

ESCOLA DE NEGÓCIOS
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA
DOUTORADO EM ECONOMIA DO DESENVOLVIMENTO

RENATA PEREIRA CARDOSO

**ENSAIOS EM ECONOMIA: EVIDÊNCIAS SOBRE ECONOMIA DA EDUCAÇÃO
E SOBRE A CORRUPÇÃO NO BRASIL.**

Porto Alegre
2023

PÓS-GRADUAÇÃO - *STRICTO SENSU*



Pontifícia Universidade Católica
do Rio Grande do Sul

RENATA PEREIRA CARDOSO

ENSAIOS EM ECONOMIA: EVIDÊNCIAS SOBRE ECONOMIA DA EDUCAÇÃO E
SOBRE A CORRUPÇÃO NO BRASIL.

Tese apresentada ao Programa de Pós- Graduação em Economia do Desenvolvimento, da Pontifícia Universidade Católica do Rio Grande do Sul (PUC-RS), como requisito parcial para obtenção do título de Doutora em Economia.

Área de concentração:
Desenvolvimento Econômico

Orientador: Prof. Dr. Marco Túlio
Aniceto França

Porto Alegre

2023

Ficha Catalográfica

C268e Cardoso, Renata Pereira

Ensaaios em Economia : Evidências Sobre Economia da Educação e Sobre a Corrupção no Brasil / Renata Pereira Cardoso. – 2023.

130.

Tese (Doutorado) – Programa de Pós-Graduação em Economia do Desenvolvimento, PUCRS.

Orientador: Prof. Dr. Marco Túlio Aniceto França.

1. Evasão Escolar. 2. Desigualdade de Gênero. 3. Educação de Jovens e Adultos. 4. Corrupção. I. França, Marco Túlio Aniceto.
II. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da PUCRS
com os dados fornecidos pelo(a) autor(a).

Bibliotecária responsável: Clarissa Jesinska Selbach CRB-10/2051

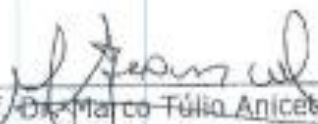
Renata Pereira Cardoso

"ENSAIOS EM ECONOMIA: EVIDÊNCIAS SOBRE ECONOMIA DA EDUCAÇÃO E SOBRE A CORRUPÇÃO NO BRASIL"

Tese apresentada como requisito parcial para a obtenção do grau de Doutor em Economia, pelo Programa de Pós-Graduação em Economia da Escola de Negócios da Pontifícia Universidade Católica do Rio Grande do Sul.

Aprovado em 27 de outubro de 2023, pela Banca Examinadora.

BANCA EXAMINADORA:


Prof. Dr. Marco Túlio Aniceto França
Orientador e Presidente da sessão


Prof. Dr. Adelar Fochezatto


Prof. Dr. Guilherme Diniz Irffl


Prof.ª Dr.ª Kalinca Leia Becker

Dedico esta tese à minha filha Laura Cardoso Costa, à minha mãe Janete Machado Pereira Bandeira e ao meu amor, Valdinei Henrique da Costa.

AGRADECIMENTOS

A conclusão desta etapa da minha trajetória acadêmica, que se aproxima com a defesa da tese, representa a chegada de uma caminhada com muitos desafios. É impossível não passar um “filme” em minha mente quando me lembro do início do curso de Doutorado, em 2019, com aulas presenciais, pesquisas com os colegas no Núcleo de Pesquisas em Economia (NEPE), encontros, reuniões, cafés, com muito apoio e parceria dos professores e colegas presencialmente. Após, passamos para o novo normal no formato “*online*”, enfrentamos uma pandemia, e junto dela, muitas dores. Nesta trajetória também me tornei mãe de uma menina linda, inteligente, iluminada, que enche nossos dias de alegria. Enfim, uma vida acontece enquanto faz-se um curso de doutoramento.

Passada esta breve introdução, primeiramente agradeço aos meus pais, Janete e Júlio, que me deram a vida e me educaram com muito amor e carinho. Com eles aprendi sobre amor, amizade e a importância da família. Eles me ensinaram sobre respeito e me passaram valores que foram essenciais em cada conquista que tive. À minha mãe, em especial, reitero os agradecimentos por todo apoio e amor e por ter sido fortaleza em nossos momentos de dificuldade.

Agradeço ao meu grande amor, meu companheiro, Henrique, que me apoia incondicionalmente, que esteve comigo desde a inscrição para a seleção do programa de doutorado e que sempre está disposto a me ajudar, como todo amor, carinho, amizade e sabedoria, com um jeitinho que só ele tem. Agradeço a compreensão por todo o tempo que deixamos de estar juntos em prol dos estudos e por todas as vezes que ele tentou suprir minha ausência com nossa filha.

À minha filha Laura, que me fez conhecer o amor incondicional. A ela que, com seus dois aninhos, ainda nem sabe sobre o mundo, nem sobre as coisas da vida, mas que me motiva a ser uma pessoa melhor todos os dias. Ela que me ensina tanto e me mostra que tão pouco sei. Com ela aprendi o quanto o amor pode ressignificar nossa vida, nossa família e nossos sonhos. Aprendi a ver alegria nos detalhes. Laura já é uma pessoa incrível, que me enche de orgulho a cada dia. Agradeço ainda a ela que, mesmo dormindo, foi minha companheira em muitos momentos de escrita.

Ao meu amado amigo e irmão Fellipe e ao meu padrasto José Bandeira e aos demais familiares que estiveram junto comigo em muitos momentos desta trajetória. Ao meu sogro Valdemar em memória, que me incentivava muito a estudar, com palavras e pequenos gestos. Aos meus amigos e amigas, que estiveram ao meu lado para me ouvir e apoiar.

Agradeço em especial ao meu orientador Prof. Dr. Marco Túlio Aniceto França, pela paciência e por sempre ter me apoiado e me compreendido nos momentos em que mais precisei. Agradeço muito pelos ensinamentos, pelos questionamentos certos em cada pesquisa, pela amizade e parceria que construímos ao longo dos anos de Doutorado. Admiro muito seu posicionamento profissional, sua inteligência, seu conhecimento e sua empatia. Foi um privilégio ter a oportunidade de aprender sendo sua aluna e sua orientanda.

Aos demais professores, colegas e funcionários do Programa de Pós-Graduação em Economia do Desenvolvimento da Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), em especial à funcionária Renata Santana, sempre atendendo nossas dúvidas com muita presteza. Aos colegas do Laboratório de Análises em Economia Social (LAES), por todas as contribuições recebidas em pesquisas, pela amizade e pelas trocas de conhecimento.

Por fim, à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela ajuda financeira, fundamental para a minha jornada.

RESUMO

Esta tese tem por objetivo analisar empiricamente três problemas sociais de interesse público, os dois primeiros referem-se à educação e o terceiro à corrupção no Brasil. Para tal propósito três ensaios são aqui propostos. No primeiro ensaio, avaliamos os fatores associados à evasão escolar em diferentes etapas de ensino, sob a hipótese de que meninos e meninas evadem por motivos distintos ao longo da trajetória escolar. Nossos resultados indicam que a evasão escolar para trabalhar acomete tanto os meninos quanto as meninas, sendo mais intensiva para eles e agravando-se com o passar dos anos escolares. Já para as meninas, além do trabalho, também é fator decisivo para a evasão o casamento e os filhos, fatores estes pouco relevantes para os meninos. No segundo ensaio, avaliamos empiricamente o efeito de cursar Educação de Jovens e Adultos (EJA) sobre a probabilidade de aprovação no Exame Nacional para Certificação de Competências de Jovens e Adultos (Encceja). Para a certificação em Nível Médio observou-se menor probabilidade de aprovação no Encceja dos indivíduos que cursam EJA, em todas as disciplinas sendo o efeito negativo de maior magnitude para Língua e Redação, e Matemática. Para certificação em Nível Fundamental identificou-se pior desempenho nas disciplinas de Língua e Redação, e Ciências da Natureza. Em geral os resultados indicam desempenho inferior do grupo participante da política de EJA. Por último, no terceiro ensaio, fazemos uma análise sobre a corrupção no Brasil, a partir da estimação de um Índice de Microrregional de Corrupção (IMRC), para 220 microrregiões do Brasil. Adicionalmente, avaliamos tendências espaciais da corrupção estimada. Os resultados apontam a formação de clusters espaciais da corrupção do tipo alta incidência (*high-high*) para as regiões dos estados do Norte, e baixa incidência (*low-low*) para os do Sul. Por fim, o trabalho inova ao calcular um índice objetivo com abrangência microrregional. Ademais, indica áreas para a atuação do Estado em possíveis ações de combate à corrupção, bem como disponibiliza o IMRC, que possibilita diversas pesquisas sobre corrupção.

Palavras-chave: Economia da Educação; Evasão Escolar; Educação de Jovens e Adultos; Logit Multinomial; Propensity Score Matching; Corrupção Regional; Análise Espacial.

ABSTRACT

This thesis aims to empirically analyze three social problems of public interest, the first two refer to education and the third to corruption in Brazil. For this purpose, three tests are proposed here. In the first test, we evaluated the factors associated with school dropout at different stages of education, under the hypothesis that boys and girls drop out for different reasons throughout their school career. Our results indicate that school dropout to work affects both boys and girls, being more intense for them and getting worse as the school years go by. For girls, in addition to work, marriage and children are also decisive factors for evasion, factors that are not very relevant for boys. In the second test, we empirically evaluated the effect of studying Youth and Adult Education (EJA) on the probability of passing the National Examination for Certification of Youth and Adult Skills (Enceja). For certification at the Secondary Level, there was a lower probability of approval in Enceja for individuals taking EJA, in all subjects, with the negative effect being of greater magnitude for Languages and Writing, and Mathematics. For certification at Fundamental Level, worse performance was identified in the subjects of Languages and Writing, and Natural Sciences. In general, the results indicate lower performance of the group participating in the EJA policy. Finally, in the third essay, we analyze corruption in Brazil, based on the estimation of a Microregional Corruption Index (IMRC), for 220 microregions in Brazil. Additionally, we evaluate spatial trends in estimated corruption. The results point to the formation of spatial clusters of corruption of the high-incidence type (high-high) for the regions of the Northern states, and low incidence (low-low) for those in the South. Finally, the work innovates by calculating an index objective with micro-regional coverage. Furthermore, it indicates areas for State action in possible actions to combat corruption, as well as making available the IMRC, which enables various research on corruption.

Keywords: Economics of Education; School Dropout; Youth and Adult Education; Multinomial Logit; Propensity Score Matching; Regional Corruption; Spatial Analysis.

LISTA DE ILUSTRAÇÕES

SEGUNDO ENSAIO

Apêndice A2 Figura 1 - Distribuição Score de Propensão dos participantes do Encceja, por tipo de certificação e por disciplina, pré e pós-pareamento Certificação Ensino Médio 72

Apêndice A2 Figura 2 - Distribuição Score de Propensão participantes do Encceja, por tipo de certificação e por disciplina, pré e pós-pareamento Certificação Ensino Fundamental 73

TERCEIRO ENSAIO

Figura 1 - Mapa da Transparência 2015 e 2016 105

Figura 2 - Mapas Quantílicos - distribuição espacial da corrupção 110

Figura 3 – Índice de Associação Espacial Local (LISA)..... 114

Gráfico 1 - Número de contas julgadas irregulares registrados no Cadirreg para os anos de 2015 e 2016, por Estados..... 102

Gráfico 2 - Número de processos de corrupção, por Estados, registrados na base do Ministério Público Federal para os anos de 2015 e 2016..... 104

Gráfico 3 - Índice de Moran Global a partir do Índice Microrregional de Corrupção 113

LISTA DE TABELAS

PRIMEIRO ENSAIO

Tabela 1 - Renda em salários-mínimos	27
Tabela 2 - Estatísticas descritivas Idade e Reprovação	27
Tabela 3 - Estatísticas descritivas da variável evasão	28
Tabela 4 - Estatísticas descritivas das variáveis trabalho e casamento e filhos.	29
Tabela 5 - Falta de interesse e oferta de escola	30
Tabela 6 - Escolaridade da mãe e do pai	31
Tabela 7 - Resultados do modelo Logit Multinomial.....	31
Apêndice A 1 - Tabela 1 - Estimativas com os controles referentes aos Estados da Federação.	41

SEGUNDO ENSAIO

Tabela 1 - Estatísticas descritivas da amostra	57
Tabela 2 - Distribuição geográfica dos participantes do exame.....	59
Tabela 3 - Estatísticas descritivas de aprovação por disciplina.....	60
Tabela 4 - Certificação Ensino Médio - Efeito médio de ter cursado EJA sobre a aprovação nas provas de Linguagens e Redação, Matemática, Ciências da Natureza e Ciências Humanas (ATT).....	62
Tabela 5 - Certificação Ensino Fundamental - Efeito médio de ter cursado EJA sobre a aprovação nas provas de Linguagens e Redação, Matemática, Ciências da Natureza e Ciências Humanas (ATT).....	64
Apêndice A2 - Tabela 1 Diferença de médias, antes e depois do pareamento, entre os grupos de tratamento e controle - Linguagens e Redação (Ensino Médio).....	74
Apêndice A2 - Tabela 2 Diferença de médias, antes e depois do pareamento, entre os grupos de tratamento e controle - Ciências Naturais (Ensino Médio).....	75
Apêndice A2 - Tabela 3 Diferença de médias, antes e depois do pareamento, entre os grupos de tratamento e controle - Ciências Humanas (Ensino Médio).....	76

Apêndice A2 - Tabela 4 Diferença de médias, antes e depois do pareamento, entre os grupos de tratamento e controle – Matemática (Ensino Médio)	77
Apêndice A2 - Tabela 5 Diferença de médias, antes e depois do pareamento, entre os grupos de tratamento e controle – Linguagens e Redação (Ensino Fundamental).....	78
Apêndice A2 - Tabela 6 Diferença de médias, antes e depois do pareamento, entre os grupos de tratamento e controle – Ciências Naturais (Ensino Fundamental)	79
Apêndice A2 - Tabela 7 Diferença de médias, antes e depois do pareamento, entre os grupos de tratamento e controle – Ciências Humanas (Ensino Fundamental).....	80
Apêndice A2 - Tabela 8 Diferença de médias, antes e depois do pareamento, entre os grupos de tratamento e controle – Matemática (Ensino Fundamental).....	81
Apêndice A2 - Tabela 9 Resultados estimados para a probabilidade de receber o tratamento - Probit (Linguagens e redação - Ensino Médio).....	82
Apêndice A2 - Tabela 10 Resultados estimados para a probabilidade de receber o tratamento - Probit (Ciências da Natureza - Ensino Médio).....	83
Apêndice A2 - Tabela 11 Resultados estimados para a probabilidade de receber o tratamento - Probit (Ciências Humanas - Ensino Médio).....	84
Apêndice A2 - Tabela 12 Resultados estimados para a probabilidade de receber o tratamento - Probit (Matemática - Ensino Médio).....	85
Apêndice A2 - Tabela 13 Resultados estimados para a probabilidade de receber o tratamento - Probit (Linguagens e Redação - Ensino Fundamental)	86
Apêndice A2 - Tabela 14 Resultados estimados para a probabilidade de receber o tratamento - Probit (Ciências da Natureza - Ensino Fundamental)	87
Apêndice A2 - Tabela 15 Resultados estimados para a probabilidade de receber o tratamento - Probit (Ciências Humanas - Ensino Fundamental).....	88
Apêndice A2 - Tabela 16 Resultados estimados para a probabilidade de receber o tratamento - Probit (Ciências Humanas - Matemática).....	89

TERCEIRO ENSAIO

Tabela 1 - Índice de Corrupção por estados brasileiros.....	110
Tabela 2 - Autocorrelação Global	112

SUMÁRIO

1. APRESENTAÇÃO	13
2. EVASÃO ESCOLAR: AS MENINAS EVADEM PARA CUIDAR DA FAMÍLIA E OS MENINOS PARA TRABALHAR? UMA ANÁLISE POR ETAPA DE ENSINO.....	16
2.1 INTRODUÇÃO.....	18
2.2 FATORES ASSOCIADOS À EVASÃO ESCOLAR.....	20
2.3 QUAIS AS DIFERENÇAS ENTRE OS MOTIVOS DA EVASÃO ESCOLAR DE MENINAS E MENINOS?.....	23
2.4 ESTRATÉGIA EMPÍRICA	25
2.5 ANÁLISE EXPLORATÓRIA DOS DADOS.....	26
2.6 RESULTADOS	31
2.7 CONSIDERAÇÕES FINAIS	35
REFERÊNCIAS	37
3. AVALIAÇÃO DE POLÍTICAS PÚBLICAS: UMA ANÁLISE DO DESEMPENHO DOS ESTUDANTES DA EJA NO EXAME NACIONAL PARA CERTIFICAÇÃO DE COMPETÊNCIAS DE JOVENS E ADULTOS.	42
3.1 INTRODUÇÃO.....	44
3.2 CONTEXTO HISTÓRICO DA EDUCAÇÃO DE JOVENS E ADULTOS NO BRASIL 46	
3.3 A AVALIAÇÃO DA EDUCAÇÃO DE JOVENS E ADULTOS	47
3.4 ESTRATÉGIA EMPÍRICA	50
3.4.1 PROPENSITY SCORE MATCHING (PSM).....	50
3.4.2 ANÁLISE DE SENSIBILIDADE.....	52
3.4.3 COARSENEDED EXACT MATCHING (CEM).....	53
3.4.4 ENTROPY BALANCING (EB)	54
3.5 ANÁLISE EXPLORATÓRIA DOS DADOS	55
3.6 RESULTADOS E DISCUSSÃO	61
3.6.1 RESULTADOS DOS TESTES DE LIMITES DE ROSENBAUM	66
3.7 CONSIDERAÇÕES FINAIS	67

REFERÊNCIAS	69
4. ONDE ESTÁ A CORRUPÇÃO NO BRASIL? UMA ANÁLISE SOB A ÓTICA DO ÍNDICE MICRORREGIONAL DE CORRUPÇÃO.....	90
4.1 INTRODUÇÃO.....	92
4.2 MENSURAÇÃO OBJETIVA DA CORRUPÇÃO.....	94
4.3 COMPORTAMENTO ESPACIAL DA CORRUPÇÃO	97
4.4 BASE DE DADOS.....	100
4.4.1 CADASTRO DE CONTAS JULGADAS IRREGULARES – CADIRREG – DO TRIBUNAL DE CONTAS DA UNIÃO.....	100
4.4.2 REGISTROS DE PROCESSOS JUDICIAIS DE CORRUPÇÃO NAS PROCURADORIAS REGIONAIS DO MINISTÉRIO PÚBLICO FEDERAL (MPF).	103
4.4.3 ÍNDICE DE TRANSPARÊNCIA MUNICIPAL - MINISTÉRIO PÚBLICO FEDERAL. 105	
4.5 METODOLOGIA EMPÍRICA	106
4.5.1 CÁLCULOS DO ÍNDICE MICRORREGIONAL DE CORRUPÇÃO.....	106
4.5.2 METODOLOGIA PARA ANÁLISE ESPACIAL DA CORRUPÇÃO	107
4.6 RESULTADOS E DISCUSSÕES.....	109
4.6.1 ÍNDICE DE MORAN GLOBAL	112
4.6.2 INDICADOR LOCAL DE ASSOCIAÇÃO ESPACIAL (LISA).....	113
4.7 CONSIDERAÇÕES FINAIS	114
REFERÊNCIAS	117
5. CONSIDERAÇÕES FINAIS	126

1. APRESENTAÇÃO

Nesta pesquisa temos por objetivo analisar empiricamente três problemas sociais. As áreas de análises compreendem questões de interesse público, tanto por dois ensaios aqui apresentados estarem relacionados à educação, quanto pelo terceiro relacionar-se ao tema da corrupção. Em geral nosso objetivo é produzir evidências para contribuir com a literatura e com gestores públicos, tendo por finalidade qualificar programas de educação, além de produzir evidência sobre a corrupção, com indícios sobre sua localização e seu comportamento espacial no Brasil.

Quanto ao primeiro tema, a educação no Brasil necessita de políticas que tenham por objetivo sua qualificação. Dentre os problemas relacionados à educação, destaca-se a baixa qualidade escolar, demonstrada pelo baixo desempenho dos estudantes brasileiros em exames de larga escala, tais como o Sistema de Avaliação da Educação Básica (Saeb). A baixa qualidade, dentre outras consequências tem implicação na evasão estudantil. No Brasil, 10 milhões de pessoas de 14 a 29 anos não completaram alguma das etapas da educação básica (Educação Infantil, o Ensino Fundamental, ou o Ensino Médio), seja por nunca terem frequentado a escola, ou por dela terem evadido, conforme dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) de 2019.

Assim, o primeiro ensaio trata da evasão escolar no Brasil. Este problema traz consequências à saúde mental dos jovens, com implicações no uso de álcool e drogas ilícitas (MASTORCI et al., 2022). Há ainda na literatura, indícios da relação entre a evasão escolar e o ingresso de jovens no mundo do crime (DRAGONE et al., 2021; GERLINGER e HIPPEL, 2020; ANDERSON, 2014).

Sob a hipótese aventada pela literatura de que meninas e meninos evadem por motivos distintos (MOREIRA, 2021, SHAHIDUL e KARIN, 2015), o objetivo proposto na pesquisa é avaliar os fatores associados à evasão escolar de meninos e meninas em diferentes etapas de ensino (1º e 4º ano do ensino fundamental; 5º e 8º ano do ensino fundamental; 1º e 3º ano do ensino médio), ao longo da trajetória escolar. A metodologia empregada é a estimação de um Logit Multinomial e os microdados utilizados são do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) coletado em questionário de inscrição para o Exame Nacional para Certificação de Competências de Jovens e Adultos (Encceja), no ano de 2019, portanto, a análise refere-se a este público em específico.

Para tratar do problema da evasão escolar e da ausência à escola, existe no Brasil a política pública de Educação de Jovens e Adultos (EJA), que tem como público-alvo aqueles que não acessaram ou não continuaram os estudos no ensino fundamental ou no ensino médio

regular em idade apropriada. A EJA se diferencia do ensino regular, pois possibilita a conclusão das etapas de ensino em períodos menores que o ensino padrão.

Nesse contexto, a partir dos Microdados do Encceja, para o ano de 2019, nosso objetivo de pesquisa no segundo ensaio é avaliar empiricamente o efeito de cursar Educação de Jovens e Adultos (EJA) sobre a aprovação no Exame Nacional para Certificação de Competências de Jovens e Adultos (Encceja). O Encceja é um exame de larga escala, aplicado para mensurar o conhecimento e as habilidades de estudantes que não concluíram o ensino fundamental ou médio em idade apropriada. Caso o indivíduo aprove no exame, recebe certificação para o nível de escolaridade almejado, qual seja fundamental ou médio.

Há evidência na literatura de que os estudantes da EJA possuem desempenho inferior que os estudantes do ensino regular tanto no Enem (DA SILVA ET AL. 2020; SILVA 2022), quanto no mercado de trabalho (OLIVA ET AL. 2014). Os métodos utilizados na pesquisa são típicos de avaliação de políticas públicas e contemplam o *Propensity Score Matching*, o *Coarsened Exact Matching*, e o *Entropy Balancing*, além da combinação entre os métodos. Nosso grupo de comparação com os estudantes da EJA são indivíduos que estão evadidos do sistema escolar e que não estavam estudando no momento da realização do Encceja.

Adicionalmente, no terceiro ensaio, considerando o problema público da corrupção, apresentamos uma proposta de indicador de corrupção para o Brasil. A literatura brasileira nos últimos anos vem buscando estimar um indicador de corrupção com dados objetivos (BOLL 2010; CARRARO et al. 2015; SOUSA 2018; COSTA et al. 2020; GARCIA 2022). A partir dessas pesquisas, propusemos um indicador de corrupção de abrangência inédita, para 220 microrregiões brasileiras, o que até então só havia sido estimado por outros pesquisadores em nível dos 27 estados da federação.

Os dados utilizados são do Cadastro de contas julgadas irregulares – do Tribunal de Contas da União; Processos judiciais de corrupção nas Procuradorias Regionais do Ministério Público Federal (MPF); Índice de transparência Municipal – do Ministério Público Federal. A metodologia adotada foi a de *Rescaling* e os anos de análise são 2015 e 2016, em decorrência da disponibilidade dos dados. Após a estimação do índice, avaliamos a existência de correlação espacial da corrupção a partir dos Índices de Moran Local e Global.

Como contribuição de pesquisa nossos principais resultados sobre o primeiro ensaio indicam que meninos e meninas são acometidos pela evasão escolar de forma distinta, sendo elas mais afetadas por questões relacionadas aos cuidados domésticos, enquanto os meninos preponderantemente evadem por questões do mercado de trabalho. Assim, em termos de

políticas públicas, indica-se que a evasão seja tratada de maneira distinta entre os gêneros e entre as etapas de ensino escolar.

Em relação à avaliação sobre a EJA, identifica-se que a aprovação desses estudantes no Encceja é inferior ao de indivíduos que não estão estudando, com efeito negativo de maior magnitude nas disciplinas de Linguagens e Redação, e matemática. Tal resultado indica a importância de repensar o desenho desta política, considerando as peculiaridades do seu público-alvo, da forma que o curso é ofertado e as suas necessidades de aprendizado.

Por fim, nosso indicador de corrupção aponta que é importante estudar a corrupção de forma desagregada, considerando as diferentes regiões de incidência e a interação entre essas regiões. A partir dos resultados encontrados, espera-se que novas pesquisas sobre os impactos da corrupção no Brasil possam ser realizadas, além de um indicativo aos órgãos de controle e fiscalização sobre como ocorre o comportamento espacial deste fenômeno que corrói o erário público brasileiro.

Assim, com os três ensaios, buscamos contribuir com a agenda de pesquisas relacionadas a políticas públicas e suas avaliações. Assim, buscamos fornecer evidências para pesquisas futuras, assim como para que gestores públicos e órgãos de controle possam atuar no sentido de promover o bem-estar da população.

2. EVASÃO ESCOLAR: AS MENINAS EVADEM PARA CUIDAR DA FAMÍLIA E OS MENINOS PARA TRABALHAR? UMA ANÁLISE POR ETAPA DE ENSINO.

Resumo

O objetivo desta pesquisa é avaliar os fatores associados à evasão escolar nas diferentes etapas de ensino, sob a hipótese de que meninos e meninas evadem por motivos distintos ao longo da trajetória escolar. A base de dados utilizada foi a do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) coletado em questionário de inscrição para o Exame Nacional para Certificação de Competências de Jovens e Adultos (ENCCEJA). O ano de análise é o de 2019, que contém informações de 2,9 milhões de participantes. O modelo econométrico adotado foi um Logit Multinomial, estimado para identificar as razões de risco de evasão em três momentos distintos, quais sejam: do 1º ao 4º ano do ensino fundamental, do 5º ao 8º ano do ensino fundamental e do 1º ao 3º ano do ensino médio. Os principais resultados confirmam a hipótese de que os fatores associados à evasão variam por gênero e ao longo da trajetória escolar. Identificamos que a evasão escolar para trabalhar acomete tanto os meninos que as meninas, sendo mais intensiva para eles e agravando-se com o passar dos anos escolares. Já para as meninas, além do trabalho, também é fator decisivo para a evasão o casamento e os filhos, fatores estes pouco relevantes para os meninos.

Palavras-chave: Evasão Escolar; Desigualdade de Gênero; Logit Multinomial.

Abstract

The objective of this research is to evaluate the factors associated with school dropout at different stages of education, under the assumption that boys and girls drop out for different reasons throughout their school trajectory. The database used was that of the National Institute of Educational Studies and Research Anísio Teixeira (INEP) collected in a registration questionnaire for the National Examination for the Certification of Youth and Adult Skills (ENCCEJA). The year of analysis is 2019, which contains information on 2.9 million participants. The econometric model adopted was a Multinomial Logit, estimated to identify the dropout risk ratios at three different times, namely: from the 1st to the 4th year of elementary school, from the 5th to the 8th year of elementary school and from the 1st to the

3rd year of elementary school. high school. The main results confirm the hypothesis that the factors associated with dropout vary by gender and along the school path. We identified that dropping out of school to work affects both boys and girls, being more intensive for them and getting worse as the school years go by. For girls, in addition to work, marriage and children are also decisive factors for dropout, factors that are not very important for boys.

Keywords: School Dropout; Gender Inequality; Multinomial Logit.

JEL: C25; I21; I24.

2.1 INTRODUÇÃO

No Brasil, a partir da Constituição de 1988, a educação passou a ser um direito social que deve ser prestado pelo Estado a todos os cidadãos. Os princípios constitucionais relacionados à oferta de educação são, dentre outros: a gratuidade do ensino público em estabelecimentos oficiais; a garantia de padrão de qualidade; e a garantia do direito à educação e à aprendizagem ao longo da vida (BRASIL, 1988).

Há também dois importantes instrumentos jurídicos que trazem diretrizes à Educação brasileira: a Lei de Diretrizes e Bases da Educação Nacional (LDB), que define e regulariza a organização da educação brasileira com base nos princípios constitucionais; e o Plano Nacional de Educação (PNE), de duração decenal, que estabelece diretrizes, metas e estratégias para a política educacional.

O PNE 2014-2024 possui 20 metas relacionadas à oferta, à qualidade e à ampliação do ensino e escolarização da sociedade brasileira. Entre essas metas, destaca-se a de 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 de 2024, erradicar o analfabetismo absoluto e reduzir em 50% (cinquenta por cento) a taxa de analfabetismo funcional.

Embora exista diretrizes e normativas legais em favor da educação, a evasão escolar segue sendo um grave problema público no Brasil. Dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) de 2019 indicam que 20% (10 milhões) das pessoas de 14 a 29 anos não completaram alguma das etapas da educação básica (Educação Infantil, o Ensino Fundamental, ou o Ensino Médio), seja por terem abandonado a escola, seja por nunca a terem frequentado.

O abandono escolar traz consequências à saúde dos jovens. Ainda que de difícil identificação entre causas e consequências, a literatura aponta forte relação entre o abandono escolar e problemas de saúde mental e a comportamentos de risco, tais como o uso de bebidas alcoólicas, nicotina, maconha, cocaína e outras substâncias ilícitas (MASTORCI et al., 2022). Ademais, há evidências da relação entre o abandono escolar e o ingresso de jovens na criminalidade (DRAGONE et al., 2021; GERLINGER e HIPPEL, 2020; ANDERSON, 2014).

Além das consequências à saúde, a ausência dos jovens na escola implica em prejuízos econômicos e sociais. De Barros (2021) estimou que a evasão escolar tem impacto na vida de mais de 550 mil jovens brasileiros a cada ano, tendo um custo financeiro de R\$ 220 bilhões, referente às coortes que não concluem a educação básica. Este valor representa 3,3% do PIB Nacional.

A literatura aponta que a evasão de meninas e meninos ocorre por motivos distintos. Os fatores associados à evasão escolar entre meninas e meninos divergem tanto por fatores físicos e biológicos, tais como gravidez e a menstruação (MOREIRA, 2021), quanto por questões sociais e religiosas que entendem que as meninas devem assumir um papel de procriação e cuidados domésticos, enquanto os meninos devem prover o sustento da família (SHAHIDUL e KARIN, 2015).

No Brasil, as meninas são mais acometidas pelo abandono escolar. Segundo dados da PNAD 2015, o número de meninas e mulheres de 15 a 29 anos que não completaram o ensino médio, não estudavam e não exerciam atividade remunerada era de 1,7 milhões, elas representam mais que o dobro dos meninos e homens em tal situação, que totalizaram 800 mil no ano de análise. A ausência de meninas na escola custa aos países entre 15 e 30 trilhões de dólares em produtividade, conforme estimado por Wodon et al. (2018).

O objetivo desta pesquisa é identificar os fatores associados à evasão escolar de meninas e meninos, avaliando as razões de chances da evasão por etapa de ensino. A amostra estudada contempla indivíduos que participaram do Exame Nacional de Certificação (ENCCEJA), no ano de 2019. O modelo econométrico adotado será um Logit Multinomial, em que serão avaliadas as razões de chance de evasão em três períodos, quais sejam: período entre o 1º e 4º ano do ensino fundamental; período entre o 5º e o 8º ano do ensino fundamental; e entre o 1º e o 3º ano do ensino médio. Essas evidências podem ser relevantes para a proposição de políticas que visem mitigar tal problemática.

A base de dados utilizada é inédita, segundo o nosso conhecimento, para a análise da evasão escolar. Os microdados do Exame Nacional de Certificação (ENCCEJA) são disponibilizados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). Este exame tem por objetivo fornecer certificação para aqueles que evadiram-se da escola, tanto para nível fundamental como para nível médio. O questionário com dados socioeconômicos é preenchido pelos participantes de forma *online*, como uma das etapas do procedimento de inscrição. Os microdados do ENCCEJA são anuais, do tipo *cross section* e não há informação de identificação dos participantes. A população da base de dados é composta por um total de 2.973.376 indivíduos. O ano de análise é 2019, pois é o último ano com informações completas sobre os participantes do exame. Trata-se do período pré-pandêmico (Covid-19).

A base de dados do ENCCEJA se diferencia das demais, pois contém informações que possibilitam identificar os fatores que motivaram a evasão escolar de cada indivíduo, bem como o ano de ensino em que a evasão ocorreu. Assim, esta base permite mitigar a

dificuldade que normalmente encontra-se ao utilizar as bases de dados educacionais para avaliar a evasão escolar, pois quando o estudante abandona os estudos, a escola perde o vínculo com o estudante e por consequência, as informações sobre as causas da evasão dificilmente são identificadas.

Todavia, deve ser considerado que embora a amostra seja composta por um grande número de indivíduos, todos se encontram em distorção idade-série. Isso ocorre, pois o regramento do ENCCEJA define que para participar do exame de certificação é necessário ter 15 anos completos na data de realização do exame para fazer as provas do ensino fundamental, e 18 anos completos na data de realização do exame para fazer as provas do ensino médio.

O artigo está organizado em seis seções. Na próxima seção é apresentada a literatura sobre evasão escolar e sobre os seus determinantes, distinguindo ainda a evasão escolar de meninos e meninas. Na seção seguinte é apresentada a metodologia econométrica adotada. Na seção quatro é realizada uma análise exploratória dos dados, na seção cinco são apresentados os resultados. A última seção é reservada às considerações finais.

2.2 FATORES ASSOCIADOS À EVASÃO ESCOLAR

Em estudo de caso britânico que acompanhou 405 jovens, Tyerman (1968) identificou que questões familiares, tais como a negligência dos pais nos cuidados com os filhos, e as circunstâncias sociais em que os estudantes estavam inseridos eram as principais razões para a evasão escolar. Ademais, o autor observou a prevalência da evasão no grupo de meninos. Reid (1985) identificou como fator preponderante para a evasão escolar questões relacionadas à escola, como *bullying*, currículo e ensino deficiente são fatores que favorecem a evasão escolar. A base de dados contou com 384 estudantes, dos quais 128 evadiram. O Autor destaca ainda que todos os casos analisados individualmente continham aspectos sociais, psicológicos e institucionais em maior ou menor grau.

