisão sobre a coorte estudada. Isso porque os resultados das avaliações externas demoram em média 12 meses para serem divulgados, o que inviabiliza a formulação de planos de recuperação para a coorte. No entanto, os dados da Prova Bimestral na SME-RJ são rapidamente processados e devolvidos para os diretores e professores da rede. Isso permite identificar de forma rápida os alunos em maior situação de risco. O segundo passo seria a adoção de um programa eficaz para recuperar o aprendizado dos alunos em situação de risco. Os resultados do modelo final apresentado na Tabela 38 reforçam o alto poder preditivo das Provas Bimestrais quando estimamos o risco de repetência nos anos finais do ensino fundamental. Acreditamos que a formulação e avaliação de impacto desses programas de recuperação devem ser realizadas com desenhos robustos de pesquisa.

Os resultados preliminares sugerem que o Programa Ginásio Carioca tem afeito protetor na trajetória escolar dos alunos, entretanto, para assegurar tais resultados, sugerimos um estudo quase experimental, com maior robustez metodológica, no qual se compare grupos similares, e assim possa aumentar a validade interna, minimizando o risco de possíveis vieses de seleção.

Uma limitação que precisamos destacar mais uma vez sobre este estudo é o desconhecimento da trajetória escolar pregressa da nossa coorte. Não sabemos o que aconteceu com nossos alunos antes de 2012. Não sabemos como foram as trajetórias dos anos iniciais e até mesmo do próprio 6º ano. Os dados permitem apenas observar que 59,3% deles estavam dentro da idade ideal (11 anos) em 2012. Portanto, quase 40% da nossa coorte pode ter tido trajetória irregular desde o início do ensino fundamental, fazendo com que este fato influenciasse na sua trajetória dos anos finais, sem que nós pudéssemos captar

neste estudo. Por isso, sugerimos para novos estudos uma análise de sobrevivência para uma coorte desde o 1º ano do ensino fundamental⁷¹.

Outro dado descritivo relevante é a taxa de alunos que saem da SME-RJ ao longo do segundo segmento – um pouco mais de 20% dos alunos na coorte estudada. É importante que estudos futuros analisem especificamente esse fenômeno e busque identificar fatores associados à saída dos alunos do sistema. Um bom sistema de ensino deve apresentar boa aderência dos estudantes. Por isso, sugerimos um estudo mais detalhado das movimentações de saída da Rede, pois dependendo do tipo de movimentação, ações diferentes devem ser aplicadas.

Outra sugestão que podemos deixar aqui é fazer o estudo da repetência através da análise de sobrevivência, entretanto considerando que a unidade de análise possa passar por diversas repetências⁷², embora tal ideia tenha sido considerada inicialmente, por parcimônia, optamos em não segui-la neste estudo, justamente por entender esse tipo de investigação exigiria mais em relação à experiência prévia da autora.

E por fim, indicamos a necessidade de replicação da dissertação para outras coortes da rede pública. A ideia aqui é comparar os resultados obtidos com a replicação e observar se há regularidade nos achados. Cabe reforçar que esse é o primeiro estudo que conhecemos que utiliza a análise de sobrevivência com dados do SGA da SME-RJ.

⁷¹ Uma análise futura irá refazer as análises para essa mesma coorte, no entanto, selecionando apenas os alunos na idade ideal e verificar se as variáveis se modificariam para este grupo, entretanto, por questões de tempo, tivemos que deixar esta ideia para um próximo estudo.

⁷² O processo investigado de multiepisódio é quando a unidade de análise pode ter experimentado o evento de interesse mais de uma vez (BERNARDI, 2006, p. 41)

REFERÊNCIAS

AFONSO, Almerindo Janela. **Avaliação educacional - regulação e emancipação**. 4. ed. São Paulo: Cortez, 2009. 152p.

ALLISON, Paul D. **Event history and survival analysis**. 2ª ed. Series: Quantitative Applications in the Social Sciences, n. 46. EUA: Sage Publications, 2014. 93p.

BACCHETTO, João Galvão. O Pisa e o custo da repetência no Fundeb. **Ensaio: Avaliação e Políticas Públicas em Educação**, Rio de Janeiro, v.24, n. 91, p. 424-444, abr./jun. 2016. Disponível em: <<http://www.scielo.br/pdf/ensaio/v24n91/1809-4465-ensaio-24-91-0424.pdf>>. Acesso em: 15 abr. 2018.

BARTHOLO, Tiago Lisboa. **Segregação escolar na rede municipal do Rio de Janeiro: Causas e consequências**. Tese (Doutorado em Educação) – Faculdade de Educação, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2014a. Disponível em: <<http://www.educacao.ufrj.br/ttiagobartholo.pdf>>. Acesso em: 02 jul. 2018.

BARTHOLO, T. L. Segregação escolar na cidade do Rio de Janeiro: análise da movimentação de estudantes. *Estudos em Avaliação Educacional (Online)*, v. 25, p. 242, 2014b.

BARTHOLO, Tiago Lisboa; COSTA, Marcio da. Turnos e segregação escolar: discutindo as desigualdades intraescolares. **Cadernos de Pesquisa**, São Paulo, v. 44, n. 153, p. 670-692, Sept. 2014. Disponível em:< http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-15742014000300010&lng=en&nrm=iso>. Acesso em: 02 jul. 2018.

BARTHOLO, Tiago Lisboa; COSTA, Marcio da. Evidence of a school composition effect in Rio de Janeiro public schools. **Ensaio: Avaliação e Políticas Públicas em Educação**. Rio de Janeiro, v.24, n. 92, p. 498-521, jul./set. 2016. Disponível em: <<http://www.scielo.br/pdf/ensaio/v24n92/1809-4465-ensaio-24-92-0498.pdf>>. Acesso em 15 jul. 2018.

BERNARDI, Fabrizio. **Análisis de la Historia de Acontecimientos**. Colección Cuadernos Metodológicos, n. 38. Madrid: Centro de Investigaciones Sociológicas, 2006. 168p.

BISINOTO, Cynthia; MARINHO, Claisy; ALMEIDA, Leandro. Contribuições da psicologia escolar à promoção do sucesso acadêmico na educação superior. In: Universidade do Minho (Org.). **Anais do I Seminário Internacional “Contributos da Psicologia em Contextos Educativos”**, p. 102-116, Braga, 2010. Disponível em: <<https://repositorium.sdum.uminho.pt/bitstream/1822/11460/1/010.pdf>>. Acesso em: 15 abr. 2018.

BOMENY, Helena. Alunos do 6º na da rede municipal do Rio terão uma só professora por quatro dias na semana. **Jornal Extra**, Rio de Janeiro, 25 nov. 2014. Entrevista concedida a Bruno Alfano. Disponível em: <<https://extra.globo.com/noticias/educacao/alunos-do-6-ano-da-rede-municipal-do-rio-terao-uma-so-professora-por-quatro-dias-na-semana-14651540.html>>. Acesso em: 05 jul. 2018.

BOUZADA, Vanessa Catarina Duarte. **O programa Escolas do Amanhã: Mais um programa interrompido?** Monografia (Graduação) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2014. Disponível em:
<<http://pantheon.ufrj.br/bitstream/11422/2764/1/VBouzada.pdf>>. Acesso em: 05 jul. 2018.

BRASIL. **Constituição da República Federativa do Brasil**. Promulgada em 05 de outubro de 1988. Brasília, DF. 1988. Disponível em:
<http://www.planalto.gov.br/ccivil_03/constituicao/constituicaocompilado.htm>. Acesso em: 08 abr. 2018.

BRASIL. **Lei de Diretrizes e Bases da Educação Nacional**. Lei nº 9.394, 20 de dezembro de 1996. Brasília, DF. 1996. Disponível em:
<http://www.planalto.gov.br/ccivil_03/Leis/L9394.htm>. Acesso em: 05 abr. 2018.

BRASIL. Ministério da Educação. Secretaria de Articulação com os Sistemas de Ensino. **Planejando a próxima década: Conhecendo as 20 metas do Plano Nacional de Educação**. Brasília, DF: MEC/SASE, 2014. Disponível em:
<http://pne.mec.gov.br/images/pdf/pne_conhecendo_20_metas.pdf>. Acesso em: 05 abr. 2018.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Plano Nacional de Educação PNE 2014-2024: Linha de Base**. Brasília, DF: Inep, 2015. Disponível em: <http://portal.inep.gov.br/informacao-da-publicacao/-/asset_publisher/6JYIsGMAMkW1/document/id/493812>. Acesso em: 05 abr. 2018.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Relatório do 1º ciclo de monitoramento das metas do PNE: biênio 2014- 2016**. Brasília, DF: Inep, 2016. Disponível em
<http://download.inep.gov.br/outras_acoes/estudos_pne/2016/relatorio_pne_2014_a_2016.pdf>. Acesso em: 05 abr. 2018.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Relatório do 2º ciclo de monitoramento das metas do Plano Nacional de Educação - 2018**. Brasília, DF: Inep, 2018. Disponível em:
<<http://portal.inep.gov.br/documents/186968/485745/RELAT%C3%93RIO+DO+2o.+CICLO+DE+MONITORAMENTO+DAS+METAS+DO+PNE+-+2018/9a039877-34a5-4e6a-bcfd-ce93936d7e60?version=1.14>>. Acesso em: 15 jul. 2018.

BRUEL, Ana Lorena; BARTHOLO, Thiago Lisboa. Desigualdade de oportunidades educacionais na rede pública municipal do Rio de Janeiro: transição entre os segmentos do ensino fundamental. **Revista Brasileira de Educação**, v. 17, n. 50, 303-328, 2012. Disponível em:
<<http://www.scielo.br/pdf/rbedu/v17n50/v17n50a04.pdf>>. Acesso em: 05 jul. 2018.

BRUNS, Barbara; EVANS, David; LUQUE, Javier. **Achieving Word-Class Education in Brasil: The next agenda**. Directions in Development, The World Bank, Washington, 190p., 2012. Disponível em:

<<http://documents.worldbank.org/curated/pt/993851468014439962/pdf/656590REPLACE%0A0hieving0World0Class0.pdf>>. Acesso em: 14 abr. 2018.

CALAFATE, Vitor Ribeiro Laufer. **Mecanismos de Segmentação no Interior da Escola: a experiência das políticas de correção de fluxo no município do Rio de Janeiro**. Dissertação (Mestrado em Educação) – Faculdade de Educação, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2017. Disponível em: <<http://www.educacao.ufrj.br/ppge/disserta%C3%A7%C3%B5es2017/dVitorCalafate.pdf>>. Acesso em 05 jul. 2018.

CANO, Ignácio. **Introdução à Avaliação de Programas Sociais**. 3ed. Rio de Janeiro: Editora FGV, 2006.

CASTRO, C. M. M. V. Avaliação do aprendizado no ciclo de alfabetização de alunos da rede municipal do Rio de Janeiro. Dissertação de Mestrado. Escola Nacional de Ciências Estatísticas IBGE, 2016.

CHAM, Heining *et al.* Effect of Retention in Elementary Grades on Grade 9 Motivation for Educational Attainment. **Journal of school psychology**, v. 53, n. 1, p. 7-24, 11/14 2015. ISSN 0022-4405/1873-3506. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4312584/>>. Acesso em: 15 abr. 2018.

CHRISTOVÃO, Ana Carolina. **Compensação Educacional no Rio de Janeiro: avaliando o programa Escolas do Amanhã**. Tese (Doutorado em Educação) – Faculdade de Educação, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2018. Disponível em: <<http://www.educacao.ufrj.br/ppge/teses2018/tAnaCarolinaChristovao.pdf>>. Acesso em 05 jul. 2018.

COLOSIMO, Enrico Antônio; GIOLO, Suely Ruiz. **Análise de sobrevivência aplicada**. São Paulo: Edgard Blücher, 2006. 370p.

CORREA, Erisson Viana; BONAMINO, Alicia; SOARES, Tufi Machado. Evidências do efeito da repetência nos primeiros anos escolares. **Estudos em Avaliação Educacional**. São Paulo, v. 25, n. 59, p. 242-269, set./dez. 2014. Disponível em: <<http://publicacoes.fcc.org.br/ojs/index.php/eae/article/view/2862/2859>>. Acesso em: 14 abr. 2018.

COSTA, Marcio da; KOSLINSKI, Mariane Campelo. Quase-mercado oculto: a disputa por escolas “comuns” no Rio de Janeiro. **Cadernos de Pesquisa**. São Paulo, v. 41, n.142, p. 246-266, Abr. 2011. Disponível em: <<http://www.scielo.br/pdf/cp/v41n142/v41n142a13.pdf>>. Acesso em: 05 ju. 2018.

_____. Escolha, estratégia e competição por escolas públicas. **Pró-Posições**. Campinas, v.23, n. 2, p.195-213, ago. 2012. Disponível em: <<http://www.scielo.br/pdf/pp/v23n2/a13v23n2.pdf>>. Acesso em: 05 jul. 2018.

CRAHAY, Marcel. Qual pedagogia para os alunos em dificuldade escolar? **Cadernos de Pesquisa**, São Paulo, v. 37, n. 130, p. 181-208, jan./abr. 2007. Disponível em: <<http://www.scielo.br/pdf/cp/v37n130/09.pdf>>. Acesso em: 18 abr. 2018.

CRUZ, Priscila. Brasil possui quase 2,5 milhões de crianças e adolescentes fora da escola, diz estudo. **G1**, Rio de Janeiro, 05 abr. 2017. Disponível em: <<https://g1.globo.com/educacao/noticia/brasil-possui-quase-25-milhoes-de-criancas-e-adolescentes-fora-da-escola-diz-estudo.ghtml>>. Acesso em 15 jul. 2018.

FAJARDO, Vanessa. Brasil gasta R\$16 bilhões com reprovação de 3 milhões de alunos em 2016, aponta levantamento. **G1**, Rio de Janeiro, 02 mar. 2018. Disponível em: <<https://g1.globo.com/educacao/noticia/brasil-gasta-r-16-bilhoes-com-reprovacao-de-3-milhoes-de-alunos-em-2016-aponta-levantamento.ghtml>> Acesso em: 14 abr. 2018.

FERNANDES, Reynaldo. Índice de Desenvolvimento da Educação Básica (IDEB). **Textos para discussão**, n.26, 2. ed. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), maio, 2007. Disponível em: <<http://portal.inep.gov.br/documents/186968/485287/%C3%8Dndice+de+Desenvolvimento+da+Educa%C3%A7%C3%A3o+B%C3%A1sica+%28Ideb%29/26bf6631-44bf-46b0-9518-4dc3c310888b?version=1.4>>. Acesso em: 18 abr. 2018.

FERRÃO, Maria Eugénia; BELTRÃO, Kaizô Iwakami; SANTOS, Denis Paulo dos. O impacto de políticas de não-repetência sobre o aprendizado dos alunos da 4ª série. **Pesquisa e Planejamento Econômico**, v. 32, n.3, p. 495-513, dez. 2002. Disponível em: <<http://ppe.ipea.gov.br/index.php/ppe/article/viewFile/141/76>> . Acesso em: 05 jul. 2018.

FLETCHER, Phillip R. A Repetência no ensino de 1º grau: um problema negligenciado da educação brasileira. **Revista Brasileira de Administração da Educação**, v. 3, n. 1, 1985. Disponível em: <<http://seer.ufrgs.br/index.php/rbpaee/issue/view/issue/3190/421>>. Acesso em: 05 jul. 2018.

FLETCHER, Phillip R.; RIBEIRO, Sérgio Costa. **Modeling Education System Performance with Demographic Data**: an introduction to the PROFUXO Model. Paris: Unesco, 1989. Disponível em: <<http://www.sergiocostaribeiro.ifcs.ufrj.br/artigos/1989%20Modeling%20education%20system.pdf>>. Acesso em: 05 jul. 2018.

FULGENCIO, Paulo Cesar. **Glossário Vade Mecum**: administração pública, ciências contábeis, direito, economia, meio ambiente: 14.000 termos e definições. Rio de Janeiro: Mauad X, 2007.

GARCIA, Teise de Oliveira Guaranha; ADRIÃO, Theresa; BORGHI, Raquel. **A nova gestão pública e o contexto brasileiro**. In: MARTINS, A. M. (Org.). Instituições educacionais: políticas, gestão e práticas profissionais. Santos: Universitária Leopoldianum; 2009. 164p.

GAWRYSZEWSKI, Bruno. O programa “Escolas do Amanhã” e a escola pública para a classe trabalhadora. **Revista Contemporânea de Educação**, vol. 10, n. 20, p. 382-404, dez/2015.

ISSN 1809-5747. Disponível em:

<<https://revistas.ufrj.br/index.php/rce/article/view/1828/2288>> Acesso em: 05 jul. 2018.

GERTLER, Paul J.; MARTINEZ, Sebastian; PREMAND, Patrick; RAWLINGS, Laura B.; VERMEERSCH, Christel M. J. **Impact Evaluation in Practice**. 2ª ed. Washington: The World Bank, 2016. Disponível em:

<<http://documents.worldbank.org/curated/pt/698441474029568469/pdf/108270-PUB-Box396299B-PUBLIC-PUBDATE-9-13-16.pdf>>. Acesso em: 05 jul. 2018.

GOMES-NETO, João Batista; HANUSHEK, Eric A. The causes and effects of grade repetition. In: BIRDSALL, Nancy; SABOT, Richard H. (Eds.). **Opportunity forgone: education in Brazil**. Washington, DC: Inter-American development Bank, 1994. p. 425-460. Disponível em: <<http://hanushek.stanford.edu/sites/default/files/publications/Gomes-Neto%20Hanushek%201994%20EDCC%2043%281%29.pdf>>. Acesso em: 15 abr. 2018.

GORARD, S. Who is eligible for free school meals? Characterising free school meals as a measure of disadvantage in England, *British Educational Research Journal*, 38:6, 2012, p. 1003-1017, DOI: <https://doi.org/10.1080/01411926.2011.608118>. Acesso em: 15 abr. 2018.

GORARD, Stephen; SEE, Beng Huat. **Overcoming Disadvantage in Education**. London: Routledge Falmer, 2013.

HOLMES, C. Thomas.; MATTHEWS, Kenneth M. The Effects of nonpromotion on elementary and junior high school pupils: a meta-analysis. **Review of Educational Research**, v.54, n.2, p.225-236, 1984.

HORTA NETO, João Luiz. **Avaliação externa de escolas e sistemas: questões presentes no debate sobre o tema**. Brasília: Revista Brasileira de Estudos Pedagógicos, v. 91, n. 227, p. 84-104, jan./abr. 2010. Disponível em:

<<http://rbep.inep.gov.br/index.php/rbep/article/view/604/585>> Acesso em: 06 abr. 2018.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Geografia da Educação Brasileira 2001**. Brasília, Inep, 2002. Disponível em:

<<http://www.portalinep.gov.br/download/texto/me002457.pdf>>. Acesso em: 05 jul. 2018.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Sinopse Estatística da Educação Básica 2015**. Brasília, Inep, 2016. Disponível em:

<http://download.inep.gov.br/informacoes_estatisticas/sinopses_estatisticas/sinopses_educacao_basica/sinopse_estatistica_educacao_basica_2017.zip>. Acesso em: 05 jul. 2018.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Indicadores de Fluxo Escolar da Educação Básica** (apresentação). Seminário 10 Anos da Metodologia de Coleta de Dados Individualizados dos Censos Educacionais. Brasília, Inep/DEED, 2017. Disponível em:

<http://download.inep.gov.br/educacao_basica/censo_escolar/apresentacao/2017/apresentacao_indicadores_de_fluxo_escolar_da_educacao_basica.pdf>. Acesso em: 05 jul. 2018.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Sinopse Estatística da Educação Básica 2017**. Brasília, Inep, 2018. Disponível em: <<http://inep.gov.br/web/guest/sinopses-estatisticas-da-educacao-basica>>. Acesso em: 02 jul. 2018.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Estimativas de fluxo escolar a partir do acompanhamento longitudinal dos registros de aluno do Censo Escolar do período 2007-2016** (Nota técnica nº. 8/2017/CGCQT/DEED). 27 jun. 2018. Disponível em: <http://download.inep.gov.br/informacoes_estatisticas/indicadores_educacionais/2014/escola_complexidade_gestao/nota_tecnica_indicador_escola_complexidade_gestao.pdf>. Acesso em: 05 jul. 2018.