Montmarquette et al. (2007) desenvolveram um modelo econométrico sobre preferência por escolaridade. O modelo separa os estudantes em dois grupos, o dos alunos que preferem estudar e o dos alunos com maior probabilidade de ingressar no mercado de trabalho (evadir). A base de dados utilizada foi a *Statistics Canada School Leavers Survey*, que contém dados de alunos do ensino médio para o ano de 1991 com coleta posterior no ano de 1995. Os principais resultados apontam que ser do sexo feminino, frequentar escola privada e morar com pais instruídos está ligado à forte preferência pela escola. Em relação à desistência da escola, foi identificado que a decisão é afetada pela idade legal para poder trabalhar e pelo

alto salário oferecido. Adicionalmente, foi identificado que trabalhar menos de quinze horas por semana não é prejudicial aos estudos.

Li (2007) utilizou um modelo de riscos proporcionais para analisar o momento da decisão da evasão escolar no ensino médio para os Estados Unidos, utilizando a base de dados do *High School and Beyond*, com dados de 1980 e 1982. Os principais resultados indicam que os alunos que poderiam abandonar o ensino médio têm 114% a mais de chance de evadir do que aqueles que não são legalmente elegíveis ao abandono, conforme a legislação estadual (idade). Ademais, o aumento de um aluno na relação aluno/professor aumenta o risco de evasão em 1,3%, já um aumento de mil dólares nas despesas por aluno reduz o risco de evasão em 8,1%.

Na literatura brasileira, Leon e Menezes-Filho (2002), realizaram pesquisa econométrica sobre os determinantes da reprovação, avanço e evasão escolar com dados em painel da Pesquisa Mensal de Emprego (PME), do IBGE, período 1984-1997. Os autores apontam que a reprovação é um dos principais motivos da evasão escolar. Ademais, foi identificado que a evasão varia a depender do ano, no ensino fundamental após a conclusão do 4º ano, mais de 95% dos alunos, à época, permanecem em sala de aula, já para o ensino médio, após a conclusão do 8º ano, 90% permanecem cursando o ensino médio.

Em pesquisa sobre as causas da evasão escolar, Queiroz (2006) realizou em uma escola pública de Cuiabá, no Mato Grosso, entrevista com os atores relacionados à evasão, quais sejam: a escola, a família e o estudante que evadiu. O ano letivo de análise foi o 5º ano do ensino fundamental. A autora cita fatores internos relacionados à evasão, tais como a precariedade de recursos nas escolas; e fatores externos, que compreendem o trabalho, as condições de aprendizagem, e as condições familiares (escolaridade e negligência com a vida escolar dos filhos). Ademais, identifica que há fatores como a violência, as drogas, e a defasagem de aprendizagem trazida dos anos anteriores também influenciam a evasão.

Vasconcelos e Rocha (2006), a partir de um modelo de variáveis instrumentais, estimaram a influência do comportamento dos amigos sobre a decisão dos jovens evadirem (*peer effect*). A base de dados utilizada foi a da Pesquisa Socioeconômica das Comunidades de Baixa Renda (PSCBR), realizada em 51 comunidades do Rio de Janeiro no período de 1998 a 2000. Os autores encontraram efeitos significativos da influência dos amigos na decisão de os jovens evadirem da escola. Ademais, foi identificado que os amigos influenciam mais que as variáveis familiares e do ambiente.

Em pesquisa realizada com os Suplementos da Educação da Pesquisa Nacional por Amostra de Domicílios (PNAD) 2004 e 2006, Neri (2008) observou que, a evasão escolar

ocorre por quatro principais motivos: falta de interesse (40,3%); necessidade de trabalhar devido à falta de renda (27,1%); falta de escolas (10,8%); e outros motivos (21,73%). Adicionalmente, o autor aponta o desconhecimento dos prêmios¹ pela educação, além da necessidade de adolescentes pobres necessitarem suprir a sua renda, irem para o mercado de trabalho e abandonarem a escola.

As principais causas da evasão escolar apresentadas por Batista et al. (2009) referem-se à dificuldade de os estudantes se concentrarem nas aulas, em decorrência do cansaço físico ocasionado pelo trabalho (52%) e aos filhos e a gravidez (20%). Ademais, 70% dos estudantes relataram que possuem dificuldade em Matemática. Os autores realizaram estudo de caso com 25 alunos de uma escola estadual em Ilhéus, na Bahia. Os alunos entrevistados foram aqueles que em algum momento da vida abandonaram a escola, mas que retornaram à instituição.

Castelar et al. (2012) estudaram acerca da evasão escolar no estado do Ceará. Os pesquisadores estimaram um modelo econométrico *Probit* ordenado a partir dos dados de 521 escolas públicas de ensino médio do Estado nos anos, 2008, 2009 e 2010. Os principais resultados da pesquisa indicam como fatores de aumento da evasão: a repetência escolar, o tamanho do PIB do município e da criminalidade (analisada pela quantidade de óbitos por diversos tipos de agressão), bem como a localização da escola ser em região metropolitana. Como atenuantes para a evasão, foram identificados o número de estudantes em idade escolar correta, o número de docentes na escola, assim como o transporte escolar de alunos da zona rural para a escola.

Com o objetivo de avaliar a evasão e a repetência escolar (Shirasu e Arraes, 2016) utilizaram dados do estado do Ceará do Sistema de Avaliação Permanente da Educação Básica (SPAECE) no período de 2008 a 2011 e do Censo Escolar de 2008 para estimar um modelo logístico multinível. Os principais resultados indicaram que têm influência sobre a evasão e retenção escolar fatores como repetências anteriores, distorção idade-série e nível educacional dos pais. Os autores analisaram também que a evasão escolar para os beneficiários do programa Bolsa Família era cerca de 10% menor. Para capturar o interesse do aluno em estudar, foi verificada a realização das tarefas de casa, nos casos em que os estudantes a faziam, a chance de evadir era 16% menor.

A partir de um modelo de regressão hierárquico logístico longitudinal, Gonçalves et al. (2016) identificaram os fatores associados à probabilidade de ocorrência da evasão entre o

¹ Os prêmios pela educação se referem ao impacto relativo da educação sobre salários e empregabilidade.

4º e o 8º ano do ensino fundamental, em certas escolas dos estados de Pernambuco, Sergipe, Pará, Rondônia, Mato Grosso do Sul e Goiás. Os alunos foram acompanhados no período de 1999 a 2003, ano em que deveriam concluir o ensino fundamental. As bases de dados utilizadas foram das escolas, com dados sobre o fluxo escolar do estudante e do Censo Escolar. Em relação aos alunos, os principais fatores relacionados à evasão foram a repetência anterior, ser do sexo masculino e ser de baixa renda. Ademais, foi identificada maior probabilidade de evadir após cursar o 7º ano do ensino fundamental que nos anos iniciais. Para os autores tal resultado sugere relação entre evasão e mercado de trabalho, visto que ao final do ensino fundamental o aluno tem idade adequada, o que possibilita sua inserção nesse mercado.

2.3 QUAIS AS DIFERENÇAS ENTRE OS MOTIVOS DA EVASÃO ESCOLAR DE MENINAS E MENINOS?

A literatura aponta que existem diferenças nos fatores associados à evasão escolar de meninas e meninos. Shahidul e Karim (2015), a partir de uma revisão sistemática da literatura, propõem um modelo conceitual sobre os determinantes da evasão escolar de meninas. Neste modelo, a evasão delas é consequência do baixo aprendizado e baixo desempenho escolar, estando relacionada a: fatores econômicos, referente a maiores investimentos dos pais nos meninos; fatores domésticos, que referem-se a um maior envolvimento delas no trabalho doméstico; fatores do nível escolar, que compreendem atividades extracurriculares, presença de professora na escola, absenteísmo; instalações femininas nas escolas; distância escolar da residência; e por fim, fatores culturais, tais como o casamento precoce e a gravidez, e crenças culturais/religiosas de que as meninas tem um papel social distinto dos meninos.

Em relação à gravidez na adolescência, parte da literatura sobre evasão escolar dedica-se a avaliar esta temática. Berg e Nelson (2016), identificaram que, no Texas, Estados Unidos, ter idade acima da faixa etária recomendada, no 9º ano, é um dos mais poderosos moduladores do risco de gravidez e consequente abandono escolar. Sarker et al. (2019) tratam da questão do casamento e gravidez precoces para a evasão de meninas, ressaltando que existe crença social de que homens são os membros da casa que devem ganhar dinheiro, por isso é sábio educar os meninos de uma família. Já as meninas que devem cuidar de assuntos familiares, não há necessidade de educá-las.

Cruz et. al (2021), avaliaram o impacto da gravidez na adolescência sobre a evasão escolar, a metodologia adotada foi a de redes bayesianas. A base de dados teve origem nos dados do Projeto Casa do Adolescente, que contempla 19 unidades distribuídas em 18 cidades

do estado de São Paulo, além de entrevistas com especialistas (médicos, psicólogos, professores e pedagogos). Os principais resultados indicam uma forte relação de causa e efeito entre a gravidez na adolescência e a evasão escolar, sendo que os efeitos são potencializados pela vulnerabilidade econômica. Sobre o perfil dessas adolescentes, foi identificado que são meninas brancas com mais de 15 anos e que possuem renda familiar inferior a R\$ 780,00 por mês. Ademais, foi identificado que filhas de mães que ficaram grávidas na adolescência tem mais chance de também passarem por uma gravidez na adolescência.

Em uma revisão sistemática da literatura sobre evasão escolar por meio do uso de mineração de dados, Kumar (2017) identificou que fatores individuais como o gênero, o casamento, a ausência de instalações básicas na instituição de ensino que sejam diferentes para meninos e meninas, o desempenho escolar, a estrutura familiar, a qualificação e a ocupação dos pais, a qualidade do ensino são os fatores mais relevantes para a predição da evasão escolar.

Outra questão que vem sendo discutida recentemente na literatura, que afeta a frequência escolar das meninas, é a pobreza menstrual. Moreira (2021) levantou dados sobre a pobreza menstrual no Brasil e identificou que 1,24 milhão de meninas (11,6% do total de alunas), não têm acesso a papel higiênico nos banheiros de suas escolas e 31,32% estudam em escolas sem esgotamento sanitário. Silva et al. (2021) ressaltam que a falta de infraestrutura adequada para a manutenção da higiene menstrual faz com que as meninas deixem de estudar. Montgomery et al. (2016) identificaram que tanto fornecer absorventes quanto promover ações de educação sobre a puberdade influenciam positivamente a frequência escolar das meninas.

Koech e Mulamba (2023) realizaram estudo para investigar o aumento da evasão escolar de meninos no Condado de Kericho (Quênia). O estudo foi realizado a partir da percepção dos professores de classe sobre os fatores que contribuem para a evasão deles. Os principais fatores que contribuíram para a evasão escolar foram: a busca de emprego para ajudar nas despesas familiares; a influência dos pares e relacionamento amoroso; a falta de informações dos pais associada à indisciplina dos meninos. Além disso, o estudo identificou que a morte dos pais, o absenteísmo e a distância da escola à casa dos meninos são fatores que contribuem para a evasão. No entanto, a falta de instalações sanitárias e banheiros adequados não tem impacto negativo nas taxas de retenção de alunos do sexo masculino nas escolas.

Paksi et al. (2023) pesquisaram sobre a relação coletiva entre fatores individuais, familiares e escolares e o risco de abandono escolar. O estudo foi realizado a partir de uma base de dados da Hungria, na qual professores e alunos responderam a um questionário online nos anos de 2018/2019 e 2019/2020. A metodologia adotada foi regressão linear múltipla e mínimos quadrados ordinários. Os resultados indicam que os fatores individuais e familiares estão mais associados ao risco de abandono escolar do que os fatores relacionados à escola. Os autores destacam ainda que os meninos estão em maior risco de abandono escolar precoce do que as meninas, mas que este dado varia entre os países. Na Bulgária e na Romênia, por exemplo, as meninas são as que possuem mais risco de evasão. Destacam também que os motivos para a evasão entre os gêneros variam, e que embora as meninas corram menos risco de abandono, mas se abandonam (por exemplo, devido ao parto) esta situação é persistente e o regresso à escola é mais difícil.

2.4 ESTRATÉGIA EMPÍRICA

O objetivo desta pesquisa é avaliar os fatores associados à evasão escolar nas diferentes etapas de ensino, sob a hipótese de que meninos e meninas evadem por motivos distintos ao longo da trajetória escolar. Para tal propósito, estimamos as razões de riscos de evasão por etapa de ensino, a partir de um Logit Multinomial. Este modelo foi utilizado, pois nossa variável de interesse (evasão) é categórica, assumindo valor zero quando o indivíduo ainda está estudando (categoria base); valor 1 quando abandonou a escola no período entre o 1º e 4º ano do ensino fundamental; valor 2 quando abandonou a escola no período entre o 5º e a 8º ano do ensino fundamental; e valor 3 quando abandonou a escola entre o 1º e o 3º ano do ensino médio.

Há uma importante característica da amostra que se faz necessário aqui destacar. O questionário do ENCCEJA pergunta aos respondentes sobre a evasão escolar, fato esse pretérito ao momento em que o respondente se depara com o questionamento dos motivos de ter evadido. Assim, é possível que ocorra viés de memória, podendo a intensidade do motivo relatado atualmente não representar exatamente o ocorrido no passado.

O modelo foi escolhido sob a hipótese de independência entre as alternativas irrelevantes (IAI). Conforme Wooldridge (2010), sob esta hipótese o logit multinomial é válido e as respostas probabilísticas são descritas por:

$$P(y = j|x) = \frac{\exp(x\beta_j)}{[1 + \sum_{h=1}^j \exp(x\beta_h)]} \quad (1)$$

na qual, para as J categorias, a variável aleatória y assume os valores $\{0, 1, \dots, J\}$. No caso desta pesquisa, as categorias são 0, 1, 2 e 3. A variável x representa as variáveis independentes do modelo (se evadiu por motivo de trabalho, casamento e filhos, se já reprovou na escola, falta de interesse em estudar, ausência de vaga na escola, escola longe de casa), ademais, foram incluídas no modelo variáveis de controle, tais como a idade do indivíduo, o estado em que reside e a escolaridade do pai e da mãe, em estudos sobre educação, existem evidências de que a educação dos filhos é influenciada pela escolaridade dos pais (CURI, NA MENEZES-FILHO, 2009; VIEIRA ET AL., 2016). Os coeficientes β são estimados associados à x , que se referem às probabilidades de pertencer a cada categoria de evasão J . Nesta pesquisa, estimamos o modelo para avaliar os riscos relativos de evadir em uma etapa de ensino, em relação àqueles que não evadiram. O modelo de risco relativo de chances é dado por:

$$\log\left[\frac{p_j(x,\beta)}{p_h(x,\beta)}\right] = x(\beta_j - \beta_h) \quad (2)$$

Pela expressão (2) observa-se que o sentido dos efeitos parciais no modelo não depende apenas de β_j , assim como sua contraparte binária ($J = 1$). Conforme Powers e Xie (2000), o risco relativo é definido como a probabilidade de ocorrência de determinado evento em um intervalo de tempo, conceito que se aproxima ao de *odds ratio*, que para melhor interpretação pode ser convertido em incremento percentual $(odds-1)\times 100$. Assim, quando as estimativas resultarem em valor superior à unidade pode-se interpretar que o risco de uma observação ser classificada na categoria em análise (evasão) relativo ao risco dessa observação permanecer no grupo de referência (não evasão) aumenta conforme se eleva o valor de uma respectiva variável com características dos indivíduos (variáveis independentes).

2.5 ANÁLISE EXPLORATÓRIA DOS DADOS

Os dados analisados nesta pesquisa contemplam informações de 2,9 milhões de participantes do ENCCEJA. Na tabela 01 a seguir são apresentadas informações da renda dos participantes do exame, ressalta-se que em 2019 o salário-mínimo (SM) era de R\$ 954,00.

Tabela 1 - Renda em salários-mínimos

	Meninas		Meninos	
	Observações	Percentual	Observações	Percentual
Sem renda	564.795	36,66%	341.317	23,83%
Até 1 SM	673.563	43,72%	577.147	40,29%
De 1 a 3 SM	286.324	18,58%	479.429	33,47%
De 3 a 6 SM	11.748	0,76%	28.618	2,00%
De 6 a 9 SM	1.758	0,11%	3.216	0,22%
De 9 a 12 SM	1.060	0,07%	1.256	0,09%
De 12 a 15 SM	779	0,05%	766	0,05%
Mais 15 SM	761	0,05%	837	0,06%
Total	1.540.788		1.432.586	

Fonte: Microdados ENCCEJA, 2019. Elaboração própria..

Observa-se que a amostra que foi objeto deste estudo é de baixa renda, segundo a tabela 01 (renda em salários-mínimos). Para o sexo masculino 97,59% da amostra ganha até 3 salários-mínimos, desse total 341.317 homens não possuem nenhuma renda. Quando avaliada a amostra feminina, a vulnerabilidade social é ainda maior, 564.795 das mulheres (36,66%) não possuem nenhuma renda, enquanto 98,96% delas possuem renda de até três salários-mínimos.

Também destacamos o quantitativo de pessoas que não trabalham e não estudam. No ano de 2019 foram 368.413 indivíduos que no momento da inscrição para o ENCCEJA, nem estudavam nem trabalhavam. Observamos que o grupo do sexo feminino representa 58% desse total (212.882), enquanto o masculino totaliza 42% (155.531). Segundo Neri (2021), esses dados para a população brasileira de 15 a 29 anos era de 23,66% em 2019.

É importante destacar que as variáveis foram construídas, a depender da forma como a questão estava presente no banco de dados, assim, para a idade (variável de controle em números) e se já reprovou de ano alguma vez (*dummy*), temos estatísticas descritivas na tabela 02 a seguir:

Tabela 2 - Estatísticas descritivas Idade e Reprovação

	Meninas				Meninos			
	Nº observações							
	1.540.789				1.432.586			
	Média	Desvio Padrão	Min	Max	Média	Desvio Padrão	Min	Max
Idade	30.98	10.70	15	94	29.30	10.2431	15	91
Já Reprovou	0.6827	0.4654	0	1	0.8051	0.3960	0	1

Fonte: Microdados ENCCEJA, 2019. Elaboração própria.

A variável idade foi utilizada como variável de controle, não tendo sido realizada qualquer interpretação relacionada à evasão, uma vez que se trata de uma informação atual do

participante e nossa variável de interesse (evasão) ocorreu no passado. Observa-se que a média de idade entre o sexo feminino e masculino é semelhante, sendo um pouco menor para os meninos, que em média tem 29 anos, enquanto elas têm em média 30 anos.

A variável se já reprovou é largamente utilizada na pesquisa sobre evasão, sendo considerada a reprovação uma preditora para o abandono escolar. Observa-se um elevado percentual de reprovação entre meninos e meninas, sendo elas com 68% e eles com 80%, percentual significativamente maior. Esse quantitativo elevado tem fundamento na própria característica da base de dados, que contém estudantes em distorção idade-série.

A variável de interesse, quando ocorreu a evasão, foi construída a partir de duas questões presentes na base de dados. A primeira refere-se a uma pergunta se o participante do exame está estudando, e quando respondido que sim, a variável evasão assume valor zero. A segunda, refere-se à pergunta de quando o estudante evadiu, tendo como resposta 1º ano do ensino fundamental (antigo primário, 1º grau); 2º ano do ensino fundamental (antigo primário, 1º grau); 3º ano do ensino fundamental (antigo primário, 1º grau); 4º ano do ensino fundamental (antigo primário, 1º grau); 5º ano do ensino fundamental (antigo ginásio, 1º grau); 6º ano do ensino fundamental (antigo ginásio, 1º grau); 7º ano do ensino fundamental (antigo ginásio, 1º grau); 8º ano do ensino fundamental (antigo ginásio, 1º grau); 1º ano do ensino médio (antigo 2º grau); 2º ano do ensino médio (antigo 2º grau); 3º ano do ensino médio (antigo 2º grau). Assim, criamos a variável categórica, que pode ser observada na tabela 03.

Tabela 3 - Estatísticas descritivas da variável evasão

	Meninas		Meninos	
	Frequência	Percentual	Frequência	Percentual
Evadiu da entre o 1º e 4º ano do ensino fundamental	100.360	6,50%	83.515	5,90%
Evadiu da entre o 5º e 8º ano do ensino fundamental	608.672	39,90%	548.774	38,70%
Evadiu da entre o 1º e 3º ano do ensino médio	812.312	53,30%	781.233	55%
Não Evadiu	4.603	0,30%	5.467	0,40%
Total	1.525.947		1.418.989	

Fonte: elaboração própria com uso do software STATA.

Em nossa amostra, obtivemos um número reduzido de indivíduos considerados como não evadidos (10.070), pois quem respondeu que estava estudando, mas que em algum momento de sua trajetória escolar já evadiu, foi considerado como evadido. Observa-se que a

maior concentração de evasão ocorre no ensino médio. Estudantes que estejam regulares (sem distorção idade-série) devem estar matriculados no ensino médio com idade entre 15 e 17 anos, idade em que os custos de oportunidade entre estudar e trabalhar é bastante presente, pois nesta idade o trabalho é legalmente permitido.

As variáveis independentes do modelo foram construídas a partir de questões do banco de dados sobre a intensidade com que determinados fatores afetam a decisão de o estudante evadir. O valor zero significa que o fator não apresenta nenhuma influência na decisão, enquanto o valor cinco representa muita influência. Os valores 1, 2, 3 e 4 são graduais e indicam menor influência quando próximo de zero e mais influência quando próximo de cinco. Na tabela 04 a seguir são apresentadas as estatísticas descritivas.

Tabela 4 - Estatísticas descritivas das variáveis trabalho e casamento e filhos.

	Falta de tempo para estudar por causa do trabalho		Casamento e Filhos	
	Meninas	Meninos	Meninas	Meninos
Grau de influência sobre a decisão de evasão escolar	Percentual	Percentual	Percentual	Percentual
0	24,48%	15,78%	26,01%	45,91%
1	5,70%	4,66%	4,99%	7,83%
2	7,85%	7,21%	6,20%	8,14%
3	12,66%	13,06%	10,05%	10,46%
4	11,36%	13,21%	9,99%	7,40%
5	37,94%	46,09%	42,75%	20,27%

Fonte: Microdados ENCCEJA, 2019. Elaboração própria.

As duas variáveis apresentadas na tabela 04 são consideradas as principais desta pesquisa, pois buscamos identificar como a sociedade tem se organizado, no que se refere às decisões entre escola, trabalho e tarefas de cuidado familiar, em relação ao gênero, para fornecer subsídios à promoção de políticas públicas. O número de observações de meninas e meninos respondentes a esta pergunta é de respectivamente 1.540.787 e 1.432.586. Na instrução para as respostas é informado que quando a variável não exerce nenhuma influência sobre a decisão da evasão deve ser marcado “0”; e quando o fator exerce muita influência na decisão de evadir deve ser marcado “5”. É possível observar que os papéis assumidos na sociedade por homens e mulheres afetam a decisão da evasão escolar.

Em relação ao abandono escolar por causa do trabalho, tanto meninos quanto meninas são acometidas por este problema. Embora para os meninos este fator tenha mais peso. Observa-se que para 46% deles o trabalho teve muita influência na decisão de abandonar os estudos, enquanto 37,9% das meninas relataram o mesmo problema.

Já o abandono escolar em relação à alocação do tempo com cuidados do casamento/casa e dos filhos, enquanto 42,7% das meninas dizem que o casamento e os filhos

tiveram muita influência para a sua evasão escolar, apenas 20,2% responderam que este motivo os influenciou muito, ou seja, menos da metade do quantitativo delas. Quando avaliamos a resposta em relação a não ter nenhuma influência, o diferencial é ainda maior: para 45,9% deles, este fator não influenciou em nada sua decisão, enquanto para elas, apenas 26% não sofreram nenhuma influência do casamento e filhos sobre a decisão de evadir.

Na tabela 05 a seguir, sob a mesma lógica vista anteriormente, apresentamos a distribuição de três variáveis do modelo econométrico. A primeira delas é quando o indivíduo responde que deixou os estudos por falta de interesse, e as duas outras variáveis referem-se à oferta escolar.

Tabela 5 - Falta de interesse e oferta de escola

	Falta de interesse em estudar		Escola longe de casa		Ausência de Vaga na Escola	
	Meninas	Meninos	Meninas	Meninos	Meninas	Meninos
Influência sobre a evasão escolar	Percentual	Percentual	Percentual	Percentual	Percentual	Percentual
0	56,11%	46,32%	47,89%	48,30%	64,48%	61,97%
1	12,60%	13,35%	9,75%	10,57%	9,31%	9,76%
2	10,98%	13,83%	10,31%	11,29%	7,53%	8,49%
3	10,48%	14,32%	11,85%	12,69%	7,52%	8,81%
4	4,21%	5,35%	5,72%	5,35%	3,63%	3,79%
5	5,62%	6,83%	14,48%	11,80%	7,53%	7,18%

Fonte: Microdados ENCCEJA, 2019. Elaboração própria...
Meninas 1.540.787 observações
Meninos 1.432.586 observações

Observa-se que a falta de interesse em estudar é apontada como um fator muito determinante para a evasão apenas para 5,6% das meninas e 6,83% dos meninos. Sobre as variáveis de oferta escolar, observamos que isso não parece ser um dos determinantes da evasão, embora ter a escola longe de casa seja apontado como determinante por 14,5% das meninas e por 11,8% dos meninos. Nas variáveis em tela não houve diferença significativa entre os gêneros.

Uma vez que a evasão escolar ocorreu em momento pretérito à resposta ao questionário, entendemos que utilizar variáveis atuais da renda dos indivíduos nas estimativas do *Logit Multinomial* poderia trazer distorções aos resultados. Assim, considerando que a literatura utiliza a variável escolaridade dos pais para estimar a escolaridade e renda dos filhos², utilizamos duas proxies para identificar a renda dos estudantes à época que estavam na escola, quais sejam: a escolaridade da sua mãe e a escolaridade do seu pai.

² Reis e Ramos (2011), utilizam o nível educacional dos pais para observar os determinantes dos rendimentos futuros dos filhos.

Tabela 6 - Escolaridade da mãe e do pai

	Escolaridade do Pai		Escolaridade da Mãe	
	Meninas	Meninos	Meninas	Meninos
Número de Observações	1.216.670	1.160.864	1.378.748	1.277.311
Não Estudou	12,59%	8,77%	11,90%	8,15%
Da 1ª à 4ª ano do Ensino Fundamental	43,54%	38,36%	39,24%	32,42%
Da 5ª à 8ª ano do Ensino Fundamental	25,21%	28,74%	27,32%	29,06%
Ensino Médio	15,85%	20,21%	17,61%	23,97%
Ensino Superior	2,37%	3,34%	3,19%	5,26%
Especialização	0,44%	0,58%	0,75%	1,14%

Fonte: Microdados ENCCEJA, 2019. Elaboração própria.

Sobre a escolaridade dos pais, observa-se que para a maior parte da amostra, ou os pais não estudaram ou estudaram apenas até o ensino fundamental. Tal dado faz sentido quando analisada a própria renda do indivíduo, apresentada na tabela 01.

2.6 RESULTADOS

Na tabela 07 a seguir são apresentados os resultados da estimação do Logit Multinomial para avaliar as razões de chance de evasão de meninas e meninos no decorrer da trajetória escolar (três etapas de ensino analisadas). Ressalta-se que foram inseridas no modelo como variáveis de controle a idade e a unidade federativa dos indivíduos, que consta no apêndice A1.

Tabela 7 - Resultados do modelo Logit Multinomial

Variáveis de interesse	Ensino Fundamental 1º ao 4º ano		Ensino Fundamental 5º ao 8º ano		Ensino Médio 1ª ao 3º ano	
	Meninas	Meninos	Meninas	Meninos	Meninas	Meninos
Casamento e Filhos	1,1860*** (0,0307)	1,0535* (0,0290)	1,3025*** (0,0335)	1,0887* (0,0297)	1,2807*** (0,0329)	1,0963** (0,0299)
Falta de tempo por causa do trabalho	1,2032*** (0,0315)	1,2311*** (0,0276)	1,2177*** (0,0316)	1,3116*** (0,0288)	1,2953*** (0,0336)	1,4165*** (0,0311)
Já reprovou na escola	1,2463* (0,1490)	1,0485 (0,1455)	1,6135*** (0,1916)	1,4332* (0,1974)	1,6588*** (0,1969)	1,4554** (0,2003)
Escola longe de casa	1,0781** (0,0336)	1,0415 (0,0288)	0,9691 (0,0300)	0,9363** (0,0255)	0,9636 (0,0299)	0,9186* (0,0250)
Ausência de Vaga	0,9123* (0,0301)	0,9307** (0,0261)	0,8843*** (0,0290)	0,8874*** (0,0245)	0,8623*** (0,0282)	0,8758*** (0,0241)
Falta de interesse em estudar	0,9908 (0,0320)	1,1803*** (0,0330)	0,9979 (0,0319)	1,1924*** (0,0329)	0,9271* (0,0296)	1,0874* (0,0299)
Escolaridade do Pai	0,8437** (0,0506)	0,7985*** (0,0380)	0,9526 (0,0564)	0,8697** (0,0405)	1,0552 (0,0625)	0,9893 (0,0460)
Escolaridade da Mãe	0,9040* (0,0549)	0,8815* (0,0416)	1,0225 (0,0614)	0,9934 (0,0459)	1,2233** (0,0734)	1,1516 (0,0532)

Controle UF	Sim	Sim	Sim	Sim	Sim	Sim
Nº Observações Meninas	333.105					
Nº Observações Meninos	334.110					
Log Likelihood estimativa Meninas	-274811.96					
Log Likelihood estimativa Meninos	-269234.70					
Erro-padrão entre parênteses. Nível de significância *p<0,01 , **p<0,05, ***p<0,01.						

Fonte: Microdados ENCCEJA, 2019. Elaboração própria.

Sobre a evasão escolar em virtude do casamento e filhos, os resultados são significativamente distintos entre meninas e meninos. Enquanto elas têm um risco de evasão de 18% (1ª etapa), 30% (2ª etapa) e 28% (3ª etapa), os meninos apresentam risco de 5%, 8% e 9%, respectivamente. Observa-se que o risco de evasão causada por casamento e filhos aumenta significativamente quando as meninas entram na idade reprodutiva (a partir dos 10 anos de idade)³. Uma possível explicação para este resultado é a cultura de que as meninas são responsáveis pelos serviços de cuidado da casa e família, enquanto os meninos são responsáveis pelo sustento familiar. Um dado que corrobora esta afirmação é o tempo despendido entre homens e mulheres nas tarefas domésticas, conforme apresentado na pesquisa de Uhr et al. (2020), as mulheres trabalham em média 37 horas por semana no mercado de trabalho e 20 horas semanais são despendidas nas tarefas domésticas, totalizando 57 horas totais. Já os homens, trabalham em média 42 horas semanais no mercado de trabalho e gastam 6 horas semanais nas tarefas domésticas, totalizando 48 horas.

Ao analisar o papel cultural da mulher de cuidadora e do homem de provedor, esperar-se-ia que o efeito da evasão para as meninas causado por conta do casamento e filhos fosse de certa forma compensado pelos motivos de evasão dos meninos por conta do mercado de trabalho. Todavia, os resultados apontam uma sobrecarga das meninas, visto que, mesmo que em menor proporção que os meninos, elas estão expostas a um alto risco de evasão devido à necessidade de trabalhar (variável falta de tempo por causa do trabalho) e também por conta do casamento e filhos, principalmente na adolescência, momento em que o risco passa para 30% para elas (a partir da 5ª série).

O problema da gravidez na adolescência é uma questão discutida na literatura e relacionada à evasão de meninas. Há pesquisas que apontam a gravidez como fato preditor da evasão (LIMA E COLS., 2004; LOSS e SAPIRO, 2005), enquanto há evidências de que a gravidez precoce ocorre após a evasão escolar (SABROZA, LEAL, SOUZA JR., e GAMA, 2004). De toda forma, é um problema que acomete de forma contundente as meninas, Castro

³ para informações sobre a idade da primeira menstruação (menarca) das meninas no Brasil ver PIKANÇO, Marilúcia Rocha de Almeida. A idade da menarca da menina brasileira: Os fatores sócio-econômicos e as diferenças regionais. 1995.

et al. (2004) apontam que, quando da ocorrência da gravidez na adolescência, as meninas abandonam a escola pois além de terem que cuidar dos filhos, muitas vezes necessitam entrar no mercado de trabalho.

Em termos de necessidade de proposição de políticas públicas, cabe aqui destacar a vulnerabilidade a que as meninas brasileiras estão expostas. Além de elas serem a maioria a não estar na escola, quando analisado em relação aos meninos, a gravidez na adolescência é uma das causas dessa ausência escolar feminina. Segundo De Matos (2022), a prevenção da gravidez precoce está vinculada às atividades familiares e educativas, que resultarão no conhecimento sobre métodos contraceptivos, além do fornecimento de informações seguras sobre sexualidade.

A problemática extrapola a evasão escolar e tem influências sobre o mercado de trabalho ocupado pelas mulheres. Sobre esse tema, Fernandes et al. (2022) identificaram aumento nas probabilidades de as mulheres, após o nascimento do primeiro filho, estarem em posições de desocupação, inatividade e informalidade. Assim, após o nascimento de um bebê, a probabilidade de a mulher estar inativa é quase três vezes maior (2,7) do que a probabilidade de estar ocupada no setor formal.

Embora a literatura aponte a relação entre repetência/reprovação e evasão, tais como Castelar et al. (2012), Shirasu e Arraes (2016), e Gonçalves et al. (2016), nossos resultados trazem pontos interessantes à literatura, uma vez que identificamos os efeitos da repetência sobre a evasão ao longo da trajetória escolar de meninas e meninos.

No caso das meninas, o efeito de reprovar sobre a evasão é progressivo ao longo do tempo. Observa-se que reprovar na primeira etapa do ensino fundamental eleva em 24% o risco de evadir, já na segunda etapa o risco vai para 61%. No ensino médio a magnitude é ainda maior, alcançando 66% de aumento no risco de evasão. Conforme apontado na literatura, um fator que diferencia o fluxo da evasão de meninas e meninos é a pobreza menstrual, que pode levá-las a sucessivas faltas, reprovações e por consequência à evasão (MOREIRA, 2021).

Para os meninos, a reprovação nos primeiros anos da vida escolar não foi significativa. Já na segunda etapa do ensino fundamental, período em que atingem a idade para ocupar postos de trabalho na condição de menor aprendiz (14 anos)⁴, o risco de evadir após a reprovação é de 43% e no ensino médio de 45%. Tal resultado pode estar relacionado,

⁴ Idade prevista na Consolidação das Leis do Trabalho (CLT).

dentre outros fatores, a uma maior exposição dos meninos ao mercado de trabalho conforme apontado por Shahidul e Karin (2015).

A reprovação é um problema que possui impacto sobre a probabilidade de evasão, sendo considerada a principal determinante do abandono escolar (LEON e MENEZES-FILHO, 2002). Conforme Rebelo (2009) a chance de evasão é significativamente maior se o aluno passar pelo processo de reprovação mais do que uma vez. Do ponto de vista pedagógico a reprovação e a conseqüente repetência são consideradas ineficazes, com potenciais prejuízos ao desenvolvimento dos estudantes, com efeitos negativos sobre a aprendizagem, personalidade e comportamento dos estudantes (REBELO, 2009)

Sobre a distância entre a escola e a residência, esta é um fator associado pela literatura à evasão escolar (SHAHIDUL e KARIM, 2015; KOECH e MULAMBA 2023). Nossos resultados apontam que este é um fator significativo ao aumento de risco de evasão apenas na primeira etapa escolar, e para as meninas. Para os meninos os resultados não foram significativos.

A oferta de vagas também não foi um fator que diminui a probabilidade de evasão. Na tabela 05 pode-se observar que este foi um fator pouco indicado como motivador para a evasão. Apenas 5% dos respondentes indicaram que este seria um fator com forte impacto sobre a evasão. Ademais, conforme dados do Observatório do Plano Nacional de Educação, o Ensino Fundamental e o Ensino Médio estão quase universalizados. Em 2020, a meta 2 de universalizar o Ensino Fundamental de 9 anos para toda a população de 6 a 14 anos, em 2020 a referida meta encontrava-se cumprida para 98% do público-alvo⁵. Já para o ensino médio, a meta 03 era de universalizar, até 2016, o atendimento escolar para toda a população de 15 a 17 anos⁶. Em 2020 a referida meta encontrava-se cumprida para 94,5% do público-alvo.

Em relação à falta de interesse em estudar, os resultados são antagônicos para meninas e meninos. Enquanto para elas este não é um fator associado à evasão, para os meninos, a falta de interesse pelos estudos corrobora a evasão nas duas primeiras etapas de ensino, representando um aumento do risco em 18% na primeira etapa do ensino fundamental e 19% na segunda. No ensino médio esta variável não foi significativa. Neri (2009) aponta que a falta de interesse em estudar pode ser causada pelo desconhecimento dos potenciais prêmios oferecidos pela educação.

A variável que remete à evasão por falta de interesse em estudar deve ser analisada com cautela. É possível que esses indivíduos estejam avocando para si uma culpa que não

⁵ Disponível em [OPNE – Meta 02 | Ensino Fundamental \(observatoriodopne.org.br\)](https://observatoriodopne.org.br/OPNE-Meta-02-Ensino-Fundamental)

⁶ Disponível em [OPNE – Meta 03 | Ensino Médio \(observatoriodopne.org.br\)](https://observatoriodopne.org.br/OPNE-Meta-03-Ensino-Medio)

lhes pertença. É difícil ter “interesse em estudar”, quando as condições socioeconômicas são escassas. Um dado importante a ser considerado nesta análise é o apontado por Neri (2009), que aponta a falta de interesse em estudar maior entre indivíduos de menor renda.

No que se refere à escolaridade dos pais, a literatura observa que este é um fator relevante à escolaridade e renda dos filhos (REIS E RAMOS, 2011). Nos resultados desta pesquisa a escolaridade do pai possui efeito para a presença dos filhos à escola, nos anos iniciais do ensino fundamental (1º ao 4º ano) para meninos e meninas, e efeito positivo apenas para os meninos (5º ao 8º ano). No ensino médio a escolaridade do pai não é significativa.

Em relação à escolaridade da mãe, esta é significativa para a permanência escolar nos anos iniciais do ensino fundamental. Nos demais anos de ensino não apresenta significância estatística. Há exceção, para as meninas, no ensino médio, no qual o aumento da escolaridade da mãe mostrou-se como um maior fator de risco à evasão. Este resultado, não apresenta coerência com a literatura, que possui evidência de quanto maior a escolaridade dos pais, maior a dos seus filhos (BARROS ET AL, 2006; GHANEM JR, 2010; BOUALAPHET e GOTO, 2020). Duas possíveis causas podem ser aventadas a esse respeito, a primeira trata-se da amostra específica desta pesquisa, de baixa escolaridade, e de baixa renda, é possível que os efeitos de escolaridade dos pais para esta amostra sejam diversos. Outro fator é que a escolaridade dos pais é respondida pelos filhos, fato esse que pode apresentar algum viés de memória (PEREIRA E KOIFMAN, 1999).

2.7 CONSIDERAÇÕES FINAIS

De forma geral, com esta pesquisa, buscamos identificar como a sociedade tem se organizado, no que se refere às decisões entre escola, trabalho e tarefas de cuidado familiar, sob a perspectiva de gênero. Ao analisar o papel cultural da mulher de cuidadora e do homem de provedor, esperar-se-ia que o efeito da evasão para meninas causado por conta do casamento e filhos fosse de certa forma compensado pelos motivos de evasão dos meninos por conta do mercado de trabalho. Todavia, os resultados apontam uma sobrecarga das meninas, visto que, mesmo em menor proporção que os meninos, elas estão expostas a um alto risco de evasão devido à necessidade de trabalhar (variável falta de tempo por causa do trabalho), além do elevado risco de evasão associada ao casamento e aos filhos.

Assim, identificamos que a necessidade de trabalho em detrimento da educação acomete tanto meninos quanto meninas, porém a magnitude do risco de evasão para trabalhar é maior para eles, chegando a 41% durante o ensino médio. Em relação à etapa de ensino, as

razões de riscos de evadir em virtude do trabalho aumentam tanto da primeira para a segunda etapa do ensino fundamental, quanto do ensino fundamental para o ensino médio.

No que se refere à evasão por motivo de casamento e filhos, este fator acomete preponderantemente as meninas. O maior risco para elas é do 5º ao 8º ano do ensino fundamental, momento em que as razões de riscos são aumentadas em 30%. Para os meninos, casamento e filhos aumentam em 5%, 8% e 9%, nas respectivas etapas de ensino, o risco de evasão.

Já ter reprovado em algum ano, aumenta mais os riscos de evasão para meninas do que para os meninos. Para elas, os resultados chegam a um aumento de 66% no risco de evasão, na etapa do ensino médio, enquanto para eles, o risco no ensino médio é de 45%.

Além de trazer subsídios para políticas públicas que tenham por objetivo tratar a evasão escolar, nossos resultados indicam que estudos sobre esta temática devem ser abordados sob a perspectiva de gênero, pois os fatores são distintos para meninas e meninos e análises que não considerem este quesito podem apresentar viés e os resultados serem limitados.

Ainda no que se refere à proposição de políticas públicas, destaca-se que é de extrema relevância a adoção de políticas que visem educar meninos e meninas com a finalidade de prevenir a gravidez na adolescência e o casamento precoce, tendo em vista as graves consequências discutidas nesta pesquisa. Outra questão necessária é a ampliação da oferta de vagas em creches públicas, para possibilitar às mães a conclusão dos estudos, a permanência ou reingresso no mercado de trabalho após o nascimento dos filhos.

Destaca-se como limitação deste estudo a ausência de variáveis relevantes na base de dados, que podem ter prejudicado as análises, tais como a ausência de informações sobre a cor dos estudantes, fator este relevante para a análise da evasão escolar, visto que autodeclarados pretos tem menor escolaridade média que os brancos, conforme dados da PNAD 2022. Ademais, cabe considerar ainda, a possibilidade de existir viés de memória nos respondentes ao questionário, pois os indivíduos estão respondendo atualmente sobre a sua evasão, fato esse ocorrido no passado.

REFERÊNCIAS

- ANAZAWA, Leandro, GUEDES, Marcelo Sanchez, KOMATSU, Bruno Kawaoka, AND MENEZES-FILHO, Naercio Aquino. "A Loteria da Vida: Examinando a Relação entre a Educação da Mãe e a Escolaridade do Jovem com Dados Longitudinais do Brasil". *Inspere - Centro de Políticas Públicas, Policy Paper 22*. 2016.
- ANDERSON, D. Mark. In school and out of trouble? The minimum dropout age and juvenile crime. *Review of Economics and Statistics*, v. 96, n. 2, p. 318-331, 2014.
- BARROS, Ricardo Paes de Organizador; FOGUEL, Miguel Nathan Organizador; ULYSSEA, Gabriel Organizador. *Desigualdade de renda no Brasil: uma análise da queda recente*. 2006.
- BATISTA, Santos Dias; SOUZA, Alesxsandra Matos; OLIVEIRA, Júlia Mara da Silva. A evasão escolar no ensino médio: um estudo de caso. *Revista Profissão Docente, UNIUBE. Uberaba/MG*, v. 9, n. 19, 2009.
- BERG, Nathan; NELSON, Teresa D. Pregnancy and dropout: Effects of family, neighborhood, and high school characteristics on girls' fertility and dropout status. *Population Research and Policy Review*, v. 35, p. 757-789, 2016.
- BRASIL. Constituição (1988). *Constituição da República Federativa do Brasil*. Brasília, DF: Senado Federal: Centro Gráfico, 1988.
- BOUALAPHET, Kikeo; GOTO, Hideaki. Determinants of school dropout in Lao People's Democratic Republic: A survival analysis. *Journal of International Development*, v. 32, n. 6, p. 961-975, 2020.
- CASTRO, M. G., ABRAMOVAY, M., e Silva, L. B. *Juventudes e sexualidades*. Brasília: UNESCO Brasil. 2004.
- CURI, Andréa Zaitune; MENEZES-FILHO, Naércio Aquino. A relação entre educação pré-primária, salários, escolaridade e proficiência escolar no Brasil. *Estudos Econômicos (São Paulo)*, v. 39, p. 811-850, 2009.
- CRUZ, Emerson et al. The impact of teenage pregnancy on school dropout in Brazil: a Bayesian network approach. *BMC Public Health*, 2021.
- DE BARROS, Ricardo Paes. *Consequências da violação do direito à educação*. Editora Autografia, 2021.
- DE CARVALHO CASTELAR, Pablo Urano; MONTEIRO, Vitor Borges; LAVOR, Daniel Campos. Um estudo sobre as causas de abandono escolar nas escolas públicas de Ensino Médio no Estado do Ceará. *GOVERNO DO ESTADO DO CEARÁ*, v. 1, p. 33, 2012.
- DE MATOS BEZERRA, Thiago; MATOS, Cintia Chagas. Impactos da gravidez na adolescência no Brasil. *Research, Society and Development*, v. 11, n. 5, p. e39111528381-e39111528381, 2022.
- DRAGONE, Davide; MIGALI, Giuseppe; ZUCHELLI, Eugenio. *High school dropout and the intergenerational transmission of crime*. Bonn: IZA Institute of Labour Economics 2021.

FERNANDES MATOS LIMA, Isabela; VERZOLA VAZ, Daniela; LEDI GONCALVES, Solange. O trade-off entre participação no mercado de trabalho e maternidade revisitado: uma análise de painel para o período 2012-2019. Available at SSRN 4219970, 2022.

GONÇALVES, Maria Elizete; RIOS-NETO, Eduardo LG; CÉSAR, Cibele Comini. Evasão no ensino fundamental brasileiro: identificação e análise dos principais determinantes. Anais do Encontro Nacional de Estudos Populacionais, p. 1-20, 2016.

GERLINGER, Julie; HIPPEL, John R. Schools and neighborhood crime: The effects of dropouts and high-performing schools on juvenile crime. *The Social Science Journal*, 2020.

KOECH, Evaline Chelangat; MULAMBULA, S. M. Class teachers' perceptions on the influence of the selected factors on boy child drop-out in public mixed day secondary schools in Kericho County, Kenya. *International Journal of Management Studies and Social Science Research*.

KUMAR, Mukesh; SINGH, A. J.; HANDA, Disha. Literature survey on educational dropout prediction. *International Journal of Education and Management Engineering*, v. 7, n. 2, p. 8, 2017.

LI, Mingliang. Bayesian proportional hazard analysis of the timing of high school dropout decisions. *Econometric Reviews*, v. 26, n. 5, p. 529-556, 2007.

LIMA, C. T. B., FELICIANO, K. V. DE O., CARVALHO, M. F. S., SOUZA, A. P. P., MENABÓ, J. B. C., RAMOS, L. S., CASSUNDÉ, L. F., e KOVÁCS, M. H. Percepções e práticas de adolescentes grávidas e de familiares em relação à gestação. *Revista Brasileira de Saúde Materno-Infantil*, 4, 71-83 2004.

LOSS, M. A., e SAPIRO, C. M. Processos psíquicos do engravidamento na adolescência em contexto de periferia: Impasses e possibilidades. *Psicologia USP*, 16(4), 69-98. 2005.

MASTORCI, Francesca et al. Gender Differences for Health Indicators in a Sample of School Dropout Adolescents: A Pilot Study. *International Journal of Environmental Research and Public Health*, v. 19, n. 13, p. 7852, 2022.

MONTEIRO, Vitor Borges; ARRUDA, Elano Ferreira. O impacto da violência urbana nos indicadores de evasão escolar na Região Metropolitana de Fortaleza. *Anais do I Circuito de Debates acadêmicos*, 2011.

MONTEIRO, Vitor Borges; ARRUDA, Elano Ferreira. O impacto da violência urbana nos indicadores de evasão escolar na Região Metropolitana de Fortaleza. *Anais do I Circuito de Debates acadêmicos*, 2011.

MONTMARQUETTE, Claude; VIENNOT-BRIOT, Nathalie; DAGENNAIS, Marcel. Dropout, school performance, and working while in school. *The Review of Economics and Statistics*, v. 89, n. 4, p. 752-760, 2007.

MOREIRA, Luisa Prado Afonso. Pobreza menstrual no Brasil: diagnóstico e alternativas. 2021. Tese de Doutorado.

MONTGOMERY, Paul et al. Menstruation and the cycle of poverty: a cluster quasi-randomised control trial of sanitary pad and puberty education provision in Uganda. *Plos one*, v. 11, n. 12, p. e0166122, 2016.

NERI, M. *Motivos da evasão escolar*. São Paulo: Instituto Unibanco, 2008. Mimeo.

NERI, M. C. *Juventudes, educação e trabalho: Impactos da pandemia nos nem-nem*. Rio de Janeiro: FGV Social, 2021.

NERI, M.; THOMAS, M. The effects of idiosyncratic shocks to father's income on child labor, school drop-outs and repetition rates in Brazil. *Anais do XXII Encontro Brasileiro de Econometria, SBE*, 2000.

LEON, Fernanda Leite Lopez de; MENEZES-FILHO, Naércio Aquino. Reprovação, avanço e evasão escolar no Brasil. *Pesquisa e Planejamento Econômico*, Rio de Janeiro, v. 32, n. 3, p. 417-451, 2002. 2002.

QUEIROZ, Lucileide Domingos. Um estudo sobre a evasão escolar: para se pensar na inclusão escolar. *Rev Bras Estudos Pedag*, v. 64, n. 147, p. 38-69, 2006.

REID, Ken. *Truancy and school absenteeism*. Hodder and Stoughton, 1985.

REIS, Mauricio Cortez; RAMOS, Lauro. Escolaridade dos pais, desempenho no mercado de trabalho e desigualdade de rendimentos. *Revista Brasileira de Economia*, v. 65, p. 177-205, 2011.

PAKSI, Borbála; SZÉLL, Krisztián; FEHÉRVÁRI, Anikó. Empirical Testing of a Multidimensional Model of School Dropout Risk. *Social Sciences*, v. 12, n. 2, p. 50, 2023.

PEREIRA, Rosângela A.; KOIFMAN, Sérgio. Uso do questionário de frequência na avaliação do consumo alimentar progressivo. *Revista de Saúde Pública*, v. 33, p. 610-621, 1999.

POWERS, D. A.; XIE, Y. *Statistical Methods for Categorical Data Analysis*. San Diego: Academic, 2000. 305 p.

REBELO, J. A. S. (2009). Efeitos da retenção escolar, segundo os estudos científicos, e orientações para uma intervenção eficaz: Uma revisão. *Revista Portuguesa De Pedagogia*, (43-1), p. 27-52.

SABROZA, A. R., LEAL, M. C., SOUZA JR., P. R., e GAMA, S. G. N. Algumas repercussões emocionais negativas da gravidez precoce em adolescentes do município do Rio de Janeiro (1999-2001). *Cadernos de Saúde Pública*, 2004.

SARKER, Md Nazirul Islam; WU, Min; HOSSIN, Md Altab. Economic effect of school dropout in Bangladesh. *International journal of information and education technology*, v. 9, n. 2, p. 136-142, 2019.

SHAHIDUL, S. M.; KARIM, A. H. M. Z. Factors contributing to school dropout among the girls: A review of literature. *European Journal of research and reflection in educational sciences*, v. 3, n. 2, 2015.

SHIRASU, Maitê Rimekká; ARRAES, Ronaldo de Albuquerque. Determinantes da evasão e repetência escolar. Encontro Nacional de Economia, v. 43, 2016.

SILVA, João Victor Ferreira da; LOPES, Yoanna Danielly Victor. A pobreza menstrual como fator de violação de direitos humanos: um olhar para adolescentes em ambiente escolar. 2022.

TYERMAN, Maurice J. Truancy. University of London Press, 1968.

UHR, Daniel de Abreu Pereira et al. Alocação do tempo entre os gêneros e o mercado de trabalho: uma análise entre casados e solteiros para o Brasil. Nova Economia, v. 29, p. 1041-1063, 2020.

WODON Q, MALE C, MONTENEGRO C, NGUYEN H, ONAGORUWA A. The Cost of Not Educating Girls: Educating Girls and Ending Child Marriage: A Priority For Africa. Washington, DC: World Bank; 2018.

WOOLDRIDGE, Jeffrey M. Econometric analysis of cross section and panel data. MIT press, 2010.

VASCONCELLOS, I. C.; ROCHA, L. F. Interação social e evasão escolar nas favelas do Rio de Janeiro: um problema de identificação. XXXIV Encontro Nacional de Economia, 2006.

VIEIRA, CATERINA SOTO, CABANAS, PEDRO, MENEZES-FILHO, NAERCIO, AND KOMATSU, BRUNO KAWAOKA. "Como as Mudanças no trabalho e Renda dos Pais afetam as Escolhas entre Estudo e Trabalho dos Jovens?" Pesquisa e Planejamento Econômico. 2016.

APÊNDICE A 1

Apêndice A1 Tabela 1 - Estimativas com os controles referentes aos Estados da Federação.

Estados da Federação	Ensino Fundamental 1º ao 4º ano		Ensino Fundamental 5º ao 8º ano		Ensino Médio 1ª ao 3º ano	
	Meninas	Meninos	Meninas	Meninos	Meninas	Meninos
Rondônia	5,60** (2,97277)	2,0088* (0,7165)	3,4323 (1,8118)	1,3685 (0,4812)	2,9604* (1,5625)	1,2917 (0,4539)
Acre	5,600 (0,5145)	1,1400 (1,4410)	0,6627 (0,2774)	5,551 (7,0209)	0,6924 (0,2896)	7,2861 (9,2209)
Amazonas	2,930 (1,5611)	1,8641 (0,8249)	2,7142 (1,4342)	1,4048 (0,6147)	2,3720 (1,2532)	1,4178 (0,6201)
Roraima	0,9801 (0,5298)	2,8571 (2,0886)	0,7770 (0,4119)	1,7821 (1,2920)	0,9072 (0,4804)	3,0057 (2,1771)
Pará	1,2070 (0,3443)	0,6828 (0,1540)	0,9728 (0,2727)	0,5191* (0,1137)	1,2782 (0,3580)	0,6903 (0,1509)
Amapá	0,8750 (0,3530)	0,5744 (0,1764)	0,4916 (0,1941)	0,3358*** (0,0994)	0,6149 (0,2423)	0,3900* (0,1152)
Tocantins	0,7432 (0,2992)	0,7474 (0,2267)	0,7577 (0,2982)	0,5134 (0,1511)	1,1212 (0,4406)	0,8226 (0,2414)
Maranhão	1,3817 (0,7418)	0,3977* (0,1321)	1,3532 (0,7170)	0,3739** (0,1195)	1,8929 (1,0022)	0,6383 (0,2034)
Piauí	0,9861 (0,5308)	0,6038 (0,2111)	1,1865 (0,6275)	0,4550 (0,1547)	1,6101 (0,8508)	0,6762 (0,2294)
Ceará	1,3699 (0,4670)	1,0908 (0,3437)	1,5373 (0,5155)	0,9277 (0,2868)	2,1769* (0,7295)	1,5141 (0,4676)
Rio Grande do Norte	0,8209 (0,4446)	1,0599 (0,6464)	0,8624 (0,4563)	1,0452 (0,6261)	1,2305 (0,6502)	1,6408 (0,9818)
Paraíba	2,484 (1,8258)	0,8878 (0,4273)	2,3410 (1,7039)	0,9643 (0,4573)	3,4801 (2,5314)	1,4200 (0,6726)
Pernambuco	0,6885 (0,3135)	1,0493 (0,4689)	0,8943 (0,3962)	1,0359 (0,4533)	0,9499 (0,4205)	1,0516 (0,4598)
Alagoas	0,8227 (0,2921)	0,8622 (0,3132)	0,6616 (0,2296)	0,7415 (0,2618)	0,7042 (0,2442)	0,8580 (0,3026)
Sergipe	1,0510 (0,4224)	0,5353 (0,1531)	0,9951 (0,3906)	0,4412** (0,1203)	1,0356 (0,4061)	0,4658* (0,1267)
Bahia	0,7142 (0,1760)	0,4705*** (0,0942)	0,6796 (0,1630)	0,3770*** (0,0724)	1,1451 (0,2741)	0,5948 (0,1267)
Minas Gerais	5,6359*** (1,8962)	2,3832* (0,6254)	5,0650*** (1,6911)	2,1187 (0,5668)	6,1205*** (2,0429)	3,1333*** (0,8081)
Espírito Santo	2,5239 (0,9987)	0,9325 (0,2749)	2,1108 (0,8266)	0,9194 (0,2636)	2,8543* (1,1170)	1,3772 (0,3943)
Rio de Janeiro	2,5239 (0,9987)	1,3295 (0,2882)	1,9077* (0,4469)	1,2378 (0,2619)	2,7087*** (0,6341)	1,9142* (0,4046)
São Paulo	1,9701 (0,4691)	2,0025** (0,4361)	3,8567*** (0,9132)	1,8140* (0,3872)	4,8143*** (1,1394)	2,8481*** (0,6074)
Paraná	2,5239 (0,9987)	2,9881*** (0,6965)	4,2972*** (1,1351)	2,8063*** (0,6422)	3,5711*** (0,9460)	2,2228*** (0,5084)
Santa Catarina	1,9701 (0,4691)	3,1347*** (0,9302)	10,5637*** (5,0467)	2,5664* (0,7514)	9,6021*** (4,5868)	2,4936* (0,7298)
Rio Grande do Sul	3,8979*** (0,9351)	1,3942 (0,2867)	2,2395*** (0,3055)	1,4799 (0,2963)	2,1150* (0,4817)	1,4670 (0,2934)
Mato Grosso do Sul	5,7731*** (1,5415)	1,0593 (0,2635)	1,1570 (0,3055)	0,9020 (0,2191)	1,1926 (0,3147)	0,9459 (0,2295)
Mato Grosso	14,6244*** (7,0141)	1,2807 (0,5322)	1,5087** (0,6257)	0,9752 (0,3990)	2,0758 (0,8603)	1,6313 (0,6665)
Goiás	1,9414* (0,6795)	1,4370 (0,4294)	2,5632 (0,8867)	1,5650 (0,4591)	3,2424** (1,1212)	2,0798* (0,6096)
Distrito Federal	<i>omitted</i>	<i>omitted</i>	<i>omitted</i>	<i>omitted</i>	<i>omitted</i>	<i>omitted</i>

Nº Observações Meninas 333.105

Nº Observações Meninos 334.110

Log Likelihood estimativa Meninas -274811.96

Log Likelihood estimativa Meninos -269234.70

Erro-padrão entre parênteses. Nível de significância *p<0,01 , **p<0,05, ***p<0,01.

Fonte: elaboração própria

3. AVALIAÇÃO DE POLÍTICAS PÚBLICAS: UMA ANÁLISE DO DESEMPENHO DOS ESTUDANTES DA EJA NO EXAME NACIONAL PARA CERTIFICAÇÃO DE COMPETÊNCIAS DE JOVENS E ADULTOS.

Resumo

O objetivo desta pesquisa é avaliar empiricamente o efeito de cursar Educação de Jovens e Adultos (EJA) sobre a probabilidade de aprovação no Exame Nacional para Certificação de Competências de Jovens e Adultos (Encceja). A avaliação foi realizada para os dois tipos de certificação, Ensino Fundamental e Ensino Médio, assim como para as quatro provas distintas aplicadas no exame, quais sejam, Ciências Humanas, Ciências da Natureza, Linguagens e Redação, e Matemática. A fonte de dados são os microdados do Encceja, divulgados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), para o ano de 2019. Com o intuito de conferir robustez às estimativas, três metodologias econométricas foram aplicadas. A primeira foi o *Propensity Score Matching*, a segunda o *Coarsened Exact Matching*, e a terceira o *Entropy Balancing*, além da combinação entre os métodos. A partir dos resultados estimados para o Ensino Médio observou-se menor probabilidade de aprovação no Encceja para os indivíduos que cursam EJA, em todas as disciplinas analisadas, sendo o efeito negativo de maior magnitude para Linguagens e Redação, e Matemática. Para certificação em Nível Fundamental identificou-se pior desempenho para as disciplinas de Linguagens e Redação, e Ciências da Natureza. Em geral os resultados indicam desempenho inferior do grupo participante da política de EJA.

Palavras-chave: Educação de Jovens e Adultos; Economia da Educação; *Propensity Score Matching*; *Coarsened Exact Matching*; *Entropy Balancing*.

Abstract

The objective of this research is to empirically evaluate the effect of attending Youth and Adult Education (EJA) on the probability of passing the National Examination for the Certification of Youth and Adult Skills (Encceja). The evaluation was carried out for the two types of certification, Elementary School and High School, as well as for the four different tests applied in the exam, namely, Human Sciences, Natural Sciences, Languages and Writing, and Mathematics. The data source is the Encceja microdata, released by the National Institute of Educational Studies and Research Anísio Teixeira (INEP), for the year 2019. In order to provide robustness to the estimates, three econometric methodologies were

applied. The first was the Propensity Score Matching, the second the Coarsened Exact Matching, and the third the Entropy Balancing, in addition to the combination between the methods. Based on the estimated results for High School, a lower probability of passing the Enceja was observed for individuals were studying EJA, in all subjects analyzed, with the negative effect being of greater magnitude for Languages and Writing, and Mathematics. For certification at the Fundamental Level, worse results was identified for Languages and Writing, and Natural Sciences. In general, the results indicate a lower performance of the group participating in the Eja policy.

Keywords: Youth and Adult Education; Economics of Education; Propensity Score Matching; Coarsened Exact Matching; Entropy Balancing.

JEL: A20; C50; J18.

3.1 INTRODUÇÃO

A Educação de Jovens e Adultos (EJA) é uma modalidade de ensino que tem como público-alvo aqueles que não acessaram ou não continuaram os estudos no ensino fundamental ou no ensino médio regular em idade apropriada (BRASIL, 1996). Uma das principais diferenças entre o Ensino Regular e a EJA é o tempo de duração para sua conclusão. No Ensino Regular o nível fundamental é composto por nove anos e o nível médio por três anos. Na EJA cada ano é cursado em seis meses. Instituições de ensino públicas e privadas ofertam cursos de EJA. Não é necessária a realização de qualquer prova adicional de certificação, caso o estudante seja aprovado no próprio curso de EJA recebe a certificação referente à etapa concluída.

Há no Brasil um número elevado da população que não frequentou a escola ou que dela se evadiu. De acordo com os dados da Pesquisa Nacional por Amostra de Domicílios Contínua (PNAD) 2019, dentre a população de 14 a 29 anos (50 milhões de pessoas), 20% não concluíram alguma etapa da educação básica, tanto por motivo de abandono, quanto pelo fato de nunca terem frequentado o ambiente escolar.

Em 2019, a taxa de analfabetismo das pessoas de 15 anos ou mais de idade foi de 6,6%, o que representa 11 milhões de analfabetos (INEP, 2019). Tal número é menor do que o número de 2018, que foi de 6,8%, uma diminuição de aproximadamente 200 mil pessoas. Em relação aos brasileiros com 25 anos ou mais de idade que concluíram a etapa de educação básica obrigatória (até o ensino médio) a proporção foi de 47,4% em 2018 para 48,8% em 2019 (INEP, 2019). No Plano Nacional de Educação – PNE 2014-2024 estão previstas as metas de elevar a escolaridade média da população de 18 a 29 anos e erradicar o analfabetismo.

Entre as estratégias para atingir tais metas está a oferta gratuita de Educação de Jovens e Adultos (EJA), assim como a garantia do acesso gratuito a exames de certificação, tais como o Exame Nacional para Certificação de Competências de Jovens e Adultos (Encceja).

Embora o Encceja seja uma política relacionada à EJA, as duas não se confundem. O Encceja representa um exame de certificação de larga escala, que adota a Teoria da Resposta ao Item (TRI), aplica instrumentos dedicados a aferir o aprendizado, e dispõe de questionário socioeconômico preenchido pelos participantes; enquanto a EJA é uma modalidade de ensino que proporciona a conclusão em tempo menor que o Ensino Regular.

Observa-se que as matrículas da EJA vêm diminuindo nos últimos anos: tendo passado de 3.598.716 em 2017, para 3.002.749 em 2020, uma redução de aproximadamente 15% (INEP, 2021), embora de 2018 para 2019 tenha ocorrido migração de aproximadamente 500 mil estudantes do Ensino Regular para a EJA (CENSO ESCOLAR, 2020).

Por outro lado, a participação de jovens e adultos no Encceja vem aumentando, em 2014 o total de inscritos foi de 148.881, nos anos de 2015 e 2016 o exame não foi aplicado. Em 2017, um total de 1,6 milhão de brasileiros participou do exame, seguido por 1,7 milhão em 2018. Já em 2019 houve expressivo aumento da participação, que contou com um total de 2,9 milhões de inscritos (INEP, 2019).

Embora a EJA e o ENCCEJA tenham naturezas distintas, do ponto de vista da elevação da escolaridade da população e da busca por certificação de jovens e adultos para o ensino fundamental e médio as políticas são semelhantes. Assim, da mesma forma que se observa rivalidade entre as matrículas no Ensino Regular e na EJA (TAVARES et al. 2014), é possível que esteja ocorrendo uma migração dos estudantes dos cursos de Educação de Jovens e Adultos para busca por certificação via ENCCEJA, este fenômeno é apontado por Ventura e Oliveira (2020) como desescolarização da EJA.

Assim, considerando a relevância da política de educação de jovens e adultos, esta pesquisa tem por objetivo avaliar o desempenho dos estudantes de EJA no ENCCEJA em relação a indivíduos que não cursaram EJA e que não estão estudando, sob a hipótese de que os estudantes da EJA seriam mais escolarizados do que os indivíduos que evadiram do sistema escolar e não estão estudando. O ano analisado será o ano de 2019, pois o ano de 2020 foi um ano atípico devido à Pandemia COVID-19, que teve impacto em diversas áreas incluindo a educação e seus formatos de ensino.

A metodologia adotada para a estimação econométrica é um *Propensity Score Matching*, a partir dos microdados do ENCCEJA, divulgados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Os resultados da análise têm por objetivo demonstrar se a política Educação de Jovens e Adultos apresenta-se como um fator favorável à certificação no ENCCEJA.

Após esta introdução é apresentado o referencial teórico, e em seguida, passa-se a apresentação da base de dados a ser utilizada e as estatísticas descritivas. Na seção seguinte é descrita a estratégia empírica, seguida da discussão dos resultados. Por fim serão apresentadas as considerações finais.

3.2 CONTEXTO HISTÓRICO DA EDUCAÇÃO DE JOVENS E ADULTOS NO BRASIL

No Brasil, até meados da década de 1940 a educação de jovens e adultos praticamente não se diferenciava como política ou metodologia pedagógica específica. Embora, a partir da década de 1930, com a Era Vargas e a criação do Ministério da Educação em 1930, o Estado tenha assumido um papel maior na implementação de políticas públicas, entre elas as políticas educacionais, conforme descrito por Haddad & Di Pierro (2000).

Os esforços empreendidos desde 1930 a partir de regulamentação, destinação de recursos e instituição de programas, para a educação de jovens e adultos, fizeram com que os índices de analfabetismo de pessoas acima de cinco anos de idade caíssem de 72% (censo de 1920) para 46,7% no censo 1960 (HADDAD E DI PIERRO, 2000).

À época, havia preocupação com a prática pedagógica aplicada a população de jovens e adultos, uma vez que a condição desses estudantes diferia da condição de alunos em idade regular. Em consequência dessa distinção, acreditava-se que práticas pedagógicas alternativas deveriam ser adotadas aos jovens e adultos.

Nesse contexto, janeiro de 1964, foi instituído o Programa Nacional de Alfabetização do Ministério da Educação e Cultura, pelo Decreto nº 53.465, no Art. 1º da norma foi previsto que a metodologia de ensino a ser adotada seria o uso do “Sistema Paulo Freire”. De acordo com Manfredi (1978 p. 15):

“o Sistema de Educação proposto por Paulo Freire distinguia-se dos que o precederam pelas seguintes características: incluía uma técnica de alfabetização que possibilitava o domínio da leitura e da escrita no prazo mínimo de 40 horas e permitia que os conteúdos culturais inerentes ao processo de alfabetização se voltassem para o exame crítico de problemas concretos (sociais, políticos, econômicos) vividos pelos alfabetizandos. Particularmente, Este segundo aspecto inovador do método garantia, pelo menos potencialmente, a possibilidade de a educação vir a ser empregada como um dentre os diversos mecanismos de ampliação da participação política.”

Em que pese a instituição do “Sistema Paulo Freire”, com o advento do período militar a partir de março de 1964, houve uma mudança na condução das políticas educacionais, no qual a repressão a movimentos de educação e cultura popular passaram a ocorrer. Nesse período, o Programa Nacional de Alfabetização foi interrompido, teve seus materiais apreendidos, além da prisão dos seus dirigentes (HADDAD & DI PIERRO, 2000).

Em 1967 foi instituído pela Lei nº 5.379, de 15 de dezembro, o Movimento Brasileiro de Alfabetização – Mobral. No desenho do programa, estavam previstos, dentre

outros, objetivos relacionados à obrigatoriedade do ensino, para a idade de 7 a 14 anos; à ampliação da escolaridade até a 6ª série; à assistência educativa aos analfabetos com idade entre 10 a 14 anos e alfabetização e educação continuada para os analfabetos com 15 anos ou mais.

Em conformidade com a exposição de Haddad (2000) há uma marca de preconceito acerca da Educação de Jovens e Adultos, que é considerado como um campo de “segunda linha” tanto entre os professores, quanto entre os trabalhadores das secretarias de educação e escolas. O autor observa que o preconceito existe até mesmo entre os próprios alunos. Conforme Leony (2013) tal preconceito tem raízes no programa Mobral, onde as práticas pedagógicas foram abandonadas e o enfoque foi nos índices educacionais brasileiros. Ainda a palavra “mobral” é utilizada popularmente com tom pejorativo.

3.3 A AVALIAÇÃO DA EDUCAÇÃO DE JOVENS E ADULTOS

A avaliação da qualidade e dos resultados da educação de jovens e adultos é um tema pouco explorado pela literatura brasileira (SILVA, 2022). Sabe-se que para além do desempenho dos estudantes auferido em provas, a educação de jovens e adultos visa promover o senso crítico e desenvolver habilidades e competências técnicas necessárias à vida em sociedade (TELES E SOARES, 2016). Todavia, é importante avaliar os resultados alcançados pela educação de jovens e adultos e fornecer subsídios aos gestores desta política pública.