ITAÚ SOCIAL. **Avaliação de Impacto do Programa 6º Ano Experimental - Secretaria Municipal de Educação Rio de Janeiro**. São Paulo, 2017. Não publicado.

JIMERSON, Shane et al. A prospective longitudinal study of the correlates and consequences of early grade retention. **Journal of School Psychology**, v. 35, p. 3-25, 1997. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0022440596000337>>. Acesso em: 15 abr.2018.

JACKSON, Gregg B. The research evidence on the effects of grade retention. **Review of Education Research**, vol 45, p.613-635, December, 1975.

JIMERSON, Shane R.; CARLSON, Elizabeth; ROTERT, Monique; EGELAND, Byron; SROUFE, L. Alan. A perspective, longitudinal study of the correlates and consequences of early grade retention. **Journal of School Psychology**, vol.35, p.3-25, march, 1997.

KLAPPROTH, Florian; SCHALTZ, Paule. Who is retained in school, and when? Survival analysis of predictors of grade retention in Luxembourgish secondary school. **European Journal of Psychology of Education**, vol. 30, n.1, p. 119-136, Mar. 2015. Disponível em: <https://www.researchgate.net/publication/268223001_Who_is_retained_in_school_and_when_Survival_analysis_of_predictors_of_grade_retention_in_Luxembourgish_secondary_school>. Acesso em 14 abr. 2018.

KLAPPROTH, F. et al. Short-term and medium-term effects of grade retention in secondary school on academic achievement and psychosocial outcome variables. **Learning and Individual Differences**, v. 50, p. 182-194, 2016/08/01/ 2016. ISSN 1041-6080. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1041608016301583>>. Acesso em: 15 abr. 2018.

KLEIN, Ruben; RIBEIRO, Sérgio da Costa. A pedagogia da repetência ao longo das décadas. *Ensaio: Avaliações e Políticas Públicas em Educação*, Rio de Janeiro, v. 3, n. 20, p. 55-61, jul./set. 1995. Disponível em: <<http://www.sergiocostaribeiro.ifcs.ufrj.br/artigos/1995%20Pedagogia%20da%20repetencia%20ao%20longo%20das%20decadas.pdf>>. Acesso em: 15 abr. 2018.

KOCH, Zenir Maria; HANFF, Beatriz B. C.; LEMOS, Glayce. A reprovação escolar: um desafio para as políticas educacionais. Anais do XXIII Simpósio Brasileiro de Políticas e Administração da Educação. Porto Alegre: 2007. Disponível em: <https://www.anpae.org.br/congressos_antigos/simposio2007/469.pdf> Acesso em: 17 abr. 2018.

LIMA JR, Paulo; SILVEIRA, Fernando Lang da; OSTERMANN, Fernanda. Análise de sobrevivência aplicada ao estudo do fluxo escolar nos cursos de graduação de física: um exemplo de uma universidade brasileira. **Revista Brasileira de Ensino de Física**, São Paulo, v. 34, n.1, 2012. Disponível em: <<http://www.scielo.br/pdf/rbef/v34n1/v34n1a14.pdf>> Acesso em: 15 abr. 2018. ISSN 1806-9126

LINS, Rodrigo; FIGUEIREDO FILHO, Dalson; ROCHA, Enivaldo. Não temos tempo a perder: uma introdução à análise de sobrevivência. **Revista Política Hoje**, Pernambuco, v. 26, n. 1, p. 279-298, 2017. Disponível em: <https://www.researchgate.net/profile/Dalson_Figueiredo2/publication/316202648_Nao_temos_tempo_a_perder_uma_introducao_a_analise_de_sobrevivencia/links/58f610144585158514b53195/Nao-temos-tempo-a-perder-uma-introducao-a-analise-de-sobrevivencia.pdf>. Acesso em: 23 abr. 2018. ISSN 0104-7094

MARINO, L. L. **Análise de dados como característica longitudinal em educação: um estudo sobre o município do Rio de Janeiro**. Dissertação de Mestrado. Escola Nacional de Ciências Estatísticas IBGE, 2014.

MICHAELIS: Dicionário Brasileiro da Língua Portuguesa. Disponível em: <<http://michaelis.uol.com.br/moderno-portugues/>>. Acesso em: 13 Jun. 2018

MORAES, Caroline da Luz; SPINDOLA, Luisa Kaufman. O currículo de história na reforma da Secretaria Municipal de Educação do Rio de Janeiro. **Educ. Pesqui.**, São Paulo, v. 43, n. 4, p. 1147-1162, Dec. 2017. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1517-97022017000401147&lng=en&nrm=iso>. Acesso em: 02 jul. 2018.

OBSERVATÓRIO DO PNE. **Metas do PNE – Meta 2**. Disponível em: <<http://www.observatoriodopne.org.br/metas-pne/2-ensino-fundamental/indicadores>>. Acesso em 05 abr. 2018.

OLIVEIRA, Romualdo Portela de; ARAUJO, Gilda Cardoso de. Qualidade do ensino: uma nova dimensão da luta pelo direito à educação. **Revista Brasileira Educação**: Rio de Janeiro, n. 28, p. 5-23, Jan/Abr, 2005. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-24782005000100002&lng=en&nrm=iso>. Acesso em: 05 abr. 2018.

OLIVEIRA, Luís Felipe Batista de; SOARES, Sergei Suarez Dillon. **Determinantes da repetência escolar no Brasil: uma análise de painel dos censos escolares entre 2007 e 2010**. 2012. Disponível em: <http://www.ipea.gov.br/portal/images/stories/PDFs/TDs/td_1706.pdf>. Acesso em: 15 abr. 2018.

OUCHANA, Deborah. Qual o futuro do Projeto GENTE? **Revista Educação**. São Paulo, 29/04/2014. Disponível em: <<http://www.revistaeducacao.com.br/qual-o-futuro-do-projeto-gente/>>. Acesso em: 05 jul. 2018.

OOD, Johan H. L. **Event History Analysis**. In: Encyclopedia of quality of life and well-being research. Springer, 2014. Disponível em: <https://www.researchgate.net/publication/281968023_Event_History_Analysis>. Acesso em: 19 mar. 2018.

PATTO, Maria Helena Souza. **A produção do fracasso escolar: história de submissão e rebeldia**. 2. ed. São Paulo: Casa do Psicólogo, 1999. 332p.

PEREGRINO, Mônica. **Trajetórias desiguais: um estudo sobre os processos de escolarização pública de jovens pobres**. Rio de Janeiro: Garamond, 2010. 328p.

PRADO, Mauricio Almeida. O modelo gerencial da educação: contribuições da experiência da Inglaterra ao debate brasileiro. **Revista do Serviço Público**: Brasília, v. 62, n. 3, p. 261-279, Jul/Set, 2011. Disponível em: <<https://revista.enap.gov.br/index.php/RSP/article/view/72/71>>. Acesso em: 23 abr. 2018.

RIANI, Juliana de Licena Ruas; SILVA, Vania Candida da; SOARES, Tufi Machado. Repetir ou progredir? Uma análise da repetência nas escolas públicas de Minas Gerais. **Educação e Pesquisa**, São Paulo, v. 38, n. 3, p. 623-636, sep. 2012. Disponível em: <<http://www.revistas.usp.br/ep/article/view/47898>>. Acesso em: 14 abr. 2018.

RIBEIRO, Sérgio Costa. A pedagogia da repetência. **Estudos Avançados**, São Paulo, v.5, n. 12, p. 07-21, Mai/Ago. 1991. ISSN 1806-9592. Disponível em: <<http://www.revistas.usp.br/eav/article/view/8604>>. Acesso em 15 abr. 2018.

RIBEIRO, Vanda Mendes *et al.* Crenças de professores sobre reprovação escolar. Belo Horizonte: **Educação em Revista**, vol 34, 2018. ISSN 1982-6621. Disponível em: <<http://educacaoemrevistaufmg.com.br/?artigo=crencas-de-professores-sobre-reprovacao-escolar>>. Acesso em: 15 abr. 2018.

RIO DE JANEIRO. **Plano Estratégico da Prefeitura do Rio de Janeiro (2009-2012)**. Rio de Janeiro, 2009. Disponível em: <http://www.rio.rj.gov.br/dlstatic/10112/6616925/4178940/planejamento_estrategico_site_01.pdf>. Acesso em: 05 jul. 2018.

RIO DE JANEIRO. **Plano Estratégico da Prefeitura do Rio de Janeiro (2013-2016)**. Rio de Janeiro, 2013. Disponível em: <http://www.rio.rj.gov.br/dlstatic/10112/2116763/4104304/planejamento_estrategico_1316.pdf>. Acesso em: 05 jul. 2018.

RIO DE JANEIRO (Município). **Educação**. Cadernos de Políticas Públicas Rio de Janeiro. v. 1, s/d. 54 p.. Rio de Janeiro, 2016. Disponível em: Acesso em:

<<http://www.rio.rj.gov.br/dlstatic/10112/4379008/4130264/EducacaoPORT1211.pdf>>. Acesso em: 05 jul. 2018.

RIO DE JANEIRO: SECRETARIA MUNICIPAL DE EDUCAÇÃO. **Experiências: Programa Escolas do Amanhã**. Rio de Janeiro, 2016, 132p. Disponível em: <<http://unesdoc.unesco.org/images/0024/002442/244284por.pdf>>. Acesso em 02 jul. 2018.

SANTOS, Jairo Campos dos. O gerencialismo no novo modelo de educação pública da cidade do Rio de Janeiro (2009-2012): origens, implantação, resultados e percepções. Tese (Doutorado em Educação) – Faculdade de Educação, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2014. Disponível em: <<http://www.educacao.ufrj.br/ppge/Teses2014/tjairocampos.pdf>>. Acesso em: 02 jul. 2018.

SANTOS, Ronny Alex Libório dos. **As possíveis relações entre a adaptação discente dos alunos do 6º ano e o fracasso escolar**. Dissertação (Mestrado Profissional em Gestão e Avaliação da Educação Pública) – Faculdade de Educação, Universidade Federal de Juiz de Fora, Minas Gerais, 2016. Disponível em: <<http://www.mestrado.caedufjf.net/wp-content/uploads/2016/12/Rony.pdf>>. Acesso: 06 jul. 2018.

SANTOS, Daniel Domingues dos; SCORZAFAVE, Luiz Guilherme; NICOLELLA, Alexandre C.; SANT'ANNA, Elder Generozo. Mais é menos? O impacto do projeto 6º ano experimental – SME/RJ. **Estudos em Avaliação Educacional**, v. 28, n. 69, p. 718-147, set./dez. 2017. E-ISSN 1984-932X. Disponível em: <<http://publicacoes.fcc.org.br/ojs/index.php/eae/article/view/3795/3467>>. Acesso em: 05 jul. 2018.

SCHIEFELBEIN, Ernesto. Repeating: An Overlooked Problem of Latin American Education. **Comparative Education Review**. v.19, n.3, p. 468-487, 1975.

SCHWARTZMAN, Simon. **O Plano Nacional de Educação e a Crise do País**. Instituto Alfa e Beto. 14 abr. 2016. Disponível em: <<http://www.alfaebeto.org.br/o-plano-nacional-de-educacao-e-a-cri-se-do-pais/>>. Acesso em: 05 abr. 2018.

SECRETARIA MUNICIPAL DE EDUCAÇÃO. Secretaria de Educação promove encontro de todos os Ginásios Cariocas. Rio de Janeiro, 16 dez. 2016. Disponível em: <<http://www.rio.rj.gov.br/web/sme/exibeconteudo?id=6612676>>. Acesso em 05 jul. 2018.

SHIRASU, Maitê Rimekká; ARRAES, Ronaldo de Albuquerque e. **Determinantes da evasão e repetência escolar**. ANPEC – Associação Nacional dos Centros de Pós-Graduação em Economia. 2016. Disponível em: <https://www.anpec.org.br/encontro/2015/submissao/files_l/i12-85f3c3774c3d65741cb278e01e61db39.pdf>. Acesso em: 15 abr. 2018.

SILVA, Wagner Ferraz Coelho. Transferência Escolar: os desafios do aluno na nova escola e o papel dos recursos didático. **Anais do Congresso de Pesquisa e Extensão e da Semana de Ciências Sociais da UEMG/Barbacena**, vol. 2, n.2, 2015. ISSN: 2525-6823. Disponível em:

<<http://revista.uemg.br/index.php/anaisbarbacena/article/view/1582/934>>. Acesso em: 15 jun. 2018.

SILVA, Nelson do Valle; HASENBALG, Carlos. Recursos familiares e transições educacionais. **Caderno de Saúde Pública**, vol. 18, p. 67-76). Rio de Janeiro, 2002.

SOARES, José Francisco. Melhoria do desempenho cognitivo dos alunos do ensino fundamental. **Cadernos de Pesquisa**, São Paulo, v. 37, n. 130, p. 135-160, 2007. Disponível em: <<http://www.scielo.br/pdf/cp/v37n130/07.pdf>> Acesso em: 07 abr. 2018.

SOARES, Sergei Suarez Dillon. A repetência no contexto internacional: o que dizem os dados de avaliações das quais o Brasil não participa? **Texto para discussão**, nº. 1300. Brasília: Instituto de Pesquisa Econômica Aplicada, 2007. Disponível em: <http://repositorio.ipea.gov.br/bitstream/11058/1392/1/TD_1300.pdf>. Acesso em: 05 jul. 2018.

SOARES, Sergei Suarez Dillon. **A repetência no contexto internacional: o que dizem os dados de avaliações das quais o Brasil não participa**. Brasília: Instituto de Pesquisa Econômica Aplicada, 2007. (Texto para Discussão, n. 1300).

SOUZA, Adilson Severo de. **Os desafios da ampliação da jornada escolar em duas escolas da favela da Rocinha – Rio de Janeiro (RJ)**. Dissertação (Mestrado em Educação) – Faculdade de Educação, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2015. Disponível em: <<http://www.educacao.ufrj.br/ppge/dissertacoes2015/dadilsonseverodesouza.pdf>>. Acesso em 05 jul. 2018.

SOUZA, André Portela; OLIVA, Bruno; PONCZEK, Vladimir. Os determinantes do fluxo escolar entre o ensino fundamental e o ensino médio no Brasil. **Texto para discussão** (Escola de Economia de São Paulo) nº 286, C-micro, nº 06. Maio, 2011. Disponível em: <<http://bibliotecadigital.fgv.br/dspace/bitstream/handle/10438/10000/TD%20286%20-%20C-Micro%2006%20-%20Andr%C3%A9%20Portela.pdf?sequence=1&isAllowed=y>>. Acesso em: 05 jul. 2018.

TAVARES JR, Fernando; FARIA, V. B.; LIMA, M. A. de. Indicadores de fluxo escolar e políticas educacionais: avaliação das últimas décadas. **Estudos em Avaliação Educacional**, v. 23, n. 52, p. 48-67, maio 2012. Disponível em: <<http://publicacoes.fcc.org.br/ojs/index.php/eae/article/view/1929/1912>>. Acesso em: 18 abr. 2018.

TAVARES JR, Fernando; MONT'ALVÃO, Arnaldo; NEUBERT, Luiz Flávio. Rendimento escolar e seus determinantes sociais no Brasil. **Revista Brasileira de Sociologia**, vol. 03, n. 06, jul-dez, 2015.

TAYLOR, C. The Reliability of Free School Meal Eligibility as a Measure of Socio-Economic Disadvantage: Evidence from the Millennium Cohort Study in Wales, **British Journal of Educational Studies**, 66:1, p. 29-51, 2018. DOI: [10.1080/00071005.2017.1330464](https://doi.org/10.1080/00071005.2017.1330464)

TEIXEIRA DE FREITAS, Mario Augusto. Dispersão demográfica e escolaridade. **Revista Brasileira de Estatística**, v. 1, n. 3, p. 497-527, 1940a. Disponível em: <https://biblioteca.ibge.gov.br/visualizacao/periodicos/760/rbe_1940_v1_n3.pdf>. Acesso em: 05 jul. 2018.

TEIXEIRA DE FREITAS, Mario Augusto. A evasão escolar no ensino primário brasileiro. **Revista Brasileira de Estatística**, v. 1, n. 4, p. 697-722, 1940b. Disponível em: <https://biblioteca.ibge.gov.br/visualizacao/periodicos/760/rbe_1940_v1_n4.pdf>. Acesso em: 05 jul. 2018.

TEIXEIRA DE FREITAS, Mario Augusto. A Escolaridade média no ensino primário brasileiro. **Revista Brasileira de Estatística**, v. 8, n. 30-31, p. 395-474, 1947. Disponível em: <https://biblioteca.ibge.gov.br/visualizacao/periodicos/760/rbe_1947_v8_n30_n31.pdf>. Acesso em: 05 jul. 2018.

TEODORO, António. **A educação em tempos de globalização neoliberal: os novos modos de regulação das políticas educacionais**. Brasília: Liber Livro, 2011. 176 p.

TERRA. **A escola de onde os alunos não querem sair**. 17 nov. 2016. Disponível em: <<https://www.terra.com.br/noticias/a-escola-de-onde-os-alunos-nao-querem-sair,678ee412328ba6598b8f9f70e1d522bajqrmdy2k.html>>. Acesso em: 05 jul. 2018.

UNESCO, Institute for Statistics. **Opportunities lost: The impact of grade repetition and early school leaving**. Global education digest. 2012. Disponível em: <http://uis.unesco.org/sites/default/files/documents/opportunities-lost-the-impact-of-grade-repetition-and-early-school-leaving-en_0.pdf>. Acesso em: 15 abr. 2018.

VIEIRA, Sofia Lerche; FARIAS, Isabel Maria Sabino de. **Política Educacional no Brasil: Introdução Histórica**. 3. ed. Brasília: Liber Livro, 2011. 220p.

WEISS, Maria Lucia Lemme. **Psicopedagogia clínica: uma visão diagnóstica dos problemas de aprendizagem escolar**. 14. ed. rev. e ampl. Rio de Janeiro: Lamparina, 2012.

ANEXOS

ANEXO A – ANÁLISE DE SOBREVIVÊNCIA – SINTAXE DO STATA

Configurando os dados para Análise de Sobrevivência

```
. stset Tempo, id(ID) failure (EventoAno)
```

```

      id: ID
      failure event: EventoAno != 0 & EventoAno < .
obs. time interval: (Tempo[_n-1], Tempo]
exit on or before: failure
```

```

161300 total obs.
      0 exclusions
```

```

161300 obs. remaining, representing
58019 subjects
16931 failures in single failure-per-subject data
161300 total analysis time at risk, at risk from t =      0
                                     earliest observed entry t =      0
                                     last observed exit t =      4
```

Descritiva da Análise de Sobrevivência com cálculo do tempo de sobrevivência

```
. stdes
```

```

      failure _d: EventoAno
analysis time _t: Tempo
      id: ID
```

Category	total	per subject			
		mean	min	median	max
no. of subjects	58019				
no. of records	161300	2.780124	1	3	4
(first) entry time		0	0	0	0
(final) exit time		2.780124	1	3	4
subjects with gap	0				
time on gap if gap	0
time at risk	161300	2.780124	1	3	4
failures	16931	.2918182	0	0	1

Calculando a Taxa de incidência

```
. stsum
```

```

      failure _d: EventoAno
analysis time _t: Tempo
      id: ID
```

	time at risk	incidence rate	no. of subjects	Survival time		
				25%	50%	75%
total	161300	.1049659	58019	3	.	.