Sobre a avaliação da EJA foi identificada a pesquisa de Silva (2022), que buscou investigar os diferenciais de desempenho entre os estudantes da EJA e do ensino médio regular, a partir da análise das notas em cada uma das cinco disciplinas das provas do ENEM (Ciências Humanas e suas tecnologias, Ciências da Natureza e suas tecnologias, Linguagens, Códigos e suas tecnologias, Matemática e suas tecnologias, Redação). A pesquisa foi realizada a partir dos microdados do ENEM e do Censo Escolar de 2018. A metodologia econométrica adotada foi o *Propensity Score Matching*. Os resultados apontam que as médias de desempenho (nota) dos estudantes do ensino regular superam as dos estudantes da EJA em todas as provas. As disciplinas em que esses diferenciais foram maiores foram em escrita (redação) e matemática, com diferença de 83,81 e 32,99 pontos a favor dos alunos do sistema regular de ensino, de um total de 1.000 pontos possíveis.

Ainda sobre a avaliação da EJA, com o intuito de verificar os diferenciais entre o de desempenho no ENEM de estudantes oriundos do Ensino Regular e da EJA, que buscavam

certificalçai, Da Silva et al. (2020), avaliaram respostas dos questionários socioeconômicos dos candidatos que prestaram o ENEM, divulgados pelo INEP. O estudo contemplou 1.329.734 candidatos oriundos do Ensino Regular e 111.326 da EJA que prestaram o Enem 2014. Os principais resultados apontam que 26% dos candidatos oriundos do Ensino Regular obtiveram a certificação, enquanto para os estudantes provenientes da EJA o percentual de certificação foi de 16%. Adicionalmente, a autora identifica que é semelhante o perfil socioeconômico de ambos os grupos que buscam certificação.

Ainda que não se refira à avaliação específica da EJA a pesquisa Oliva et al. (2014) apresenta *insights* sobre o desempenho no mercado de trabalho dos estudantes da EJA em relação aos estudantes da educação técnica e profissional (ETP). O estudo considerou indivíduos que não possuíam ensino médio completo. A base de dados utilizada para o estudo são os microdados da Pesquisa Mensal de Emprego (PME), realizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE), para o período de janeiro de 2002 a dezembro de 2013. O método adotado foi o de diferenças em diferenças com efeito fixo no nível do indivíduo. Os resultados indicam que cursar ETP tem mais impacto sobre a ocupação no mercado de trabalho do que cursar EJA. Também foram encontradas evidências de diferenças no impacto sobre os rendimentos de quem cursa ETP, chegando a 19 p.p para os homens que trabalham no setor privado.

Em pesquisa sobre a Educação de Jovens e Adultos e o Ensino Médio Regular, Tavares et al. (2014) identificaram a ocorrência de rivalidade entre as duas modalidades de ensino. A base de dados utilizada foi a da Pesquisa Mensal de Empregos (PME) de 2002 a 2010. A metodologia econométrica consiste em uma regressão com descontinuidade, que explora a descontinuidade da regra de idade para a elegibilidade ao ensino médio na modalidade EJA (17 anos à época). Os autores observaram que quando os alunos passam a ser elegíveis para a EJA, ocorre redução de dois a três pontos percentuais nas matrículas do ensino médio, o que caracteriza a rivalidade entre as duas formas de ensino. Adicionalmente, os autores apontam que indivíduos mais jovens (entre 19 e 20 anos) estão mais propensos a abandonar a modalidade de ensino médio regular e cursar EJA do que indivíduos na mesma idade que estão fora da escola. Já os estudantes de EJA com mais idade, são aqueles que abandonaram a escola no passado e retornaram para o sistema educacional.

Ventura e Oliveira (2020) analisaram a substituição do oferecimento da Educação de Jovens e Adultos, em escolas, pela oferta de certificação via Encceja. O estudo foi realizado para a rede pública estadual do estado do Rio de Janeiro. Inicialmente foi trazida como demanda potencial a quantidade de brasileiros com mais de 15 anos que são

analfabetos no Brasil (11,3 milhões), em confronto com 3,5 milhões de estudantes matriculados na EJA, em 2018. Os autores apontam que a demanda potencial da EJA nos municípios do estado do Rio de Janeiro varia entre 10% e 23% (pessoas com 18 anos ou mais) e que a oferta de EJA e números de matrículas vêm sendo reduzidas. A região metropolitana apresentou uma taxa de atendimento da demanda de 6%, enquanto as regiões com maiores porcentagens não atingiram nem 20%. Por fim é apontada como causa da desescolarização da EJA o fechamento de turmas e sucessivas reduções na oferta de vagas, junto com a adesão aos mecanismos de certificação.

As demais pesquisas identificadas na literatura tratam de identificar preponderantemente o perfil dos estudantes da EJA. Dos Santos e Santos (2021), utilizaram os microdados do Encceja nos anos de 2018 e 2019 para avaliar o perfil dos participantes do exame. Os pesquisadores destacam que os indivíduos possuem majoritariamente idade inferior a 29 anos, em especial os que buscam certificação para o ensino médio, e que há que se desenvolver uma reflexão sobre a procura de um público jovem pelo exame, pois pode estar havendo uma migração da EJA para o Encceja. Ressaltam ainda que a política pode estar sendo utilizada para correção do fluxo escolar e como uma forma de lidar com o problema da distorção idade-série, na garantia do direito à educação. Destacam ainda que os participantes possuem renda domiciliar baixa e que em geral estão inseridos no mercado de trabalho.

Ferreira e Martinelli (2016) acompanharam 67 estudantes matriculados regularmente na etapa de ensino fundamental da EJA, em cinco escolas da rede pública municipal de São Paulo. A pesquisa teve por objetivo, caracterizar o perfil do grupo de estudantes da EJA e avaliar o desempenho escolar desses estudantes. Os autores destacam que os estudantes são jovens, sendo 21% com idade entre 14 e 24 anos, 42% com idades entre 27 e 38 anos e 37% com 40 a 61 anos. Outra característica destacada é que em relação ao gênero, a maioria das mulheres participantes era casada, enquanto a maioria dos homens era jovem e solteiro. Em relação ao desempenho escolar, foi observado baixo desempenho em escrita e matemática e bom desempenho dos estudantes em leitura.

Assim, esta pesquisa se insere na literatura utilizando de forma inédita, segundo nosso conhecimento, os dados do ENCCEJA disponibilizados pelo INEP, com o objetivo de fornecer uma avaliação sobre a qualidade da educação de jovens e adultos, mensurada pela aprovação dos estudantes da modalidade EJA, no Exame de certificação.

3.4 ESTRATÉGIA EMPÍRICA

A estratégia de identificação para esta pesquisa consiste na designação dos indivíduos que compõem a amostra em dois grupos. O primeiro é o grupo de tratados, que contém os indivíduos que estavam cursando EJA no momento da realização do Enceja, nosso grupo de interesse. O segundo grupo refere-se à amostra controle, que contempla indivíduos que não estavam estudando no momento da realização do exame.

A técnica utilizada é a de pareamento, que visa a criação de um grupo de controle semelhante ao grupo de tratamento com base em determinadas características observáveis. Como a participação no curso de EJA não ocorre estritamente ao acaso é realizado também controle de características observáveis pré-pareamento.

Há evidências na literatura de que a combinação de diversos tipos de pareamento (*matching*) confere robustez aos resultados. Assim, para esta pesquisa foram combinados os três seguintes métodos: i) o *Propensity Score Matching (PSM)*, que consiste no cálculo da probabilidade de ser tratado, dadas suas características observáveis, proposto por Rosenbaum e Rubim (1993); ii) o *Coarsened Exact Matching (Cem)*, que consiste na realização de balanceamento da amostra pré-matching (Iacus, King, and Porro (2008) ; BLACKWELL et al., 2009; iii) o *Entropy Balancing (EB)*, que realiza um pré-tratamento entre as variáveis no qual é gerada uma função peso que visa garantir um melhor ajustamento entre os grupos. Nas seções seguintes as metodologias serão descritas detalhadamente.

3.4.1 PROPENSITY SCORE MATCHING (PSM)

Esta pesquisa propõe avaliar se a participação em cursos da Educação de Jovens e Adultos tem impacto sobre a aprovação no exame nacional de certificação de ensino fundamental e médio. O método a ser utilizado será o *Propensity Score Matching*, proposto por Rosenbaum e Rubim (1993). Na estimação do PSM, nosso parâmetro de interesse será o efeito médio do tratamento sobre os tratados (*Average Treatment Effect on Treated - ATT*). Para o cálculo do ATT é construído um grupo de controle que representa estatisticamente o que aconteceria com o grupo de tratados na ausência do tratamento, situação essa não observável.

Segundo Khandker et al. (2009) a variável de interesse, a *dummy* de tratamento e o vetor de características observáveis são definidas, respectivamente, por:

$$Y_i, T_i \text{ e } X_i \tag{1}$$

Para estudar o ATT, duas condições são necessárias. A primeira é a independência condicional ou seleção em observáveis, na qual o grupo de controle é exógeno à variável de resultado potencial, de forma que:

$$Y_i(0) \perp T_i | X_i. \quad (2)$$

Tal condição implica que, condicional às variáveis observáveis (vetor X), o tratamento seja aleatório entre os grupos de tratamento e controle. Assim, o resultado de um indivíduo do grupo de controle é um bom preditor para o resultado potencial de um indivíduo do grupo de tratamento, desde que possuam o mesmo vetor X' de variáveis observáveis.

A segunda condição, segundo Khandker et al. (2009), é a existência de suporte comum, ou seja, os participantes da EJA (tratados) devem ser semelhantes em características observáveis ao grupo de não participantes do EJA (controle) de forma que ao observar essas características do indivíduo, no suporte comum, não seja possível distinguir de qual grupo (tratado ou controle) o indivíduo faz parte, de forma que:

$$0 < pr[T_i = 1 | X_i] < 1 \quad (3)$$

Neste trabalho, o indivíduo é considerado tratado se cursar a modalidade de ensino de Educação de Jovens e Adultos, caso em que uma variável *dummy* assume valor 1, se o estudante não cursou EJA o valor assumido pela *dummy* é 0.

A condição de suporte comum garante que o grupo de tratados tenha observações comparativas “próximas” à distribuição do escore de propensão (HECKMAN, LALONDE E SMITH, 1999). O escore de propensão representa o cálculo da probabilidade de pertencer ao grupo de tratados, dadas as suas características observáveis e é dado por:

$$p(X) = pr[T = 1 | X] \quad (4)$$

e a hipótese de independência condicional é dada por:

$$Y_i(0) \perp T_i | p(X_i) \quad (5)$$

Logo, a propensão de cursar EJA, dadas as características observadas, será estimada a partir de um probit, e o efeito médio de cursar EJA sobre a aprovação nas diferentes provas do Encceja (ATT) será calculado pela equação:

$$ATT = E_{p(x)|T=1}\{E[Y_i(1)|T = 1, p(x)] - E[Y_i(0)|T = 0, p(x)]\} \quad (6)$$

Para esta pesquisa o tipo de pareamento adotado foi o do vizinho mais próximo (*Nearest Neighbor Matching – NNM*), em que cada participante do Encceja tratado (àqueles que cursam ou cursaram EJA) é pareado à um participante que não seja tratado mais próximo, avaliado em termos de escores de propensão. No NNM admite-se o tipo de pareamento com e sem reposição. No tipo com reposição, os indivíduos não tratados podem

ser utilizados mais de uma vez como par dos indivíduos tratados. Já no tipo sem reposição cada indivíduo é utilizado como par apenas uma vez (CALIENDO, KOPEINIG, 2005).

3.4.2 ANÁLISE DE SENSIBILIDADE

A hipótese de identificação do PSM pressupõe a ausência de características não observáveis que afetem a chance de participação dos indivíduos no programa. Assim a participação no programa depende de um vetor de características individuais (X').

Considerando que a probabilidade de receber o tratamento (π_i) seja uma função $\lambda(x_i)$ do vetor X que contém as observáveis de cada indivíduo (i). Caso dois indivíduos (i e j) que possuam as mesmas características ($X_i = X_j$) tenham probabilidade distinta de receber o tratamento ($\pi_i \neq \pi_j$) é possível que esteja ocorrendo viés por características não observáveis ou não controladas no estudo.

Não é possível testar esta hipótese, todavia a análise de sensibilidade proposta por Rosenbaum (2002) permite avaliar o quanto a ocorrência de variáveis não observadas poderia influenciar os resultados do estudo. Considerando que as chances de receber o tratamento para cada indivíduo (i, j, k, \dots) sejam dadas por $\frac{\pi_i}{1-\pi_i}$, a razão de chances entre os indivíduos é representada por um número $\gamma \geq 1$, tal que:

$$\frac{1}{\gamma} \leq \frac{\pi_i(1-\pi_j)}{\pi_j(1-\pi_i)} \geq \gamma, \quad \forall i \text{ e } j \text{ com } x_i = x_j \quad (7)$$

Quando as chances de receber o tratamento são iguais entre os indivíduos o $\gamma = 1$ e não há viés. Se γ assumir um valor maior do que um, como por exemplo, $\gamma = 2$, há indícios de que, embora os indivíduos sejam iguais em observáveis, existam variáveis não observáveis ou não controladas no estudo que podem afetar as chances de participação no programa.

Intuitivamente o teste de Rosenbaum (2002) indica que, se há sensibilidade dos resultados para valores muito próximos de $\gamma = 1$, podem existir potenciais problemas de variáveis omitidas no modelo.

No caso desta pesquisa, a ocorrência do problema implicaria que os alunos pareados no PSM seriam aparentemente semelhantes em termos das características utilizadas no pareamento, mas difeririam em alguma variável não observada ou não controlada, com implicação nas probabilidades de participação na EJA, de forma que a hipótese de identificação do PSM poderia ser questionada.

3.4.3 COARSENEDED EXACT MATCHING (CEM)

Os métodos de pareamento tradicionais são realizados por semelhança de variáveis observáveis entre os indivíduos dos grupos de tratamento e controle. Assim, buscam para cada unidade tratada, pelo menos um indivíduo semelhante no grupo de controle.

O *Coarsened Exact Matching (CEM)* é um algoritmo de pré-processamento de dados, que combina os valores exatos das variáveis observáveis de uma unidade tratada com todas as unidades de controle disponíveis na amostra. Na estimação do *CEM* são inseridos intervalos (Bins) entre os valores originais da variável X para identificar grupos substancialmente representativos dessas variáveis.

Assim, o *CEM* busca reduzir temporariamente as variáveis em grupos significativos, fazer o *matching* exato das variáveis reduzidas e então reter apenas os valores originais dos dados correspondentes (BLACKWELL et al., 2009; IACUS; KING, 2008; IACUS; KING; PORRO, 2011).

Conforme exposto por Blackwell et al. (2009) o algoritmo funciona em observância às seguintes etapas:

- i) Realiza uma cópia das características observáveis presentes no vetor X' , denotada X^* ;
- ii) Agrupa X^* conforme o algoritmo de *binning* automático do *CEM*, ou conforme definido pelo usuário;
- iii) Cria um estrato para cada observação de X^* e inclui cada observação nesse estrato;
- iv) Atribui os estratos aos dados originais (X') e elimina todas as observações que, no estrato, não contenham pelo menos uma unidade tratada e uma unidade de controle.

Após a conclusão das etapas os estratos criados constituirão as bases para o cálculo do efeito do tratamento, que pode ser estimado a partir de uma regressão linear, ou pela diferença de médias entre os grupos de tratamento e controle, ou ainda pela combinação dos seus resultados com outros métodos.

Entre as vantagens da utilização desta técnica de pré-processamento, destaca-se: que o *CEM* atende ao “*congruence principle*”, que afirma que o espaço de dados e o espaço de análise devem ser iguais; a melhora no desequilíbrio entre os grupos de tratamento e de controle; e sua eficiência para grandes conjuntos de dados. Uma limitação do *CEM* é que, devido à riqueza de covariadas normalmente presente em pesquisas, o algoritmo pode produzir poucas correspondências.

3.4.4 ENTROPY BALANCING (EB)

O *Entropy Balancing (EB)* é um método multivariado proposto por Hainmueller (2012), no qual é possível reponderar uma base de dados de forma que as distribuições dessas variáveis observem um conjunto de condições de momentos especificadas (o primeiro momento representa a média, o segundo a variância, e o terceiro assimetria, de uma distribuição de probabilidades). Neste método, aplicável a estudos com tratamento binário, os dados do grupo de controle podem ser reponderados para corresponder aos momentos covariáveis no grupo de tratamento.

A seguir, segue a formalização da reponderação da média contrafactual, momento em geral utilizado no EB, descrita por Hainmueller, Jens; XU, Yiqing, (2013):

$$E[y(0)|\widehat{T} = 1] = \frac{\sum_{\{i|T=0\}} Y_i \omega_i}{\sum_{\{i|T=0\}} \omega_i} \quad (8)$$

Na qual, ω_i é o peso atribuído para as unidades de controle e o peso utilizado na reponderação é a função

$$\min_{\omega_i} H(\omega) = \sum_{\{i|T=0\}} \omega_i \log \frac{\omega_i}{\varphi_i} \quad (9)$$

Sujeita às restrições de equilíbrio e normalização

$$\sum_{\{i|T=0\}} \omega_i c_{ri}(X_i) = m_r \quad \text{com } r \in 1, \dots, R \quad e \quad (10)$$

$$\sum_{\{i|T=0\}} \omega_i = 1 \quad e \quad (11)$$

$$\omega_i \geq 0 \quad \forall i / T = 0 \quad (12)$$

Onde $\varphi_i = \frac{1}{n_0}$ é o peso base, e $c_{ri}(X_i)$ descreve o conjunto de restrições do balanceamento impostas aos momentos das observáveis do grupo de controle reponderado.

Em síntese, deve-se escolher as covariáveis que serão incluídas no procedimento, e para cada uma delas são especificadas um conjunto de restrições de equilíbrio (equação 9) para que os momentos das distribuições das covariadas sejam reponderadas entre os grupos de tratamento e controle. Assim, no EB os pesos são estimados previamente às análises de avaliação entre os grupos, diferentemente do que ocorre no pareamento por escore de propensão.

3.5 ANÁLISE EXPLORATÓRIA DOS DADOS

A fonte de dados para esta pesquisa são as respostas prestadas pelos participantes do Enceja aos questionários socioeconômicos do exame, assim como as variáveis de desempenho nas provas. A inscrição no Enceja é voluntária e gratuita e pode ser realizada por tipo de prova. No total são quatro provas distintas, quais sejam: Linguagens e Redação, Matemática, Ciências e Humanidades e Ciências da Natureza.

As provas são aplicadas para certificação de ensino fundamental ou médio. Para obter a certificação, o participante do exame deve atingir a nota mínima de 50% de acertos nas quatro provas e na prova de redação. Há também a possibilidade de declaração parcial de proficiência, casos em que o participante obtém aprovação em uma ou mais provas, mas não em todas elas.

Em 2019 houve um total de 2.973.386 inscritos no exame, sendo que havia sido estimado pelo órgão um total de 1.942.627. O valor total despendido pelos cofres públicos para a aplicação das provas foi de quase 140 milhões (R\$138.981.074,17), de acordo com os dados solicitados pelos autores ao INEP, por meio da Lei de Acesso à Informação. Observa-se pelos microdados do Enceja que há um elevado número de inscritos que não compareceram às provas.

Assim, considerando o elevado quantitativo de ausentes ao exame, com o intuito de obter uma amostra confiável, foram excluídos da base de dados os indivíduos que se inscreveram no exame, responderam ao questionário, mas não participaram da prova. Dessa forma, a amostra inicial de 2.973.376 restou em 1.212.606 pessoas.

O grupo de tratados é composto por indivíduos que estavam cursando EJA na data do exame. A variável para identificar tal grupo foi criada pela combinação das respostas de duas questões, a primeira é se o indivíduo cursa ou já cursou EJA, e a segunda é sobre se deixou de cursar EJA. Assim, o grupo é composto por aqueles que cursam EJA e não evadiram da modalidade. Cabe aqui destacar que a EJA é uma modalidade de ensino com período de conclusão mais célere que o ensino regular, todavia EJA e Enceja não se confundem e um não é requisito para certificação no outro. Ressalta-se que, se o indivíduo está cursando EJA e busca certificar-se via Enceja, pode se inferir que há intenção de abandonar o curso de EJA caso a certificação por meio do exame ocorra.

O grupo de controle contempla indivíduos que não estão estudando e que não cursam EJA. Para realizar uma análise mais fidedigna possível dos resultados da EJA, excluiu-se da base de dados àqueles participantes do Enceja que estavam cursando o Ensino Regular na

data do exame (88.139). Foram também excluídos da amostra indivíduos idosos⁷, ou seja, aqueles com 60 anos ou mais, assim a base de dados restou em 256.906 estudantes da EJA e 896.000 indivíduos que não estão estudando e não cursaram EJA, ou seja, estão fora dos sistemas de ensino.

A análise desta pesquisa contempla a aprovação por tipo de certificação (Certificação Ensino Médio e Ensino Fundamental) e em cada diferente tipo de prova (Linguagens e Redação, Matemática, Ciências e Humanidades e Ciências da Natureza), sendo avaliado o aumento médio na probabilidade de aprovação, dado que o indivíduo participa de curso na modalidade de Educação de Jovens e Adultos.

Uma das limitações deste trabalho refere-se à base de dados, pois não há informações sobre em qual escola o indivíduo está cursando EJA, o que seria relevante para controlar questões de qualidade, que podem afetar os resultados. Cita-se também a ausência na base de dados de informação sobre a cor dos indivíduos, dado este relevante para compreender desigualdades e especificidades das diferentes populações.

Diante do exposto, as variáveis consideradas para o pareamento são comumente utilizadas na literatura que estuda sobre educação e contemplam: idade, gênero, renda individual em salários mínimos, idade em que evadiu da escola, se o pai possui ensino médio, se a mãe possui ensino médio, além dos 27 estados da federação.

As características dos grupos de estudantes da EJA (tratados) e daqueles que não estão estudando (controles) são apresentadas na Tabela 1 a seguir.

⁷ Estatuto do Idoso (Lei nº 10.741, de 1º de outubro de 2003.)

Tabela 1 - Estatísticas descritivas da amostra

Variável	Certificação Ensino Fundamental				Certificação Ensino Médio			
	Cursa EJA (T) Nº=38.008		Não Cursa EJA (C) Nº=173.949		Cursa EJA (T) Nº=218.898		Não Cursa EJA (C) Nº=722.051	
	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão
Idade	30,81	11,12	33,00	11,60	29,89	9,46	30,33	10,04
Feminino	0,55	0,49	0,54	0,49	0,53	0,49	0,51	0,49
Renda Individual (SM)	4,30	5,58	4,52	5,79	3,51	4,88	3,50	4,91
Ano do Ensino Evadiu da Escola	5,19	1,52	5,11	1,59	7,41	1,90	7,72	1,89
Mãe possui Ensino Médio	0,21	0,41	0,18	0,38	0,29	0,45	0,28	0,45
Pai possui Ensino Médio	0,15	0,36	0,13	0,33	0,22	0,42	0,22	0,41

Fonte: Elaboração própria

A tabela 1 contém estatísticas descritivas dos indivíduos que compõem a amostra deste estudo e que cursam EJA, assim como dos indivíduos que não estão estudando. Também foram consideradas as diferentes provas aplicadas, tendo em vista a evidência na literatura de que o desempenho dos estudantes da EJA varia a depender da prova realizada (SILVA, 2022; FERREIRA E MARTINELLI, 2016). Assim, há um total de oito grupos distintos analisados.

Em relação ao grupo que busca certificação para o ensino fundamental (CEF), observa-se que a proporção de estudantes da EJA é de 18%, enquanto os indivíduos que cursam EJA e buscam certificação para o ensino médio (Ceme) totalizam 23%.

Ao analisar a idade dos estudantes da EJA (CEF) observa-se que a idade média é de 31 anos, enquanto aqueles que não estudam (CEF) possuem 33 anos em média. Já o grupo que busca Ceme é mais jovem, sendo sua idade média de 30 anos para ambos os grupos, pois a diferença fica por conta de o desvio padrão ser maior entre aqueles que não estudam, a diferença de desvio padrão pode estar sendo influenciada pelo maior tamanho da amostra.

A questão dos estudantes da EJA serem jovens já foi apontada pela pesquisa de Dos Santos e Santos (2021). Um dado que corrobora tal fato é a questão da rivalidade entre ensino regular e a educação de jovens e adultos, tendo sido observado por Tavares et al. (2014) que, no momento em que os estudantes do ensino regular atingem a idade mínima para cursar EJA, ocorre aumento na probabilidade de migração entre as modalidades.

Em relação ao gênero, a proporção de mulheres estudantes da EJA é maior tanto para a certificação em nível fundamental quanto para nível médio, sendo elas 55% do primeiro grupo e 53% do segundo. No grupo dos indivíduos que não estudam o gênero feminino representa 54% dos que buscam CEF e 51% dos que buscam Ceme. Ferreira e Martinelli

(2016), em estudo realizado sobre a EJA, identificaram que a maior parte das estudantes é do gênero feminino.

A variável renda foi gerada por número de salários mínimos, variando este de nenhuma renda (0) até 15 salários mínimos. É conhecido pela literatura que escolaridade e renda possuem correlação positiva. A literatura aponta ainda que os estudantes da EJA são indivíduos de baixa renda (DOS SANTOS E SANTOS; 2021). Em termos médios, a renda dos que buscam CEF é maior do que os indivíduos que buscam Ceme, todavia o desvio padrão daquele grupo também é maior.

A variável referente ao ano em que evadiu contempla o ano em que o estudante evadiu e assume valor zero quando não frequentou a escola, valor um quando evadiu no primeiro ano do ensino fundamental, e assim sucessivamente, até o terceiro ano do ensino médio. Uma vez que foram excluídos da amostra aqueles indivíduos que ainda estão estudando no ensino regular, a amostra desta pesquisa trata de pessoas evadidas do sistema escolar.

Em que pese esta pesquisa não trate da evasão escolar, cabe aqui ressaltar que a amostra (participantes do Encceja) contempla indivíduos em distorção idade-série, pelas próprias regras de participação no exame (limites de idade). Assim, uma variável utilizada no pareamento e considerada relevante é o momento em que o indivíduo evadiu da escola.

Assim, ao analisar o ano em que os indivíduos evadem, observa-se que o grupo que busca CEF evadiu em média no quinto ano, enquanto os indivíduos do grupo CEF evadiram em média após o sétimo ano. Como a avaliação é em termos de média, pode-se inferir que aqueles que evadiram do ensino regular antes do sétimo ano e não cursaram EJA, já realizaram o exame de certificação. Na literatura sobre economia da educação é apontado por Leon e Menezes-Filho (2002) que há um efeito diploma, caracterizado pela manutenção do aluno na escola, até a conclusão do ciclo escolar.

Em estudos sobre a educação, há também evidências de que a escolaridade dos pais é relevante para a educação dos filhos (CURI, MENEZES-FILHO, 2009; VIEIRA et al., 2016; ANAZAWA et al., 2016). A variável sobre escolaridade representa uma *dummy*, que assume valor 1 quando os pais possuem ensino médio ou nível maior de ensino e 0, caso a escolaridade seja menor do que o ensino médio.

Observou-se que tanto as mães como os pais dos indivíduos da amostra possuem baixa escolaridade. A escolaridade dos pais está correlacionada diretamente com a escolaridade dos filhos da amostra. A menor proporção de mães e pais com ensino médio completo está entre os evadidos que buscam certificação para EF, sendo de 18% as mães que

possuem ensino médio e 13% dos pais. Uma maior quantidade média de mães possui ensino médio no grupo de estudantes da EJA, são 21%, enquanto os pais do mesmo grupo são 15%. No grupo que busca Ceme, 29% das mães dos estudantes da EJA possuem ensino médio, enquanto no grupo dos que não estudam o percentual é de 28%. Para a escolaridade dos pais, ambos os grupos tem 22% com ensino médio.

Nas estimações do PSM foram considerados cada um dos 27 estados da federação, todavia, para fins de apresentação dos dados, na tabela 2, a seguir, os estados são compilados nas cinco regiões do Brasil, para fins de uma melhor análise da distribuição dos participantes do exame.

Tabela 2 - Distribuição geográfica dos participantes do exame

Variável	Certificação Ensino Fundamental		Certificação Ensino Médio	
	Cursa EJA (T) Nº=38.008	Não Cursa EJA (C) Nº=173.949	Cursa EJA (T) Nº=218.898	Não Cursa EJA (C) Nº=722.051
	Média	Média	Média	Média
Norte	0,12	0,12	0,09	0,09
Nordeste	0,10	0,11	0,13	0,15
Centro-oeste	0,28	0,36	0,37	0,40
Sudeste	0,14	0,14	0,13	0,11
Sul	0,36	0,27	0,28	0,25

Fonte: Elaboração própria

Conforme dados estimados pelo Instituto Brasileiro de Geografia e Estatística, em 2021 as cinco regiões brasileiras contemplavam a seguinte distribuição da população: Norte 9% (18.906.962); Nordeste 27% (57.667.842); Centro-oeste 8% (16.707.336); Sudeste 42% (89.632.912); Sul 14% (30.402.587).

Em relação a distribuição pelas regiões do país, a amostra não guarda relação com a distribuição da população. Pode-se verificar que o maior percentual de inscritos para CEF, que cursam EJA é na Região sul (36%). Já para os que não estão estudando e buscam CEF a maior concentração é no centro-oeste. Para Ceme a distribuição é concentrada no Centro-Oeste.

Norte e Nordeste são as regiões do país com menores taxas de escolarização da população e com maiores taxas de evasão escolar entre os jovens, conforme dados da PNAD 2022, todavia, não são os locais com maior incidência de participantes da EJA e do Enceja. Tal fato pode indicar a existência de potencial público alvo que não está acessando ao programa. Um fator que pode explicar a distribuição dos participantes é a questão da oferta do exame. Em análise ao edital do Enceja 2019, observou-se que esta prova não chega a todos os municípios. De um total de 5.568 o exame é aplicado em 611 municípios. Em relação a distribuição nas regiões, Norte recebe o exame em 16% dos seus municípios, Nordeste em

apenas 5,9%, Centro-Oeste em 13,5%, sudeste e sul aplicam o exame em 13% dos seus municípios.

Outra possível causa da distribuição dos participantes pode ser as próprias condições de baixa escolaridade e desigualdade socioeconômica dessas pessoas, pois o norte e nordeste são, igualmente, as regiões com maiores índices de pobreza. Assim, é possível que a política de educação de jovens e adultos não esteja atingindo os mais necessitados para o programa. Uma possível explicação para não acessarem ao Enceja é o fato de a inscrição só ser realizada via internet, assim a amostra em análise já estaria passando por algum tipo de “seleção” ao conseguir se inscrever e realizar o exame.

As variáveis de interesse desta pesquisa são a aprovação em cada uma das quatro provas aplicadas no Enceja. Embora as variáveis possuam os mesmos nomes para o ensino fundamental e médio, os componentes das disciplinas diferem, conforme pode ser analisado a seguir:

- Ensino fundamental: i) Linguagens e Redação, composta pelas matérias de Língua Portuguesa com redação, Inglês, Espanhol, Artes e Educação Física; ii) Ciências da Natureza: Ciências; iii) Ciências Humanas: História, Geografia, Filosofia e Sociologia; e iv) Matemática: Matemática

- Ensino médio: i) Linguagens e Redação, composta pelas matérias Linguagens, Códigos e suas Tecnologias e Redação Língua Portuguesa com redação, Inglês, Espanhol, Artes e Educação Física; ii) Ciências da Natureza: Ciências da Natureza e suas Tecnologias Química, Física e Biologia; iii) Ciências Humanas: Ciências Humanas e suas Tecnologias História, Geografia, Filosofia e Sociologia; e iv) Matemática: Matemática e suas Tecnologias Matemática.

Na tabela 3 a seguir podem ser observadas as estatísticas descritivas da aprovação dos grupos analisados, nas diferentes provas.

Tabela 3 - Estatísticas descritivas de aprovação por disciplina

Variável	Certificação Ensino Fundamental				Certificação Ensino Médio			
	Cursa EJA (T) Nº=38.008		Não Cursa EJA (C) Nº=173.949		Cursa EJA (T) Nº=218.898		Não Cursa EJA (C) Nº=722.051	
	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão
Ciências da Natureza	0,84	0,35	0,86	0,34	0,94	0,24	0,94	0,24
Ciências Humanas	0,77	0,41	0,79	0,40	0,84	0,36	0,85	0,35
Linguagens e Redação	0,73	0,44	0,73	0,44	0,77	0,41	0,78	0,41
Matemática	0,65	0,47	0,64	0,47	0,53	0,49	0,55	0,49

Fonte: Elaboração própria

Em geral observa-se que a maior parte dos indivíduos que prestaram as provas obteve aprovação em cada matéria. Em ordem de aprovação, a disciplina com maior percentual é Ciências da Natureza, seguida por Ciências Humanas. Em terceiro fica Linguagens e Redação (que inclui prova escrita) e por fim Matemática. Esse comportamento foi identificado para todos os grupos analisados.

Para CEF, em Linguagens e Redação, 73% dos participantes do exame obtiveram aprovação, sem diferença de médias entre os grupos. Já em Ciências da Natureza e Ciências Humanas foram identificados menores percentuais entre os estudantes da EJA do que entre os que não estão estudando, ainda que em termos absolutos o percentual seja elevado, 84% e 86% respectivamente.

A disciplina de Matemática é a disciplina com o menor percentual de aprovação entre os grupos, todavia foi a que apresentou maior percentual para os estudantes da EJA (65%), ainda que a diferença tenha sido de um ponto percentual (p.p.), pois os indivíduos que não estão estudando obtiveram 64% de aprovação na prova de Matemática. Observa-se que o desvio padrão também se manteve semelhante entre os grupos. Na seção seguinte são apresentados os resultados e discussão.

3.6 RESULTADOS E DISCUSSÃO

A Educação de Jovens e Adultos é uma política relevante no Brasil. Em 2019 o número de matriculados na modalidade foi de 3.273.668 de indivíduos conforme Censo Escolar daquele ano. Se por um lado a EJA possui um elevado número de matriculados, raros estudos buscam averiguar a sua qualidade em termos de desempenho dos estudantes, conforme apontado por Silva (2022).

Em que pese à proposta desta pesquisa seja avaliar o desempenho dos estudantes da EJA no Enceja, cabe destacar que os resultados devem ser analisados com cautela, tendo em vista que não se sabe a duração do período que o participante do exame estava cursando a educação de jovens e adultos, ou seja, expostos à política.

Ademais, ressalta-se que, embora o desempenho de estudantes da EJA no Enceja seja um indicador da qualidade dessa política de jovens e adultos, de acordo com o Parecer 11/2000 da Comissão de Educação Básica, do Conselho Nacional de Educação, a EJA possui três funções básicas: a função reparadora, que contempla a possibilidade de acesso ao ensino de qualidade a todos que não tiveram esse direito em idade própria; a função equalizadora, a qual possibilita o retorno ao sistema educacional de segmentos específicos da sociedade como donas de casa, migrantes, trabalhadores rurais, aposentados e encarcerados que tiveram sua

escolaridade interrompida por diversos motivos como evasão, repetência ou outras circunstâncias desfavoráveis. Assim, o desempenho dos estudantes da EJA no Enceja não possui o condão de avaliar a política como um todo. Postas estas considerações, passa-se a discussão dos resultados.