Função de Sobrevivência para cada tempo

```
. sts list
```

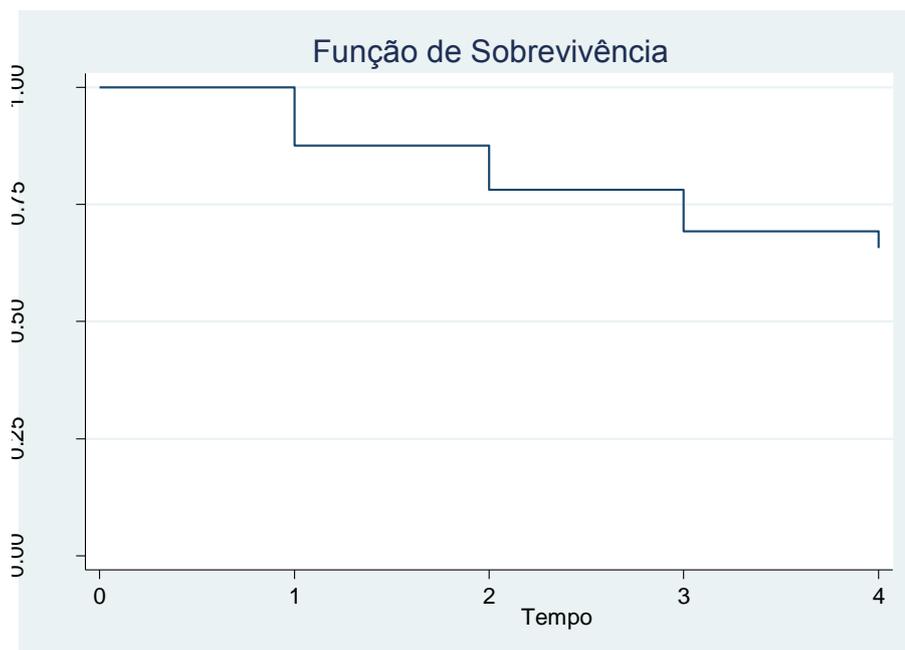
```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
1	58019	7208	9371	0.8758	0.0014	0.8731	0.8784
2	41440	4446	3287	0.7818	0.0018	0.7782	0.7853
3	33707	3816	1757	0.6933	0.0021	0.6892	0.6974
4	28134	1461	2.7e+04	0.6573	0.0022	0.6530	0.6616

Gráfico – Função de Sobrevivência estimada por Kaplan-Meier

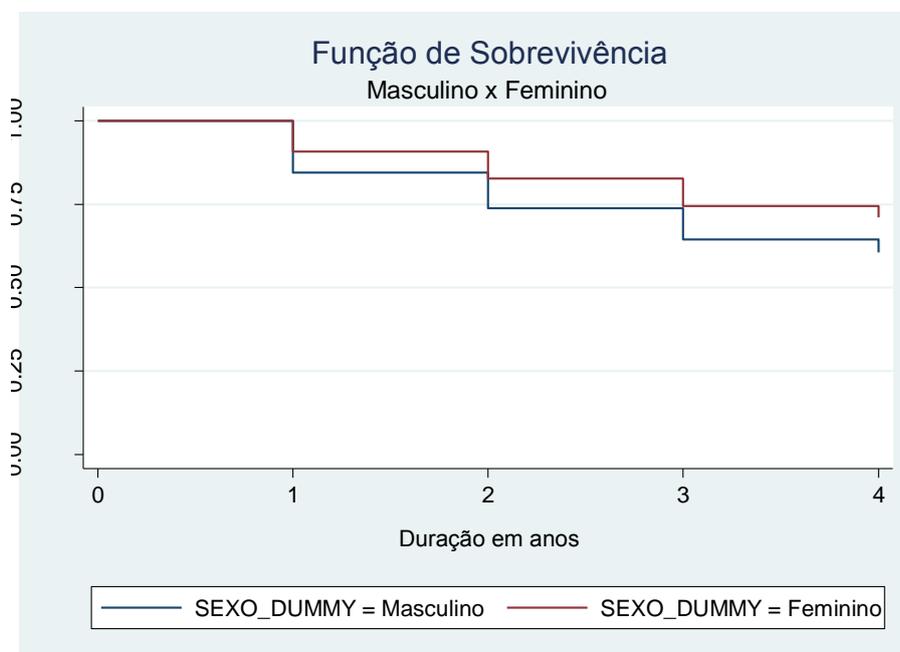
```
. sts graph
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```



ANEXO B – GRÁFICOS COMPARATIVOS DAS FUNÇÕES DE SOBREVIVÊNCIA E SINTAXE DO STATA

- Sexo



```
. sts list, by (Sexo Nao_Branco)
```

```
failure _d: EventoAno
analysis time _t: Tempo
id: ID
```

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
Masculino Branco							
1	10108	1236	1547	0.8777	0.0033	0.8712	0.8840
2	7325	761	546	0.7865	0.0043	0.7780	0.7948
3	6018	684	304	0.6971	0.0050	0.6873	0.7068
4	5030	263	4767	0.6607	0.0052	0.6504	0.6708
Masculino Não Branco							
1	18572	3165	3296	0.8296	0.0028	0.8241	0.8349
2	12111	1692	996	0.7137	0.0035	0.7067	0.7205
3	9423	1300	504	0.6152	0.0040	0.6074	0.6229
4	7619	499	7120	0.5749	0.0041	0.5669	0.5829
Feminino Branco							
1	9952	717	1422	0.9280	0.0026	0.9227	0.9329
2	7813	627	617	0.8535	0.0037	0.8460	0.8606
3	6569	545	337	0.7827	0.0045	0.7737	0.7913
4	5687	216	5471	0.7529	0.0047	0.7435	0.7621
Feminino Não Branco							
1	16883	1720	2664	0.8981	0.0023	0.8935	0.9026
2	12499	1189	972	0.8127	0.0032	0.8064	0.8188
3	10338	1111	529	0.7253	0.0038	0.7179	0.7326
4	8698	422	8276	0.6902	0.0039	0.6824	0.6978

```
. sts test SEXO_DUMMY
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

Log-rank test for equality of survivor functions

SEXO_DUMMY	Events observed	Events expected
Masculino	10046	8403.08
Feminino	6885	8527.92
Total	16931	16931.00

chi2(1) = 718.38
Pr>chi2 = 0.0000

```
. sts test SEXO_DUMMY, wilcoxon
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

Wilcoxon (Breslow) test for equality of survivor functions

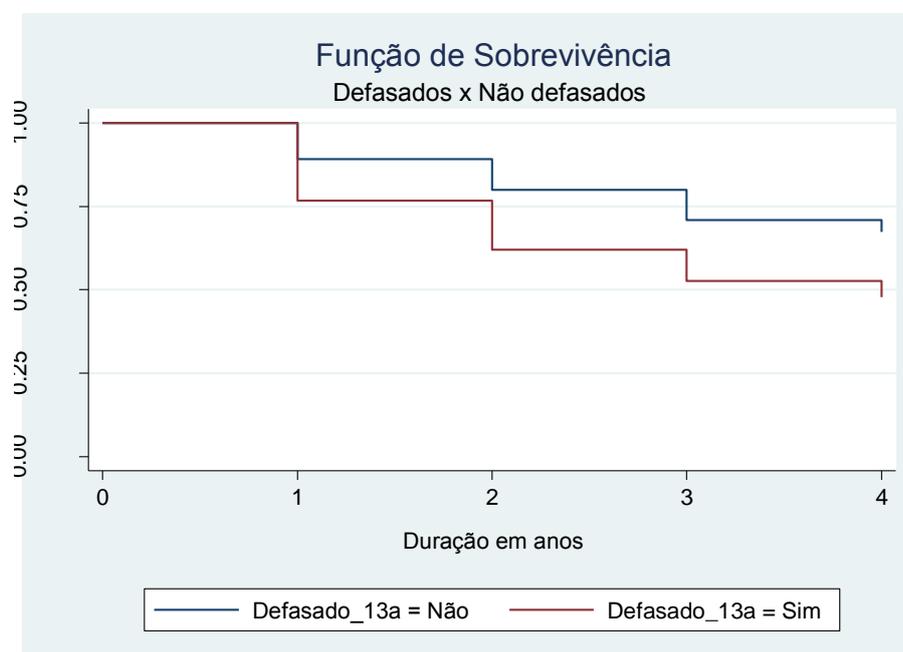
SEXO_DUMMY	Events observed	Events expected	Sum of ranks
Masculino	10046	8403.08	79893798
Feminino	6885	8527.92	-79893798
Total	16931	16931.00	0

chi2(1) = 774.47
Pr>chi2 = 0.0000

```
. sts graph, by( SEXO_DUMMY) title (Função de Sobrevivência) subtitle (Masculino x Feminino) xtit
> le (Duração em anos, margin (medium))
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

- Não defasados x defasados



```
. sts list, by( Defasado)
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

	Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
Não	1	50368	5432	5349	0.8922	0.0014	0.8894	0.8948
	2	39587	4091	2496	0.8000	0.0018	0.7963	0.8035
	3	33000	3708	1627	0.7101	0.0021	0.7058	0.7143
	4	27665	1419	2.6e+04	0.6736	0.0022	0.6692	0.6780
Sim	1	7651	1776	4022	0.7679	0.0048	0.7583	0.7772
	2	1853	355	791	0.6208	0.0080	0.6048	0.6363
	3	707	108	130	0.5259	0.0108	0.5045	0.5469
	4	469	42	427	0.4788	0.0120	0.4550	0.5022

```
. sts test Defasado
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

Log-rank test for equality of survivor functions

Defasado	Events observed	Events expected
Não	14650	15677.28
Sim	2281	1253.72
Total	16931	16931.00

```
chi2(1) = 1072.01
Pr>chi2 = 0.0000
```

```
. sts test Defasado, wilcoxon
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

Wilcoxon (Breslow) test for equality of survivor functions

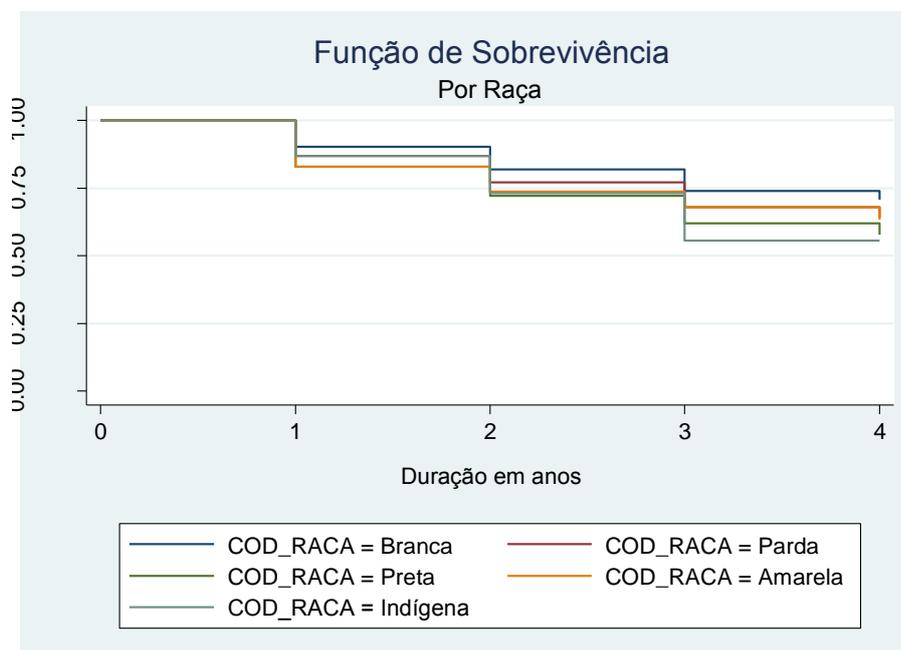
Defasado	Events observed	Events expected	Sum of ranks
Não	14650	15677.28	-55804961
Sim	2281	1253.72	55804961
Total	16931	16931.00	0

```
chi2(1) = 1104.03
Pr>chi2 = 0.0000
```

```
. sts graph, by ( Defasado)
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

- Raças



```
. sts list, by( COD_RACA)
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

Time	Beg. Total	Net Fail	Survivor Lost	Std. Function	Std. Error	[95% Conf. Int.]	
Branca							
1	20060	1953	2969	0.9026	0.0021	0.8985	0.9067
2	15138	1388	1163	0.8199	0.0028	0.8142	0.8254
3	12587	1229	641	0.7398	0.0034	0.7332	0.7463
4	10717	479	1.0e+04	0.7068	0.0035	0.6998	0.7136
Parda							
1	28585	3718	4731	0.8699	0.0020	0.8660	0.8738
2	20136	2302	1586	0.7705	0.0026	0.7653	0.7756
3	16248	1915	866	0.6797	0.0030	0.6737	0.6856
4	13467	732	1.3e+04	0.6427	0.0032	0.6365	0.6489
Preta							
1	6681	1138	1190	0.8297	0.0046	0.8204	0.8385
2	4353	563	374	0.7224	0.0058	0.7108	0.7336
3	3416	481	164	0.6206	0.0066	0.6076	0.6334
4	2771	186	2585	0.5790	0.0068	0.5655	0.5922
Amarela							
1	105	18	23	0.8286	0.0368	0.7418	0.8883
2	64	7	6	0.7379	0.0460	0.6349	0.8161
3	51	4	3	0.6801	0.0507	0.5694	0.7680
4	44	3	41	0.6337	0.0538	0.5183	0.7286
Indígena							
1	84	11	16	0.8690	0.0368	0.7760	0.9252
2	57	9	2	0.7318	0.0522	0.6136	0.8191
3	46	11	0	0.5568	0.0608	0.4301	0.6661
4	35	0	35	0.5568	0.0608	0.4301	0.6661

```
. sts test COD_RACA
```

```
      failure _d: EventoAno
analysis time _t: Tempo
              id: ID
```

Log-rank test for equality of survivor functions

COD_RACA	Events observed	Events expected
Branca	5049	6068.08
Parada	8667	8209.30
Preta	2368	1818.33
Amarela	32	27.82
Indígena	31	23.46
Total	16147	16147.00

```
      chi2(4) = 411.75
      Pr>chi2 = 0.0000
```

```
. sts test COD_RACA, wilcoxon
```

```
      failure _d: EventoAno
analysis time _t: Tempo
              id: ID
```

Wilcoxon (Breslow) test for equality of survivor functions

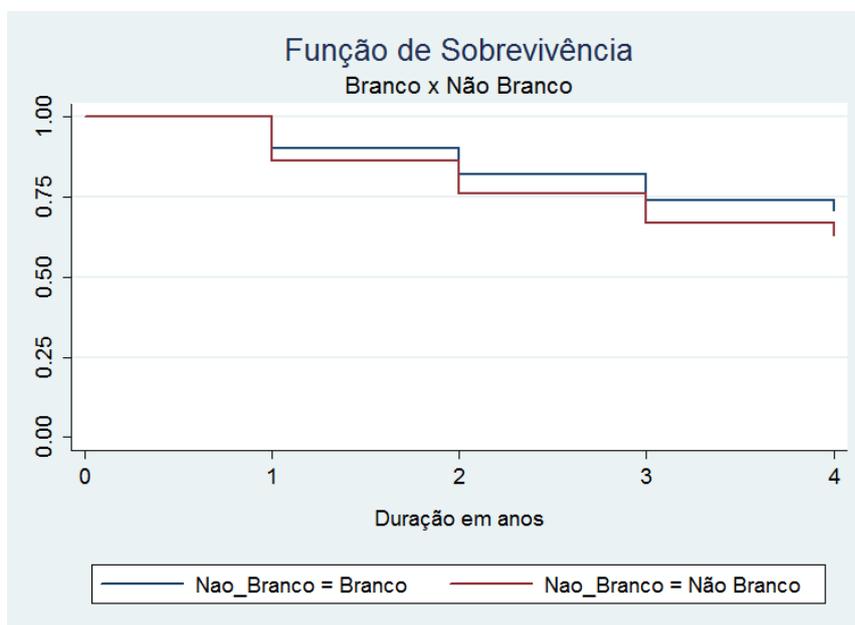
COD_RACA	Events observed	Events expected	Sum of ranks
Branca	5049	6068.08	-46318885
Parada	8667	8209.30	20218640
Preta	2368	1818.33	25560631
Amarela	32	27.82	249554
Indígena	31	23.46	290060
Total	16147	16147.00	0

```
      chi2(4) = 427.95
      Pr>chi2 = 0.0000
```

```
. sts graph, by( COD_RACA) title (Função de Sobrevivência) subtitle (Por Raça) xtitle (Duração em
> anos, margin (medium))
```

```
      failure _d: EventoAno
analysis time _t: Tempo
              id: ID
```

- Branco x Não-Branco



```
. sts list, by (Nao_Branco)
```

```
      failure _d: EventoAno
analysis time _t: Tempo
      id: ID
```

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
Branco							
1	20060	1953	2969	0.9026	0.0021	0.8985	0.9067
2	15138	1388	1163	0.8199	0.0028	0.8142	0.8254
3	12587	1229	641	0.7398	0.0034	0.7332	0.7463
4	10717	479	1.0e+04	0.7068	0.0035	0.6998	0.7136
Não Branco							
1	35455	4885	5960	0.8622	0.0018	0.8586	0.8658
2	24610	2881	1968	0.7613	0.0024	0.7566	0.7659
3	19761	2411	1033	0.6684	0.0027	0.6630	0.6738
4	16317	921	1.5e+04	0.6307	0.0029	0.6250	0.6363

```
. sts test Nao_Branco
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

Log-rank test for equality of survivor functions

Nao_Branco	Events observed	Events expected
Branco	5049	6068.08
Não Branco	11098	10078.92
Total	16147	16147.00

chi2(1) = 308.37
Pr>chi2 = 0.0000

```
. sts test Nao_Branco, wilcoxon
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

Wilcoxon (Breslow) test for equality of survivor functions

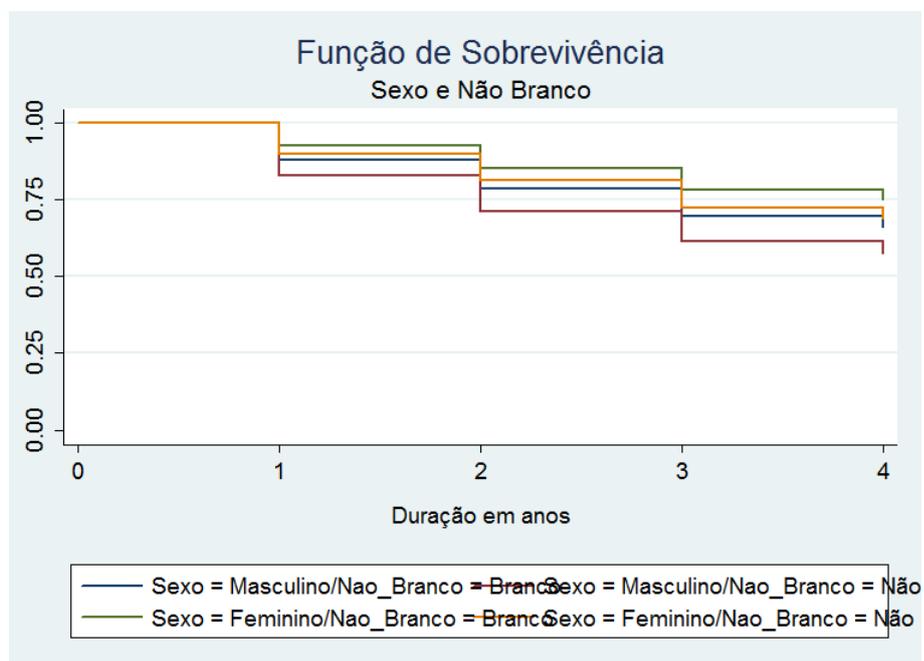
Nao_Branco	Events observed	Events expected	Sum of ranks
Branco	5049	6068.08	-46318885
Não Branco	11098	10078.92	46318885
Total	16147	16147.00	0

chi2(1) = 319.28
Pr>chi2 = 0.0000

```
. sts graph, by( RACA_NAOBRANCO) title (Função de Sobrevida) subtitle (Branco x Não Branco) x
> title (Duração em anos, margin (medium))
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

- Sexo e Não-Branco



```
. sts list, by (Sexo Nao_Branco)
```

```
      failure _d: EventoAno
analysis time _t: Tempo
      id: ID
```

	Beg.	Net		Survivor	Std.	[95% Conf. Int.]	
Time	Total	Fail	Lost	Function	Error		
Masculino Branco							
1	10108	1236	1547	0.8777	0.0033	0.8712	0.8840
2	7325	761	546	0.7865	0.0043	0.7780	0.7948
3	6018	684	304	0.6971	0.0050	0.6873	0.7068
4	5030	263	4767	0.6607	0.0052	0.6504	0.6708
Masculino Não Branco							
1	18572	3165	3296	0.8296	0.0028	0.8241	0.8349
2	12111	1692	996	0.7137	0.0035	0.7067	0.7205
3	9423	1300	504	0.6152	0.0040	0.6074	0.6229
4	7619	499	7120	0.5749	0.0041	0.5669	0.5829
Feminino Branco							
1	9952	717	1422	0.9280	0.0026	0.9227	0.9329
2	7813	627	617	0.8535	0.0037	0.8460	0.8606
3	6569	545	337	0.7827	0.0045	0.7737	0.7913
4	5687	216	5471	0.7529	0.0047	0.7435	0.7621
Feminino Não Branco							
1	16883	1720	2664	0.8981	0.0023	0.8935	0.9026
2	12499	1189	972	0.8127	0.0032	0.8064	0.8188
3	10338	1111	529	0.7253	0.0038	0.7179	0.7326
4	8698	422	8276	0.6902	0.0039	0.6824	0.6978