O PSM foi estimado por um modelo *probit*, cujas estimativas constam no apêndice A 2. As covariadas utilizadas para o pareamento foram descritas nas tabelas 1, 2 e 3 da seção anterior. Ao todo foram estimados 64 modelos econométricos, uma vez que a análise ocorreu pelos dois tipos de certificação, pelas quatro diferentes provas aplicadas (Linguagens e Redação, Matemática, Ciências da Natureza e Ciências Humanas), e, como medida de robustez, oito metodologias distintas foram adotadas.

No PSM o método de pareamento adotado foi o de vizinho mais próximo (*nearest-neighbor matching*), considerando um vizinho mais próximo, três vizinhos mais próximos e cinco vizinhos mais próximos. Os gráficos da distribuição dos escores de propensão antes e depois do pareamento podem ser observados no Apêndice A2, assim como os resultados do teste de diferença de média antes e depois do pareamento por grupo de controles e tratados.

Conforme já descrito anteriormente, nas seções 3.2 e 3.3, foram também utilizados os métodos de pré-processamento da base de dados CEM e EB. Além da combinação do CEM com EB e do CEM com o PSM, conforme apontado por Iacus, King e Porro (2009) o CEM combinado com outros métodos pode conferir maior confiabilidade aos resultados.

Os resultados constam descritos nas tabelas 4 e 5 a seguir, referentes às análises sobre o ensino médio e ensino fundamental, respectivamente.

Tabela 4 - Certificação Ensino Médio - Efeito médio de ter cursado EJA sobre a aprovação nas provas de Linguagens e Redação, Matemática, Ciências da Natureza e Ciências Humanas (ATT).

Disciplina	NN (1)	NN (1) *	NN (3)	NN (5)	CEM	EB	CEM+EB	CEM+PSM
Linguagens e Redação	-3,37*** (0.0039)	-1,33*** (0.0016)	-2,12*** (0.0025)	-1,53*** (0.0021)	-3,67*** (0.0046)	0.06 (0.0063)	-3,56*** (0.0049)	-2,84*** (0.0070)
Ciências da Natureza	-0,85*** (0.0033)	-0,17* (0.0009)	-0,31* (0.0015)	-0,22* (0.0012)	-1,26*** (0.0034)	-0.02 (0.0052)	-1,31*** (0.0040)	-1,46*** (0.0048)
Ciências Humanas	-1,54*** (0.0033)	-0,44*** (0.0014)	-0,89*** (0.0021)	-0,45* (0.0018)	-0,64** (0.0028)	0.10 (0.0045)	-0,06* (0.0032)	-0,06** (0.0038)
Matemática	-2,37*** (0.0046)	-0,50*** (0.0019)	-1,39*** (0.0029)	-1,11*** (0.0024)	-1,18** (0.0059)	-0.67 (0.0048)	-0.11* (0.0060)	-0.05 (0.0096)

Fonte: Elaboração própria Observações: 1. *, ** e *** representam confiança de 90%, 95% e 99%, respectivamente.

2. NN (1) = nearest neighbor – o primeiro vizinho mais próximo; NN (1) SR = o primeiro vizinho mais próximo sem reposição; NN (3) = os três primeiros vizinhos mais próximos; NN (5) os cinco primeiros vizinhos mais próximos; CEM = Coarsened Exact Matching; e EB = Entropy Balancing.

3. T e C representam o quantitativo de indivíduos nos grupos de Tratamento e Controle, respectivamente e foram expressos pois houve perda de observações nos procedimentos CEM e EB.

As estimações para a certificação em nível médio indicam, em geral, diminuição na probabilidade média de aprovação dos estudantes da EJA, em relação a indivíduos que não

estão estudando. Nossos resultados corroboram estudos anteriores que trataram de auferir o desempenho de estudantes de EJA no ensino médio e identificaram desempenho inferior deste grupo, tais como Andriola (2014).

Os resultados com mais magnitude referem-se às disciplinas de Linguagens e Redação, com as estimativas apontando uma redução (entre -1,33% a -3,67%) nas chances médias de aprovação para os estudantes da EJA, com um nível de confiança de 99%, nos diversos tipos de metodologia utilizados. Igualmente observou-se menor a probabilidade de aprovação dos estudantes da EJA na prova de Matemática, (com uma variação de -0,5% à -2,37%), nos diversos tipos de metodologia adotados, variando para esses resultados o grau de confiabilidade, entre 90% e 99%. Lima et al. (2019) e Silva et al. (2020), chegaram a conclusões semelhantes às deste estudo, pois verificaram habilidade inferior dos estudantes da EJA em conhecimentos de leitura, escrita e matemática, quando comparados aos estudantes do ensino médio regular, em análise realizada sobre a prova do ENEM.

Em relação às disciplinas de Ciências da Natureza e de Ciências Humanas, observa-se redução média aprovação para estudantes da EJA, nos diversos tipos de metodologias adotadas, com efeitos significativos, porém a magnitude desse efeito é menor que os efeitos em Linguagens e Redação, e Matemática. Para Ciências da Natureza o coeficiente varia de -0,17 à -1,46, com 90% e 99% de confiança, respectivamente. Para Ciências Humanas, observa-se que o efeito varia de -0,06% à -1,54% com 90% e 99% de confiança, respectivamente.

Em pesquisa sobre o desempenho dos estudantes da EJA na prova do ENEM, Silva (2022) identificou resultados semelhantes aos encontrados por esta pesquisa, indicando maior qualidade no sistema regular de ensino na produção de conhecimento comparativamente à modalidade de ensino EJA. Todavia, no caso desta pesquisa o grupo de comparação refere-se à estudantes que não estão estudando.

Na tabela 5 a seguir são apresentados os resultados das estimativas realizadas para as provas de certificação em nível fundamental, os procedimentos adotados foram os mesmos para a análise em nível de ensino médio, já citados anteriormente.

Tabela 5 - Certificação Ensino Fundamental - Efeito médio de ter cursado EJA sobre a aprovação nas provas de Linguagens e Redação, Matemática, Ciências da Natureza e Ciências Humanas (ATT).

Disciplina	NN (1)	NN (1) *	NN (3)	NN (5)	CEM	EB	CEM+EB	CEM+PSM
Linguagens e Redação	-2,36*** (0.0061)	-1,73*** (0.0041)	-1,87*** (0.0044)	-1,47*** (0.0039)	-1,29 (0.0109)	0,19 (0.0117)	-0,08 (0.0123)	-0,02 (0.0156)
Ciências da Natureza	-0,92* (0.0049)	-0,90*** (0.0034)	-0,88*** (0.0036)	-0,64** (0.0032)	0,04 (0.0183)	-0,04 (0.0185)	0,03 (0.1422)	-2,52** (0.0144)
Ciências Humanas	-0,64 (0.0059)	-0,19 (0.0040)	0,05 (0.0042)	-0,01 (0.0038)	-1,46 (0.0183)	-1,68 (0.0176)	-1,47 (0.0199)	0,00 (0.0278)
Matemática	1,05 (0.0068)	0,35 (0.0046)	0,09 (0.0048)	0,07 (0.0043)	0,01 (0.0416)	0,06 (0.0221)	0,00 (0.0413)	0,04 (0.0612)

Fonte: Elaboração própria.

Observações: 1. *, ** e *** representam confiança de 90%, 95% e 99%, respectivamente.

2. NN (1) = *nearest neighbor* – o primeiro vizinho mais próximo; NN (1) * = o primeiro vizinho mais próximo com reposição; NN (3) = os três primeiros vizinhos mais próximos; NN (5) os cinco primeiros vizinhos mais próximos; CEM = *Coarsened Exact Matching*; e EB = *Entropy Balancing*.

Observa-se na tabela 5 que para a certificação em nível fundamental os resultados apresentam-se significantes para as disciplinas de Linguagens e Redação e Ciências da Natureza. Cabe destacar que o quantitativo de indivíduos que buscam certificação para Ensino Fundamental é expressivamente menor em relação aos que buscam para Ensino Médio. Nos procedimentos de *CEM* e *EB* observou-se considerável redução da amostra após os modelos serem estimados, o que pode ser observado abaixo de cada resultado, na tabela 4. Tal fato pode prejudicar a confiabilidade dos resultados, pela própria definição de que os métodos devem ser aplicados a grandes bases de dados. Entretanto, para fins de robustez, os resultados estimados foram mantidos na tabela 5.

Sobre a disciplina de Linguagens e Redação, para certificação em nível fundamental, o efeito médio de cursar EJA é negativo sobre a aprovação em todos os métodos de pareamento adotados, o que indica certa robustez dos resultados. A magnitude do efeito, que é significativo a um nível de confiança de 99%, varia de 1,28% a 1,77%. Observa-se que a disciplina de Linguagens e Redação é a disciplina com maior número de participantes entre as quatro provas analisadas da amostra (103.016).

Considerando que para a aprovação em Linguagens e Redação é necessário obter 50% na nota na prova objetiva, além de 50% na prova de redação (escrita) nossos resultados podem ir ao encontro do estudo de Ferreira e Martinelli (2016), que identificaram baixo desempenho em escrita dos estudantes da EJA.

Resultado semelhante aos supracitados foi encontrado no desempenho na matéria de Ciências da Natureza, observou-se um efeito médio negativo sobre a aprovação na disciplina, para aqueles que cursam EJA. Para a certificação em Ciências Humanas e Matemática as diferenças não foram significativas em nenhum dos métodos estimados.

Ressalta-se que raros estudos dedicaram-se a avaliar o desempenho de estudantes da EJA. Assim, estudos adicionais devem ser elaborados a fim de identificar possíveis causas para tais diferenças de desempenho entre as disciplinas e os diferentes tipos de certificação, bem como para descobrir os mecanismos que levam à variação no desempenho entre as provas.

Em geral, em nossas análises foi possível observar que as chances de aprovação para os estudantes da EJA são reduzidas, em relação aos indivíduos que não estão estudando. Uma hipótese para explicar tais resultados pode ser relacionada à pedagogia da EJA. Na literatura que trata sobre a educação de jovens e adultos há evidência de que os professores que atuam com jovens e adultos devem ter uma capacitação específica para lidar com esse público (LOPES & SOUZA, 2005).

Todavia, conforme a pesquisa realizada por Da Silveira (2019) os materiais e metodologias aplicados à EJA não são adequados ao público alvo do programa, tendo sido identificada a utilização de materiais infantilizados, o mesmo utilizado para a educação básica. A autora destaca para a necessidade de reavaliar os caminhos da EJA assim como a metodologia de ensino e proposta pedagógica, que deve ser diferente do ensino regular.

Pode-se supor também a existência do custo de oportunidade de cursar EJA, quando se tem a intenção de certificar-se via Enceja, ou seja, quando se busca estritamente a certificação, e não os conhecimentos que podem ser adquiridos no curso de EJA. Há no próprio sítio do INEP materiais e apostilas voltadas especificamente para o estudo do Enceja, além de existirem também cursos *on-line* comercializados pelo mercado privado, que podem implicar em uma melhor preparação para o exame, mesmo para aqueles que não estão no ensino regular.

Ao considerar que tanto o EJA quanto o Enceja possuem como um de seus objetivos principais a certificação em nível fundamental e médio, pode-se inferir que são políticas que concorrem entre si, conforme abordado por Ventura e Oliveira (2020). Embora exista toda uma questão de inserção social, indicada pela literatura, relacionada à oferta dos cursos de EJA. Corrobora a hipótese de concorrência o fato de 34% dos inscritos para o Enceja de nível fundamental e 36% para o nível médio já terem cursado EJA. Depreende-se então que tais inscritos podem ter intenção de abandonar os cursos de EJA se aprovados no Enceja, visto que a finalidade máxima de ambos é a certificação.

3.6.1 RESULTADOS DOS TESTES DE LIMITES DE ROSENBAUM

A tabela 6 mostra os resultados do teste de sensibilidade conhecido por limites de Rosenbaum. Conforme Diprete e Gangl (2004), após a estimação do ATT, sob a hipótese de independência condicional, simula-se a existência de uma variável que afeta as chances de participação do indivíduo na política, condicional a suas características X (viés de seleção). À medida que o impacto potencial da variável sobre a probabilidade de participação na política torna-se mais forte, o intervalo de confiança dos efeitos estimados é elevado (p-valor eleva-se).

Tabela 6 Análise de Sensibilidade estimativa *Propensity Score Matching*

Ensino Médio								
	Linguagens e Redação		Ciências da Natureza		Ciências Humanas		Matemática	
γ	p-crit+	p-crit-	p-crit+	p-crit-	p-crit+	p-crit-	p-crit+	p-crit-
1	0,0290**	0,0290**	0,0501*	0,0501*	0,0508*	0,0508*	0,0074***	0,0074***
1.1	0,000***	0,0453	0,000***	0,1320	0,000***	0,0499	0,0801*	0,000***
1.2	0,000***	0,000***	0,000***	0,000***	0,000***	0,000***	0,000***	0,000***
1.3	0,000***	0,000***	0,000***	0,000***	0,000***	0,000***	0,000***	0,000***
1.4	0,000***	0,000***	0,000***	0,000***	0,000***	0,000***	0,000***	0,000***
1.5	0,000***	0,000***	0,000***	0,000***	0,000***	0,000***	0,000***	0,000***

Ensino Fundamental								
	Linguagens e Redação		Ciências da Natureza		Ciências Humanas		Matemática	
γ	p-crit+	p-crit-	p-crit+	p-crit-	p-crit+	p-crit-	p-crit+	p-crit-
1	0,1115	0,1115	0,1115	0,1115	0,0235**	0,0235**	0,0004***	0,0004***
1.1	0,000***	0,000***	0,0069	0,000***	0,000***	0,0002	0,000***	0,000***
1.2	0,000***	0,000***	0,000***	0,000***	0,000***	0,000***	0,000***	0,000***
1.3	0,000***	0,000***	0,000***	0,000***	0,000***	0,000***	0,000***	0,000***
1.4	0,000***	0,000***	0,000***	0,000***	0,000***	0,000***	0,000***	0,000***
1.5	0,000***	0,000***	0,000***	0,000***	0,000***	0,000***	0,000***	0,000***

Fonte: Elaboração própria.

Observações: 1. *, ** e *** representam confiança de 90%, 95% e 99%, respectivamente.

2. Teste estimado a partir do comando “*mhbounds*” no software Stata - H0: não há efeito de variável não controlada nos resultados estimados.

3. Teste realizado após estimação do *Propensity Score Matching* com um vizinho mais próximo sem reposição.

Os resultados da Tabela 6 indicam que a robustez ao viés de seleção varia entre as variáveis de resultado. Em termos gerais, observa-se que o efeito do tratamento parece ser estável mesmo diante da possível presença de viés de seleção, uma vez que a variável demonstra maior robustez à medida que o valor do nível crítico γ aumenta.

Em relação aos resultados para a certificação em nível médio, para a prova de Ciências da Natureza, foi identificado o nível crítico de γ de 1,1, significando que se a presença de variáveis não observadas podem levar a uma diferença na probabilidade de participar do curso de EJA por um fator de 1,1, de forma que os resultados sobre esta disciplina podem ser questionados.

Já para a certificação em nível fundamental, observa-se que para as provas das disciplinas de Linguagens e Redação e Ciências da natureza o γ crítico é a unidade, embora com o aumento da variação de γ os resultados apresentem robustez. Nas disciplinas de Ciências Humanas e Matemáticas os resultados são robustos a possíveis vieses.

De acordo com Diprete e Gangl (2004), é relevante observar que esses resultados são considerados os cenários mais desfavoráveis. Importa destacar ainda que um valor de 1,10, por exemplo, não implica de imediato a ausência de um efeito da política aqui analisada. Segundo os autores, isso indica apenas que, se houver viés de seleção que leve as probabilidades de participação a diferirem entre os grupos de tratamento e controle em um fator de 1,10, o intervalo de confiança para o efeito do tratamento conterá zero.

Embora a análise de sensibilidade de Rosenbaum apresente resultados no cenário mais desfavorável, ela ilustra a magnitude necessária para a influência de uma variável não observada para questionar as conclusões derivadas dos métodos de pareamento.

3.7 CONSIDERAÇÕES FINAIS

Esta pesquisa teve por objetivo avaliar empiricamente o efeito de cursar Educação de Jovens e Adultos (EJA) sobre a probabilidade de aprovação desses estudantes no Exame Nacional para Certificação de Competências de Jovens e Adultos (Enceja). Embora a Educação de Jovens e Adultos seja uma política pública de longa data, poucas pesquisas dedicaram-se a avaliar os seus resultados quantitativamente.

Para conferir robustez às estimativas utilizamos diversos tipos de procedimentos econométricos de pareamento e de preparação da base de dados. Ademais, foi realizado teste de sensibilidade para possível viés de seleção na participação na EJA, o qual indicou que, em geral, a possibilidade da ocorrência de variáveis não controladas no presente estudo não alteraria significativamente os resultados estimados. A partir das estimativas, foi possível observar menor probabilidade dos indivíduos que cursam EJA obterem certificação via Enceja. Tal fato indica um pior desempenho do grupo participante da política de EJA.

No Ensino Médio identificou-se menor probabilidade de aprovação no Enceja para os indivíduos que cursam EJA, em todas as disciplinas analisadas, sendo o efeito negativo de maior magnitude para Linguagens e Redação, e Matemática.

Na certificação em Nível Fundamental verificou-se desempenho inferior dos estudantes da EJA nas disciplinas de Linguagens e Redação, e Ciências da Natureza. As

estimativas para a aprovação nas disciplinas de Ciências Humanas e Matemática não apresentaram coeficientes estatisticamente significativos.

Em geral os resultados indicam desempenho inferior do grupo participante da política de EJA e que esse desempenho varia a depender do tipo de certificação que os estudantes estão buscando (Nível Fundamental ou Nível Médio), e entre as diferentes disciplinas.

Pode-se inferir que os estudantes da EJA que realizaram o Encceja têm intenção de evadir da EJA, uma vez que são políticas semelhantes no que se refere à certificação. Entretanto, ainda que a EJA certifique em menos tempo que o ensino regular, o Encceja é um processo mais célere, bastando realizar o exame e ser aprovado para obter a certificação.

Ressalta-se ainda que, nas análises descritivas identificou-se que a Encceja está longe de ser uma política universalizada e que apenas 11% dos municípios brasileiros realizaram a prova em 2019. Adiciona-se que a oferta é maior na região sul, sudeste e centro-oeste, seguido pela região norte. A região do nordeste possui aplicação da prova em apenas 5,9% dos seus municípios. Assim, destacamos que a amostra estudada contempla parte dos estudantes da EJA, ou seja, aqueles que realizaram o Encceja.

Por todo o exposto, em termos de proposição de políticas públicas é necessário adotar meios para qualificar o ensino da EJA e levar em conta sua interação com o Encceja, verificando questões de competição e rivalidade entre as duas políticas. Deve-se ainda serem consideradas questões já apontadas pela literatura, tais como metodologia de ensino da EJA, pedagogia, qualificação dos professores, tendo em vista as especificidades de jovens e adultos que retornam ao ambiente escolar.

REFERÊNCIAS

ANAZAWA, Leandro, GUEDES, Marcelo Sanchez, KOMATSU, Bruno Kawaoka, AND MENEZES-FILHO, Naercio Aquino. "A Loteria da Vida: Examinando a Relação entre a Educação da Mãe e a Escolaridade do Jovem com Dados Longitudinais do Brasil". Insper - Centro de Políticas Públicas, Policy Paper 22. 2016.

ANDRIOLA, Wagner Bandeira. Avaliação diagnóstica da Educação de Jovens e Adultos (EJA) no Brasil. Ensaio: Avaliação e Políticas Públicas em Educação. v. 22, n. 82, p. 171-196, 2014.

CALIENDO, M., KOPEINIG, S. "Some practical guidance for the implementation of propensity score matching". *IZA Discussion Papers*, n.1588, IZA.2005.

CATELLI JR., Roberto; GISI, B., Serrao, LFS. "Encceja: cenário de disputas na EJA". *Rev. Bras. Estud. Pedagog.*, Brasília, v. 94, n. 238, p. 721-744. 2013.

CONZATTI, FBK; DAVOGLIO, TR. "Estado de Conhecimento da Pesquisa Acadêmica sobre o Aluno Adulto da Educação de Jovens e Adultos". *Revista Educação Por Escrito*, v. 7(1), p.5973. 2016.

CURI, Andréa Zaitune; MENEZES-FILHO, Naércio Aquino. A relação entre educação pré-primária, salários, escolaridade e proficiência escolar no Brasil. *Estudos Econômicos (São Paulo)*, v. 39, p. 811-850, 2009.

DA SILVA, Ilza Ferreira; GOMES ELLIOT, LIGIA; SANTOS FONTANIVE, NILMA. Certificação do Ensino Médio via ENEM: como foi o desempenho dos candidatos. *Meta: Avaliação*, v. 12, n. 36, 2020.

DA SILVEIRA MOTA, Rosangela. Aprendizagem do adulto e correspondentes metodologias. 2009. Tese de Doutorado. 2009.

DOS SANTOS, Robson; SANTOS, Priscila Pereira. O ENCCEJA e a Educação de Jovens e Adultos. *Cadernos de Estudos e Pesquisas em Políticas Educacionais*, v. 3, n. 4, p. 37-37, 2021.

DIPRETE, T.; GANGL, M. "Assessing bias in the estimation of causal effects: Rosenbaum bounds on matching estimators and instrumental variables estimation with imperfect instruments." *Sociological Methodology*, v. 34, n. 1, p. 271-310. 2004.

FERREIRA, A; MARTINELLI, SC. 2016 "Estudantes da Educação de Jovens e Adultos: considerações sobre o perfil e desempenho escolar". *Educação: teoria e prática*, v. 26 (52), p. 312-331, ago. 2016.

IACUS, Stefano M; KING, Gary; PORRO, Giuseppe. cem : Software for Coarsened Exact Matching. *Journal of Statistical Software*, v. 30, n. 9, 2009.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA - INEP. Relatório Divulgação Censo Escolar 2020, 2021.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE. Pesquisa Nacional por Amostra de Domicílios Contínua – PNAD Contínua 2019.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE. Síntese de Indicadores Sociais 2018.

HADDAD, Sérgio. O estado da arte das pesquisas em educação de jovens e adultos no Brasil: a produção discente da pós-graduação em educação no período 1986-1998. 2000.

HADDAD, Sérgio; DI PIERRO, Maria Clara. Escolarização de jovens e adultos. Revista brasileira de educação, n. 14, p. 108-130, 2000.

HAINMUELLER, Jens. Entropy Balancing for Causal Effects : A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies. *Political Analysis*, v. 20, p. 25–46, 2012.

HAINMUELLER, Jens; XU, Yiqing. Ebalance: A Stata package for entropy balancing. *Journal of Statistical Software*, v. 54, n. 7, 2013.

HECKMAN, J., LALONDE, R, e SMITH, J. 1999. “The Economics and Econometrics of Active Labor Market Programs.” *In Handbook of Labor Economics*, vol. 3, ed. Orley Ashenfelter and David Card, 1865–2097.

KHANDKER, Shahidur, GAYATRI B. Koolwal, and HUSSAIN Samad. “Handbook on impact evaluation: quantitative methods and practices”. The World Bank. 2009.

LOPES, Selva Paraguassu; SOUSA, Luzia Silva. EJA: uma educação possível ou mera utopia. *Revista Alfabetização Solidária (Alfasol)*, v. 5, p. 75-80, 2005.

LAZARETTI, Lauana Rossetto; FRANÇA, Marco Túlio Aniceto. A forma de ingresso importa? Uma análise do desempenho dos alunos do ensino médio na rede federal. XXIV Encontro de Economia da Região Sul, 2021, Brasil., 2021.

LEON, Fernanda Leite Lopez de; NA MENEZES-FILHO, Naércio Aquino. Reprovação, avanço e evasão escolar no Brasil. 2002.

LEONCY, Christiane. Mulheres na EJA: questões de identidade e gênero. Mestrado em Educação)–Universidade Estadual de Campinas, Campinas, SP, 2013.

MANFREDI, Silvia Maria. Política: educação popular. São Paulo: Símbolo, 1978

OLIVA, Bruno et al. Requalificação e mercado de trabalho: impactos do EJA e da educação técnica e profissional. São Paulo School of Economics Working Paper, v. 4, p. 8-36, 2014.

ROSENBAUM, P. “Observational studies”. Nova York: Springer Verlag, 375p. 2000.

ROSENBAUM, Paul R. *Observational Studies*. Second Edition. [S.l.]: Springer Series in Statistics, 2002.

ROSENBAUM, P. R., Rubin, D. B. “The central role of the propensity score in observational studies for causal effects”. *Biometrika*, v. 70, n. 1, p. 41–55. 1983.

SILVA, Maria da Conceição. Desigualdade educacional no ensino público: educação de jovens e adultos no ENEM. 2022. 51f. Dissertação (Mestrado em Economia do Setor Público) - Faculdade de Economia, Administração, Atuária e Contabilidade - FEAAC, Programa de Economia Profissional - PEP, Universidade Federal do Ceará - UFC, Fortaleza (CE), 2022.

TAVARES, Priscilla Albuquerque; SOUZA, André Portela de; PONCZEK, Vladimir Pinheiro. Uma análise dos fatores associados à frequência ao ensino médio na educação de jovens e adultos (EJA) no Brasil. 2014.

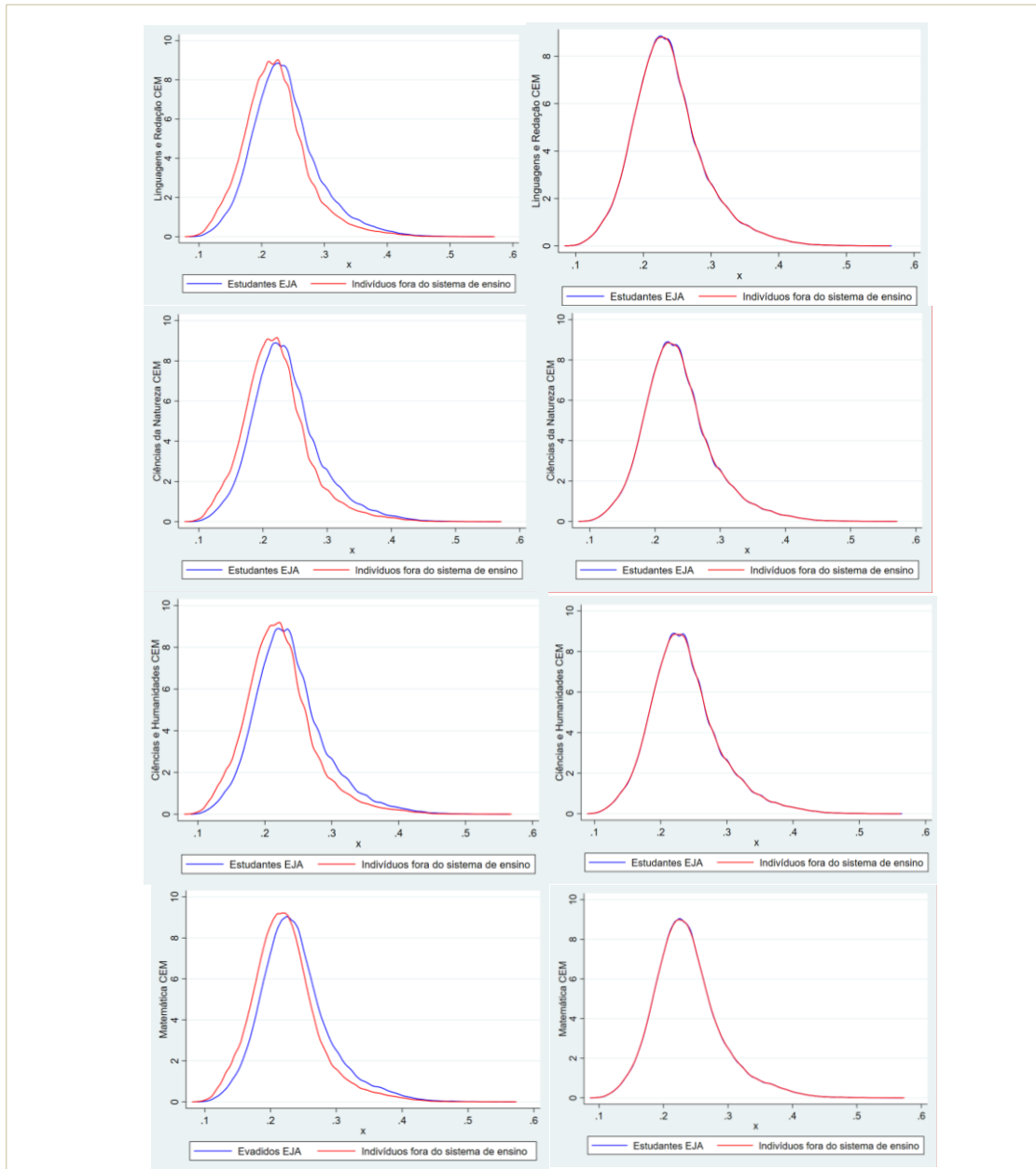
TELES, Damares Araujo; SOARES, Maria Perpétua do Socorro Beserra. Educação de jovens e adultos: desafios e possibilidades na alfabetização. 2016.

VENTURA, Jaqueline Pereira; OLIVEIRA, Francisco Gilson. A travessia “do EJA” ao Enceja: Será o mercado da educação não formal o novo rumo da EJA no Brasil?. Revista Internacional de Educação de Jovens e Adultos, v. 3, n. 5, p. 80-97, 2020.

VIEIRA, Caterina Soto et al. Como as mudanças no trabalho e na renda dos pais afetam as escolhas entre estudo e trabalho dos jovens?. 2016.

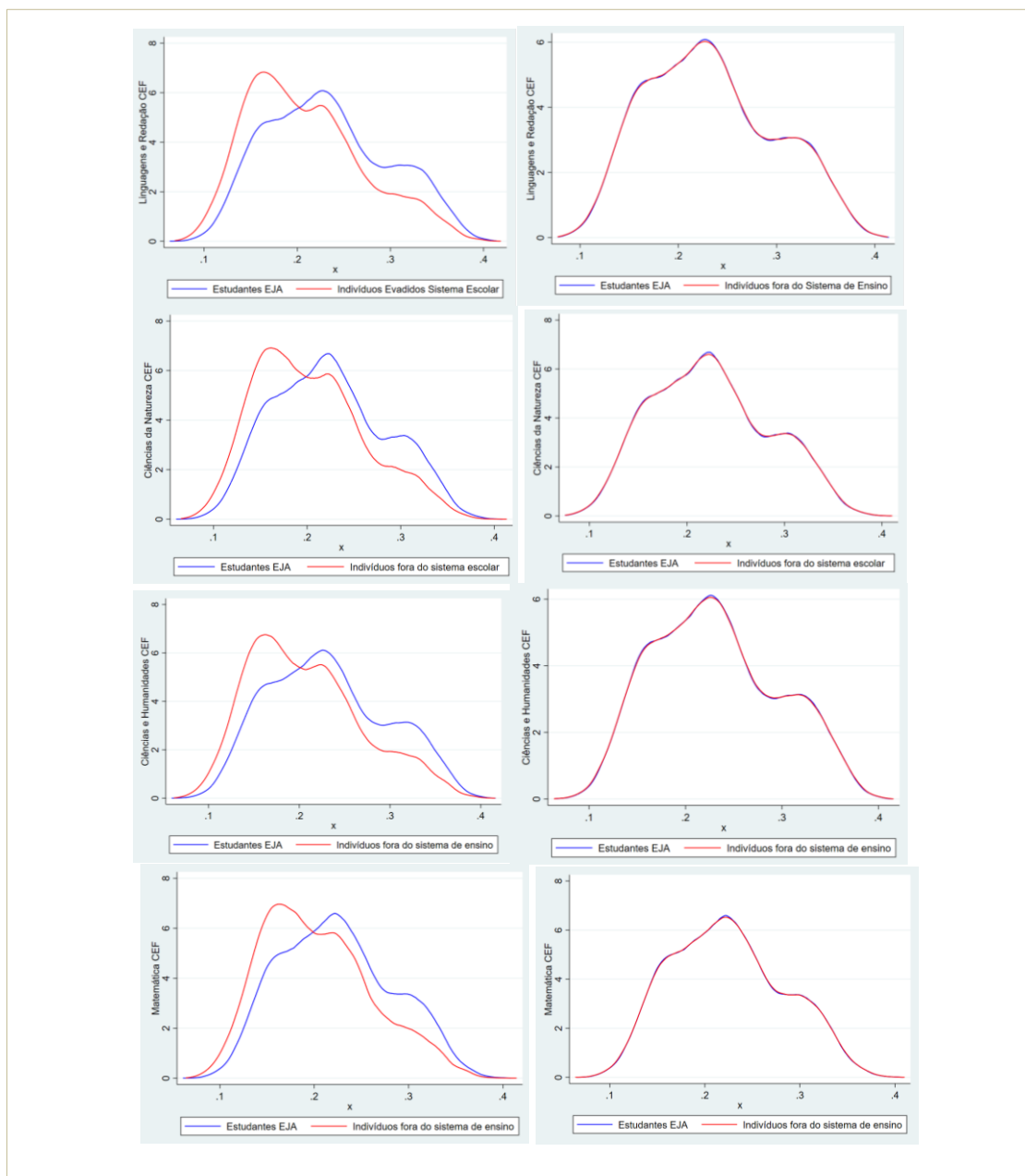
APÊNDICE A2

Apêndice A2 Figura 1 - Distribuição Score de Propensão dos participantes do Enceja, por tipo de certificação e por disciplina, pré e pós-pareamento Certificação Ensino Médio



Fonte: Elaboração Própria

Apêndice A2 Figura 2 - Distribuição Score de Propensão participantes do Encceja, por tipo de certificação e por disciplina, pré e pós-pareamento Certificação Ensino Fundamental



Fonte: Elaboração Própria

Apêndice A2 - Tabela 1 Diferença de médias, antes e depois do pareamento, entre os grupos de tratamento e controle - Linguagens e Redação (Ensino Médio)

Variável	Antes Matching			Depois Matching		
	Tratado	Controle	Diferença	Tratado	Controle	Diferença
Idade	30,06300	30,46200	-0,39900 ***	30,06300	29,98800	0,07500 *
Feminino	0,51887	0,48978	0,02909 ***	0,51887	0,51937	-0,00050
Renda	1,41510	1,49850	-0,08340 ***	1,41510	1,38640	0,02870 ***
Quando Evadiu	7,36110	7,70100	-0,33990 ***	7,36110	7,36300	-0,00190
Mãe Possui E.M.	0,26784	0,26388	0,00396 ***	0,26784	0,26054	0,00730 ***
Pai Possui E.M.	0,14827	0,15073	-0,00246 **	0,14827	0,14304	0,00523 ***
Rondônia	0,01295	0,00790	0,00505 ***	0,01295	0,01233	0,00062
Acre	0,00613	0,00440	0,00173 ***	0,00613	0,00567	0,00046
Amazonas	0,00844	0,00842	0,00002	0,00844	0,00755	0,00089 **
Roraima	0,00445	0,00288	0,00157 ***	0,00445	0,00349	0,00096 ***
Pará	0,03302	0,03881	-0,00579 ***	0,03302	0,03269	0,00033
Amapá	0,00845	0,00709	0,00136 ***	0,00845	0,00751	0,00094 ***
Tocantins	0,01056	0,00982	0,00074 **	0,01056	0,01010	0,00046
Maranhão	0,00725	0,00945	-0,00220 ***	0,00725	0,00698	0,00027
Piauí	0,00781	0,00751	0,00030	0,00781	0,00709	0,00072 **
Ceará	0,02953	0,01646	0,01307 ***	0,02953	0,02945	0,00008
Rio Grande do Norte	0,00464	0,00526	-0,00062 ***	0,00464	0,00416	0,00048 *
Paraíba	0,00792	0,00704	0,00088 ***	0,00792	0,00706	0,00086 **
Pernambuco	0,00541	0,00598	-0,00057 **	0,00541	0,00468	0,00073 **
Alagoas	0,01058	0,01143	-0,00085 **	0,01058	0,01036	0,00022
Sergipe	0,01278	0,01191	0,00087 **	0,01278	0,01268	0,00010
Bahia	0,03545	0,06083	-0,02538 ***	0,03545	0,03585	-0,00040
Minas Gerais	0,07700	0,10093	-0,02393 ***	0,07700	0,07847	-0,00147
Espírito Santo	0,01942	0,01586	0,00356 ***	0,01942	0,01865	0,00077
Rio de Janeiro	0,10920	0,11010	-0,00090	0,10920	0,11094	-0,00174
São Paulo	0,15986	0,17659	-0,01673 ***	0,15986	0,16139	-0,00153
Paraná	0,14244	0,12394	0,01850 ***	0,14244	0,14486	-0,00242 *
Santa Catarina	0,06343	0,05555	0,00788 ***	0,06343	0,06362	-0,00019
Rio Grande do Sul	0,10359	0,09206	0,01153 ***	0,10359	0,10457	-0,00098
Mato Grosso do Sul	0,03861	0,03606	0,00255 ***	0,03861	0,03877	-0,00016
Mato Grosso	0,00995	0,00937	0,00058 *	0,00995	0,00883	0,00112 ***
Goiás	0,03326	0,03779	-0,00453 ***	0,03326	0,03367	-0,00041
Distrito Federal	0,03786	0,02656	0,01130 ***	0,03786	0,03856	-0,00070

Fonte: Elaborada pelos autores.

Notas: 1. A hipótese nula do teste de diferença de médias é H_0 : Diferença = 0;

2. , * , ** e *** representam confiança de 90%, 95% e 99%, respectivamente.

Apêndice A2 - Tabela 2 Diferença de médias, antes e depois do pareamento, entre os grupos de tratamento e controle - Ciências Naturais (Ensino Médio)

Variável	Antes Matching			Depois Matching		
	Tratado	Controle	Diferença	Tratado	Controle	Diferença
Idade	29,91200	30,24700	-0,33500 ***	29,91200	29,82900	0,08300 *
Feminino	0,53296	0,50133	0,03163 ***	0,53296	0,53429	-0,00133
Renda	1,39210	1,47780	-0,08570 ***	1,39210	1,36480	0,02730 ***
Quando Evadiu	7,38110	7,72970	-0,34860 ***	7,38110	7,38190	-0,00080
Mãe Possui E.M.	0,26734	0,26538	0,00196	0,26734	0,25985	0,00749 ***
Pai Possui E.M.	0,14798	0,15170	-0,00372 ***	0,14798	0,14167	0,00631 ***
Rondônia	0,01253	0,00751	0,00502 ***	0,01253	0,01166	0,00087 **
Acre	0,00581	0,00404	0,00177 ***	0,00581	0,00561	0,00020 **
Amazonas	0,00828	0,00822	0,00006	0,00828	0,00768	0,00060
Roraima	0,00433	0,00282	0,00151 ***	0,00433	0,00406	0,00027 **
Pará	0,03474	0,04070	-0,00596 ***	0,03474	0,03457	0,00017
Amapá	0,00817	0,00726	0,00091 ***	0,00817	0,00731	0,00086 **
Tocantins	0,00952	0,00894	0,00058 *	0,00952	0,00916	0,00036
Maranhão	0,00771	0,00979	-0,00208 ***	0,00771	0,00700	0,00071
Piauí	0,00800	0,00758	0,00042	0,00800	0,00734	0,00066 **
Ceará	0,02996	0,01661	0,01335 ***	0,02996	0,02919	0,00077
Rio Grande do Norte	0,00483	0,00546	-0,00063 **	0,00483	0,00431	0,00052 **
Paraíba	0,00829	0,00714	0,00115 ***	0,00829	0,00773	0,00056
Pernambuco	0,00582	0,00622	-0,00040	0,00582	0,00485	0,00097 **
Alagoas	0,01091	0,01156	-0,00065 *	0,01091	0,01057	0,00034
Sergipe	0,01325	0,01216	0,00109 ***	0,01325	0,01334	-0,00009
Bahia	0,03673	0,06291	-0,02618 ***	0,03673	0,03703	-0,00030
Minas Gerais	0,07833	0,10135	-0,02302 ***	0,07833	0,07997	-0,00164
Espírito Santo	0,01940	0,01544	0,00396 ***	0,01940	0,01852	0,00088
Rio de Janeiro	0,11122	0,11220	-0,00098	0,11122	0,11325	-0,00203 **
São Paulo	0,16221	0,17765	-0,01544 ***	0,16221	0,16457	-0,00236
Paraná	0,14240	0,12387	0,01853 ***	0,14240	0,14399	-0,00159
Santa Catarina	0,06222	0,05464	0,00758 ***	0,06222	0,06199	0,00023
Rio Grande do Sul	0,10177	0,09040	0,01137 ***	0,10177	0,10319	-0,00142
Mato Grosso do Sul	0,03349	0,03234	0,00115 *	0,03349	0,03305	0,00044
Mato Grosso	0,00992	0,00926	0,00066 **	0,00992	0,00966	0,00026
Goiás	0,03367	0,03799	-0,00432 ***	0,03367	0,03372	-0,00005
Distrito Federal	0,03650	0,02595	0,01055 ***	0,03650	0,03669	-0,00019

Fonte: Elaborada pelos autores.

Notas: 1. A hipótese nula do teste de diferença de médias é H_0 : Diferença = 0;

2. , * , ** e *** representam confiança de 90%, 95% e 99%, respectivamente.

Apêndice A2 - Tabela 3 Diferença de médias, antes e depois do pareamento, entre os grupos de tratamento e controle - Ciências Humanas (Ensino Médio)

Variável	Antes Matching			Depois Matching		
	Tratado	Controle	Diferença	Tratado	Controle	Diferença
Idade	29,87200	30,19100	-0,31900 ***	29,87200	29,79900	0,07300 *
Feminino	0,53320	0,50255	0,03065 ***	0,53320	0,53463	-0,00143
Renda	1,39100	1,47510	-0,08410 ***	1,39100	1,36640	0,02460 ***
Quando Evadiu	7,37160	7,72310	-0,35150 ***	7,37160	7,36940	0,00220
Mãe Possui E.M.	0,26810	0,26574	0,00236	0,26810	0,26004	0,00806 ***
Pai Possui E.M.	0,14856	0,15171	-0,00315 ***	0,14856	0,14251	0,00605 ***
Rondônia	0,01313	0,00792	0,00521 ***	0,01313	0,01247	0,00066
Acre	0,00622	0,00440	0,00182 ***	0,00622	0,00569	0,00053 *
Amazonas	0,00853	0,00857	-0,00004	0,00853	0,00801	0,00052
Roraima	0,00447	0,00298	0,00149 ***	0,00447	0,00377	0,00070 ***
Pará	0,03413	0,03990	-0,00577 ***	0,03413	0,03384	0,00029
Amapá	0,00859	0,00738	0,00121 ***	0,00859	0,00759	0,00100 ***
Tocantins	0,01006	0,00946	0,00060 *	0,01006	0,00987	0,00019
Maranhão	0,00773	0,00971	-0,00198 ***	0,00773	0,00732	0,00041
Piauí	0,00790	0,00755	0,00035	0,00790	0,00755	0,00035
Ceará	0,02877	0,01626	0,01251 ***	0,02877	0,02777	0,00100
Rio Grande do Norte	0,00480	0,00533	-0,00053 **	0,00480	0,00416	0,00064 **
Paraíba	0,00824	0,00709	0,00115 ***	0,00824	0,00764	0,00060
Pernambuco	0,00570	0,00615	-0,00045 *	0,00570	0,00451	0,00119 ***
Alagoas	0,01097	0,01157	-0,00060 *	0,01097	0,01054	0,00043
Sergipe	0,01311	0,01257	0,00054	0,01311	0,01255	0,00056
Bahia	0,03693	0,06274	-0,02581 ***	0,03693	0,03766	-0,00073
Minas Gerais	0,07633	0,09930	-0,02297 ***	0,07633	0,07707	-0,00074
Espírito Santo	0,01872	0,01537	0,00335 ***	0,01872	0,01868	0,00004
Rio de Janeiro	0,11073	0,11207	-0,00134	0,11073	0,11365	-0,00292 **
São Paulo	0,15996	0,17576	-0,01580 ***	0,15996	0,16144	-0,00148
Paraná	0,14455	0,12510	0,01945 ***	0,14455	0,14559	-0,00104
Santa Catarina	0,06205	0,05448	0,00757 ***	0,06205	0,06161	0,00044
Rio Grande do Sul	0,10144	0,09054	0,01090 ***	0,10144	0,10319	-0,00175
Mato Grosso do Sul	0,03568	0,03416	0,00152 **	0,03568	0,03576	-0,00008
Mato Grosso	0,01028	0,00950	0,00078 **	0,01028	0,01014	0,00014
Goiás	0,03318	0,03759	-0,00441 ***	0,03318	0,03362	-0,00044
Distrito Federal	0,03781	0,02656	0,01125 ***	0,03781	0,03832	-0,00051

Fonte: Elaborada pelos autores.

Notas: 1. A hipótese nula do teste de diferença de médias é H_0 : Diferença = 0;

2. , * , ** e *** representam confiança de 90%, 95% e 99%, respectivamente.

Apêndice A2 - Tabela 4 Diferença de médias, antes e depois do pareamento, entre os grupos de tratamento e controle – Matemática (Ensino Médio)

Variável	Antes Matching			Depois Matching		
	Tratado	Controle	Diferença	Tratado	Controle	Diferença
Idade	30,22800	30,60500	-0,37700 ***	30,22800	30,15500	0,07300 **
Feminino	0,54312	0,51252	0,03060 ***	0,54312	0,54488	-0,00176
Renda	1,39510	1,48390	-0,08880 ***	1,39510	1,36810	0,02700 ***
Quando Evadiu	7,36320	7,69920	-0,33600 ***	7,36320	7,36210	0,00110
Mãe Possui E.M.	0,26302	0,25934	0,00368 ***	0,26302	0,25628	0,00674 ***
Pai Possui E.M.	0,14503	0,14761	-0,00258 **	0,14503	0,13950	0,00553 ***
Rondônia	0,01207	0,00734	0,00473 ***	0,01207	0,01119	0,00088 *
Acre	0,00573	0,00412	0,00161 ***	0,00573	0,00514	0,00059
Amazonas	0,00822	0,00838	-0,00016	0,00822	0,00799	0,00023
Roraima	0,00454	0,00294	0,00160 ***	0,00454	0,00389	0,00065
Pará	0,03451	0,04004	-0,00553 ***	0,03451	0,03435	0,00016
Amapá	0,00869	0,00740	0,00129 ***	0,00869	0,00785	0,00084 **
Tocantins	0,01054	0,00972	0,00082 ***	0,01054	0,01004	0,00050
Maranhão	0,00769	0,00946	-0,00177 ***	0,00769	0,00723	0,00046 **
Piauí	0,00809	0,00744	0,00065 **	0,00809	0,00726	0,00083 *
Ceará	0,03090	0,01673	0,01417 ***	0,03090	0,02990	0,00100
Rio Grande do Norte	0,00471	0,00519	-0,00048 **	0,00471	0,00412	0,00059 *
Paraíba	0,00808	0,00704	0,00104 ***	0,00808	0,00766	0,00042
Pernambuco	0,00560	0,00598	-0,00038	0,00560	0,00503	0,00057 ***
Alagoas	0,01106	0,01208	-0,00102 ***	0,01106	0,01067	0,00039
Sergipe	0,01274	0,01205	0,00069 **	0,01274	0,01264	0,00010
Bahia	0,03726	0,06245	-0,02519 ***	0,03726	0,03770	-0,00044
Minas Gerais	0,07824	0,10157	-0,02333 ***	0,07824	0,07910	-0,00086
Espírito Santo	0,01968	0,01557	0,00411 ***	0,01968	0,01959	0,00009
Rio de Janeiro	0,10978	0,10976	0,00002	0,10978	0,11233	-0,00255
São Paulo	0,16087	0,17720	-0,01633 ***	0,16087	0,16218	-0,00131
Paraná	0,13703	0,12374	0,01329 ***	0,13703	0,13864	-0,00161
Santa Catarina	0,05880	0,05260	0,00620 ***	0,05880	0,05838	0,00042
Rio Grande do Sul	0,10342	0,09100	0,01242 ***	0,10342	0,10459	-0,00117
Mato Grosso do Sul	0,03939	0,03598	0,00341 ***	0,03939	0,04009	-0,00070
Mato Grosso	0,00957	0,00899	0,00058 *	0,00957	0,00934	0,00023
Goiás	0,03418	0,03809	-0,00391 ***	0,03418	0,03417	0,00001
Distrito Federal	0,03859	0,02713	0,01146 ***	0,03859	0,03891	-0,00032

Fonte: Elaborada pelos autores.

Notas: 1. A hipótese nula do teste de diferença de médias é H_0 : Diferença = 0;

2. , * , ** e *** representam confiança de 90%, 95% e 99%, respectivamente.

Apêndice A2 - Tabela 5 Diferença de médias, antes e depois do pareamento, entre os grupos de tratamento e controle – Linguagens e Redação (Ensino Fundamental)

Variável	Antes Matching			Depois Matching		
	Tratado	Controle	Diferença	Tratado	Controle	Diferença
Idade	31,13500	33,96400	-2,82900 ***	31,13500	31,09800	0,03700
Feminino	0,54165	0,52743	0,01422 ***	0,54165	0,54119	0,00046
Renda	1,25880	1,38020	-0,12140 ***	1,25880	1,23320	0,02560 **
Quando Evadiu	5,16630	5,05490	0,11140 ***	5,16630	5,20350	-0,03720 **
Mãe Possui E.M.	0,19100	0,15295	0,03805 ***	0,19100	0,17407	0,01693 ***
Pai Possui E.M.	0,09296	0,07777	0,01519 ***	0,09296	0,08261	0,01035 ***
Rondônia	0,01793	0,01537	0,00256 ***	0,01793	0,01620	0,00173
Acre	0,00763	0,00536	0,00227 ***	0,00763	0,00472	0,00291 ***
Amazonas	0,01235	0,01052	0,00183 **	0,01235	0,00994	0,00241 **
Roraima	0,00431	0,00369	0,00062	0,00431	0,00340	0,00091
Pará	0,04357	0,04702	-0,00345 **	0,04357	0,04403	-0,00046
Amapá	0,01371	0,01042	0,00329 ***	0,01371	0,01298	0,00073
Tocantins	0,00758	0,01200	-0,00442 ***	0,00758	0,00672	0,00086
Maranhão	0,00640	0,00777	-0,00137 **	0,00640	0,00409	0,00231 ***
Piauí	0,00985	0,00735	0,00250 ***	0,00985	0,00812	0,00173 *
Ceará	0,01580	0,00964	0,00616 ***	0,01580	0,01407	0,00173
Rio Grande do Norte	0,00413	0,00345	0,00068	0,00413	0,00200	0,00213 ***
Paraíba	0,00536	0,00398	0,00138 ***	0,00536	0,00350	0,00186 ***
Pernambuco	0,00499	0,00446	0,00053	0,00499	0,00318	0,00181 ***
Alagoas	0,01003	0,01066	-0,00063	0,01003	0,00776	0,00227 **
Sergipe	0,01257	0,01183	0,00074	0,01257	0,01171	0,00086
Bahia	0,02742	0,04570	-0,01828 ***	0,02742	0,02701	0,00041
Minas Gerais	0,06763	0,10391	-0,03628 ***	0,06763	0,06936	-0,00173
Espírito Santo	0,01793	0,01818	-0,00025	0,01793	0,01620	0,00173
Rio de Janeiro	0,07589	0,08728	-0,01139 ***	0,07589	0,07866	-0,00277
São Paulo	0,11906	0,14753	-0,02847 ***	0,11906	0,12514	-0,00608 *
Paraná	0,20571	0,13232	0,07339 ***	0,20571	0,21551	-0,00980 **
Santa Catarina	0,06037	0,04639	0,01398 ***	0,06037	0,06255	-0,00218
Rio Grande do Sul	0,10726	0,11971	-0,01245 ***	0,10726	0,11075	-0,00349
Mato Grosso do Sul	0,05029	0,05597	-0,00568 ***	0,05029	0,05061	-0,00032
Mato Grosso	0,00867	0,00958	-0,00091	0,00867	0,00803	0,00064
Goiás	0,02583	0,03119	-0,00536 ***	0,02583	0,02528	0,00055
Distrito Federal	0,05774	0,03872	0,01902 ***	0,05774	0,05846	-0,00072

Fonte: Elaborada pelos autores.

Notas: 1. A hipótese nula do teste de diferença de médias é H_0 : Diferença = 0;

2. , * , ** e *** representam confiança de 90%, 95% e 99%, respectivamente.

Apêndice A2 - Tabela 6 Diferença de médias, antes e depois do pareamento, entre os grupos de tratamento e controle – Ciências Naturais (Ensino Fundamental)

Variável	Antes Matching			Depois Matching		
	Tratado	Controle	Diferença	Tratado	Controle	Diferença
Idade	30,89400	33,67200	-2,77800 ***	30,89400	30,75200	0,14200
Feminino	0,54554	0,52912	0,01642 ***	0,54554	0,54702	-0,00148
Renda	1,23830	1,36380	-0,12550 ***	1,23830	1,21460	0,02370 **
Quando Evadiu	5,18320	5,08220	0,10100 ***	5,18320	5,21620	-0,03300 ***
Mãe Possui E.M.	0,18993	0,15541	0,03452 ***	0,18993	0,17581	0,01412 *
Pai Possui E.M.	0,09379	0,07823	0,01556 ***	0,09379	0,08326	0,01053 *
Rondônia	0,01609	0,01361	0,00248 ***	0,01609	0,01447	0,00162
Acre	0,00773	0,00511	0,00262 ***	0,00773	0,00526	0,00247 ***
Amazonas	0,01230	0,01019	0,00211 ***	0,01230	0,00945	0,00285
Roraima	0,00438	0,00386	0,00052	0,00438	0,00295	0,00143
Pará	0,04753	0,05034	-0,00281	0,04753	0,04689	0,00064
Amapá	0,01481	0,01077	0,00404 ***	0,01481	0,01481	0,00000
Tocantins	0,00763	0,01189	-0,00426 ***	0,00763	0,00576	0,00187
Maranhão	0,00684	0,00834	-0,00150 **	0,00684	0,00467	0,00217
Piauí	0,01019	0,00747	0,00272 ***	0,01019	0,00969	0,00050 **
Ceará	0,01693	0,00998	0,00695 ***	0,01693	0,01324	0,00369
Rio Grande do Norte	0,00448	0,00348	0,00100 **	0,00448	0,00315	0,00133
Paraíba	0,00581	0,00408	0,00173 ***	0,00581	0,00428	0,00153 *
Pernambuco	0,00531	0,00469	0,00062	0,00531	0,00349	0,00182
Alagoas	0,01102	0,01125	-0,00023	0,01102	0,01038	0,00064
Sergipe	0,01294	0,01227	0,00067	0,01294	0,01225	0,00069
Bahia	0,02938	0,04836	-0,01898 ***	0,02938	0,02997	-0,00059
Minas Gerais	0,06904	0,10376	-0,03472 ***	0,06904	0,07159	-0,00255
Espírito Santo	0,01919	0,01793	0,00126	0,01919	0,01880	0,00039 *
Rio de Janeiro	0,07897	0,08929	-0,01032 ***	0,07897	0,08316	-0,00419 **
São Paulo	0,12188	0,14739	-0,02551 ***	0,12188	0,12631	-0,00443
Paraná	0,19348	0,12941	0,06407 ***	0,19348	0,20081	-0,00733
Santa Catarina	0,05757	0,04597	0,01160 ***	0,05757	0,05796	-0,00039
Rio Grande do Sul	0,10722	0,11757	-0,01035 ***	0,10722	0,11032	-0,00310
Mato Grosso do Sul	0,04822	0,05347	-0,00525 ***	0,04822	0,04729	0,00093
Mato Grosso	0,00866	0,00926	-0,00060	0,00866	0,00669	0,00197
Goiás	0,02795	0,03204	-0,00409 ***	0,02795	0,02982	-0,00187
Distrito Federal	0,05447	0,03825	0,01622 ***	0,05447	0,05654	-0,00207

Fonte: Elaborada pelos autores.

Notas: 1. A hipótese nula do teste de diferença de médias é H_0 : Diferença = 0;

2. , * , ** e *** representam confiança de 90%, 95% e 99%, respectivamente.

Apêndice A2 - Tabela 7 Diferença de médias, antes e depois do pareamento, entre os grupos de tratamento e controle – Ciências Humanas (Ensino Fundamental)

Variável	Antes Matching			Depois Matching		
	Tratado	Controle	Diferença	Tratado	Controle	Diferença
Idade	30,85400	33,69700	-2,84300 ***	30,85400	30,77500	0,07900
Feminino	0,54712	0,53357	0,01355 ***	0,54712	0,55187	-0,00475
Renda	1,24140	1,36350	-0,12210 ***	1,24140	1,21460	0,02680 **
Quando Evadiu	5,18060	5,08320	0,09740 ***	5,18060	5,21270	-0,03210 **
Mãe Possui E.M.	0,19157	0,15492	0,03665 ***	0,19157	0,17198	0,01959 ***
Pai Possui E.M.	0,09346	0,07861	0,01485 ***	0,09346	0,08265	0,01081 ***
Rondônia	0,01803	0,01501	0,00302 ***	0,01803	0,01866	-0,00063
Acre	0,00785	0,00535	0,00250 ***	0,00785	0,00460	0,00325 ***
Amazonas	0,01246	0,01093	0,00153 *	0,01246	0,01125	0,00121
Roraima	0,00417	0,00386	0,00031	0,00417	0,00257	0,00160 ***
Pará	0,04590	0,04845	-0,00255	0,04590	0,04649	-0,00059
Amapá	0,01352	0,01060	0,00292 ***	0,01352	0,01110	0,00242 **
Tocantins	0,00693	0,01184	-0,00491 ***	0,00693	0,00557	0,00136 *
Maranhão	0,00635	0,00793	-0,00158 **	0,00635	0,00548	0,00087
Piauí	0,01003	0,00734	0,00269 ***	0,01003	0,00873	0,00130
Ceará	0,01600	0,00975	0,00625 ***	0,01600	0,01386	0,00214 *
Rio Grande do Norte	0,00417	0,00355	0,00062	0,00417	0,00228	0,00189 ***
Paraíba	0,00562	0,00405	0,00157 ***	0,00562	0,00364	0,00198 ***
Pernambuco	0,00538	0,00457	0,00081	0,00538	0,00383	0,00155 **
Alagoas	0,01071	0,01106	-0,00035	0,01071	0,00887	0,00184 *
Sergipe	0,01285	0,01202	0,00083	0,01285	0,01275	0,00010
Bahia	0,02831	0,04745	-0,01914 ***	0,02831	0,02758	0,00073
Minas Gerais	0,06747	0,10282	-0,03535 ***	0,06747	0,06854	-0,00107
Espírito Santo	0,01837	0,01773	0,00064	0,01837	0,01629	0,00208
Rio de Janeiro	0,07794	0,08993	-0,01199 ***	0,07794	0,08245	-0,00451 *
São Paulo	0,12065	0,14846	-0,02781 ***	0,12065	0,12627	-0,00562 *
Paraná	0,20572	0,13213	0,07359 ***	0,20572	0,21260	-0,00688 *
Santa Catarina	0,06011	0,04568	0,01443 ***	0,06011	0,06238	-0,00227
Rio Grande do Sul	0,10252	0,11619	-0,01367 ***	0,10252	0,10620	-0,00368
Mato Grosso do Sul	0,04804	0,05353	-0,00549 ***	0,04804	0,04779	0,00025
Mato Grosso	0,00868	0,00997	-0,00129 *	0,00868	0,00780	0,00088
Goiás	0,02685	0,03186	-0,00501 ***	0,02685	0,02700	-0,00015
Distrito Federal	0,05536	0,03793	0,01743 ***	0,05536	0,05540	-0,00004

Fonte: Elaborada pelos autores.

Notas: 1. A hipótese nula do teste de diferença de médias é H_0 : Diferença = 0;

2. , *, ** e *** representam confiança de 90%, 95% e 99%, respectivamente.

Apêndice A2 - Tabela 8 Diferença de médias, antes e depois do pareamento, entre os grupos de tratamento e controle – Matemática (Ensino Fundamental)

Variável	Antes Matching			Depois Matching		
	Tratado	Controle	Diferença	Tratado	Controle	Diferença
Idade	31,26900	34,00200	-2,73300 ***	31,26900	31,11700	0,15200
Feminino	0,55104	0,53364	0,01740 ***	0,55104	0,55045	0,00059
Renda	1,24110	1,37320	-0,13210 ***	1,24110	1,21420	0,02690 *
Quando Evadiu	5,16360	5,06520	0,09840 ***	5,16360	5,20870	-0,04510 **
Mãe Possui E.M.	0,18742	0,15143	0,03599 ***	0,18742	0,18013	0,00729 ***
Pai Possui E.M.	0,09381	0,07601	0,01780 ***	0,09381	0,08880	0,00501 ***
Rondônia	0,01696	0,01405	0,00291 ***	0,01696	0,01551	0,00145
Acre	0,00749	0,00508	0,00241 ***	0,00749	0,00525	0,00224 ***
Amazonas	0,01239	0,00987	0,00252 ***	0,01239	0,01069	0,00170 ***
Roraima	0,00467	0,00378	0,00089 *	0,00467	0,00369	0,00098 **
Pará	0,04734	0,04904	-0,00170	0,04734	0,04890	-0,00156
Amapá	0,01405	0,01081	0,00324 ***	0,01405	0,01463	-0,00058
Tocantins	0,00768	0,01166	-0,00398 ***	0,00768	0,00671	0,00097 **
Maranhão	0,00651	0,00809	-0,00158 **	0,00651	0,00588	0,00063 ***
Piauí	0,01055	0,00754	0,00301 ***	0,01055	0,00826	0,00229
Ceará	0,01701	0,01003	0,00698 ***	0,01701	0,01555	0,00146 ***
Rio Grande do Norte	0,00452	0,00350	0,00102 **	0,00452	0,00350	0,00102 **
Paraíba	0,00578	0,00401	0,00177 ***	0,00578	0,00442	0,00136 **
Pernambuco	0,00535	0,00476	0,00059	0,00535	0,00452	0,00083 ***
Alagoas	0,01060	0,01099	-0,00039	0,01060	0,00962	0,00098
Sergipe	0,01283	0,01206	0,00077	0,01283	0,01288	-0,00005
Bahia	0,02887	0,04740	-0,01853 ***	0,02887	0,02805	0,00082
Minas Gerais	0,07023	0,10540	-0,03517 ***	0,07023	0,07038	-0,00015
Espírito Santo	0,01930	0,01807	0,00123	0,01930	0,01711	0,00219
Rio de Janeiro	0,07923	0,08893	-0,00970 ***	0,07923	0,08525	-0,00602
São Paulo	0,12258	0,14792	-0,02534 ***	0,12258	0,12628	-0,00370
Paraná	0,18902	0,12912	0,05990 ***	0,18902	0,19316	-0,00414 *
Santa Catarina	0,05735	0,04536	0,01199 ***	0,05735	0,05837	-0,00102
Rio Grande do Sul	0,10912	0,11882	-0,00970 ***	0,10912	0,10980	-0,00068
Mato Grosso do Sul	0,04899	0,05432	-0,00533 ***	0,04899	0,05138	-0,00239
Mato Grosso	0,00865	0,00946	-0,00081	0,00865	0,00739	0,00126 **
Goiás	0,02780	0,03195	-0,00415 ***	0,02780	0,02741	0,00039
Distrito Federal	0,05512	0,03798	0,01714 ***	0,05512	0,05541	-0,00029

Fonte: Elaborada pelos autores.

Notas: 1. A hipótese nula do teste de diferença de médias é H_0 : Diferença = 0;

2. , *, ** e *** representam confiança de 90%, 95% e 99%, respectivamente.

Apêndice A2 - Tabela 9 Resultados estimados para a probabilidade de receber o tratamento - Probit (Linguagens e redação - Ensino Médio)

Variável	Parâmetro Estimado	Erro Padrão
Idade	-0,00467 ***	0,00022
Feminino	0,06196 ***	0,00399
Renda	-0,02003 ***	0,00147
Quando Evadiu	-0,05858 ***	0,00103
Mãe Possui E.M.	0,04835 ***	0,00606
Pai Possui E.M.	-0,03215 ***	0,00732
Rondônia	0,07102 ***	0,02171
Acre	-0,02572 ***	0,02852
Amazonas	-0,20662	0,02347
Roraima	0,05326 ***	0,03343
Pará	-0,29371	0,01478
Amapá	-0,11299 ***	0,02421
Tocantins	-0,16780 ***	0,02175
Maranhão	-0,35131 ***	0,02381
Piauí	-0,17302 ***	0,02436
Ceará	0,14650 ***	0,01668
Rio Grande do Norte	-0,26793 ***	0,02940
Paraíba	-0,13754 ***	0,02457
Pernambuco	-0,26028 ***	0,02769
Alagoas	-0,26104 ***	0,02125
Sergipe	-0,18551 ***	0,02032
Bahia	-0,50996 ***	0,01403
Minas Gerais	-0,35052 ***	0,01254
Espírito Santo	-0,07673 ***	0,01794
Rio de Janeiro	-0,19280 ***	0,01217
São Paulo	-0,23849 ***	0,01172
Paraná	-0,12000 ***	0,01195
Santa Catarina	-0,11397 ***	0,01332
Rio Grande do Sul	-0,12819 ***	0,01234
Mato Grosso do Sul	-0,16334 ***	0,01459
Mato Grosso	-0,15871 ***	0,02223
Goiás	-0,26813 ***	0,01480
Distrito Federal	<i>Omitted</i>	
Constante	0,01966	0,01555
Número de observações	528.713	

Fonte: Elaborada pelos autores.

Notas: 1. *, ** e *** representam confiança de 90%, 95% e 99%, respectivamente.

Apêndice A2 - Tabela 10 Resultados estimados para a probabilidade de receber o tratamento - Probit (Ciências da Natureza - Ensino Médio)

Variável	Parâmetro Estimado	Erro Padrão
Idade	-0,00434 ***	0,00022
Feminino	0,06574 ***	0,00408
Renda	-0,02086 ***	0,00151
Quando Evadiu	-0,05951 ***	0,00105
Mãe Possui E.M.	0,04627 ***	0,00619
Pai Possui E.M.	-0,03231 ***	0,00748
Rondônia	0,08885 ***	0,02265
Acre	0,00036 ***	0,03013
Amazonas	-0,19654 ***	0,02431
Roraima	0,05568 ***	0,03465
Pará	-0,28507	0,01508
Amapá	-0,14025 ***	0,02494
Tocantins	-0,16822 ***	0,02321
Maranhão	-0,32928 ***	0,02396
Piauí	-0,15813 ***	0,02483
Ceará	0,15696 ***	0,01711
Rio Grande do Norte	-0,25811 ***	0,02966
Paraíba	-0,10919 ***	0,02489
Pernambuco	-0,23432 ***	0,02773
Alagoas	-0,24277 ***	0,02166
Sergipe	-0,16886 ***	0,02066
Bahia	-0,50031 ***	0,01439
Minas Gerais	-0,33517 ***	0,01296
Espírito Santo	-0,05349 ***	0,01854
Rio de Janeiro	-0,18585 ***	0,01258
São Paulo	-0,22643 ***	0,01214
Paraná	-0,11284 ***	0,01238
Santa Catarina	-0,10809 ***	0,01381
Rio Grande do Sul	-0,12074 ***	0,01280
Mato Grosso do Sul	-0,17730 ***	0,01551
Mato Grosso	-0,14524 ***	0,02289
Goiás	-0,25742 ***	0,01524
Distrito Federal	<i>Omitted</i>	
Constante	0,00265 **	0,01605
Número de observações	506.453	

Fonte: Elaborada pelos autores.

Notas: 1. *, ** e *** representam confiança de 90%, 95% e 99%, respectivamente.

Apêndice A2 - Tabela 11 Resultados estimados para a probabilidade de receber o tratamento - Probit (Ciências Humanas - Ensino Médio)

Variável	Parâmetro Estimado	Erro Padrão
Idade	-0,00425 ***	0,00022
Feminino	0,06350 ***	0,00405
Renda	-0,02119 ***	0,00150
Quando Evadiu	-0,05989 ***	0,00105
Mãe Possui E.M.	0,04634 ***	0,00615
Pai Possui E.M.	-0,02987 ***	0,00742
Rondônia	0,07818 ***	0,02200
Acre	-0,01541 ***	0,02891
Amazonas	-0,21037	0,02375
Roraima	0,03268 ***	0,03374
Pará	-0,28972	0,01492
Amapá	-0,12634 ***	0,02436
Tocantins	-0,17403 ***	0,02249
Maranhão	-0,32988 ***	0,02375
Piauí	-0,16959 ***	0,02469
Ceará	0,13810 ***	0,01706
Rio Grande do Norte	-0,25471 ***	0,02960
Paraíba	-0,11677 ***	0,02473
Pernambuco	-0,24614 ***	0,02767
Alagoas	-0,24758 ***	0,02140
Sergipe	-0,20001 ***	0,02037
Bahia	-0,50280 ***	0,01416
Minas Gerais	-0,34608 ***	0,01278
Espírito Santo	-0,08058 ***	0,01845
Rio de Janeiro	-0,19437 ***	0,01236
São Paulo	-0,23490 ***	0,01192
Paraná	-0,11652 ***	0,01214
Santa Catarina	-0,11475 ***	0,01360
Rio Grande do Sul	-0,13030 ***	0,01258
Mato Grosso do Sul	-0,17838 ***	0,01508
Mato Grosso	-0,14485 ***	0,02240
Goiás	-0,26631 ***	0,01507
Distrito Federal	<i>Omitted</i>	
Constante	0,01450	0,01584
Número de observações	511.990	

Fonte: Elaborada pelos autores.

Notas: 1. *, ** e *** representam confiança de 90%, 95% e 99%, respectivamente.