```
. sts test Nao_Branco Sexo
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

Log-rank test for equality of survivor functions

Nao_Br~o, Sexo	Events observed	Events expected
Branco, Masculino	2944	2969.43
Branco, Feminino	2105	3098.65
Não Branco, Masculino	6656	5043.22
Não Branco, Feminino	4442	5035.70
Total	16147	16147.00

```
chi2(3) = 1019.25
Pr>chi2 = 0.0000
```

```
. sts test Nao_Branco Sexo, wilcoxon
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

Wilcoxon (Breslow) test for equality of survivor functions

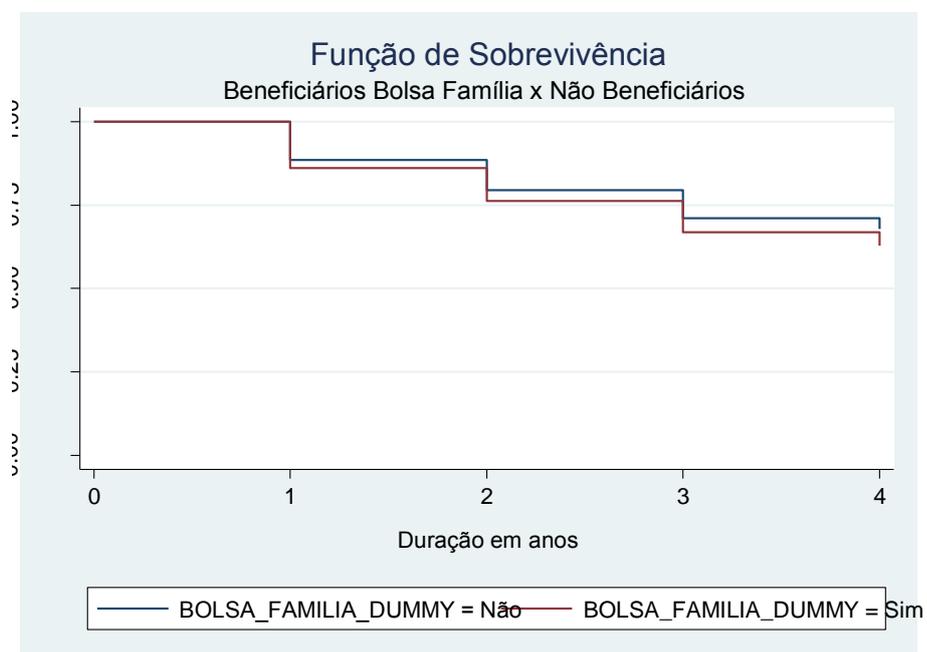
Nao_Br~o, Sexo	Events observed	Events expected	Sum of ranks
Branco, Masculino	2944	2969.43	-1235707
Branco, Feminino	2105	3098.65	-45083178
Não Branco, Masculino	6656	5043.22	74837442
Não Branco, Feminino	4442	5035.70	-28518557
Total	16147	16147.00	0

```
chi2(3) = 1074.67
Pr>chi2 = 0.0000
```

```
. sts graph, by(Sexo Nao_Branco) title (Função de Sobrevivência) subtitle (Sexo e Não Branco) xti
> tle (Duração em anos, margin (medium))
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

- Beneficiários do Programa Bolsa Família x Não beneficiários



```
. sts list, by( BOLSA_FAMILIA_DUMMY)
```

```
      failure _d: EventoAno
analysis time _t: Tempo
      id: ID
```

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
Não							
1	34861	3978	7056	0.8859	0.0017	0.8825	0.8892
2	23827	2458	2176	0.7945	0.0023	0.7899	0.7990
3	19193	2032	1068	0.7104	0.0027	0.7050	0.7157
4	16093	735	1.5e+04	0.6779	0.0028	0.6723	0.6835
Sim							
1	23158	3230	2315	0.8605	0.0023	0.8560	0.8649
2	17613	1988	1111	0.7634	0.0029	0.7577	0.7690
3	14514	1784	689	0.6696	0.0033	0.6631	0.6759
4	12041	726	1.1e+04	0.6292	0.0034	0.6225	0.6358

```
. sts test BOLSA_FAMILIA_DUMMY

      failure _d: EventoAno
analysis time _t: Tempo
      id: ID
```

Log-rank test for equality of survivor functions

BOLSA_FAMI~Y	Events observed	Events expected
Não	9203	9895.87
Sim	7728	7035.13
Total	16931	16931.00

chi2(1) = 131.43
Pr>chi2 = 0.0000

```
. sts test BOLSA_FAMILIA_DUMMY, wilcoxon

      failure _d: EventoAno
analysis time _t: Tempo
      id: ID
```

Wilcoxon (Breslow) test for equality of survivor functions

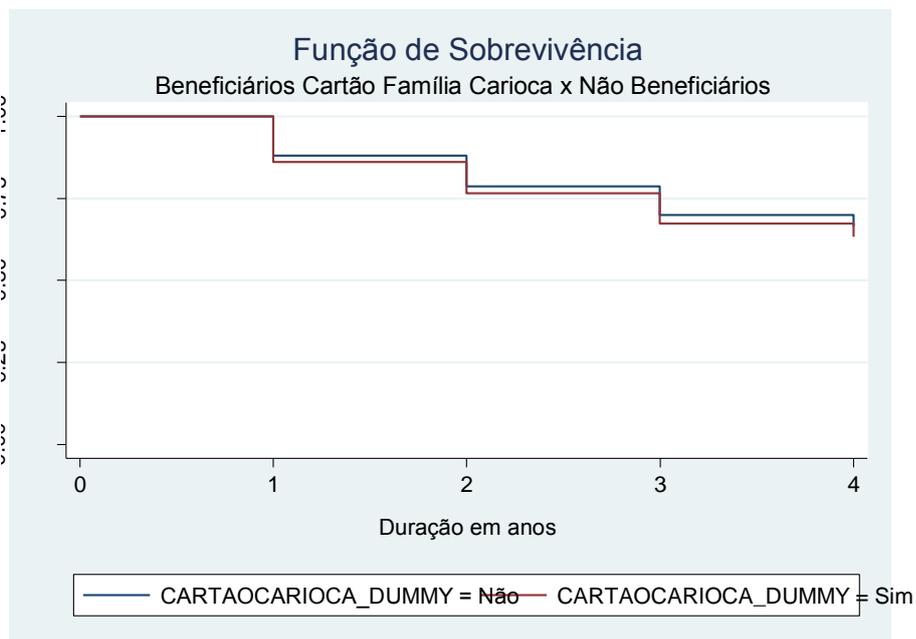
BOLSA_FAMI~Y	Events observed	Events expected	Sum of ranks
Não	9203	9895.87	-32135075
Sim	7728	7035.13	32135075
Total	16931	16931.00	0

chi2(1) = 129.52
Pr>chi2 = 0.0000

```
. sts graph, by( BOLSA_FAMILIA_DUMMY) title (Função de Sobrevivência) subtitle (Beneficiários Bol
> sa Família x Não Beneficiários) xtitle (Duração em anos, margin (medium))
```

```
      failure _d: EventoAno
analysis time _t: Tempo
      id: ID
```

- Beneficiários do Programa Cartão Família Carioca x Não beneficiários



```
. sts list, by( CARTAOARIOCA_DUMMY)
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
Não							
1	43803	5245	8066	0.8803	0.0016	0.8772	0.8833
2	30492	3230	2624	0.7870	0.0021	0.7829	0.7911
3	24638	2719	1310	0.7002	0.0024	0.6954	0.7049
4	20609	1011	2.0e+04	0.6658	0.0025	0.6608	0.6708
Sim							
1	14216	1963	1305	0.8619	0.0029	0.8561	0.8675
2	10948	1216	663	0.7662	0.0036	0.7589	0.7732
3	9069	1097	447	0.6735	0.0041	0.6653	0.6816
4	7525	450	7075	0.6332	0.0043	0.6247	0.6416

```
. sts test CARTAOCARIOCA_DUMMY
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

Log-rank test for equality of survivor functions

CARTAOCARI~Y	Events observed	Events expected
Não	12205	12572.80
Sim	4726	4358.20
Total	16931	16931.00

```
chi2(1) = 47.03
```

```
Pr>chi2 = 0.0000
```

```
. sts test CARTAOCARIOCA_DUMMY, wilcoxon
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

Wilcoxon (Breslow) test for equality of survivor functions

CARTAOCARI~Y	Events observed	Events expected	Sum of ranks
Não	12205	12572.80	-17174151
Sim	4726	4358.20	17174151
Total	16931	16931.00	0

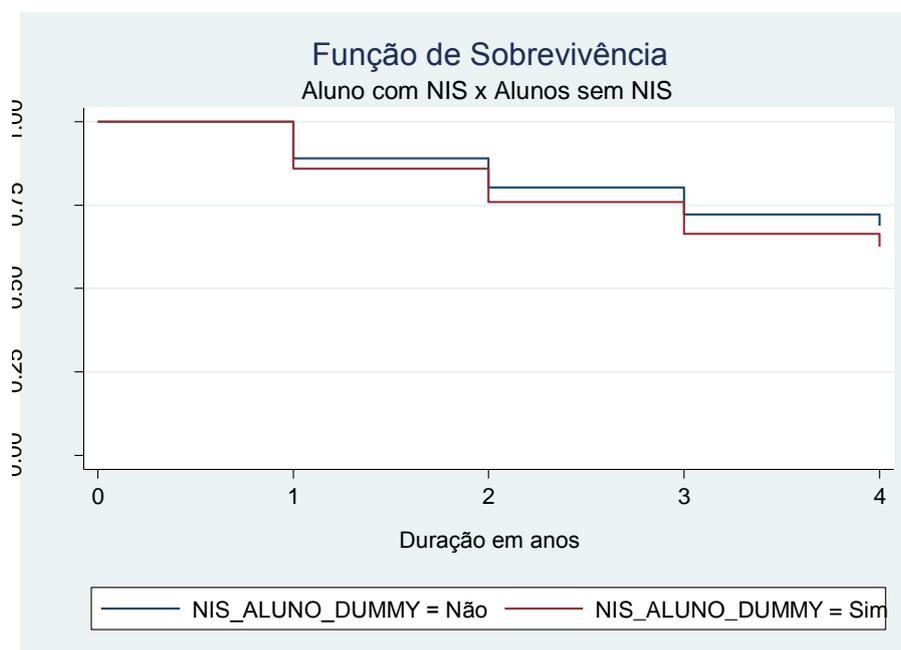
```
chi2(1) = 47.37
```

```
Pr>chi2 = 0.0000
```

```
. sts graph, by( CARTAOCARIOCA_DUMMY) title (Função de Sobrevida) subtitle (Beneficiários Car
> tão Família Carioca x Não Beneficiários) xtitle (Duração em anos, margin (medium))
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

- Alunos com NIS x Alunos sem NIS



```
. sts list, by( NIS_ALUNO_DUMMY)
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
Não							
1	30123	3280	5872	0.8911	0.0018	0.8875	0.8946
2	20971	2060	1857	0.8036	0.0024	0.7987	0.8083
3	17054	1740	925	0.7216	0.0029	0.7159	0.7272
4	14389	653	1.4e+04	0.6888	0.0030	0.6829	0.6947
Sim							
1	27896	3928	3499	0.8592	0.0021	0.8551	0.8632
2	20469	2386	1430	0.7590	0.0027	0.7538	0.7642
3	16653	2076	832	0.6644	0.0030	0.6584	0.6703
4	13745	808	1.3e+04	0.6254	0.0032	0.6191	0.6315

```
. sts test NIS_ALUNO_DUMMY
```

```
      failure _d: EventoAno
analysis time _t: Tempo
              id: ID
```

Log-rank test for equality of survivor functions

NIS_ALUNO_~Y	Events	Events
	observed	expected
Não	7733	8670.19
Sim	9198	8260.81
Total	16931	16931.00

chi2(1) = 233.60
Pr>chi2 = 0.0000

```
. sts test NIS_ALUNO_DUMMY, wilcoxon
```

```
      failure _d: EventoAno
analysis time _t: Tempo
              id: ID
```

Wilcoxon (Breslow) test for equality of survivor functions

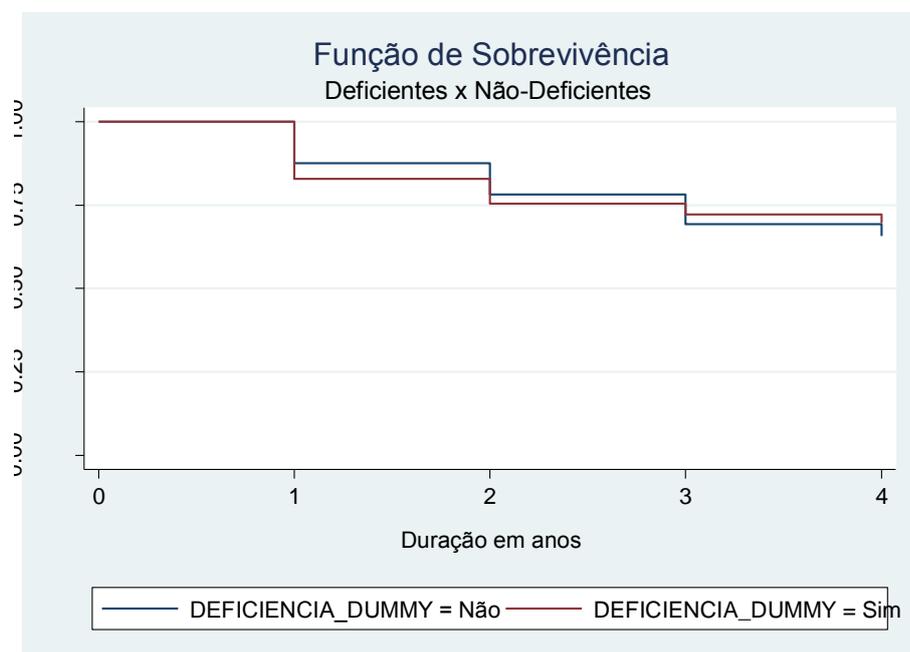
NIS_ALUNO_~Y	Events	Events	Sum of
	observed	expected	ranks
Não	7733	8670.19	-43773641
Sim	9198	8260.81	43773641
Total	16931	16931.00	0

chi2(1) = 232.45
Pr>chi2 = 0.0000

```
. sts graph, by( NIS_ALUNO_DUMMY) title (Função de Sobrevivência) subtitle (Aluno com NIS x Aluno
> s sem NIS) xtitle (Duração em anos, margin (medium))
```

```
      failure _d: EventoAno
analysis time _t: Tempo
              id: ID
```

- Deficientes x Não Deficientes (NÃO SIGNIFICATIVO)



```
. sts list, by( DEFICIENCIA_DUMMY)
```

```
      failure _d: EventoAno
analysis time _t: Tempo
      id: ID
```

	Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
Não								
	1	57567	7131	9270	0.8761	0.0014	0.8734	0.8788
	2	41166	4421	3252	0.7820	0.0018	0.7785	0.7856
	3	33493	3807	1743	0.6931	0.0021	0.6890	0.6972
	4	27943	1454	2.6e+04	0.6571	0.0022	0.6528	0.6614
Sim								
	1	452	77	101	0.8296	0.0177	0.7917	0.8613
	2	274	25	35	0.7539	0.0216	0.7085	0.7933
	3	214	9	14	0.7222	0.0231	0.6739	0.7647
	4	191	7	184	0.6958	0.0244	0.6452	0.7406

```
. sts test DEFICIENCIA_DUMMY
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

Log-rank test for equality of survivor functions

DEFICIENCIA~Y	Events observed	Events expected
Não	16813	16811.30
Sim	118	119.70
Total	16931	16931.00

chi2(1) = 0.03
Pr>chi2 = 0.8688

```
. sts test DEFICIENCIA_DUMMY, wilcoxon
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

Wilcoxon (Breslow) test for equality of survivor functions

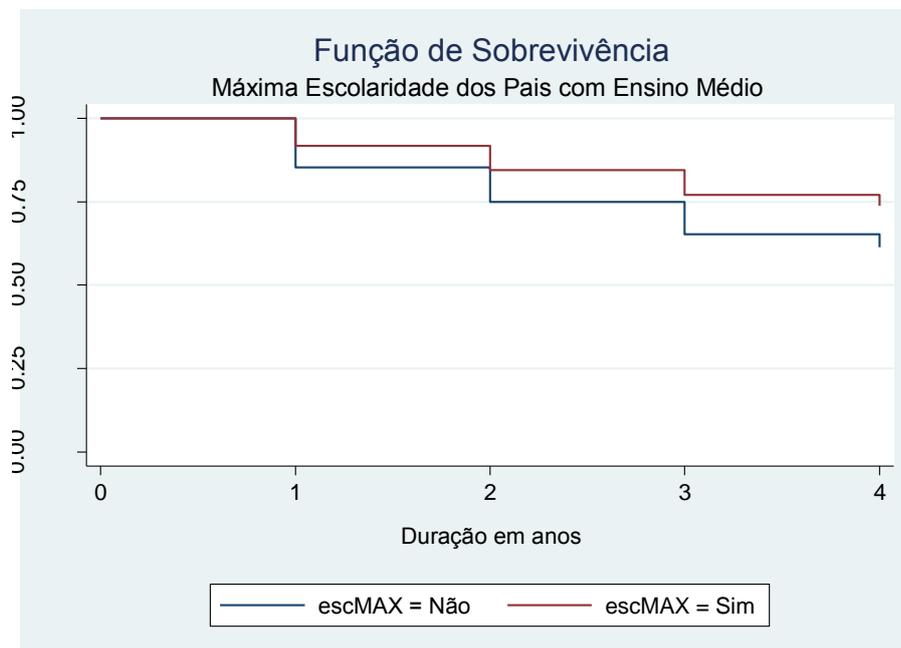
DEFICIENCIA~Y	Events observed	Events expected	Sum of ranks
Não	16813	16811.30	-431869
Sim	118	119.70	431869
Total	16931	16931.00	0

chi2(1) = 0.77
Pr>chi2 = 0.3787

```
. sts graph, by( DEFICIENCIA_DUMMY) title (Função de Sobrevivência) subtitle (Deficientes x Não-D
> eficientes) xtitle (Duração em anos, margin (medium))
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

- Máxima Escolaridade dos Pais Ensino Médio completo



```
. sts list, by(escMAX)
```

```
      failure _d: EventoAno
analysis time _t: Tempo
      id: ID
```

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
Não							
1	32333	4727	5327	0.8538	0.0020	0.8499	0.8576
2	22279	2705	1738	0.7501	0.0025	0.7451	0.7551
3	17836	2318	930	0.6526	0.0029	0.6469	0.6583
4	14588	849	1.4e+04	0.6147	0.0030	0.6087	0.6206
Sim							
1	18089	1492	2448	0.9175	0.0020	0.9134	0.9214
2	14149	1106	1051	0.8458	0.0028	0.8402	0.8512
3	11992	1062	584	0.7709	0.0034	0.7642	0.7774
4	10346	443	9903	0.7379	0.0036	0.7308	0.7448

```
. sts test escMAX

      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

Log-rank test for equality of survivor functions

escMAX	Events observed	Events expected
Não	10599	9095.70
Sim	4103	5606.30
Total	14702	14702.00

chi2(1) = 733.07
Pr>chi2 = 0.0000

```
. sts test escMAX, wilcoxon

      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

Wilcoxon (Breslow) test for equality of survivor functions

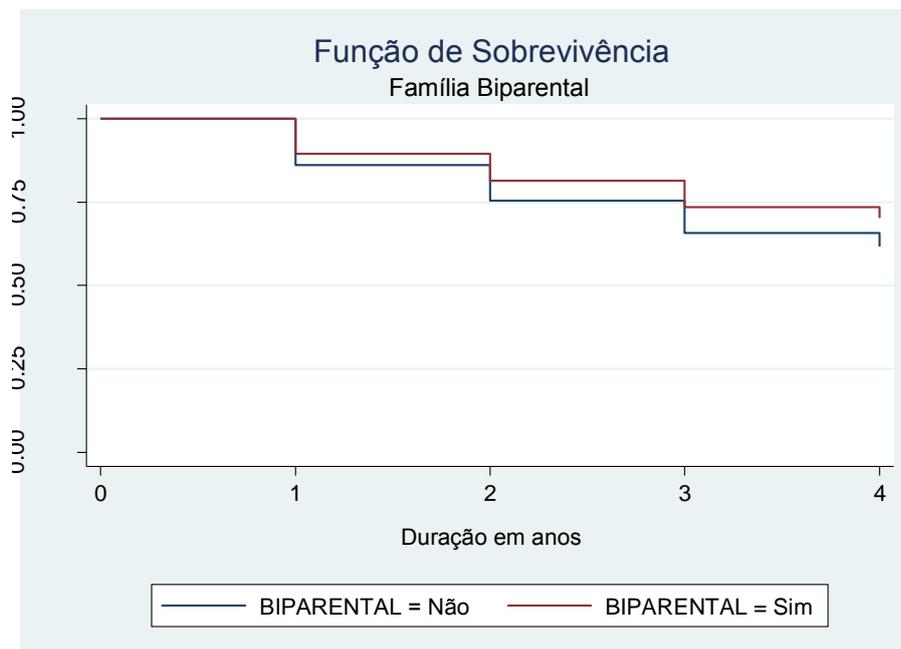
escMAX	Events observed	Events expected	Sum of ranks
Não	10599	9095.70	62075232
Sim	4103	5606.30	-62075232
Total	14702	14702.00	0

chi2(1) = 756.98
Pr>chi2 = 0.0000

```
. sts graph, by( escMAX) title (Função de Sobrevivência) subtitle (Máxima Escolaridade dos Pais c
> om Ensino Médio) xtitle (Duração em anos, margin (medium))
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

- Família Biparental



```
. sts list, by( BIPARENTAL)
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

Time	Beg. Total	Net Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
Não							
1	32407	4510	6330	0.8608	0.0019	0.8570	0.8646
2	21567	2675	2076	0.7541	0.0026	0.7490	0.7590
3	16816	2162	1092	0.6571	0.0030	0.6513	0.6629
4	13562	825	1.3e+04	0.6171	0.0031	0.6110	0.6232
Sim							
1	25612	2698	3041	0.8947	0.0019	0.8908	0.8984
2	19873	1771	1211	0.8149	0.0025	0.8099	0.8198
3	16891	1654	665	0.7351	0.0029	0.7293	0.7408
4	14572	636	1.4e+04	0.7030	0.0031	0.6970	0.7090

```
. sts test BIPARENTAL
```

```
      failure _d: EventoAno
analysis time _t: Tempo
              id: ID
```

Log-rank test for equality of survivor functions

BIPARENTAL	Events observed	Events expected
Não	10172	8947.99
Sim	6759	7983.01
Total	16931	16931.00

```
      chi2(1) =    400.70
      Pr>chi2 =     0.0000
```

```
. sts test BIPARENTAL, wilcoxon
```

```
      failure _d: EventoAno
analysis time _t: Tempo
              id: ID
```

Wilcoxon (Breslow) test for equality of survivor functions

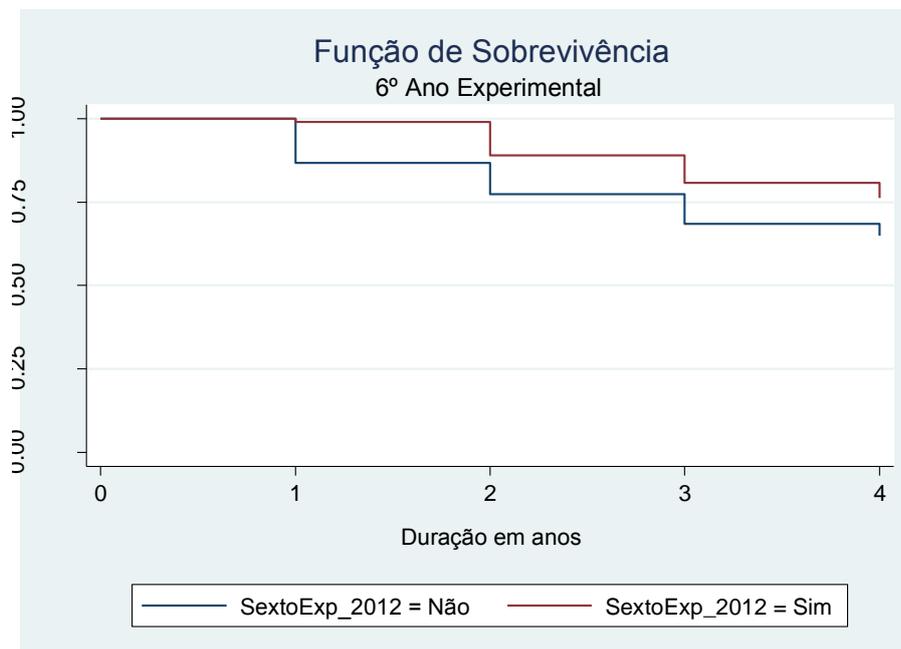
BIPARENTAL	Events observed	Events expected	Sum of ranks
Não	10172	8947.99	55142298
Sim	6759	7983.01	-55142298
Total	16931	16931.00	0

```
      chi2(1) =    371.92
      Pr>chi2 =     0.0000
```

```
. sts graph, by( BIPARENTAL) title (Função de Sobrevivência) subtitle (Família Biparental) xtitle
> (Duração em anos, margin (medium))
```

```
      failure _d: EventoAno
analysis time _t: Tempo
              id: ID
```

- 6º Ano Experimental 2012



```
. sts list, by( SextoExp_2012)
```

```

      failure _d: EventoAno
analysis time _t: Tempo
           id: ID

```

Time	Beg. Total	Net Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
Não							
1	54326	7174	9000	0.8679	0.0015	0.8651	0.8708
2	38152	4115	3049	0.7743	0.0019	0.7706	0.7780
3	30988	3564	1618	0.6853	0.0022	0.6810	0.6895
4	25806	1331	2.4e+04	0.6499	0.0023	0.6454	0.6544
Sim							
1	3693	34	371	0.9908	0.0016	0.9871	0.9934
2	3288	331	238	0.8911	0.0054	0.8800	0.9011
3	2719	252	139	0.8085	0.0070	0.7944	0.8217
4	2328	130	2198	0.7633	0.0076	0.7480	0.7779

```
. sts test SextoExp_2012
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

Log-rank test for equality of survivor functions

SextoEx~2012	Events observed	Events expected
Não	16184	15690.72
Sim	747	1240.28
Total	16931	16931.00

chi2(1) = 238.02
Pr>chi2 = 0.0000

```
. sts test SextoExp_2012, wilcoxon
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

Wilcoxon (Breslow) test for equality of survivor functions

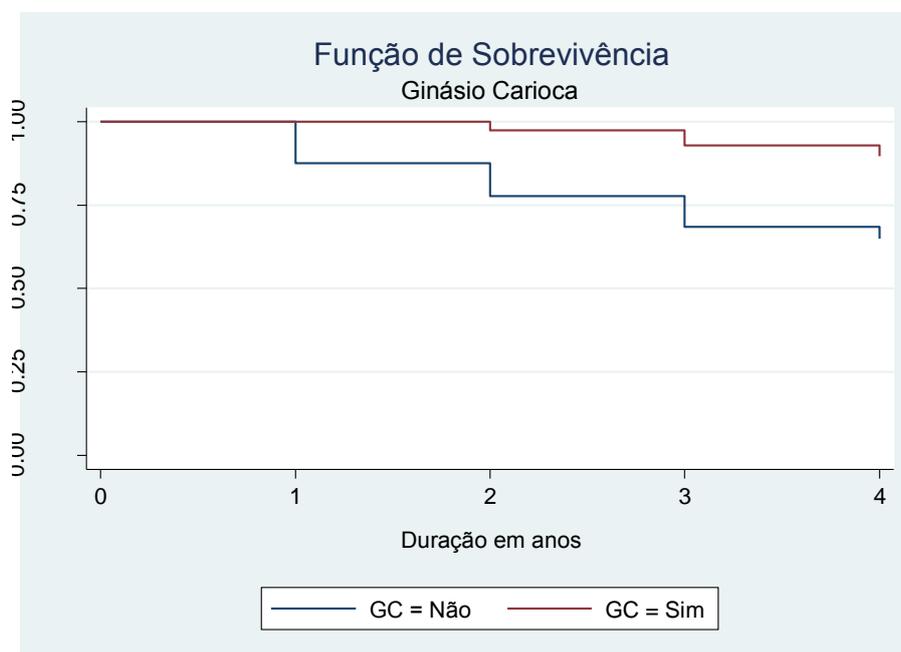
SextoEx~2012	Events observed	Events expected	Sum of ranks
Não	16184	15690.72	27173634
Sim	747	1240.28	-27173634
Total	16931	16931.00	0

chi2(1) = 346.23
Pr>chi2 = 0.0000

```
. sts graph, by( SextoExp_2012) title (Função de Sobrevivência) subtitle (6° Ano Experimental) xt
> itle (Duração em anos, margin (medium))
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

- Ginásio Carioca



```
. sts list, by (GC)
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

Time	Beg. Total	Net Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
Não							
1	58000	7208	1.2e+04	0.8757	0.0014	0.8730	0.8784
2	38992	4382	3030	0.7773	0.0019	0.7736	0.7809
3	31580	3718	1594	0.6858	0.0022	0.6815	0.6900
4	26268	1395	2.5e+04	0.6494	0.0023	0.6449	0.6538
Sim							
1	19	0	-2429	1.0000	.	.	.
2	2448	64	257	0.9739	0.0032	0.9667	0.9795
3	2127	98	163	0.9290	0.0054	0.9176	0.9388
4	1866	66	1800	0.8961	0.0065	0.8825	0.9082

```
. sts test GC
```

```
      failure _d: EventoAno
analysis time _t: Tempo
              id: ID
```

Log-rank test for equality of survivor functions

GC	Events observed	Events expected
Não	16703	16328.30
Sim	228	602.70
Total	16931	16931.00

chi2(1) = 276.09
Pr>chi2 = 0.0000

```
. sts test GC, wilcoxon
```

```
      failure _d: EventoAno
analysis time _t: Tempo
              id: ID
```

Wilcoxon (Breslow) test for equality of survivor functions

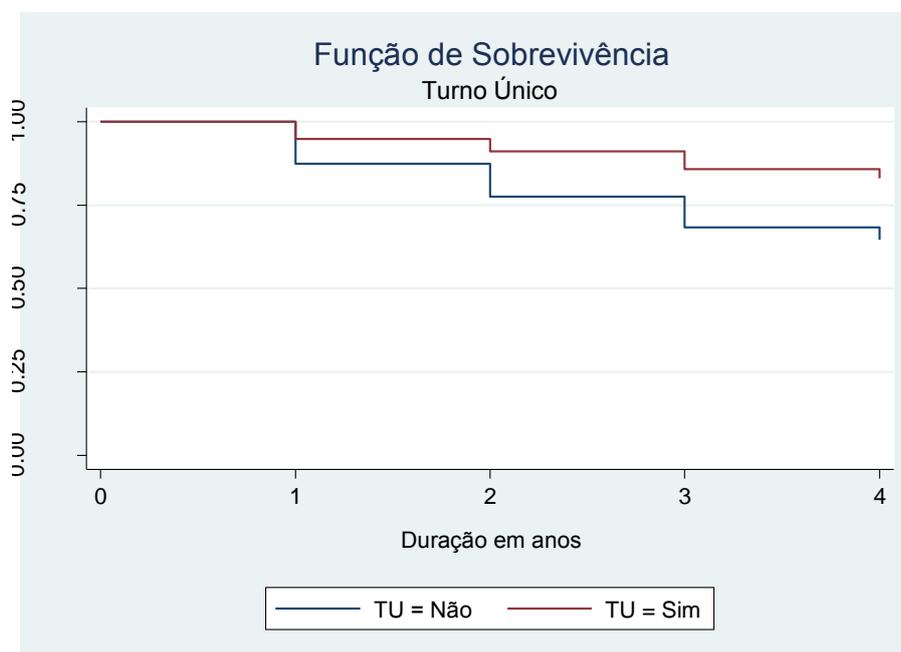
GC	Events observed	Events expected	Sum of ranks
Não	16703	16328.30	14051328
Sim	228	602.70	-14051328
Total	16931	16931.00	0

chi2(1) = 289.92
Pr>chi2 = 0.0000

```
. sts graph, by(GC) title (Função de Sobrevivência) subtitle (Ginásio Carioca) xtitle (Duração em
> anos, margin (medium))
```

```
      failure _d: EventoAno
analysis time _t: Tempo
              id: ID
```

- Turno Único



```
. sts list, by (TU)
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

	Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
Não								
	1	56870	7148	1.1e+04	0.8743	0.0014	0.8716	0.8770
	2	38250	4321	3097	0.7755	0.0019	0.7718	0.7792
	3	30832	3649	1515	0.6838	0.0022	0.6795	0.6880
	4	25668	1381	2.4e+04	0.6470	0.0023	0.6425	0.6514
Sim								
	1	1149	60	-2101	0.9478	0.0066	0.9333	0.9592
	2	3190	125	190	0.9106	0.0071	0.8957	0.9236
	3	2875	167	242	0.8577	0.0078	0.8417	0.8723
	4	2466	80	2386	0.8299	0.0081	0.8133	0.8452

```
. sts test TU
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

Log-rank test for equality of survivor functions

TU	Events observed	Events expected
Não	16499	15992.47
Sim	432	938.53
Total	16931	16931.00

```
chi2(1) = 329.22
Pr>chi2 = 0.0000
```

```
. sts test TU, wilcoxon
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

Wilcoxon (Breslow) test for equality of survivor functions

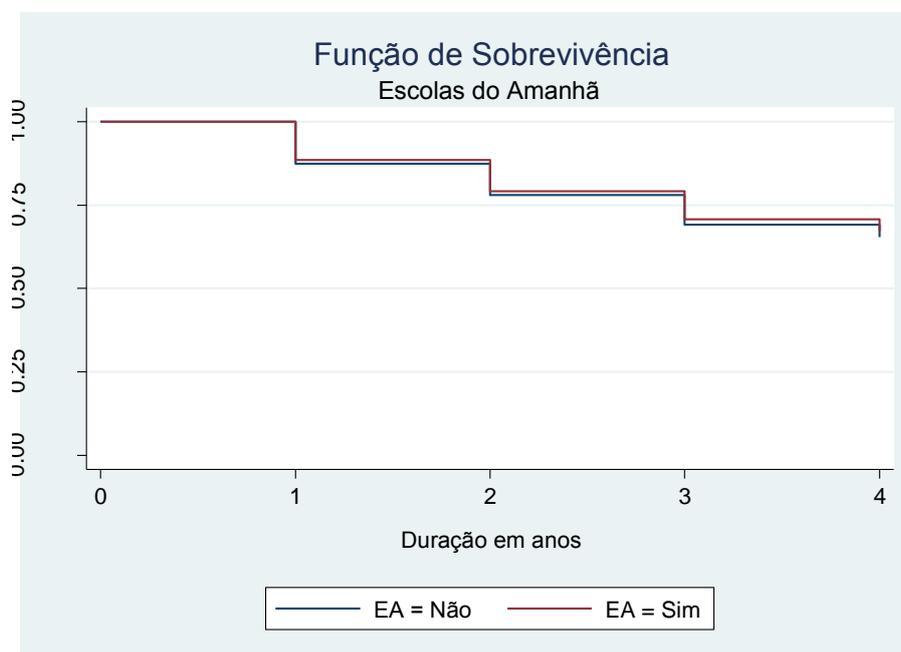
TU	Events observed	Events expected	Sum of ranks
Não	16499	15992.47	20497629
Sim	432	938.53	-20497629
Total	16931	16931.00	0

```
chi2(1) = 327.11
Pr>chi2 = 0.0000
```

```
. sts graph, by(TU) title (Função de Sobrevida) subtitle (Turno Único) xtitle (Duração em ano)
> s, margin (medium)
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

- Escolas do Amanhã



```
. sts list, by (EA)
```

```
      failure _d: EventoAno
analysis time _t: Tempo
      id: ID
```

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
Não							
1	49315	6213	6819	0.8740	0.0015	0.8711	0.8769
2	36283	3903	2797	0.7800	0.0019	0.7761	0.7838
3	29583	3371	1449	0.6911	0.0022	0.6867	0.6955
4	24763	1296	2.3e+04	0.6549	0.0023	0.6503	0.6595
Sim							
1	8704	995	2552	0.8857	0.0034	0.8788	0.8922
2	5157	543	490	0.7924	0.0049	0.7827	0.8018
3	4124	445	308	0.7069	0.0058	0.6954	0.7181
4	3371	165	3206	0.6723	0.0061	0.6602	0.6841

```
. sts test EA
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

Log-rank test for equality of survivor functions

EA	Events observed	Events expected
Não	14783	14654.44
Sim	2148	2276.56
Total	16931	16931.00

chi2(1) = 9.46
Pr>chi2 = 0.0021

```
. sts test EA, wilcoxon
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

Wilcoxon (Breslow) test for equality of survivor functions

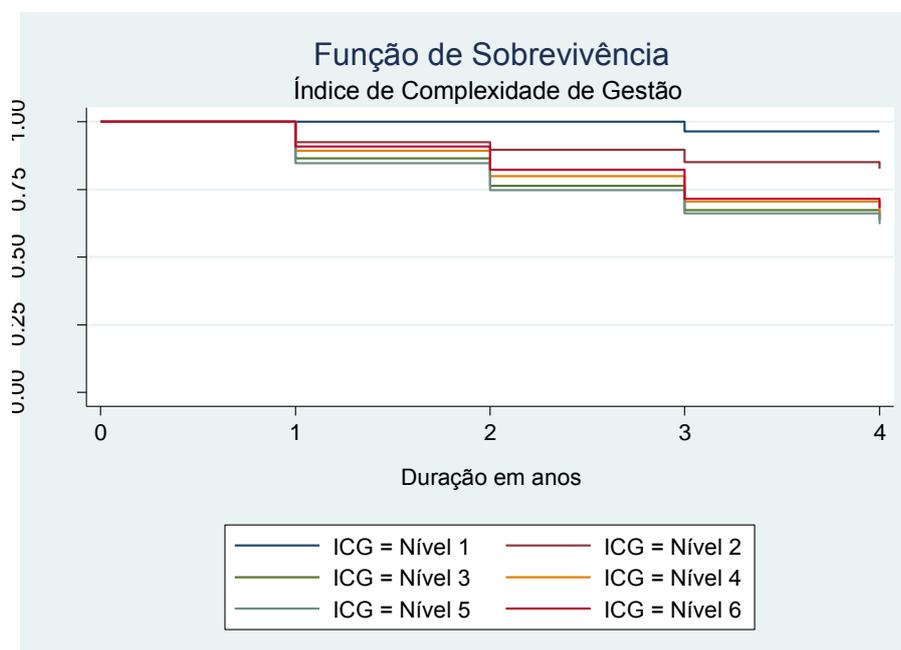
EA	Events observed	Events expected	Sum of ranks
Não	14783	14654.44	6456119
Sim	2148	2276.56	-6456119
Total	16931	16931.00	0

chi2(1) = 10.47
Pr>chi2 = 0.0012

```
. sts graph, by(EA) title (Função de Sobrevida) subtitle (Escolas do Amanhã) xtitle (Duração
> em anos, margin (medium))
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

- ICG