Apêndice A2 - Tabela 12 Resultados estimados para a probabilidade de receber o tratamento - Probit (Matemática - Ensino Médio)

Variável	Parâmetro Estimado	Erro Padrão
Idade	-0,00443 ***	0,00021
Feminino	0,06468 ***	0,00385
Renda	-0,02093 ***	0,00143
Quando Evadiu	-0,05756 ***	0,00099
Mãe Possui E.M.	0,04719 ***	0,00587
Pai Possui E.M.	-0,03219 ***	0,00711
Rondônia	0,07369 ***	0,02148
Acre	-0,02600 ***	0,02829
Amazonas	-0,21916	0,02277
Roraima	0,05181 ***	0,03195
Pará	-0,28584	0,01406
Amapá	-0,12008 ***	0,02299
Tocantins	-0,16420 ***	0,02099
Maranhão	-0,31716 ***	0,02262
Piauí	-0,14610 ***	0,02332
Ceará	0,16322 ***	0,01587
Rio Grande do Norte	-0,24986 ***	0,02830
Paraíba	-0,12587 ***	0,02355
Pernambuco	-0,24194 ***	0,02646
Alagoas	-0,26679 ***	0,02008
Sergipe	-0,19240 ***	0,01952
Bahia	-0,49554 ***	0,01335
Minas Gerais	-0,34535 ***	0,01199
Espírito Santo	-0,05961 ***	0,01726
Rio de Janeiro	-0,18821 ***	0,01165
São Paulo	-0,23742 ***	0,01121
Paraná	-0,14211 ***	0,01146
Santa Catarina	-0,12766 ***	0,01292
Rio Grande do Sul	-0,12340 ***	0,01182
Mato Grosso do Sul	-0,15248 ***	0,01397
Mato Grosso	-0,15727 ***	0,02173
Goiás	-0,25761 ***	0,01414
Distrito Federal	<i>Omitted</i>	
Constante	0,00381	0,01497
Número de observações	567.875	

Fonte: Elaborada pelos autores.

Notas: 1. *, ** e *** representam confiança de 90%, 95% e 99%, respectivamente.

Apêndice A2 - Tabela 13 Resultados estimados para a probabilidade de receber o tratamento - Probit (Linguagens e Redação - Ensino Fundamental)

Variável	Parâmetro Estimado	Erro Padrão
Idade	-0,01101 ***	0,00046
Feminino	0,06453 ***	0,00920
Renda	-0,01404 ***	0,00355
Quando Evadiu	-0,00903 ***	0,00305
Mãe Possui E.M.	0,07787 ***	0,01572
Pai Possui E.M.	-0,05010 **	0,02048
Rondônia	-0,09877 ***	0,03953
Acre	-0,01105 *	0,05827
Amazonas	-0,13145	0,04579
Roraima	-0,11694 **	0,07215
Pará	-0,25841	0,02892
Amapá	-0,10045 **	0,04473
Tocantins	-0,46981	0,04995
Maranhão	-0,29834 ***	0,05648
Piauí	-0,01384 ***	0,05147
Ceará	0,08780 ***	0,04403
Rio Grande do Norte	-0,12143 **	0,07417
Paraíba	-0,04486	0,06730
Pernambuco	-0,12726	0,06702
Alagoas	-0,25073 *	0,04779
Sergipe	-0,22079 ***	0,04455
Bahia	-0,52560 ***	0,03118
Minas Gerais	-0,39099 ***	0,02552
Espírito Santo	-0,20359 ***	0,03855
Rio de Janeiro	-0,26687 ***	0,02544
São Paulo	-0,25335 ***	0,02378
Paraná	0,06181 ***	0,02286
Santa Catarina	-0,01695	0,02783
Rio Grande do Sul	-0,25012 ***	0,02408
Mato Grosso do Sul	-0,24878 ***	0,02783
Mato Grosso	-0,22148 ***	0,05054
Goiás	-0,28227 ***	0,03315
Distrito Federal	<i>Omitted</i>	
Constante	-0,22607 ***	0,03155
Número de observações	103.016	

Fonte: Elaborada pelos autores.

Notas: 1. *, ** e *** representam confiança de 90%, 95% e 99%, respectivamente.

Apêndice A2 - Tabela 14 Resultados estimados para a probabilidade de receber o tratamento - Probit (Ciências da Natureza - Ensino Fundamental)

Variável	Parâmetro Estimado	Erro Padrão
Idade	-0,01125 ***	0,00048
Feminino	0,06957 ***	0,00949
Renda	-0,01497 ***	0,00367
Quando Evadiu	-0,01088 ***	0,00315
Mãe Possui E.M.	0,05289 ***	0,01622
Pai Possui E.M.	-0,02365	0,02111
Rondônia	-0,06530	0,04279
Acre	0,05402	0,06069
Amazonas	-0,08493 *	0,04774
Roraima	-0,10742	0,07367
Pará	-0,22286 ***	0,02955
Amapá	-0,04949	0,04532
Tocantins	-0,43704	0,05174
Maranhão	-0,28240 ***	0,05676
Piauí	0,01823 ***	0,05270
Ceará	0,13454	0,04470
Rio Grande do Norte	-0,04480 **	0,07489
Paraíba	0,01505	0,06784
Pernambuco	-0,09425	0,06752
Alagoas	-0,20580 *	0,04803
Sergipe	-0,19973 ***	0,04560
Bahia	-0,49090 ***	0,03187
Minas Gerais	-0,35193 ***	0,02659
Espírito Santo	-0,13106 ***	0,03954
Rio de Janeiro	-0,22877 ***	0,02642
São Paulo	-0,21239 ***	0,02483
Paraná	0,06462 ***	0,02404
Santa Catarina	-0,01345	0,02920
Rio Grande do Sul	-0,21622	0,02520
Mato Grosso do Sul	-0,22413 ***	0,02929
Mato Grosso	-0,17709 ***	0,05270
Goiás	-0,22610 ***	0,03388
Distrito Federal	<i>Omitted</i>	
Constante	-0,25876 ***	0,03283
Número de observações	97.860	

Fonte: Elaborada pelos autores.

Notas: 1. *, ** e *** representam confiança de 90%, 95% e 99%, respectivamente.

Apêndice A2 - Tabela 15 Resultados estimados para a probabilidade de receber o tratamento - Probit (Ciências Humanas - Ensino Fundamental)

Variável	Parâmetro Estimado	Erro Padrão
Idade	-0,01146 ***	0,00048
Feminino	0,06325 ***	0,00949
Renda	-0,01499 ***	0,00368
Quando Evadiu	-0,01209 ***	0,00315
Mãe Possui E.M.	0,06971 ***	0,01615
Pai Possui E.M.	-0,04516 **	0,02105
Rondônia	-0,06695	0,04105
Acre	0,02114	0,05979
Amazonas	-0,12998 **	0,04695
Roraima	-0,15413 **	0,07447
Pará	-0,23189 ***	0,02974
Amapá	-0,10614 **	0,04625
Tocantins	-0,49952 ***	0,05291
Maranhão	-0,30476 ***	0,05825
Piauí	0,01361	0,05299
Ceará	0,10446	0,04534
Rio Grande do Norte	-0,11415	0,07592
Paraíba	-0,00645	0,06841
Pernambuco	-0,07829	0,06756
Alagoas	-0,22054 *	0,04838
Sergipe	-0,20414 ***	0,04575
Bahia	-0,51301 ***	0,03206
Minas Gerais	-0,37023 ***	0,02662
Espírito Santo	-0,16152 ***	0,03990
Rio de Janeiro	-0,25037 ***	0,02638
São Paulo	-0,23027 ***	0,02479
Paraná	0,07801 ***	0,02388
Santa Catarina	0,00500	0,02901
Rio Grande do Sul	-0,24429 ***	0,02526
Mato Grosso do Sul	-0,23776 ***	0,02924
Mato Grosso	-0,22970 ***	0,05191
Goiás	-0,25635 ***	0,03407
Distrito Federal	<i>Omitted</i>	
Constante	-0,21763 ***	0,03271
Número de observações	97.222	

Fonte: Elaborada pelos autores.

Notas: 1. *, ** e *** representam confiança de 90%, 95% e 99%, respectivamente.

Apêndice A2 - Tabela 16 Resultados estimados para a probabilidade de receber o tratamento
- Probit (Ciências Humanas - Matemática)

Variável	Parâmetro Estimado	Erro Padrão
Idade	-0,01081 ***	0,00047
Feminino	0,07114 ***	0,00944
Renda	-0,01750 ***	0,00365
Quando Evadiu	-0,01079 ***	0,00312
Mãe Possui E.M.	0,05467 ***	0,01627
Pai Possui E.M.	-0,00546	0,02115
Rondônia	-0,06801	0,04182
Acre	0,02040	0,06081
Amazonas	-0,07678	0,04763
Roraima	-0,07131	0,07228
Pará	-0,22341 ***	0,02942
Amapá	-0,09589 **	0,04549
Tocantins	-0,43705 ***	0,05149
Maranhão	-0,30326 ***	0,05740
Piauí	0,02117 **	0,05186
Ceará	0,12036	0,04431
Rio Grande do Norte	-0,06174	0,07415
Paraíba	0,00940 *	0,06771
Pernambuco	-0,11527 ***	0,06683
Alagoas	-0,22693 ***	0,04833
Sergipe	-0,20752 ***	0,04553
Bahia	-0,50390 ***	0,03181
Minas Gerais	-0,36566 ***	0,02630
Espírito Santo	-0,14610 ***	0,03918
Rio de Janeiro	-0,23895 ***	0,02622
São Paulo	-0,22602 ***	0,02461
Paraná	0,03918	0,02389
Santa Catarina	-0,02131	0,02904
Rio Grande do Sul	-0,22430 ***	0,02495
Mato Grosso do Sul	-0,23459 ***	0,02897
Mato Grosso	-0,20396 ***	0,05217
Goiás	-0,24084 ***	0,03367
Distrito Federal	<i>Omitted</i>	0,03260
Constante	-0,25641 ***	
Número de observações	99.105	

Fonte: Elaborada pelos autores.

Notas: 1. *, ** e *** representam confiança de 90%, 95% e 99%, respectivamente.

4. ONDE ESTÁ A CORRUPÇÃO NO BRASIL? UMA ANÁLISE SOB A ÓTICA DO ÍNDICE MICRORREGIONAL DE CORRUPÇÃO.

Resumo

O objetivo deste trabalho é calcular um Índice de Microrregional de Corrupção (IMRC), para 220 regiões do Brasil. A metodologia adotada foi o *Rescaling* (normalização das variáveis, unidas em um índice por média simples). Para avaliar tendências espaciais da corrupção, utilizou-se os Índices de Moran Global e Local (*LISA*). Os dados, ponderados pela população de cada município, referem-se aos anos de 2015 e 2016 e são oriundos de três fontes: (i) Análise de 27.975 processos judiciais de corrupção, nas 220 Procuradorias Regionais do Ministério Público Federal, com jurisdição nos 5.570 municípios brasileiros; (ii) Análise de 2.312 contas irregulares do Cadastro de Contas Julgadas Irregulares – *Cadirreg* – do Tribunal de Contas da União; e (iii) Análise do Índice de transparência Municipal do MPF, calculado para todos os municípios da federação. O IMRC varia em entre 0 e 1, sendo um indicador de menos corrupção valores próximos a zero, e mais corrupção valores próximos a um. Os resultados apontam a formação de clusters espaciais da corrupção do tipo alta incidência (*high-high*) para as regiões dos estados do Norte, e baixa incidência (*low-low*) para os do Sul. Para o Brasil o IMRC médio reduziu e passou de 0,23 em 2015 para 0,21 em 2016. Por fim, o trabalho inova ao calcular um índice com abrangência microrregional, considerando variáveis relevantes para indicar corrupção. Ademais, indica áreas para a atuação do Estado em ações de combate à corrupção, bem como disponibiliza o IMRC, que possibilita diversas pesquisas sobre corrupção.

Palavras Chave: Corrupção Regional, Análise Espacial, Combate à Corrupção.

JEL: D73 R10 R30

Abstract

The objective of this work is to calculate a Microregional Corruption Index (IMCR), for 220 regions os Brazil. The methodology adopted for *Rescaling* (normalization of variables, joined in an index by simple average). To assess spatial trends in corruption, using the Global and Local Moran Indexes (*LISA*). The data, weighted by the population of each location, is located in the years 2015 and 2016 and comes from three sources: (i) Analysis of 27,975 judicial corruption cases, in the 220 Regional Prosecutorships of the Federal Public Ministry, with jurisdiction over 5,570 Brazilian municipalities; (ii) Analysis of 2,312 irregular accounts from the Register of Judged Irregular Accounts - *Cadirreg* - from the Federal Audit Court; and (iii) Analysis of the Municipal Transparency Index of the MPF, distribution to all municipalities in the federation. The IMRC varies between 0 and 1, being an indicator of less corruption values close to zero, and more

corruption values close to one. The results point to the formation of spatial clusters of corruption of the high incidence (high-high) type for the regions of the northern states, and low incidence (low-low) for those in the south. At the Brazilian level, the average IMRC decreased to 0.23 in 2015 to 0.21 in 2016. Finally, the work innovates when calculating an index with regional coverage, considering relevant variables to indicate corruption. Furthermore, it indicates areas for the State to act in actions to combat corruption, as well as providing the IMRC, which enables research on corruption.

Keywords: Regional Corruption, Spatial Analysis, Combating Corruption.

JEL: D73 R10 R30

4.1 INTRODUÇÃO

O estudo do crime sob a ótica da racionalidade econômica foi proposto por Gary Becker no trabalho "*Crime and Punishment: An economic approach*" em 1968. À luz da Teoria Econômica do Crime, os indivíduos, ao decidirem se praticam certo delito, ponderam entre os custos incorridos e os benefícios esperados pela prática, tal como um agente racional que maximiza a sua função utilidade (BECKER, 1968). Desde este trabalho pioneiro, diversos economistas vêm estudando o fenômeno social complexo denominado crime, em especial, o crime de corrupção.

O crime de corrupção recebe diversas definições que, em sua essência, são bastante próximas. De acordo com World Bank (2019), a corrupção é o poder público exercido para ganhos privados, incluindo pequenas e grandes formas de corrupção, bem como a ‘captura’ do Estado por elites e interesses privados. Define-se ainda a prática como o abuso de poder confiado a alguém para obtenção de ganho privado (TRANSPARENCY INTERNATIONAL, 2019; ONU, 2019). A corrupção se dá em espaços públicos, privados e sociais. No âmbito público relaciona-se ao aparelho do Estado, no espaço privado é gerado nas relações empresariais, enquanto no campo social ocorre em organizações não governamentais, além de outras instâncias sociais (CARDENAS et al. 2005).

Nos primórdios da discussão, a corrupção era vista como a “lubrificação” da economia, a qual proporcionaria Crescimento Econômico (LEFF, 1964; E HUNTINGTON, 1968), visto que por intermédio da prática corrupta entraves burocráticos seriam eliminados, possibilitando um ambiente mais fértil aos negócios.

Atualmente essa questão parece superada e a literatura converge para resultados empíricos que mostram a corrupção como um fator de entrave ao Crescimento Econômico (BARDHAN, 1997; MAURO, 1998; GRÜNDLER e POTRAFKE, 2019; BAKLOUTI e BOUJELBENE, 2020), reduzindo os investimentos externos (LAMBSDORFF, 2003; ZAKHAROV, 2019), diminuindo a renda (KAUFMANN; KRAAY, MASTRUZI 2003), produzindo ineficiência e injustiça (ROSE-ACKERMANN, 1975), influencia negativamente nas decisões de esforço produtivo das empresas, tendo em vista que, a corrupção diminui a eficiência econômica e a produtividade ao mesmo tempo (SOTO, 2003), além de retardar processos de inovação (DINCER, 2019)

Em virtude da complexidade e do potencial lesivo da corrupção, diversos Organismos Internacionais, tais como o Banco Mundial, a Organização das Nações Unidas - ONU, a Organização para a Cooperação e Desenvolvimento Econômico - OCDE, a

Transparência Internacional, entre outros, empenharam-se para disponibilizar medidas de mensuração da corrupção e a fomentar políticas efetivas para combater esse mal que assola diversos países.

No Brasil, não foram identificadas políticas específicas que visem combater a corrupção. As atuais ferramentas contra a corrupção ficam no âmbito das instituições com competência para tal, tais como os tribunais de contas, as controladorias, o ministério público e as polícias judiciárias. O que se verifica no país são programas de promoção à integridade e à transparência, tais como o Conselho de Transparência, Integridade e Combate à Corrupção⁸, órgão consultivo vinculado à estrutura da Controladoria-Geral da União, é um colegiado composto por representantes do governo (11 membros) e da sociedade civil (30 membros). De fato é relevante mensurar a corrupção, a fim de combatê-la.

Atualmente a principal forma de mensuração da corrupção é um índice subjetivo, denominado Índice de Percepção da Corrupção (IPC), calculado pela Transparência Internacional desde 1995. Em 2019, o Brasil passou a ocupar a posição 106º no Índice de Percepção da Corrupção (IPC), que considera 180 países, sendo classificado como um país altamente corrupto. Esta foi a 5º piora do Brasil no ranking desde 2012, ano em que passou a ser possível avaliar a série histórica.

A estimação do IPC é criticada, pois países com instituições democráticas possuem mais chance de ter corrupção revelada, fator não considerado no índice, e que influencia na percepção geral da sociedade (AVRITZER et al., 2012). Ademais, Abramo (2005) ressalta que o meio de informação que mais traz notícias sobre corrupção é o noticiário, assim é possível que ações de combate à corrupção tenham implicações na percepção de aumento no fenômeno, e o que pode estar ocorrendo de fato são ações de combate a esse tipo de crime.

Além disso, embora todos os países sejam comparados no mesmo ranking, a metodologia de mensuração diverge entre os países. Soma-se a tais desvantagens, o fato de que por ranquear os países em lista, pode ser que simples movimentos de alguns países acarretem alteração da posição de outros, sem que qualquer mudança real em termos de corrupção esteja ocorrendo.

Apesar das críticas ao índice subjetivo, é reconhecido que há dificuldade também em mensurar a corrupção com variáveis objetivas e ferramentas estatísticas, visto que, como qualquer outro crime, ela ocorre às margens da legalidade, às escondidas. Em que pese as

⁸ Decreto nº 11.528, de 16 de maio de 2023- Institui o Conselho de Transparência, Integridade e Combate à Corrupção.

dificuldades é relevante estudar este fenômeno social complexo, que possui poucas pesquisas para o Brasil.

Nesse contexto, o presente trabalho propõe a estimação do Índice Microrregional de Corrupção para o Brasil, calculado com base em fatores objetivos, com cobertura para 220 distintas regiões do país, nos anos de 2015 e 2016. O índice proposto contempla dados da Transparência Municipal e de processos judiciais de crimes de corrupção julgados nas Procuradorias Gerais do Ministério Público Federal. Considera-se ainda o Cadastro de Responsáveis com Contas Julgadas Irregulares (Cadirreg) utilizado para estimar índices objetivos de corrupção em nível estadual, propostos por Boll (2010) e Carraro et al. (2015), para o período de 1998 a 2008. A partir do índice proposto será avaliada a correlação espacial da corrupção entre as microrregiões do Brasil. Pesquisas que exploraram essa distribuição espacial apontam que existe correlação espacial da corrupção entre países vizinhos, de forma que políticas focalizadas em uma região podem acarretar diminuição da corrupção (CRUZ, 2017).

Após esta introdução a revisão de literatura sobre a mensuração objetiva da corrupção consta na seção 2. Na seção seguinte é apresentada a literatura sobre correlação espacial da corrupção. Na seção 4 a base de dados utilizada é descrita, seguida pela descrição da metodologia aplicada para a estimação do índice. Por fim, apresentamos os resultados e as considerações finais.

4.2 MENSURAÇÃO OBJETIVA DA CORRUPÇÃO

A corrupção é um crime de difícil mensuração, seja com dados objetivos ou subjetivos, pois os atos ilícitos ocorrem às margens do que se pode observar. O índice de corrupção utilizado mundialmente é de caráter subjetivo e recebe diversas críticas já citadas na introdução desta pesquisa. Assim, questiona-se como mensurar a corrupção com dados objetivos.

Na literatura brasileira, a primeira pesquisa identificada para avaliar a corrupção com dados municipais foi a de Ferraz e Finan (2008), na qual foi avaliado o efeito da corrupção nos resultados das eleições municipais para a reeleição de prefeitos. Os autores utilizaram como indicador de corrupção o resultado de Relatórios de Auditoria da Controladoria Geral da União (CGU), que sortearam aleatoriamente municípios para serem fiscalizados, com uma amostra de 493 municípios com população inferior a 450.000 habitantes. Como resultado, foi identificado que a divulgação dos resultados dos relatórios de auditoria da CGU relacionados

à corrupção resultou em uma redução média de 20% nas chances de reeleição para prefeitos indicados como corruptos.

A partir da mesma base de dados utilizada por Ferraz e Finan (2008) para estimar a corrupção no Brasil, Brollo et al. (2013) pesquisaram sobre o efeito de receitas adicionais do governo sobre a corrupção política, além de avaliar o efeito dessa corrupção sobre a qualidade dos políticos, mensurada pela média da escolaridade dos prefeitos. A metodologia adotada foi uma Regressão com Descontinuidade no modelo *Fuzzy*, gerada a partir da população municipal de corte para recebimento de recursos do fundo de participação dos municípios. Os resultados indicam que aumentos nas transferências recebidas levam ao aumento na corrupção dos municípios, à uma piora na qualidade dos candidatos, ou seja, candidatos com menos escolaridade, e ao aumento nas chances de reeleição desses políticos.

Desde então, pesquisas com dados objetivos vêm sendo realizadas para estimar a corrupção no Brasil. O principal dado objetivo utilizado nessas pesquisas é o Cadastro de contas julgadas irregulares (Cadirreg) do Tribunal de Contas da União, a partir da pesquisa de Boll (2010), que propôs para o Brasil o Índice de Corrupção Estadual (ICE). Esse indicador é composto por fatores objetivos e contempla a quantidade de processos julgados irregulares, pelo Tribunal de Contas da União, em cada Estado (Cadirreg), ponderado: pela população, pelo PIB Estadual, e por recursos federais destinados aos Estados via Lei Orçamentária Anual (LOA). Após o cálculo desses índices simples, foi calculado um índice composto, ponderado pela população e pelo PIB estadual, o índice foi calculado para o período de 1998 a 2008.

Carraro et al. (2015), utilizaram base de dados semelhante à utilizada por Boll (2010), que também foi coautor da referida pesquisa. Os dados utilizados referem-se ao período de 1998 a 2008 e contemplam o Cadastro de Responsáveis com Contas Julgadas Irregulares pelo Tribunal de Contas da União (Cadirreg) e o montante de recursos distribuídos pela Lei Orçamentária Anual, em números absolutos e em reais, todavia sem a ponderação pelo quantitativo de habitantes do município. A principal alteração em relação à pesquisa de Boll (2010) refere-se à utilização do método multivariado, a partir da Análise de Componentes Principais. Os resultados do índice apontam que os Estados brasileiros que se mostraram mais corruptos foram a Bahia, São Paulo e o Maranhão; já o Acre, Rondônia e Santa Catarina apresentaram os menores índices de corrupção.

Sousa (2018) propôs um indicador para mensurar a corrupção no estado do Ceará. A estratégia utilizada foi baseada na de Boll (2010), e considerou os dados do Cadirreg disponibilizados pelo Tribunal de Contas do Estado do Ceará – TCE/CE, além de dados socioeconômicos sobre a população e PIB dos municípios daquele estado. Assim, foram

gerados indicadores compostos objetivos para os municípios do Estado. O período de análise foi de 2002 a 2011. Entre os principais resultados, foi observado que a corrupção varia entre os municípios, o que, conforme o autor contraria os preceitos do Índice da Corrupção Percebida da Transparência Internacional, que considera a corrupção homogênea entre os países. A partir de um ranking elaborado, foi verificado que os maiores valores médios do Indicador de Corrupção Municipal foram apresentados por Juazeiro do Norte, Caucaia e Santana do Acaraú, enquanto os menores por Ipueiras, Varjota e Catarina.

Costa et al. (2020), adotaram a mesma metodologia de Boll (2010) para calcular o “indicador de irregularidades por estado” para o período de 2013 a 2019. A diferença na pesquisa fica por conta de, após o cálculo do indicador, ter sido utilizado o “Índice de Transparência dos Governos Estaduais”, estimado via Ranking Nacional da Transparência, do Ministério Público Federal, para calcular a correlação entre esses dois índices. O teste de correlação de Pearson identificou correlação negativa entre corrupção e transparência. Ademais, os autores citam que, analisar a corrupção em conjunto com ferramentas de transparência estimula o controle social sobre os governantes, com efeitos na demanda por políticas públicas e serviços eficazes. Por fim, foram também encontradas evidências estatísticas da diferença da incidência da corrupção por Estado, a partir do teste ANOVA para diferenças de médias.

Na pesquisa de Garcia (2022) foi proposto um indicador de corrupção em nível Estadual, resultado da divisão da quantidade de contas do ente federativo, julgadas irregulares pelo TCU (Cadirreg) pelo respectivo quantitativo populacional do Estado, denominado CIPM (contas irregulares por milhão de habitantes). Na visão do autor, a vantagem deste índice seria permitir comparações entre estados e ao longo do tempo, para planejar, analisar e controlar políticas públicas de combate à corrupção.

Campos et al. (2018) estimaram a probabilidade de ocorrência de corrupção municipal. A base de dados contemplou achados irregulares que constam em relatórios de Auditoria da Controladoria Geral da União, considerados graves e médios, para 60 municípios. Foram analisados 1600 relatórios, e adotado um modelo de contagem para identificar as irregularidades, além de um *logit*, para estimar os fatores associados à ocorrência dessas irregularidades (corrupção). Os resultados indicam que a corrupção está associada a um maior volume de recursos orçamentários recebidos e que está mais presente em locais onde o controle social possui fraca atuação, baixo índice de desenvolvimento municipal, mas com alta densidade demográfica. Adicionalmente foi identificado que a

corrupção possui comportamento de transbordamento entre os municípios, pois os municípios localizados em regiões onde a corrupção é comum estão associados a maiores irregularidades.

Observa-se também na literatura que, a partir do índice de corrupção criado por Boll (2010), diversas outras pesquisas buscaram avaliar os impactos da corrupção, a partir desse índice. Melo et al. (2015) avaliaram a relação entre corrupção e empreendedorismo; Matos (2018) pesquisou sobre as relações entre corrupção e endividamento dos estados; Souza (2017) analisou a correlação entre o Índice de Percepção da Corrupção (dados subjetivos) com o índice de Boll (2010), que é de caráter objetivo; Sobral (2014) avaliou a dinâmica entre a corrupção e o crescimento econômico; entre outras pesquisas que utilizam índices de corrupção propostos para avaliar seus efeitos em diversas áreas.

Assim, nosso indicador se diferencia do que já foi proposto na literatura, pois considera dois novos componentes importantes para captar a corrupção, o primeiro refere-se aos processos de corrupção registrados nas procuradorias federais do Ministério Público Federal e o segundo, um fator de governança, que representa um fator positivo e de inibição à corrupção, também disponibilizado pelo MPF, além do Cadirreg, já utilizado na literatura. Adicionalmente, propomos a análise inédita de correlação espacial da corrupção para 220 microrregiões do país.

4.3 COMPORTAMENTO ESPACIAL DA CORRUPÇÃO

Além de propor um índice para 220 microrregiões do país, buscamos identificar a ocorrência de correlação espacial para a corrupção entre essas regiões, a partir do índice microrregional que será estimado. Na literatura há evidências de que a corrupção é correlacionada espacialmente, assim como de que a corrupção deve ser estudada dentro dos países, e não apenas de forma agregada entre países (DINCER, 2019).

Attila (2008) avaliou como a corrupção de um país pode afetar a corrupção dos países vizinhos, a partir de uma análise econométrica. A partir de um índice de corrupção proposto pelo índice de percepção da corrupção e seis fatores de governança, o autor investigou o comportamento espacial da corrupção entre os países. O modelo adotado foi os testes de autocorrelação espacial (Moran) e o Método dos Momentos Generalizados. Foi identificado que a corrupção de um país afeta o país vizinho. Assim, menor corrupção regional (medida pela média do nível de corrupção nos países vizinhos) está associada a um menor nível de corrupção nacional. Um mecanismo potencial indicado pelo autor é o nível de desenvolvimento econômico (PIB per capita), pois países menos desenvolvidos possuem

maiores índices de corrupção. Adicionalmente é citado que a governança importa para o combate à corrupção.

Cruz (2017) analisou a correlação espacial da corrupção em nível de entidade federativa do México entre 2001 e 2010. Os dados utilizados são da Transparência Mexicana, que calcula o índice de percepção da corrupção e da boa governança para o país. A metodologia de análise contemplou o cálculo do Indicador Local de Associação Espacial (LISA) e do índice de Moran. Os resultados do autor apontam que o padrão de localização da corrupção não é aleatório, bem como que a concentração de localidades com altos índices de corrupção encontra-se no centro do país. Os resultados ainda demonstram que há efeito *spillover* da corrupção entre as localidades analisadas, bem como que há correlação positiva entre o nível produtivo e a corrupção, da mesma forma que o tamanho do Estado. O autor recomenda que leve-se em consideração a correlação espacial da corrupção para a implementação de medidas de combate a esse mal.

Donfouet et al. (2015), estimaram um painel dinâmico espacial para contabilizar *spillovers* na corrupção. Primeiramente, os autores realizaram um teste de correlação espacial, a partir da construção de uma matriz de pesos espaciais com base nos cinco vizinhos mais próximos de cada país na amostra. Foi observado que, existe uma relação positiva e significativa entre os níveis percebidos de corrupção e sua defasagem espacial. O método de estimação adotado foi o Painel Espacial Dinâmico, com base nos cinco países vizinhos mais próximos. As variáveis utilizadas foram: o Índice de Percepção da Corrupção (IPC), o PIB per capita, o percentual de mulheres parlamentares, os gastos públicos com educação, o índice de liberdade econômica e o sistema legal adotado pelos países (*common law*), do banco de Indicadores de Desenvolvimento do Banco Mundial. Os resultados apontam que a corrupção não só exhibe *spillovers* espaciais, mas também, possui efeito persistente ao longo do tempo. Ademais, aumentar a renda per capita, a liberdade econômica e a porcentagem de mulheres nos parlamentos têm um efeito de longo prazo sobre os níveis percebidos de corrupção.

LV e GAO (2021) investigaram a ocorrência de autocorrelação espacial da corrupção, estimada a partir de um índice de percepção da corrupção do *Worldwide Governance Indicators*, e seu impacto ambiental em 91 países com dados dos anos de 2002 a 2012. A metodologia adotada foram os índices I de Moran e de Geary. Foi observado que a corrupção possui autocorrelação espacial entre os países analisados com impacto negativo sobre o desempenho ambiental desses países.

Becker et al. (2009) utilizaram os dados do Índice de Percepção da Corrupção do *World Bank* de 123 países para avaliar a correlação espacial da corrupção no período de 2000

a 2005. A metodologia adotada foi o *Spatial Lag Model* e os principais resultados indicam que a corrupção se dissemina entre um país e seus países vizinhos. Assim como que o efeito disseminação diminui à medida que a distância geográfica entre os países aumenta.

Sui et al. (2018), utilizando dados sobre percepção da corrupção do *World Bank*, testou a autocorrelação espacial sobre o efeito de contágio da corrupção entre 109 países, no período de 2002 a 2013, utilizando o índice de Moran. Uma vez identificada a correlação espacial, foi estimado um modelo de painel dinâmico (GMM), no qual foram avaliados choques de corrupção e os efeitos entre os países vizinhos. Os resultados apontam que a corrupção se espalha não apenas pelo limite geográfico, mas também entre países com níveis semelhantes de PIB per capita e com instituições democráticas. Em conclusão é identificado que, para o controle da corrupção, são necessários esforços conjuntos entre os países.

Borsky e Kalkschmied (2019) realizam uma análise para 1.232 regiões subnacionais de 81 países. O índice de corrupção utilizado na pesquisa contempla a percepção da corrupção a partir de entrevistas via questionário aplicado em diversos países. Para controlar um possível viés de percepção entre os países, influenciado por diferentes culturas, as autoras utilizaram controles para diferenças culturais que poderiam afetar a percepção da corrupção. Além da relação espacial da corrupção entre as regiões analisadas, as autoras destacam a importância de realizar pesquisas em níveis desagregados, uma vez que características econômicas, socioculturais e política tendem a ter impacto nos níveis de corrupção de regiões vizinhas. Especificamente sobre a correlação espacial da corrupção no Brasil, foi identificada uma pesquisa realizada neste tipo de análise.

Lopes (2021) investigou os efeitos *spillover* (transbordamento) da corrupção, a partir dos dados oriundos dos Relatórios de Fiscalização da Controladoria Geral da União, utilizados por Ferraz e Finan (2008) e Brollo et al. (2013). O estudo contemplou o período de 2001 a 2008. Os principais resultados indicam que o comportamento corrupto é difundido para além das fronteiras municipais e que é expressivo entre municípios menores, o que reforça o papel da interação social entre comunidades como um mecanismo de compartilhamento de normas e comportamentos sociais. A autora destaca que os resultados contribuem com uma nova abordagem para a expansão da literatura mais recente acerca do transbordamento da corrupção em níveis menores que os níveis estaduais.

4.4 BASE DE DADOS

Os dados utilizados para estimar o Índice Microrregional de Corrupção para o Brasil são oriundos de três fontes distintas, quais sejam: Cadastro de contas julgadas irregulares – Cadirreg – do Tribunal de Contas da União; Processos judiciais de corrupção nas Procuradorias Regionais do Ministério público Federal (MPF); Índice de transparência Municipal – do Ministério Público Federal, que serão explicitados a seguir.

Nossa principal contribuição com esta pesquisa é fornecer um Índice Microrregional De Corrupção para o Brasil, em complemento às evidências estaduais já produzidas pela literatura. O Cadirreg é o dado objetivo mais utilizado pela literatura brasileira (BOLL 2010; CARRARO et al. 2015; SOUSA 2018; COSTA et al. 2020; GARCIA 2022). Todavia pelo baixo quantitativo de processos registrados, as pesquisas que até então utilizaram esse dado, produziram índices de corrupção em nível estadual.

No intuito de suprir este problema nós propusemos a adição de um dado objetivo sobre corrupção que permita uma análise micro regionalizada, qual seja, o número de processos de corrupção registrado nas Procuradorias Federais do Ministério Público Federal. Há evidências na literatura de que é importante avaliar a corrupção dentro dos países (DINCER, 2019), bem como de a análise da corrupção deve ocorrer entre os municípios (LOPES, 2021).

Ainda com o objetivo de estimar um indicador confiável, considerando que há evidência na literatura de que a transparência é um fator inibidor da corrupção (RAUSCH E SOARES 2010; FORTINI E MOTTA 2016), consideramos também o índice de transparência municipal, calculado pelo Ministério Público Federal nos anos de 2015 e 2016. Assim, nas seções seguintes são apresentados os dados que comporão o Índice Microrregional de Corrupção proposto nesta pesquisa.

4.4.1 CADASTRO DE CONTAS JULGADAS IRREGULARES – CADIRREG – DO TRIBUNAL DE CONTAS DA UNIÃO.

Os responsáveis pela guarda ou gestão de dinheiro público estão obrigados constitucionalmente a prestarem contas da aplicação desses recursos. Nos casos em que a origem dos recursos é Federal, a instância responsável por julgar tais contas é o Tribunal de Contas da União-TCU.