```
. sts list, by (ICG)
```

```
failure _d: EventoAno
analysis time _t: Tempo
id: ID
```

Time	Beg. Total	Net Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
Nível 1							
1	0	0	-29	1.0000	.	.	.
2	29	0	1	1.0000	.	.	.
3	28	1	2	0.9643	0.0351	0.7724	0.9949
4	25	0	25	0.9643	0.0351	0.7724	0.9949
Nível 2							
1	1614	123	-1350	0.9238	0.0066	0.9097	0.9357
2	2841	83	231	0.8968	0.0070	0.8821	0.9098
3	2527	127	244	0.8517	0.0077	0.8358	0.8662
4	2156	63	2093	0.8268	0.0081	0.8102	0.8421
Nível 3							
1	31071	4174	4049	0.8657	0.0019	0.8618	0.8694
2	22848	2676	-209	0.7643	0.0025	0.7593	0.7692
3	20381	2407	127	0.6740	0.0028	0.6685	0.6795
4	17847	944	1.7e+04	0.6384	0.0029	0.6327	0.6440
Nível 4							
1	14411	1535	4474	0.8935	0.0026	0.8883	0.8984
2	8402	879	2276	0.8000	0.0038	0.7925	0.8073
3	5247	626	1242	0.7046	0.0049	0.6949	0.7140
4	3379	182	3197	0.6666	0.0054	0.6560	0.6770
Nível 5							
1	6220	948	178	0.8476	0.0046	0.8384	0.8563
2	5094	599	118	0.7479	0.0056	0.7368	0.7586
3	4377	504	-56	0.6618	0.0061	0.6497	0.6736
4	3929	234	3695	0.6224	0.0063	0.6100	0.6345
Nível 6							
1	4703	428	2049	0.9090	0.0042	0.9004	0.9169
2	2226	209	870	0.8236	0.0068	0.8099	0.8365
3	1147	151	198	0.7152	0.0101	0.6948	0.7345
4	798	38	760	0.6812	0.0110	0.6590	0.7022

```
. sts test ICG
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

Log-rank test for equality of survivor functions

ICG	Events observed	Events expected
Nível 1	1	7.58
Nível 2	396	903.36
Nível 3	10201	9545.57
Nível 4	3222	3461.27
Nível 5	2285	2018.82
Nível 6	826	994.39
Total	16931	16931.00

```
chi2(5) = 466.55
```

```
Pr>chi2 = 0.0000
```

```
. sts test ICG, wilcoxon
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

Wilcoxon (Breslow) test for equality of survivor functions

ICG	Events observed	Events expected	Sum of ranks
Nível 1	1	7.58	-238600
Nível 2	396	903.36	-20428658
Nível 3	10201	9545.57	31365652
Nível 4	3222	3461.27	-14483156
Nível 5	2285	2018.82	13471671
Nível 6	826	994.39	-9686909
Total	16931	16931.00	0

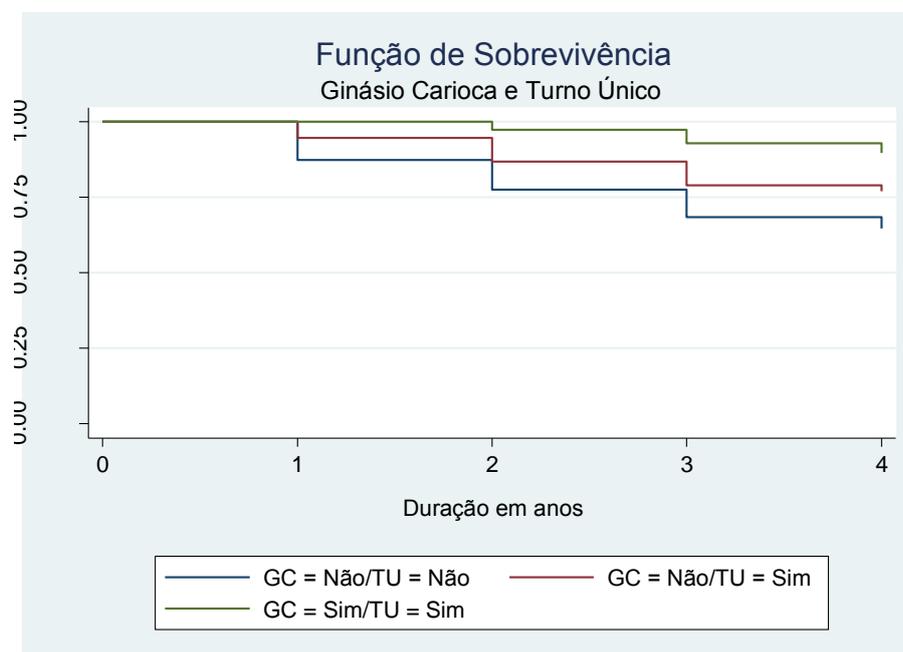
```
chi2(5) = 471.01
```

```
Pr>chi2 = 0.0000
```

```
. sts graph, by(ICG) title (Função de Sobrevivência) subtitle (Índice de Complexidade de Gestão)
> xtitle (Duração em anos, margin (medium))
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

- Ginásio Carioca e Turno Único



```
. sts list, by (GC TU)
```

```
      failure _d: EventoAno
analysis time _t: Tempo
      id: ID
```

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
Não Não							
1	56870	7148	1.1e+04	0.8743	0.0014	0.8716	0.8770
2	38250	4321	3097	0.7755	0.0019	0.7718	0.7792
3	30832	3649	1515	0.6838	0.0022	0.6795	0.6880
4	25668	1381	2.4e+04	0.6470	0.0023	0.6425	0.6514
Não Sim							
1	1130	60	328	0.9469	0.0067	0.9321	0.9585
2	742	61	-67	0.8691	0.0113	0.8450	0.8896
3	748	69	79	0.7889	0.0138	0.7603	0.8145
4	600	14	586	0.7705	0.0143	0.7409	0.7971
Sim Sim							
1	19	0	-2429	1.0000	.	.	.
2	2448	64	257	0.9739	0.0032	0.9667	0.9795
3	2127	98	163	0.9290	0.0054	0.9176	0.9388
4	1866	66	1800	0.8961	0.0065	0.8825	0.9082

```
. sts test GC TU
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

Log-rank test for equality of survivor functions

GC, TU	Events observed	Events expected
Não, Não	16499	15992.47
Não, Sim	204	335.83
Sim, Sim	228	602.70
Total	16931	16931.00

chi2(2) = 343.01
Pr>chi2 = 0.0000

```
. sts test GC TU, wilcoxon
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

Wilcoxon (Breslow) test for equality of survivor functions

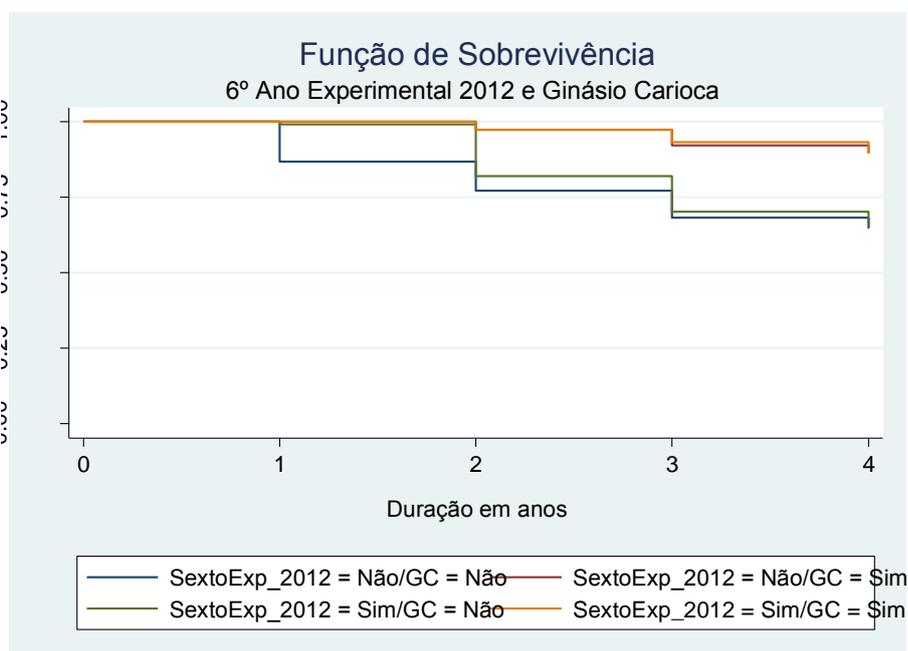
GC, TU	Events observed	Events expected	Sum of ranks
Não, Não	16499	15992.47	20497629
Não, Sim	204	335.83	-6446301
Sim, Sim	228	602.70	-14051328
Total	16931	16931.00	0

chi2(2) = 361.85
Pr>chi2 = 0.0000

```
. sts graph, by(GC TU) title (Função de Sobrevivência) subtitle (Ginásio Carioca e Turno Único) x
> title (Duração em anos, margin (medium))
```

```
      failure _d: EventoAno
analysis time _t: Tempo
           id: ID
```

- 6º Ano Experimental 2012 x Ginásio Carioca



```
. sts list, by ( SextoExp_2012 GC)
```

```

      failure _d: EventoAno
analysis time _t: Tempo
      id: ID

```

Time	Beg. Total	Net Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
Não Não							
1	54307	7174	9810	0.8679	0.0015	0.8650	0.8707
2	37323	4093	3003	0.7727	0.0019	0.7690	0.7764
3	30227	3524	1600	0.6826	0.0022	0.6783	0.6869
4	25103	1312	2.4e+04	0.6470	0.0023	0.6424	0.6514
Não Sim							
1	19	0	-810	1.0000	.	.	.
2	829	22	46	0.9735	0.0056	0.9600	0.9824
3	761	40	18	0.9223	0.0095	0.9014	0.9389
4	703	19	684	0.8974	0.0108	0.8740	0.9166
Sim Não							
1	3693	34	1990	0.9908	0.0016	0.9871	0.9934
2	1669	289	27	0.8192	0.0093	0.8002	0.8366
3	1353	194	-6	0.7018	0.0111	0.6793	0.7230
4	1165	83	1082	0.6518	0.0116	0.6285	0.6740
Sim Sim							
1	0	0	-1619	1.0000	.	.	.
2	1619	42	211	0.9741	0.0040	0.9651	0.9808
3	1366	58	145	0.9327	0.0065	0.9187	0.9444
4	1163	47	1116	0.8950	0.0083	0.8776	0.9101

Log-rank test for equality of survivor functions

Sex~2012, GC	Events	Events
	observed	expected
Não, Não	16103	15476.76
Não, Sim	81	213.96
Sim, Não	600	851.54
Sim, Sim	147	388.74
Total	16931	16931.00

chi2(3) = 376.46
Pr>chi2 = 0.0000

```
. sts test SextoExp_2012 GC, wilcoxon
```

```
failure _d: EventoAno
analysis time _t: Tempo
id: ID
```

Wilcoxon (Breslow) test for equality of survivor functions

Sex~2012, GC	Events	Events	Sum of
	observed	expected	ranks
Não, Não	16103	15476.76	32132873
Não, Sim	81	213.96	-4959239
Sim, Não	600	851.54	-18081545
Sim, Sim	147	388.74	-9092089
Total	16931	16931.00	0

chi2(3) = 492.94
Pr>chi2 = 0.0000

```
. sts graph, by( SextoExp_2012 GC) title (Função de Sobrevida) subtitle (6° Ano Experimental
> 2012 e Ginásio Carioca) xtitle (Duração em anos, margin (medium))
```

```
failure _d: EventoAno
analysis time _t: Tempo
id: ID
```

ANEXO C – MODELAGEM DA REGRESSÃO DE COX

Modelo 1: Bloco das variáveis sobre Perfil Socioeconômico

Passo 1: Testar as variáveis individualmente e comparar os Log likelihood através do Teste de Razão de Verossimilhança (TRV)

Cod. Var.	Covariável	Log	TRV	valor p
Nulo		-181082,13	-	-
Mod. 1.1	Idade_2012	-180295,91	1572,44	0,000
Mod. 1.2	Defasado	-180691,71	780,84	0,000
Mod. 1.3	Anos_Defasagem	-180812,88	538,50	0,000
Mod. 1.4	Anos_Idade_Ideal	-180345,76	1472,74	0,000
Mod. 1.5	Sexo	-180761,11	642,04	0,000
Mod. 1.6	Raça_Branca	-171856,05	18452,16	0,000
Mod. 1.7	Raça_Parda	-171970,84	18222,58	0,000
Mod. 1.8	Raça_Preta	-171910,32	18343,62	0,000
Mod. 1.9	Raça_Outras	-171995,58	18173,10	0,102
Mod. 1.10	Não_Branco	-171856,05	18452,16	0,000
Mod. 1.11	Bolsa Família	-181024,19	115,88	0,000
Mod. 1.12	CFC	-181061,58	41,10	0,000
Mod. 1.13	NIS	-180978,27	207,72	0,000
Mod. 1.14	Deficiente	-181082,12	0,02	0,876
Mod. 1.15	Biparental	-180902,39	359,48	0,000
Mod. 1.16	Máx_Escolaridade_Pais	-154931,83	52300,60	0,000

Passo 2: Como são muitas covariáveis para este bloco, optou-se em analisar antes do próximo passo algumas possíveis correlações entre variáveis do mesmo tema.

- Tema Idade:

```
. correlate Idade_2012 Defasado Anos_Defasagem Anos_Idade_Ideal
(obs=161300)
```

	Ida~2012	Defasado	Anos_D~m	Anos_I~1
Idade_2012	1.0000			
Defasado	0.7245	1.0000		
Anos_Defas~m	0.7760	0.9443	1.0000	
Anos_Idade~1	0.9708	0.7538	0.8089	1.0000

➔ Como a correlação é alta, optamos em seguir o modelo com a variável Idade_2012 e Defasado

- Tema: Raça
 - ➔ Optou-se pela variável não-branco, já que todas tem comportamento parecido sob modelo e seguindo a literatura.

- Tema: Programa Social:

```
. correlate Bolsa_Familia CFC NIS_Aluno
(obs=161300)
```

	Bolsa_~a	CFC	NIS_Al~o
Bolsa_Fami~a	1.0000		
CFC	0.6770	1.0000	
NIS_Aluno	0.8401	0.6004	1.0000

- ➔ Como a correlação com o NIS_Aluno é alta, optou-se pelas covariáveis Bolsa Família e CFC.

Passo 3: Como as variáveis pré-selecionadas, testar o modelo com todas as covariáveis que foram significativas a nível de 0,10 (mod. 1.17) como comparação para os outros modelos que serão retiradas uma a uma covariável, verificando-se que covariáveis provocam um aumento estatisticamente significativo no TRV.

Cod. Var.	Covariável	Log	TRV	valor p
Mod. 1.17	Idade_2012; Defasado; Sexo; Não_Branco; Bolsa Família; CFC; Biparental; Máx_Escolaridade_Pais	-147354,24	-	-
Mod. 1.18	Defasado; Sexo; Não_Branco; Bolsa Família; CFC; Biparental; Máx_Escolaridade_Pais	-147613,69	518,90	0,000
Mod. 1.19	Idade_2012; Sexo; Não_Branco; Bolsa Família; CFC; Biparental; Máx_Escolaridade_Pais	-147354,42	0,36	0,000
Mod. 1.20	Idade_2012; Defasado; Não_Branco; Bolsa Família; CFC; Biparental; Máx_Escolaridade_Pais	-147603,35	498,22	0,000
Mod. 1.21	Idade_2012; Defasado; Sexo; Bolsa Família; CFC; Biparental; Máx_Escolaridade_Pais	-153942,07	13175,66	0,000
Mod. 1.22	Idade_2012; Defasado; Sexo; Não_Branco; CFC; Biparental; Máx_Escolaridade_Pais	-147371,08	33,68	0,000
Mod. 1.23	Idade_2012; Defasado; Sexo; Não_Branco; Bolsa Família; Biparental; Máx_Escolaridade_Pais	-147355,18	1,88	0,000
Mod. 1.24	Idade_2012; Defasado; Sexo; Não_Branco; Bolsa Família; CFC; Máx_Escolaridade_Pais	-147472,67	236,86	0,000
Mod. 1.25	Idade_2012; Defasado; Sexo; Não_Branco; Bolsa Família; CFC; Biparental	-170706,96	46705,44	0,000

- Pelos resultados dos modelos 1.19 e 1.23, optou-se em retirar da modelo as variáveis Defasado e CFC;

Passo 4: Verificar se todas as covariáveis atendem a Suposição de Riscos Proporcionais exigida para o Modelo de Cox. Para isto, utilizar-se os resíduos padronizados de Schoenfeld

```
. estat phtest, detail
```

```
Test of proportional-hazards assumption
```

```
Time: Time
```

	rho	chi2	df	Prob>chi2
Idade_2012	-0.01801	1.57	1	0.2096
Sexo	0.03424	16.47	1	0.0000
Nao_Branco	-0.01602	3.65	1	0.0562
Bolsa_Fami~a	-0.00951	1.28	1	0.2574
Biparental	-0.02028	5.88	1	0.0153
Max_Esc_Pais	0.02439	8.40	1	0.0038
global test		39.60	6	0.0000

- As variáveis Idade_2012, Não-Branco e Bolsa Família atenderam à Suposição de Riscos Proporcionais, para as outras variáveis, deverá ser feita a interação delas com o tempo, como sugere Allison (2014, p. 44);

```
. estat phtest, detail
```

```
Test of proportional-hazards assumption
```

```
Time: Time
```

	rho	chi2	df	Prob>chi2
Idade_2012	-0.01709	1.44	1	0.2307
Sexo	0.01470	3.04	1	0.0814
Nao_Branco	-0.01615	3.70	1	0.0543
Bolsa_Fami~a	-0.00947	1.27	1	0.2594
Biparental	-0.00187	0.05	1	0.8233
Max_Esc_Pais	-0.00034	0.00	1	0.9674
Biparental_t	0.00116	0.02	1	0.8907
Max_Esc_Pa~t	-0.00140	0.03	1	0.8679
sexo_t	-0.01645	3.81	1	0.0511
global test		11.19	9	0.2629

→ Acrescentando a interação de tais variáveis, o modelo foi corrigido, com todas as variáveis atendendo à suposição de proporcionalidade de risco exigida pelo Modelo de Cox.

Conclusão: O modelo final para este bloco ficou sendo o Modelo 1.27

Cod. Var.	Covariável	Log	TRV	valor p
Mod. 1.26	Idade_2012; Sexo; Não_Branco; Bolsa Família; Biparental; Máx_Escolaridade_Pais	-147355,39	-	-
Mod. 1.27	Idade_2012; Sexo; Não_Branco; Bolsa Família; Biparental; Máx_Escolaridade_Pais; Biparental*t; Máx_Escolaridade_Pais*t; Sexo*t	-147339,34	-32,10	0,000

```
. stcox Idade_2012 Sexo Nao_Branco Bolsa_Familia Biparental Max_Esc_Pais Biparental_t Max_Esc_Pai
> s_t sexo_t
```

```
failure _d: EventoAno
analysis time _t: Tempo
id: ID
```

```
Iteration 0: log likelihood = -148694.09
Iteration 1: log likelihood = -148302.85
Iteration 2: log likelihood = -147601.97
Iteration 3: log likelihood = -147390.01
Iteration 4: log likelihood = -147343.16
Iteration 5: log likelihood = -147339.37
Iteration 6: log likelihood = -147339.34
Refining estimates:
Iteration 0: log likelihood = -147339.34
```

Cox regression -- Breslow method for ties

```
No. of subjects = 48598 Number of obs = 136705
No. of failures = 14128
Time at risk = 136705
Log likelihood = -147339.34 LR chi2(9) = 2709.49
Prob > chi2 = 0.0000
```

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
Idade_2012	1.271359	.0076464	39.92	0.000	1.25646 1.286434
Sexo	.5938452	.022886	-13.52	0.000	.5506417 .6404384
Nao_Branco	1.194435	.0217984	9.74	0.000	1.152466 1.237932
Bolsa_Familia	1.116773	.0191157	6.45	0.000	1.079928 1.154875
Biparental	.8290207	.0314331	-4.95	0.000	.7696465 .8929753
Max_Esc_Pais	.6054502	.0259179	-11.72	0.000	.5567246 .6584402
Biparental_t	.9633604	.0163057	-2.21	0.027	.9319261 .995855
Max_Esc_Pais_t	1.064041	.0197301	3.35	0.001	1.026065 1.103422
sexo_t	1.071887	.018344	4.06	0.000	1.03653 1.108451

_t	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
Idade_2012	.2400861	.0060144	39.92	0.000	.2282982	.251874
Sexo	-.5211367	.0385386	-13.52	0.000	-.596671	-.4456023
Nao_Branco	.177673	.01825	9.74	0.000	.1419037	.2134423
Bolsa_Familia	.1104433	.0171169	6.45	0.000	.0768948	.1439918
Biparental	-.1875102	.0379159	-4.95	0.000	-.2618239	-.1131964
Max_Esc_Pais	-.501783	.0428077	-11.72	0.000	-.5856846	-.4178815
Biparental_t	-.0373277	.0169259	-2.21	0.027	-.0705018	-.0041536
Max_Esc_Pais_t	.0620736	.0185426	3.35	0.001	.0257308	.0984164
sexo_t	.0694209	.0171138	4.06	0.000	.0358786	.1029633

Teste de resíduos de Schoenfeld

```
. estat phtest, detail
```

Test of proportional-hazards assumption

Time: Time

	rho	chi2	df	Prob>chi2
Idade_2012	-0.01709	1.44	1	0.2307
Sexo	0.01470	3.04	1	0.0814
Nao_Branco	-0.01615	3.70	1	0.0543
Bolsa_Familia	-0.00947	1.27	1	0.2594
Biparental	-0.00187	0.05	1	0.8233
Max_Esc_Pais	-0.00034	0.00	1	0.9674
Biparental_t	0.00116	0.02	1	0.8907
Max_Esc_Pais_t	-0.00140	0.03	1	0.8679
sexo_t	-0.01645	3.81	1	0.0511
global test		11.19	9	0.2629

Modelo 2: Bloco das variáveis sobre Desempenho Acadêmico

Passo 1: Testar as variáveis individualmente e comparar os Log likelihood através do Teste de Razão de Verossimilhança (TRV)

Cod. Var.	Covariável	Log	TRV	valor p
Nulo		-181082,13	-	-
Mod. 2.1	Conc_l	-168852,92	24458,42	0,0000
Mod. 2.2	FreqAnual	-175432,83	11298,60	0,0000
Mod. 2.3	PB_LPL_1	-157188,80	47786,66	0,0000
Mod. 2.4	PB_LPL_1e2	-171316,55	19531,16	0,0000
Mod. 2.5	PB_MAT_1	-157365,15	47433,96	0,0000
Mod. 2.6	PB_MAT_1e2	-171599,70	18964,86	0,0000
Mod. 2.7	Mov_Escola	-180979,33	205,60	0,0000
Mod. 2.8	Mov_Turma	-180914,81	334,64	0,0000

Passo 2: Como as variáveis pré-selecionadas, testar o modelo com todas as covariáveis que foram significativas a nível de 0,10 (mod. 1.17) como comparação para os outros modelos que serão retiradas uma a uma covariável, verificando-se que covariáveis provocam um aumento estatisticamente significativo no TRV.

Cod. Var.	Covariável	Log	TRV	valor p
Mod. 2.9	Conc_I; FreqAnual; PB_LPL_1; PB_LPL_1e2; PB_MAT_1; PB_MAT_1e2; Mov_Escola; Mov_Turma	-134843,14	-	-
Mod. 2.10	FreqAnual; PB_LPL_1; PB_LPL_1e2; PB_MAT_1; PB_MAT_1e2; Mov_Escola; Mov_Turma	-140002,27	10318,26	0,0000
Mod. 2.11	Conc_I; PB_LPL_1; PB_LPL_1e2; PB_MAT_1; PB_MAT_1e2; Mov_Escola; Mov_Turma	-137036,86	4387,44	0,0000
Mod. 2.12	Conc_I; FreqAnual; PB_LPL_1e2; PB_MAT_1; PB_MAT_1e2; Mov_Escola; Mov_Turma	-143185,05	16683,82	0,0000
Mod. 2.13	Conc_I; FreqAnual; PB_LPL_1; PB_MAT_1; PB_MAT_1e2; Mov_Escola; Mov_Turma	-134906,07	125,86	0,0000
Mod. 2.14	Conc_I; FreqAnual; PB_LPL_1; PB_LPL_1e2; PB_MAT_1e2; Mov_Escola; Mov_Turma	-143494,09	17301,90	0,0000
Mod. 2.15	Conc_I; FreqAnual; PB_LPL_1; PB_LPL_1e2; PB_MAT_1; Mov_Escola; Mov_Turma	-135068,71	451,14	0,0000
Mod. 2.16	Conc_I; FreqAnual; PB_LPL_1; PB_LPL_1e2; PB_MAT_1; PB_MAT_1e2; Mov_Turma	-134849,45	12,62	0,0000
Mod. 2.17	Conc_I; FreqAnual; PB_LPL_1; PB_LPL_1e2; PB_MAT_1; PB_MAT_1e2; Mov_Escola	-134854,46	22,64	0,0000

Passo 3: Verificar se todas as covariáveis atendem a Suposição de Riscos Proporcionais exigida para o Modelo de Cox. Para isto, utilizar-se os resíduos padronizados de Schoenfeld

```
. estat phtest, detail
```

```
Test of proportional-hazards assumption
```

```
Time: Time
```

	rho	chi2	df	Prob>chi2
Conc_I	0.01239	2.27	1	0.1317
FreqAnual	-0.08093	49.07	1	0.0000
PB_LPL_1	-0.02627	9.86	1	0.0017
PB_LPL_1e2	0.00859	1.06	1	0.3041
PB_MAT_1	0.00131	0.02	1	0.8768
PB_MAT_1e2	-0.00533	0.39	1	0.5313
Mov_Escola	0.01410	2.84	1	0.0921
Mov_Turma	0.01985	5.56	1	0.0184
global test		117.93	8	0.0000

→ As variáveis FreqAnual e PB_LPL_1 não atenderam à Suposição de Riscos Proporcionais, por isso, deverá ser feita a interação delas com o tempo;

```
. estat phtest, detail
```

```
Test of proportional-hazards assumption
```

```
Time: Time
```

	rho	chi2	df	Prob>chi2
Conc_I	-0.01133	1.90	1	0.1679
FreqAnual	0.15402	326.63	1	0.0000
PB_LPL_1	-0.00553	0.43	1	0.5130
PB_LPL_1e2	0.00489	0.34	1	0.5576
PB_MAT_1	-0.00003	0.00	1	0.9970
PB_MAT_1e2	-0.00976	1.32	1	0.2498
Mov_Escola	0.01031	1.52	1	0.2177
Mov_Turma	0.02375	7.96	1	0.0048
FreqAnual_t	-0.15946	361.70	1	0.0000
PB_LPL_1_t	0.00127	0.02	1	0.8825
global test		375.47	10	0.0000

→ Com essa solução, resolveu-se o problema da variável o PB_LPL_1, mas não a da FreqAnual. Como esta variável é numérica, podendo variar de 0 a 100, não é possível plotar um gráfico inteligível para a mesma, não sendo possível utilizar o método gráfico como proposto por Colosimo e Giolo (2006,p.166-167). Por não ter como comprovar a satisfação à suposição do Modelo de Cox, e por entender que, segundo à lei, os alunos com baixa frequência não pode ser aprovado (o que poderia enviesar o modelo) optou-se em retirar essa variável.

```
. estat phtest, detail
```

```
Test of proportional-hazards assumption
```

```
Time: Time
```

	rho	chi2	df	Prob>chi2
Conc_I	0.02746	10.88	1	0.0010
PB_LPL_1	0.00652	0.60	1	0.4381
PB_LPL_1e2	0.00613	0.55	1	0.4591
PB_MAT_1	-0.00430	0.26	1	0.6069
PB_MAT_1e2	-0.00704	0.69	1	0.4050
Mov_Escola	0.02138	6.55	1	0.0105
Mov_Turma	0.01247	2.20	1	0.1384
PB_LPL_1_t	-0.00639	0.54	1	0.4635
global test		30.16	8	0.0002

- Ao retirar a covariável FreqAnual, a covariável Conc_I deixou de atender à suposição, sendo necessário verificar a sua interação com o tempo.

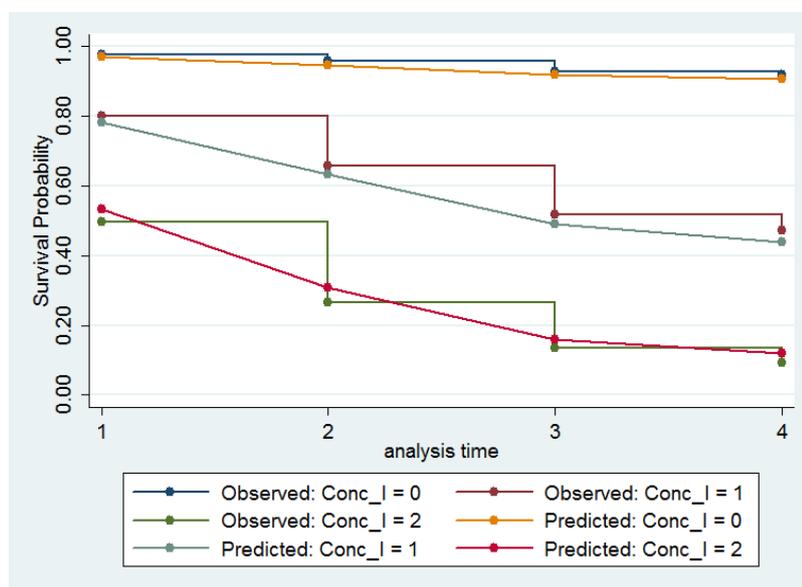
```
. estat phtest, detail
```

Test of proportional-hazards assumption

Time: Time

	rho	chi2	df	Prob>chi2
Conc_I	-0.02226	6.95	1	0.0084
PB_LPL_1	-0.00193	0.05	1	0.8155
PB_LPL_1e2	0.00507	0.38	1	0.5393
PB_MAT_1	-0.00426	0.26	1	0.6099
PB_MAT_1e2	-0.00838	0.99	1	0.3207
Mov_Escola	0.02120	6.45	1	0.0111
Mov_Turma	0.01248	2.20	1	0.1382
PB_LPL_1_t	0.00071	0.01	1	0.9321
Conc_I_t	0.02302	7.05	1	0.0079
global test		20.88	9	0.0132

- Ainda assim, não há como garantir que essa variável atende a Suposição, por isso, optou-se em utilizar o método gráfico, sugerido por Colosimo e Giolo (2006, p. 166-167)



- Como é possível analisar no gráfico, as curvas não são paralelas, demonstrando assim uma possível não proporcionalidade no tempo, sendo assim, esta variável não poderá ser utilizada no modelo.

```
. estat phtest, detail
```

```
Test of proportional-hazards assumption
```

```
Time: Time
```

	rho	chi2	df	Prob>chi2
PB_LPL_1	-0.00324	0.15	1	0.6996
PB_LPL_1e2	-0.00284	0.12	1	0.7335
PB_MAT_1	-0.00241	0.08	1	0.7729
PB_MAT_1e2	0.00226	0.07	1	0.7911
Mov_Escola	0.03311	15.46	1	0.0001
Mov_Turma	0.00667	0.63	1	0.4287
PB_LPL_1_t	0.00972	1.20	1	0.2736
global test		17.69	7	0.0135

- ➔ A covariável Mov_Escola ainda não está atendendo ao pressuposto, por isso foi feita a interação da variável com o tempo, neste modelo, a variável Mov_Escola deixou de ser significativa, por isso, optou-se em retirar a variável do modelo.

Passo 4: Verificar se os coeficientes estão apontados para a direção esperada.

- ➔ Optou-se retirar a variável PB_LPL_1 e PB_MAT_1 do modelo, pois está contradizendo o que se espera, já que sua interpretação, seria para cada ponto a mais na nota, maior é a chance de repetência.

```

. *** Modelo 2

. stcox PB_LPL_1 PB_LPL_1e2 PB_MAT_1 PB_MAT_1e2 Mov_Turma PB_LPL_1_t

      failure _d:  EventoAno
analysis time _t:  Tempo
           id:   ID

Iteration 0:  log likelihood = -148304.92
Iteration 1:  log likelihood = -143501.78
Iteration 2:  log likelihood = -143408.63
Iteration 3:  log likelihood = -143408.41
Iteration 4:  log likelihood = -143408.41
Refining estimates:
Iteration 0:  log likelihood = -143408.41

Cox regression -- Breslow method for ties

No. of subjects =          54241                Number of obs   =       147088
No. of failures =          13989
Time at risk    =          147088

LR chi2(6)      =       9793.02
Prob > chi2     =         0.0000

Log likelihood  =      -143408.41

```

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
PB_LPL_1	1.093113	.0130035	7.48	0.000	1.067921	1.118899
PB_LPL_1e2	.7389685	.0079813	-28.01	0.000	.7234899	.7547783
PB_MAT_1	1.033072	.0095028	3.54	0.000	1.014614	1.051866
PB_MAT_1e2	.7231356	.0080741	-29.03	0.000	.7074826	.7391349
Mov_Turma	1.473313	.0509311	11.21	0.000	1.376797	1.576596
PB_LPL_1_t	.9679799	.0040947	-7.69	0.000	.9599876	.9760387

➔ Com a retirada das variáveis acima, a variável PB_LPL1e2 não atendeu mais a suposição de proporcionalidade, optando-se assim por incluir a sua interação com o tempo.

Conclusão: O modelo final para este bloco ficou sendo o Modelo 2.23

Cod. Var.	Covariável	Log	TRV	valor p
Mod. 2.9	Conc_I; FreqAnual; PB_LPL_1; PB_LPL_1e2; PB_MAT_1; PB_MAT_1e2; Mov_Escola; Mov_Turma	-134843,14	-	-
Mod. 2.18	Conc_I; FreqAnual; PB_LPL_1; PB_LPL_1e2; PB_MAT_1; PB_MAT_1e2; Mov_Escola; Mov_Turma; FreqAnual_t; PB_LPL_1_t	-134734,87	216,54	0,0000
Mod. 2.19	Conc_I; PB_LPL_1; PB_LPL_1e2; PB_MAT_1; PB_MAT_1e2; Mov_Escola; Mov_Turma; PB_LPL_1_t	-137009,94	4333,60	0,0000
Mod. 2.20	Conc_I; PB_LPL_1; PB_LPL_1e2; PB_MAT_1; PB_MAT_1e2; Mov_Escola; Mov_Turma; PB_LPL_1_t; Conc_I_t	-137001,80	4317,32	0,0000
Mod. 2.21	PB_LPL_1; PB_LPL_1e2; PB_MAT_1; PB_MAT_1e2; Mov_Escola; Mov_Turma;	-143361,19	17036,10	0,0000

Modelo 3: Bloco das variáveis sobre Políticas Públicas

Passo 1: Testar as variáveis individualmente e comparar os Log likelihood através do Teste de Razão de Verossimilhança (TRV)

Modelo	Covariável	Log	TRV	valor p
Nulo	-	-181082,13	-	-
Mod. 3.1	SextoExp_2012	-180959,82	244,62	0,000
Mod. 3.2	GC	-180921,51	321,24	0,000
Mod. 3.3	TU	-180900,19	363,88	0,000
Mod. 3.4	EA	-181077,86	8,54	0,004
Mod. 3.5	ICG	-181079,14	5,98	0,014

→ Todas as covariáveis foram significativas, portanto, faremos o próximo passo com todas as variáveis.

Passo 2: Testar o modelo com todas as covariáveis que foram significativas a nível de 0,10 (mod. 3.6) como comparação para os outros modelos que serão retiradas uma a uma covariável, verificando-se que covariáveis provocam um aumento estatisticamente significativo no TRV.

Modelo	Covariável	Log	TRV	p-valor
Mod. 3.6	SextoExp_2012; GC; TU; EA; ICG	-180840,08	-	0,000
Mod. 3.7	GC; TU; EA; ICG	-180880,01	79,86	0,000
Mod. 3.8	SextoExp_2012; TU; EA; ICG	-180846,23	12,30	0,000
Mod. 3.9	SextoExp_2012; GC; EA; ICG	-180870,98	61,80	0,000
Mod. 3.10	SextoExp_2012; GC; TU; ICG	-180841,27	2,38	0,000
Mod. 3.11	SextoExp_2012; GC; TU; EA	-180845,04	9,92	0,000

→ Com o TVR do modelo 3.10, observou-se que a variável EA não altera o modelo de forma significativa, por isso, optou-se em prosseguir com esse modelo.

Passo 3: Verificar se há correlação entre as covariáveis

```
. correlate SextoExp_2012 GC TU ICG
(obs=161300)
```

	Sex~2012	GC	TU	ICG
SextoEx~2012	1.0000			
GC	0.4413	1.0000		
TU	0.3791	0.8084	1.0000	
ICG	-0.1152	-0.3034	-0.3456	1.0000

→ As variáveis GC e TU possuem correlação, por isso, há presença de multicolinearidade entre elas, optando-se em prosseguir com a variável GC, devido à importância que se é dada para tal política na rede.

Passo 4: Verificar se todas as covariáveis atendem a Suposição de Riscos Proporcionais exigida para o Modelo de Cox. Para isto, utilizar-se os resíduos padronizados de Schoenfeld⁷³

```
. estat phtest, detail
```

```
Test of proportional-hazards assumption
```

```
Time: Time
```

	rho	chi2	df	Prob>chi2
SextoEx~2012	0.14486	348.83	1	0.0000
GC	-0.00542	0.49	1	0.4826
ICG	0.02751	12.36	1	0.0004
global test		383.14	3	0.0000

→ O teste indica um problema de não proporcionalidade na variável do SextoExp2012 e ICG; para tentar solucionar, deve-se fazer a interação de tais variáveis com o tempo.

⁷³ “Se os riscos são verdadeiramente proporcionais, os resíduos de Schoenfeld devem ser não correlacionados com o tempo ou com qualquer função do tempo.” (ALLISON, 2014, p.44)

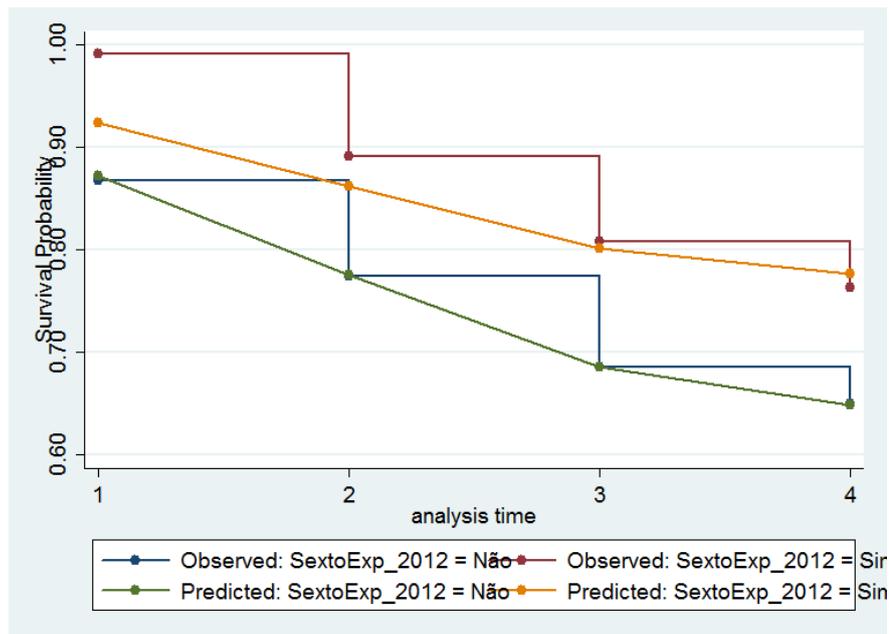
```
. estat phtest, detail
```

Test of proportional-hazards assumption

Time: Time

	rho	chi2	df	Prob>chi2
SextoEx~2012	0.10729	126.66	1	0.0000
GC	0.02402	9.98	1	0.0016
ICG	0.00104	0.02	1	0.8961
SextoExp_2~t	-0.12360	156.62	1	0.0000
ICG_t	-0.00031	0.00	1	0.9689
global test		158.89	5	0.0000

→ Com essa solução, resolveu-se o problema da variável o ICG, mas não a do SextoExp2012. Como não pode rejeitar a H0 de que essa variável não é proporcional ao longo do tempo, optou-se pelo método gráfico⁷⁴.



→ Como as curvas não são paralelas ao longo do tempo, não se pode garantir que essa variável atende à suposição, optando-se, portanto em retirar a mesma do modelo.

⁷⁴ Conforme Colosimo e Giolo (2006, p. 166-167)

```
. estat phtest, detail
```

```
Test of proportional-hazards assumption
```

```
Time: Time
```

	rho	chi2	df	Prob>chi2
GC	0.03632	22.55	1	0.0000
ICG	0.00439	0.31	1	0.5781
ICG_t	-0.00141	0.03	1	0.8570
global test		23.28	3	0.0000

→ Ao retirar a variável SextoExp_2012, a variável GC passou a não atender mais a suposição exigida para o Modelo de Cox, sendo necessário, acrescentar também a interação desta variável com o tempo.

Conclusão: O modelo final para este bloco ficou sendo o Modelo 3.15

Modelo	Covariável	Log	TRV	p-valor
Mod. 3.12	SextoExp_2012; GC; ICG;	--180874,90		
Mod. 3.13	SextoExp_2012; GC; ICG; SextoExp_2012*t; ICG*t	- 180692,77	364,26	0,000
Mod. 3.14	GC; ICG; ICG*t	-180914,64	79,48	0,000
Mod. 3.15	GC; ICG; GC*t; ICG*t	-180893,90	38,00	0,000

```
. *** MODELO 3 FINAL
```

```
. stcox GC ICG GC_t ICG_t
```

```
      failure _d:  EventoAno
analysis time _t:  Tempo
              id:  ID
```

```
Iteration 0:  log likelihood = -181082.13
Iteration 1:  log likelihood = -180916.55
Iteration 2:  log likelihood = -180894.94
Iteration 3:  log likelihood = -180893.9
Iteration 4:  log likelihood = -180893.9
Refining estimates:
Iteration 0:  log likelihood = -180893.9
```

```
Cox regression -- Breslow method for ties
```

```
No. of subjects =      58019                Number of obs   =    161300
No. of failures =      16931
Time at risk    =      161300
Log likelihood  =     -180893.9
LR chi2(4)      =      376.46
Prob > chi2     =      0.0000
```

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
GC	.0709895	.0195842	-9.59	0.000	.0413399	.1219044
ICG	.9234318	.016639	-4.42	0.000	.891389	.9566263
GC_t	1.785386	.1599725	6.47	0.000	1.497834	2.128143
ICG_t	1.034146	.0086996	3.99	0.000	1.017235	1.051339

_t	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
GC	-2.645223	.2758749	-9.59	0.000	-3.185928	-2.104518
ICG	-.0796584	.0180187	-4.42	0.000	-.1149743	-.0443424
GC_t	.5796346	.0896011	6.47	0.000	.4040197	.7552496
ICG_t	.0335763	.0084124	3.99	0.000	.0170883	.0500643

Teste de resíduos de Schoenfeld

```
. estat phtest, detail
```

```
Test of proportional-hazards assumption
```

```
Time:  Time
```

	rho	chi2	df	Prob>chi2
GC	0.00033	0.00	1	0.9663
ICG	-0.00226	0.08	1	0.7732
GC_t	-0.00021	0.00	1	0.9783
ICG_t	0.00253	0.11	1	0.7456
global test		0.11	4	0.9985

Modelo 4: Modelo 1 + Modelo 2

Passo 1: Para começar o modelo 4, optou-se em unir as versões finais do Modelo 1 e 2.

→ Ao unir os modelos, as covariáveis que perderam a significância estatística foram as de Perfil Socioeconômico e a Idade_2012 e Mov_Turma deixou de atender a suposição de proporcionalidade de risco;

```
. stcox Idade_2012 Sexo sexo_t Nao_Branco Bolsa_Familia Biparental Biparental_t Max
> Esc_Pais_t PB_LPL_1e2 PB_LPL_1e2_t PB_MAT_1e2 Mov_Turma
```

```
      failure _d:  EventoAno
analysis time _t:  Tempo
              id:  ID
```

```
Iteration 0:  log likelihood = -141369.44
Iteration 1:  log likelihood = -136676.19
Iteration 2:  log likelihood = -136514.46
Iteration 3:  log likelihood = -136511.33
Iteration 4:  log likelihood = -136511.33
Refining estimates:
Iteration 0:  log likelihood = -136511.33
```

Cox regression -- Breslow method for ties

```
No. of subjects =          47229                Number of obs   =    133773
No. of failures =          13459
Time at risk    =          133773
Log likelihood   =    -136511.33                LR chi2(13)       =    9716.22
                                                Prob > chi2       =    0.0000
```

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
Idade_2012	1.101397	.0104373	10.19	0.000	1.081129	1.122045
Sexo	.7054179	.0282443	-8.72	0.000	.6521764	.7630059
sexo_t	1.033478	.0183724	1.85	0.064	.9980891	1.070122
Nao_Branco	1.048405	.0196482	2.52	0.012	1.010594	1.087631
Bolsa_Familia	1.020923	.017942	1.18	0.239	.986356	1.056701
Biparental	.9231394	.0357689	-2.06	0.039	.8556295	.9959759
Biparental_t	.9438665	.0164125	-3.32	0.001	.9122406	.9765888
Max_Esc_Pais	.8348404	.0372867	-4.04	0.000	.7648671	.9112152
Max_Esc_Pais_t	1.006835	.0195084	0.35	0.725	.9693161	1.045806
PB_LPL_1e2	.8534074	.0090584	-14.93	0.000	.8358366	.8713476
PB_LPL_1e2_t	.9577438	.0049054	-8.43	0.000	.9481776	.9674066
PB_MAT_1e2	.7633024	.0050105	-41.15	0.000	.753545	.7731862
Mov_Turma	1.364802	.046968	9.04	0.000	1.275782	1.460033

```
. estat phtest, detail
```

```
Test of proportional-hazards assumption
```

```
Time: Time
```

	rho	chi2	df	Prob>chi2
Idade_2012	0.03698	14.54	1	0.0001
Sexo	0.00786	0.83	1	0.3608
sexo_t	-0.00764	0.79	1	0.3750
Nao_Branco	-0.01035	1.45	1	0.2288
Bolsa_Fami~a	-0.00379	0.19	1	0.6602
Biparental	-0.00181	0.04	1	0.8330
Biparental_t	0.00283	0.11	1	0.7420
Max_Esc_Pais	-0.01329	2.46	1	0.1170
Max_Esc_Pa~t	0.01577	3.44	1	0.0637
PB_LPL_1e2	-0.00914	1.04	1	0.3067
PB_LPL_1e2_t	0.01420	2.44	1	0.1183
PB_MAT_1e2	0.00896	1.04	1	0.3088
Mov_Turma	0.01808	4.52	1	0.0336
global test		31.09	13	0.0033

- ➔ Optou-se por retirar as interações que deixaram de ser significativas, como Sexo*t e Max_Escolaridade_Pais*t, uma a uma.
- ➔ Optou-se por acrescentar a variável CFC para tentar melhorar a significância do Bolsa Família.

```
. estat phtest, detail
```

```
Test of proportional-hazards assumption
```

```
Time: Time
```

	rho	chi2	df	Prob>chi2
Idade_2012	0.03792	15.30	1	0.0001
Sexo	0.01835	4.54	1	0.0332
Nao_Branco	-0.01029	1.43	1	0.2313
Bolsa_Fami~a	-0.00477	0.31	1	0.5795
CFC	0.00319	0.14	1	0.7115
Biparental	-0.00190	0.05	1	0.8248
Biparental_t	0.00302	0.12	1	0.7258
Max_Esc_Pais	0.00499	0.34	1	0.5599
PB_LPL_1e2	-0.01175	1.72	1	0.1903
PB_LPL_1e2_t	0.01592	3.04	1	0.0812
PB_MAT_1e2	0.00868	0.97	1	0.3240
Mov_Turma	0.01800	4.47	1	0.0344
global test		30.85	12	0.0021

- ➔ A variável Idade_2012 ainda está com problema na suposição de proporcionalidade, optou-se por verificar a interação dela com o tempo, o que não resolveu, optou-se

por testar a combinação com outras variáveis como Defasados e Anos_Defasagem, somente a segunda variável corrigiu.

```
. estat phtest, detail
```

```
Test of proportional-hazards assumption
```

```
Time: Time
```

	rho	chi2	df	Prob>chi2
Idade_2012	-0.01411	2.73	1	0.0982
Anos_Defasagem	0.01246	2.07	1	0.1503
Sexo	0.01786	4.30	1	0.0382
Nao_Branco	-0.00982	1.30	1	0.2536
Bolsa_Familia	-0.00381	0.20	1	0.6581
CFC	0.00330	0.15	1	0.7022
Biparental	-0.00238	0.08	1	0.7821
Biparental_t	0.00234	0.07	1	0.7857
Max_Esc_Pais	0.00357	0.17	1	0.6771
PB_LPL_1e2	-0.01638	3.34	1	0.0676
PB_LPL_1e2_t	0.01475	2.62	1	0.1052
PB_MAT_1e2	0.00848	0.93	1	0.3354
Mov_Turma	0.01489	3.06	1	0.0803
global test		15.65	13	0.2687

→ Com essa alteração, o sexo apresentou problemas na suposição de proporcionalidade, trouxe novamente a interação do sexo com o tempo.

Conclusão: O modelo final para este bloco ficou sendo o Modelo 4.5

Modelo	Covariável	Log	TRV	valor p
Mod. 4.1	Idade_2012; Sexo; Sexo*t; NãoBranco; Bolsa Família; Biparental; Biparental*t; Máx_Escolaridade_Pais; Máx_Escolaridade_Pais*t; PB_LPL_1e2; PB_MAT_1e2; Mov_Turma; PB_LPL_1*t;	-136511,33	-	-
Mod. 4.2	Idade_2012; Sexo; NãoBranco; Bolsa Família; Biparental; Biparental*t; Máx_Escolaridade_Pais; PB_LPL_1e2; PB_MAT_1e2; Mov_Turma; PB_LPL_1*t;	-136513,08	3,50	0,000
Mod. 4.3	Idade_2012; Sexo; NãoBranco; Bolsa Família; CFC; Biparental; Biparental*t; Máx_Escolaridade_Pais; PB_LPL_1e2; PB_MAT_1e2; Mov_Turma; PB_LPL_1*t;	-136509,5	-3,66	0,000
Mod. 4.4	Idade_2012; Anos_Defasagem; Sexo; NãoBranco; Bolsa Família; CFC; Biparental; Biparental*t; Máx_Escolaridade_Pais; PB_LPL_1e2; PB_MAT_1e2; Mov_Turma; PB_LPL_1*t;	-136390,89	-240,88	0,000
Mod. 4.5	Idade_2012; Anos_Defasagem; Sexo; Sexo*t; NãoBranco; Bolsa Família; CFC; Biparental; Biparental*t; Máx_Escolaridade_Pais; PB_LPL_1e2; PB_MAT_1e2; Mov_Turma; PB_LPL_1*t;	-136388,87	-244,92	0,000

```
. *** MODELO 4 FINAL

. stcox Idade_2012 Anos_Defasagem Sexo sexo_t Nao_Branco Bolsa_Familia CFC Biparental Biparen
> tal_t Max_Esc_Pais PB_LPL_1e2 PB_LPL_1e2_t PB_MAT_1e2 Mov_Turma
```

```
failure _d: EventoAno
analysis time _t: Tempo
id: ID
```

```
Iteration 0: log likelihood = -141369.44
Iteration 1: log likelihood = -136505.4
Iteration 2: log likelihood = -136389.56
Iteration 3: log likelihood = -136388.87
Iteration 4: log likelihood = -136388.87
Refining estimates:
Iteration 0: log likelihood = -136388.87
```

Cox regression -- Breslow method for ties

```
No. of subjects = 47229 Number of obs = 133773
No. of failures = 13459
Time at risk = 133773
LR chi2(14) = 9961.14
Log likelihood = -136388.87 Prob > chi2 = 0.0000
```

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
Idade_2012	1.386497	.0247635	18.30	0.000	1.338801	1.435892
Anos_Defasagem	.7579305	.0136131	-15.43	0.000	.7317134	.7850869
Sexo	.704538	.0281716	-8.76	0.000	.6514309	.7619746
sexo_t	1.036343	.0183976	2.01	0.044	1.000905	1.073037
Nao_Branco	1.041682	.0195201	2.18	0.029	1.004117	1.080652
Bolsa_Familia	1.051431	.0242102	2.18	0.029	1.005034	1.099969
CFC	.9330294	.023649	-2.73	0.006	.8878108	.9805512
Biparental	.9193236	.0356055	-2.17	0.030	.8521211	.991826
Biparental_t	.9479715	.0164807	-3.07	0.002	.916214	.9808297
Max_Esc_Pais	.854086	.0168497	-7.99	0.000	.8216915	.8877576
PB_LPL_1e2	.844355	.0089109	-16.03	0.000	.8270692	.862002
PB_LPL_1e2_t	.9636937	.0049113	-7.26	0.000	.9541155	.9733679
PB_MAT_1e2	.7636853	.0050145	-41.06	0.000	.75392	.7735771
Mov_Turma	1.383233	.0475866	9.43	0.000	1.29304	1.479717

_t	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
Idade_2012	.3267806	.0178605	18.30	0.000	.2917747	.3617865
Anos_Defasagem	-.2771636	.0179609	-15.43	0.000	-.3123663	-.2419609
Sexo	-.350213	.0399859	-8.76	0.000	-.428584	-.2718421
sexo_t	.0356987	.0177524	2.01	0.044	.0009046	.0704927
Nao_Branco	.0408363	.0187391	2.18	0.029	.0041084	.0775642
Bolsa_Familia	.0501519	.023026	2.18	0.029	.0050219	.095282
CFC	-.0693185	.0253464	-2.73	0.006	-.1189966	-.0196404
Biparental	-.0841171	.0387301	-2.17	0.030	-.1600266	-.0082076
Biparental_t	-.0534309	.0173852	-3.07	0.002	-.0875053	-.0193565
Max_Esc_Pais	-.1577234	.0197284	-7.99	0.000	-.1963903	-.1190565
PB_LPL_1e2	-.1691823	.0105536	-16.03	0.000	-.1898669	-.1484977
PB_LPL_1e2_t	-.0369818	.0050964	-7.26	0.000	-.0469705	-.0269931
PB_MAT_1e2	-.2695995	.0065662	-41.06	0.000	-.2824691	-.25673
Mov_Turma	.3244235	.0344024	9.43	0.000	.256996	.3918511

Teste de resíduos de Schoenfeld

```
. estat phtest, detail
```

```
Test of proportional-hazards assumption
```

```
Time: Time
```

	rho	chi2	df	Prob>chi2
Idade_2012	-0.01419	2.77	1	0.0963
Anos_Defasagem	0.01248	2.08	1	0.1494
Sexo	0.00830	0.93	1	0.3345
sexo_t	-0.00901	1.09	1	0.2954
Nao_Branco	-0.00983	1.31	1	0.2532
Bolsa_Familia	-0.00381	0.20	1	0.6580
CFC	0.00328	0.14	1	0.7034
Biparental	-0.00240	0.08	1	0.7796
Biparental_t	0.00221	0.07	1	0.7970
Max_Esc_Pais	0.00352	0.17	1	0.6814
PB_LPL_1e2	-0.01609	3.23	1	0.0722
PB_LPL_1e2_t	0.01558	2.94	1	0.0866
PB_MAT_1e2	0.00845	0.92	1	0.3372
Mov_Turma	0.01495	3.08	1	0.0792
global test		12.61	14	0.5572

Modelo 5: Modelo 4 + Modelo 3

Passo 1: Para começar o modelo 5, optou-se em unir as versões finais do Modelo 4 e 3.

➔ O modelo se ajustou perfeitamente, não sendo necessária nenhuma ação sob ele.

```

. *** Modelo 5 FINAL

. stcox Idade_2012 Anos_Defasagem Sexo sexo_t Nao_Branco Bolsa_Familia CFC Biparental Biparen
> tal_t Max_Esc_Pais PB_LPL_1e2 PB_LPL_1e2_t PB_MAT_1e2 Mov_Turma GC GC_t ICG ICG_t

      failure _d:  EventoAno
      analysis time _t:  Tempo
                id:  ID

Iteration 0:  log likelihood = -141369.44
Iteration 1:  log likelihood = -136476.43
Iteration 2:  log likelihood = -136341.13
Iteration 3:  log likelihood = -136339.19
Iteration 4:  log likelihood = -136339.18
Refining estimates:
Iteration 0:  log likelihood = -136339.18

Cox regression -- Breslow method for ties

No. of subjects =          47229                Number of obs   =       133773
No. of failures =          13459
Time at risk    =          133773

LR chi2(18)      =    10060.51
Prob > chi2     =         0.0000

Log likelihood   =    -136339.18

```

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
Idade_2012	1.385205	.0247381	18.25	0.000	1.337558 1.434549
Anos_Defasagem	.7584595	.0136228	-15.39	0.000	.7322238 .7856353
Sexo	.7054351	.0282044	-8.73	0.000	.652266 .7629383
sexo_t	1.035711	.0183846	1.98	0.048	1.000297 1.072378
Nao_Branco	1.04384	.0195621	2.29	0.022	1.006195 1.082894
Bolsa_Familia	1.052327	.0242347	2.21	0.027	1.005884 1.100914
CFC	.9334488	.0236638	-2.72	0.007	.8882021 .9810005
Biparental	.9177291	.0355519	-2.22	0.027	.8506282 .9901232
Biparental_t	.947493	.0164776	-3.10	0.002	.9157418 .9803451
Max_Esc_Pais	.8537572	.0168421	-8.01	0.000	.8213773 .8874137
PB_LPL_1e2	.8430146	.008916	-16.15	0.000	.8257194 .8606721
PB_LPL_1e2_t	.9646723	.0049346	-7.03	0.000	.9550489 .9743926
PB_MAT_1e2	.7663098	.0050486	-40.40	0.000	.7564784 .7762691
Mov_Turma	1.390621	.0478534	9.58	0.000	1.299923 1.487648
GC	.0969036	.0306203	-7.39	0.000	.052164 .1800151
GC_t	1.926076	.1971468	6.40	0.000	1.575967 2.353962
ICG	.9478936	.0189239	-2.68	0.007	.9115197 .985719
ICG_t	1.029916	.0096456	3.15	0.002	1.011184 1.048996

_t	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
Idade_2012	.3258482	.0178588	18.25	0.000	.2908457 .3608508
Anos_Defasagem	-.2764658	.0179612	-15.39	0.000	-.3116691 -.2412625
Sexo	-.3489405	.0399815	-8.73	0.000	-.4273029 -.2705782
sexo_t	.0350878	.0177507	1.98	0.048	.0002969 .0698786
Nao_Branco	.0429062	.0187405	2.29	0.022	.0061754 .0796369
Bolsa_Familia	.0510039	.0230296	2.21	0.027	.0058668 .0961411
CFC	-.0688691	.0253509	-2.72	0.007	-.118556 -.0191823
Biparental	-.085853	.038739	-2.22	0.027	-.1617801 -.0099259
Biparental_t	-.0539357	.0173907	-3.10	0.002	-.0880209 -.0198506
Max_Esc_Pais	-.1581084	.0197271	-8.01	0.000	-.1967727 -.119444
PB_LPL_1e2	-.170771	.0105764	-16.15	0.000	-.1915003 -.1500417
PB_LPL_1e2_t	-.0359668	.0051153	-7.03	0.000	-.0459927 -.0259409
PB_MAT_1e2	-.2661687	.0065882	-40.40	0.000	-.2790813 -.2532561
Mov_Turma	.3297506	.0344116	9.58	0.000	.2623052 .397196
GC	-2.334038	.3159872	-7.39	0.000	-2.953362 -1.714715
GC_t	.6554845	.1023567	6.40	0.000	.454869 .8561
ICG	-.053513	.0199642	-2.68	0.007	-.0926421 -.0143839
ICG_t	.0294774	.0093654	3.15	0.002	.0111215 .0478333

Teste de resíduos de Schoenfeld

```
. estat phtest, detail
```

```
Test of proportional-hazards assumption
```

```
Time: Time
```

	rho	chi2	df	Prob>chi2
Idade_2012	-0.01420	2.77	1	0.0960
Anos_Defasagem	0.01230	2.02	1	0.1554
Sexo	0.00838	0.95	1	0.3300
sexo_t	-0.00907	1.11	1	0.2925
Nao_Branco	-0.00981	1.30	1	0.2543
Bolsa_Familia	-0.00420	0.24	1	0.6255
CFC	0.00390	0.21	1	0.6506
Biparental	-0.00372	0.19	1	0.6653
Biparental_t	0.00364	0.18	1	0.6722
Max_Esc_Pais	0.00336	0.15	1	0.6947
PB_LPL_1e2	-0.01336	2.23	1	0.1353
PB_LPL_1e2_t	0.01177	1.68	1	0.1952
PB_MAT_1e2	0.00960	1.20	1	0.2737
Mov_Turma	0.01589	3.48	1	0.0620
GC	-0.00453	0.29	1	0.5892
GC_t	0.00424	0.26	1	0.6121
ICG	0.00104	0.01	1	0.9071
ICG_t	-0.00102	0.01	1	0.9078
global test		12.62	18	0.8139