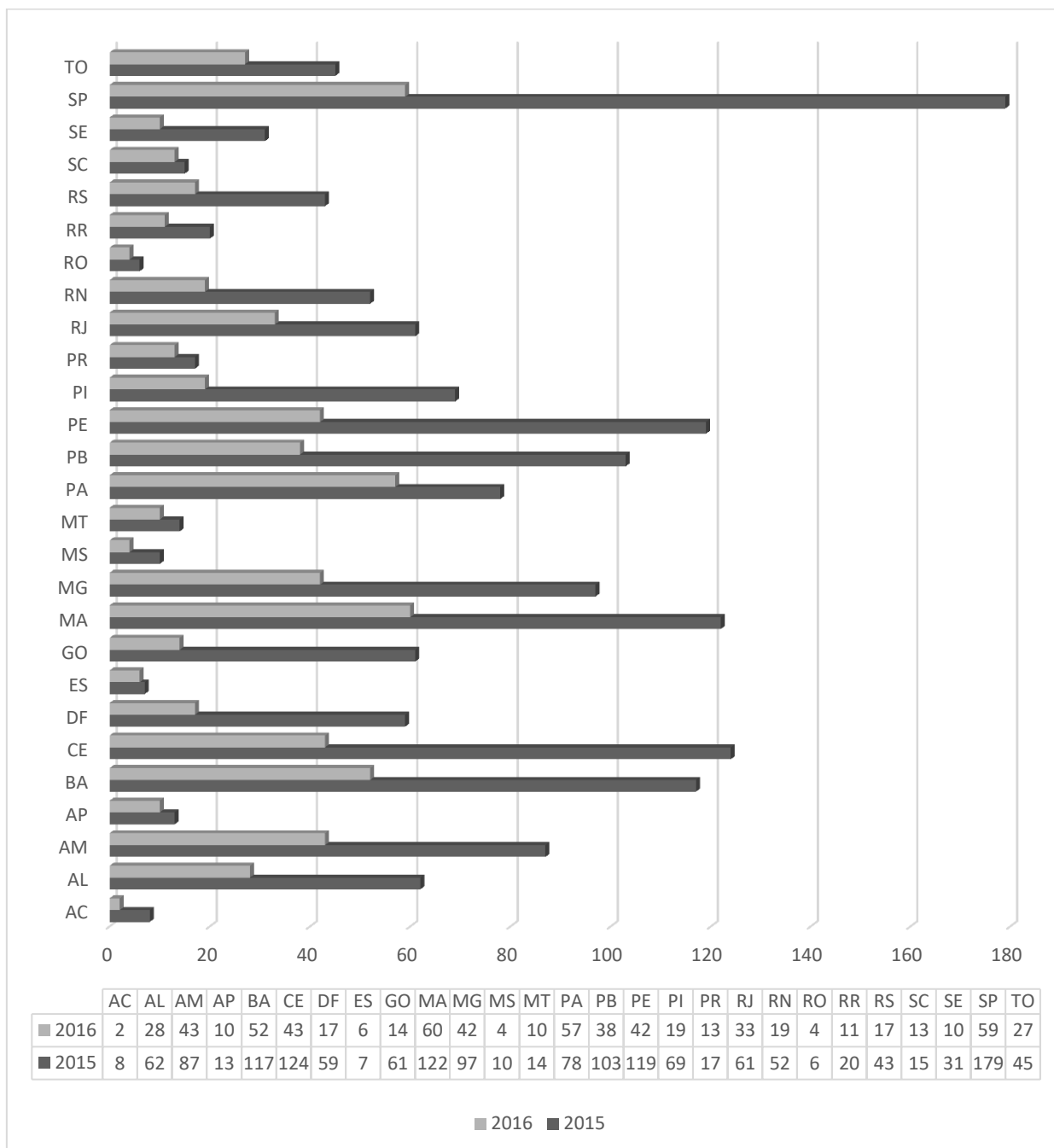
No sítio eletrônico do TCU consta registro histórico de todos os processos em trânsito julgado para pessoas físicas ou jurídicas, vivas ou falecidas, detentoras ou não de

cargo/função pública, cuja conta foi julgada como irregular pelo Tribunal. Esse banco de dados é chamado de “Cadirreg” e é possível identificar nele o quantitativo de contas julgadas irregulares em nível municipal.

As contas dos gestores são julgadas irregulares pelo TCU nos casos em que: i) ocorra omissão no dever de prestar contas; ii) exista prática de ato de gestão ilegal, ilegítimo ou antieconômico, ou infração a norma legal ou regulamentar de natureza contábil, financeira, orçamentária, operacional ou patrimonial; iii) ocorra danos ao erário decorrente de ato de gestão ilegítimo ou antieconômico; e iv) desfalque ou desvio de dinheiros, bens ou valores públicos.

O Cadirreg demonstra fragilidades na gestão dos recursos públicos, e representa um indicativo de vulnerabilidade dos órgãos à corrupção e a má aplicação do dinheiro. Este dado objetivo vem sendo utilizado pela literatura brasileira a partir do trabalho de Boll (2010). No gráfico 1 a seguir apresenta-se um resumo com o número de ocorrências registradas por Estado nos anos de 2015 e 2016, período de análise desta pesquisa.

Gráfico 1 - Número de contas julgadas irregulares registrados no Cadirreg para os anos de 2015 e 2016, por Estados.



Fonte: Elaborado pelos autores.

O estado de São Paulo concentra o maior número de contas julgadas irregulares, seguido por Maranhão, Ceará e Pernambuco. Aqueles com menor número de processos registrados são: Roraima, Acre e Espírito Santo. Destaca-se que neste quadro que, os dados ainda são brutos, sem considerar qualquer tipo de ponderação pela população.

4.4.2 REGISTROS DE PROCESSOS JUDICIAIS DE CORRUPÇÃO NAS PROCURADORIAS REGIONAIS DO MINISTÉRIO PÚBLICO FEDERAL (MPF).

O MPF é um dos principais atores no combate à corrupção no Brasil. A instituição detém legitimidade para propor ações criminais e ações por ato de improbidade administrativa contra aqueles que desviam e aplicam indevidamente recursos públicos.

Os processos judiciais relacionados à corrupção são registrados junto ao MPF, nas chamadas Procuradorias Federais. No Brasil, existem atualmente 220 Procuradorias Federais, que são responsáveis pela jurisdição de todos os 5.570 municípios. Esta divisão em 220 unidades nos permitirá a elaboração de um índice em unidades menores que a estadual.

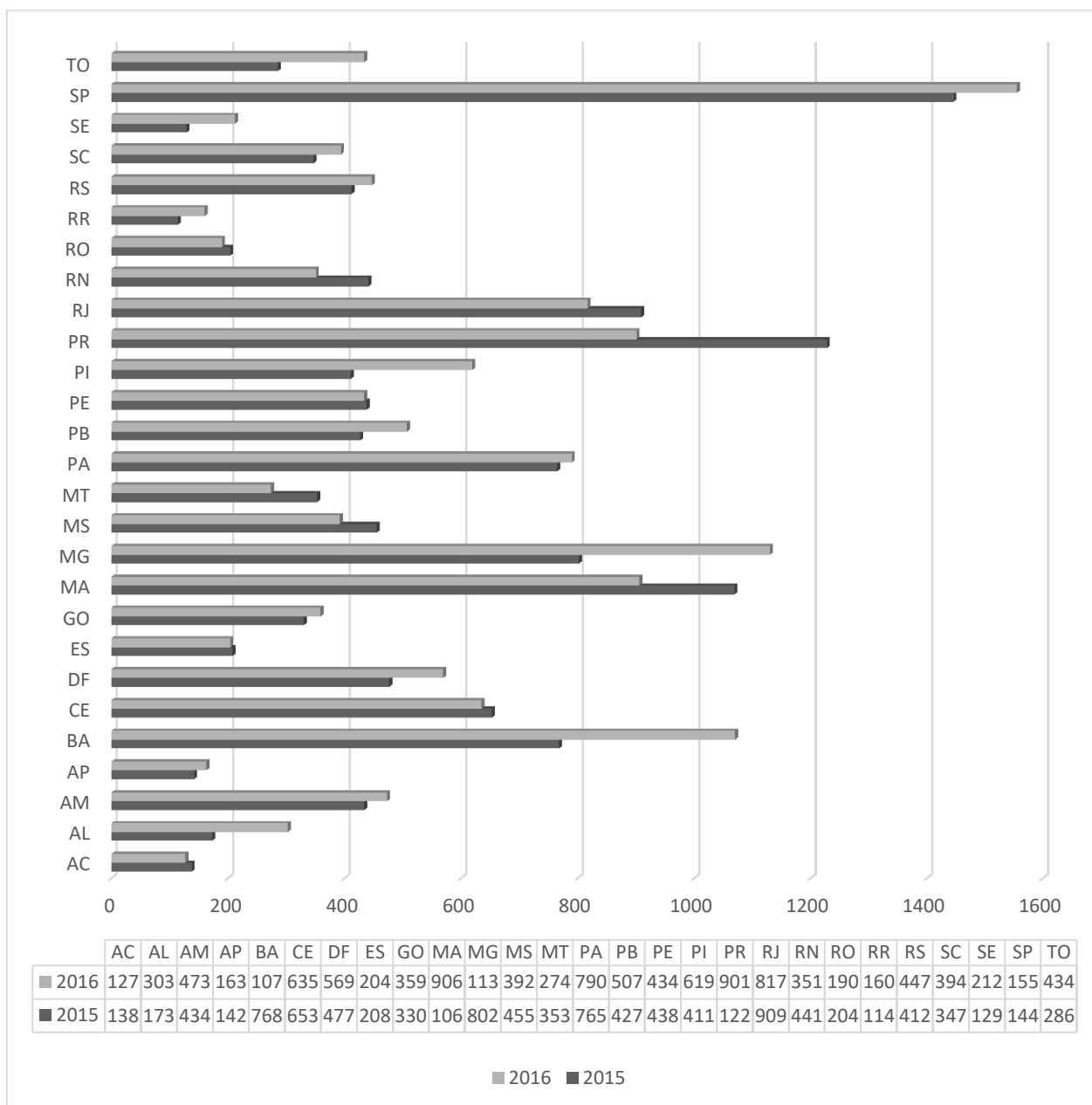
Consideramos o registro de processos por corrupção um dado semelhante ao Cadirreg, sob a hipótese de que o MPF, por ser um órgão federal, possui eficiência administrativa similar entre as 220 Procuradorias. Ademais, o combate à fraude e à corrupção no Brasil é competência de diversos órgãos, entre eles o Tribunal de Contas da União, responsável pela elaboração do Cadirreg, assim como o Ministério Público Federal, que instrui processos de corrupção e que, para os anos de 2015 e 2016 calculou o índice de transparência municipal.

Para identificar os respectivos municípios sob a jurisdição de cada Procuradoria Federal, foi realizado um levantamento pelo sítio eletrônico do Ministério Público Federal. Após a organização da base de dados, foram selecionados os processos relacionados à corrupção, classificação essa realizada pelo MPF.

Não foram identificados quais os critérios para a distribuição dos municípios entre as Procuradorias Federais, o que se verificou foi que a jurisdição está definida em regimento interno. Todavia, observa-se que a divisão está correlacionada ao tamanho populacional. A maior parte dos estados possui mais de uma Procuradoria Federal, com exceção dos estados do Amapá, Roraima e Sergipe, que possuem apenas uma.

No Gráfico 2 a seguir, podemos consultar o número de processos registrados por Estado, somando-se os quantitativos das procuradorias, e relacionados a corrupção.

Gráfico 2 - Número de processos de corrupção, por Estados, registrados na base do Ministério Público Federal para os anos de 2015 e 2016.



Fonte: Elaborado pelos autores.

Observa-se que há variação no número de processos de corrupção entre os anos analisados. Em 2016 o número de processos foi maior em 16 estados, enquanto em 2015 o número foi maior em 9 estados. Em geral, o número variou pouco entre os anos, exceto nos estados da Bahia, Maranhão, Minas Gerais e Paraná. Pesquisas adicionais devem ser realizadas para identificar as causas das variações entre os anos analisados.

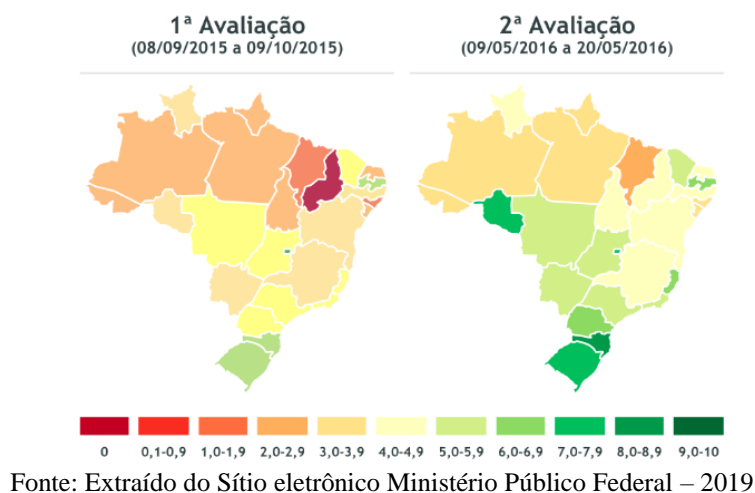
4.4.3 ÍNDICE DE TRANSPARÊNCIA MUNICIPAL - MINISTÉRIO PÚBLICO FEDERAL.

O índice de transparência municipal foi estimado nos anos de 2015 e 2016. A composição das informações acerca do Índice de transparência Municipal no Brasil foi elaborada por representantes do Ministério Público Federal (MPF), Conselho Nacional do Ministério Público (CNMP), Controladoria Geral da União (CGU), Tribunal de Contas da União (TCU), Secretaria do Tesouro Nacional (STN), Associação dos Membros dos Tribunais de Contas (ATRICON), Banco Central, entre outras instituições de controle e fiscalização (MINISTÉRIO PÚBLICO FEDERAL, 2009).

O Índice, que foi calculado para todos os municípios da federação, e avalia questões gerais de transparência, tais como página de internet, disponibilidade de relatórios, links e telefones para contato, além de questões específicas sobre divulgação de Receitas, Despesas, e Licitações e Contratos, que pode ser consultado em sua íntegra nos apêndices A3 e B3.

A Figura 1, a seguir, demonstra as duas avaliações realizadas pelo MPF, bem como o progresso ocorrido após essas avaliações.

Figura 1 - Mapa da Transparência 2015 e 2016



O MPF tem metodologia própria para ponderar o valor de cada pergunta presente no questionário sobre a corrupção, dessa forma para o índice será considerado o valor da nota dada para cada município. Esta variável foi incluída para a estimação do índice de corrupção devido ao peso do nível de transparência na percepção da corrupção, além da divulgação das informações ser uma obrigação legal dos entes federados. Ademais, a promoção da transparência é um mecanismo de prevenção à corrupção e participação democrática no país (MINISTÉRIO PÚBLICO FEDERAL, 2019).

A presença de corrupção é muitas vezes uma manifestação da falta de respeito tanto do corruptor (normalmente um cidadão privado ou empresa) quanto do corrompido (normalmente um funcionário público ou político) pelas regras que regem suas interações e, portanto, representa uma falha de governança (KAUFMANN; KRAAY; MASTRUZZI, 2003).

4.5 METODOLOGIA EMPÍRICA

4.5.1 CÁLCULOS DO ÍNDICE MICRORREGIONAL DE CORRUPÇÃO

Com o objetivo de dar o correto peso ao número de processos por procuradoria, utilizou-se uma estratégia realizada por Boll (2010) de ponderar o número de ocorrências pela população da área de cobertura de cada procuradoria. Essa metodologia foi aplicada para os processos ligados a corrupção e para as contas julgadas irregulares no Cadireg.

Para calcularmos a variável de transparência que está ao nível municipal, utilizamos uma proporção que pondera a avaliação que cada município recebeu do MPF em relação ao seu nível de transparência, segundo a população total da área de jurisdição de cada procuradoria. O valor desse somatório é subtraído de 1, assim quanto mais próximo de 1 pior é o desempenho da região da procuradoria em questão da transparência. Em termos matemáticos temos que a nota da avaliação da transparência referente à procuradoria é:

$$Aval. Procuradoria_i = 1 - \sum_{j=1}^n Aval. Município_j \cdot \left(\frac{Pop. Município_j}{Pop. Procuradoria_i} \right) \quad (1)$$

em que *Aval. Município_j* é a nota oriunda da avaliação do MPF atribuída ao município *j* pelo MPF, *Pop. Município_j* é a população do município *j* e *Pop. Procuradoria_i* é a população total dos municípios que pertencem a jurisdição da procuradoria *i*.

Após esta etapa, realizou-se a normalização das variáveis construídas, para posterior estimação do índice composto de corrupção. Essa etapa se fez necessária, pois, utilizamos as bases de dados com diferentes escalas de medida. O método de normalização adotado foi o do tipo *Re-Scaling* anual, também conhecido como normalização linear no intervalo [0,1] conforme previsto em OECD (2005, p.61), com a seguinte equação:

$$f(x_i) = \frac{x_i - \min(x_i)}{\max(x_i) - \min(x_i)}$$

(2)

em que $f(x_i)$ é o valor da variável linearmente normalizada no intervalo [0,1], (x_i) é o valor mínimo da variável x_i e (x_i) é o seu máximo, já o i é referente a umas das 220 Procuradorias Federais.

Por considerar que os componentes do indicador proposto, quais sejam: nº processos Cadirreg; nº de processos de corrupção; e notas atribuídas em avaliação sobre transparência são igualmente relevantes para mensurar a corrupção foi atribuído peso de 0,33 a cada um desses componentes. Também se considerou que adotar o mesmo peso evitaria distorções entre microrregiões que se destacam em apenas um dos componentes utilizados no item. Ademais, a estratégia de ponderar igualmente os componentes de indicadores de corrupção objetivos foi também adotada na literatura por Boll (2010). A equação se dá por:

$$\begin{aligned} \text{Índice Microrregional de Corrupção} = & \frac{1}{3} \left[\frac{\text{Processos de Corrupção}}{\text{população}} \right] \\ & + \frac{1}{3} \left[\frac{\text{Contas julgadas ilícitas}}{\text{população}} \right] + \frac{1}{3} [\text{Aval. Procuradoria}] \end{aligned} \quad (3)$$

Na subseção seguinte, será apresentada a metodologia de análise espacial a fim de compreender os resultados e as possíveis correlações espaciais da corrupção, a partir do índice que será calculado.

4.5.2 METODOLOGIA PARA ANÁLISE ESPACIAL DA CORRUPÇÃO

Na avaliação da corrupção para o Brasil é relevante que seja observado o comportamento espacial da variável de interesse. Como vemos em Donfouet et al. (2016) a corrupção apresenta uma forte correlação espacial positiva, e a fim de compreender o comportamento desta para o Brasil, realizar-se-á uma análise exploratória espacial do Índice Microrregional de Corrupção.

Para a análise de correlação espacial foi estimado o Índice de Moran, proposto por Moran (1948), que oferece uma medida de associação espacial que ocorre dentre um conjunto de dados. Essa estatística utiliza a autocovariância na forma de produto cruzado (ALMEIDA, 2012). O I de Moran está em um intervalo de [-1 a 1], sendo que valores mais próximos de 1 indicam maior correlação espacial e valores mais próximos de zero indicam menor intensidade da correlação. Em relação à magnitude, a correlação pode ser negativa ou

positiva, indicada pela inclinação da reta resultante do cálculo da estatística. Quando o índice é positivo, a correlação espacial é positiva e quando é negativa, a correlação é contrária.

Em termos matriciais, o teste global de Moran pode ser expresso da seguinte forma (ANSELIN, 1995):

$$I = \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad \text{para } j \neq i \quad (4)$$

Em que y_i e y_j são os valores do índice nas regiões i e j , respectivamente, \bar{y} = média de y e w_{ij} são os coeficientes da matriz de peso w_{ij} .

A matriz de pesos w_{ij} estimada, foi uma matriz de contiguidade, nesse caso, duas regiões foram consideradas vizinhas caso compartilhassem de uma fronteira física. Espera-se que em regiões onde existam fronteiras compartilhadas, ocorra maior interação espacial. Assim, a matriz w_{ij} assume valor 1 se i e j são contíguos e 0 se não são contíguos (ANSELIN, 1995).

A convenção de contiguidade adotada foi a matriz do tipo *queen*, na qual além das fronteiras com extensão diferente de zero, possam ser considerados os vértices como contíguos, considerando no caso desse estudo, as regiões geográficas de jurisdição do MPF. O nome da matriz é adotado em alusão ao movimento das peças do jogo de xadrez (ANSELIN, 1995).

Ressalta-se que a indicação de padrões globais de correlação espacial não necessariamente se traduz em padrões locais. Conseqüentemente, o Índice de Moran Global apresentado necessita ser complementado, uma vez que é desejável examinar os padrões espaciais em uma escala detalhada. Assim, é preciso utilizar indicadores de associação espacial que possam ser relacionados às diferentes localidades de uma variável distribuída espacialmente. Para tanto, utiliza-se o Índice de Moran Local, também chamado de Indicador Local de Associação Espacial.

Em termos gerais, o Indicador Local de Associação Espacial (LISA), ou I de Moran Local, é um parâmetro estatístico que fornece valores específicos de correlação espacial para cada Procuradoria, isto é, são calculadas n (220) computações da estatística de Moran. De acordo com Anselin (1995), o I de Moran Local apresenta a propriedade de que, o somatório dos indicadores locais, é proporcional ao I de Moran Global. Além disso, os indicadores locais permitem a identificação de situações atípicas, por meio de padrões de associação espacial e agrupamentos de valores semelhantes (“clusters”) significativos estatisticamente.

Este conjunto de informações pode ser apresentado em mapas, os quais expressam os padrões de associação espacial presentes nos dados que podem ser do tipo: (i) Alto-alto - *High-high*, significando que regiões com altos valores de determinada variável são circundados por regiões com valores semelhantes, no caso deste trabalho a incidência de corrupção; (ii) Baixo-baixo (*Low-low*), que representa regiões e regiões vizinhas com baixa incidência de corrupção; (iii) Alto-baixo (*High-low*), revela regiões de altos valores, com vizinhança de baixo valor; e (iv) Baixo-Alto (*Low-High*), indica regiões com baixo valor, que são vizinhas de regiões com valores altos (ANSELIN, 1995). Os padrões *i* e *ii* indicam associação espacial positiva, já os padrões *iii* e *iv* indicam associação espacial negativa.

Em termos matriciais, o teste local de Moran pode ser expresso da seguinte forma (ANSELIN, 1995):

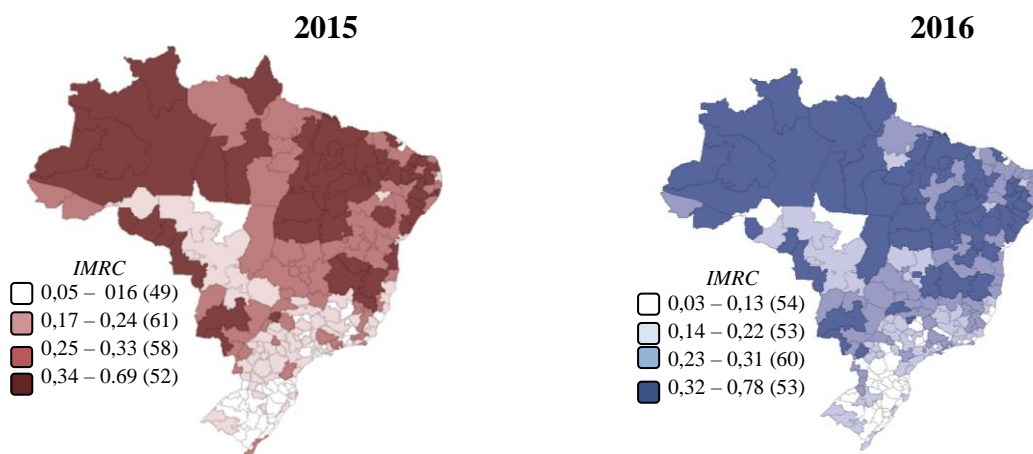
$$I_i = y_i \sum_{j=1}^n w_{ij} y_j \quad \text{para } j \neq i \quad (5)$$

em que y_i e y_j são os valores do índice nas regiões i e j , respectivamente e w_{ij} é o coeficiente da matriz de peso W descrita anteriormente. Os resultados do índice de Moran Local podem ser visualizados na próxima seção.

4.6 RESULTADOS E DISCUSSÕES

Nesta seção são apresentados os índices microrregionais de corrupção calculados para as 220 regiões que contemplam esta pesquisa. Os valores individuais para cada microregião podem ser consultados no apêndice A3, no qual consta o nome da Procuradoria Federal e os valores do Índice de Corrupção para os anos 2015 e 2016, bem como a população considerada do período. Para uma visualização mais global do índice, buscou-se construir um mapa *quantílico* dividido por *quartis*, na qual se pode observar a distribuição espacial do índice nos dois períodos estudados.

Figura 2 - Mapas *Quantílicos* - distribuição espacial da corrupção



Fonte: Elaborado pelos autores com o software GeoDa.

Observa-se pela distribuição espacial da corrupção, que as regiões com mais incidência de corrupção são as do Norte, Nordeste e a região de fronteira do Centro-Oeste, enquanto a que apresentara menor índice de corrupção é a região Sul. Como mencionado, do ano de 2015 para 2016 pode-se observar uma alteração nos *quantis* de concentração das 220 regiões. Isso se deve a uma melhora nos níveis de transparência de um ano para o outro e, por consequência, uma média menor no Índice Microrregional de Corrupção.

Os resultados da estimação do índice de corrupção para o Brasil, contendo todas as regiões das Procuradorias, nos anos de 2015 e 2016, estão descritos no apêndice A3. Para um entendimento geral, foram estimados os índices com valores em nível estadual, assim em vez de serem as procuradorias a referência, utilizamos os Estados. Este exercício foi realizado apenas com o fim de compreensão dos dados, visto que o diferencial do artigo é criar um índice ao nível das microrregiões definidas pelo Ministério Público Federal, para o Brasil. Os dados estaduais do índice são demonstrados na tabela 1 a seguir, ordenando os estados com os índices mais baixos em relação ao ano de 2015 até as unidades federativas com os índices mais altos, com as respectivas alterações para o ano de 2016.

Tabela 1 - Índice de Corrupção por estados brasileiros

Estado	IMRC 2015	Posição 2015	IMRC 2016	Posição 2016
RS	0,130730	1°	0,103033	↓ 2°
SC	0,139878	2°	0,098541	↑ 1°
RJ	0,141079	3°	0,143137	↓ 4°
SP	0,150542	4°	0,140172	↑ 3°
ES	0,176599	5°	0,163392	- 5°
PR	0,227666	6°	0,164048	- 6°

MT	0,231317	7°	0,212916	↓	8°
MG	0,237819	8°	0,237145	↓	11°
GO	0,265222	9°	0,234781	↓	10°
BA	0,267914	10°	0,262200	↓	12°
CE	0,284816	11°	0,222077	↑	9°
RO	0,285359	12°	0,191145	↑	7°
PE	0,286046	13°	0,262945	-	13°
PA	0,303767	14°	0,329212	↓	18°
DF	0,305377	15°	0,317108	↓	16°
AC	0,317623	16°	0,332078	↓	19°
MS	0,322226	17°	0,327815	-	17°
PB	0,324718	18°	0,298840	↑	15°
AL	0,337086	19°	0,334078	↓	20°
SE	0,355403	20°	0,338143	↓	21°
RN	0,360435	21°	0,275529	↑	14°
AM	0,429304	22°	0,362822	-	22°
AP	0,438819	23°	0,505854	↓	25°
PI	0,451463	24°	0,372798	↑	23°
MA	0,469185	25°	0,431807	↑	24°
RR	0,516215	26°	0,595049	↓	27°
TO	0,533527	27°	0,570458	↑	26°
Brasil	0,237809		0,218764		

Fonte: Elaborado pelos autores utilizando o software STATA.

Os dados apontam que houve uma sensível diminuição do Índice Microrregional de Corrupção - IMRC, que passou de IMRC = 0,237809 em 2015 para IMRC 0,218764 em 2016, uma variação na casa dos 8%. Em 2015, os três estados com maior índice de corrupção foram o Tocantins com IMRC igual a 0,533527, Roraima com IMRC = 0,5162 e Maranhão com IMRC = 0,4692; os menores índices ficaram nos estados do Rio Grande do Sul com IMRC = 0,1307, Santa Catarina com IMRC = 0,1399 e Rio de Janeiro com IMRC = 0,1411.

Para o ano de 2016, os estados com índices mais elevados foram Roraima com IMRC igual a 0,5950, ocupando a primeira posição, seguido por Tocantins com IMRC igual a 0,5706 e Amapá com IMRC igual a 0,5059. Já os estados com menores índices foram Santa Catarina com IMRC igual a 0,0985, Rio Grande do Sul com IMRC igual a 0,1030 e São Paulo com IMRC igual a 0,1402. Em relação ao ano anterior, os destaques positivos ficam para o estado do Rio Grande do Norte que subiu sete posições na tabela e para o estado de Rondônia que subiu cinco. Os destaques negativos ficam para o Pará que caiu quatro posições e Minas Gerais que caiu três posições no ranqueamento.

Nossos resultados contribuem com a literatura ao identificar o comportamento espacial da corrupção em 220 regiões distintas do país e lançam luz à discussão de que o Índice de Percepção da Corrupção, estimado pela Transparência Internacional pode deixar de

considerar questões relevantes entre os países. No trabalho seminal sobre a mensuração da corrupção com dados objetivos no Brasil de Boll (2010) foi identificada diferença na incidência da corrupção entre os estados brasileiros no período de 2001 a 2008. Nossos principais resultados indicam que, nos anos de 2015 e de 2016, a corrupção diverge para além das fronteiras de estados, variando entre as 220 regiões analisadas.

4.6.1 ÍNDICE DE MORAN GLOBAL

Com o fim de compreender a correlação espacial do índice de corrupção calculado, realizaram-se os testes de dependência espacial de Moran Global e Local. Segundo Almeida (2012), resumidamente, a partir do *I* de Moran pode-se realizar três inferências: i) a partir do nível de significância pode-se verificar se os dados estão aleatoriamente distribuídos ou não; ii) quando significativo, o sinal positivo do *I* de Moran indica que os dados estão concentrados em regiões, já o sinal negativo indica dispersão dos dados; e iii) a magnitude do índice fornece a intensidade da autocorrelação, sendo mais forte quanto mais próxima de 1 e menos forte quanto mais próxima de -1.

Na tabela 2 a seguir é possível observar os valores do *I* de Moran para o índice de corrupção nos anos de 2015 e 2016.

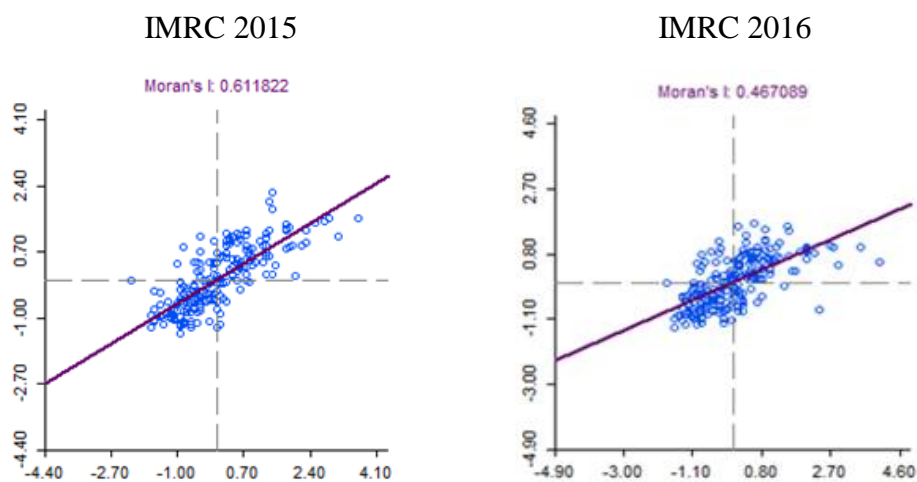
Tabela 2 – *I* de Moran - Autocorrelação Global

Variável	I de Moran	Convenção	P-valor
Índice Microrregional de Corrupção 2015	0,6118	Rainha (<i>Queen</i>)	0,00001
Índice Microrregional de Corrupção 2016	0,4670	Rainha (<i>Queen</i>)	0,00001

Fonte: Elaborado pelos autores utilizando o software STATA.

Assim, os resultados do cálculo do Índice de Moran Global são significativos e demonstram que existe correlação espacial para a corrupção, ou seja, a corrupção em determinada microrregião pode afetar a ocorrência desse fenômeno em outras microrregiões. Conforme pode ser observado no gráfico 3 a seguir.

Gráfico 3 - Índice de Moran Global a partir do Índice Microrregional de Corrupção



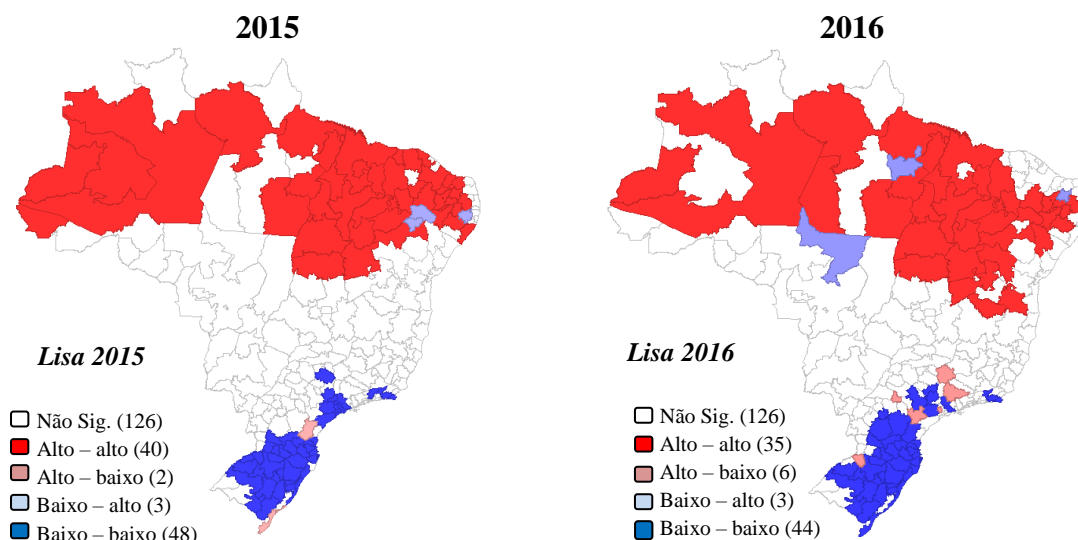
Fonte: Elaborado pelos autores com o software GeoDa.

Assim, o Índice de Moran Global para 2015 apresenta o valor de 0,6118, significativo. Já para o ano de 2016 o valor do índice é de 0,4670, também significativo. No Gráfico 3, percebe-se que o índice de correlação espacial assume valores um pouco maiores para 2016. É possível ainda observar a formação de clusters espaciais. No caso deste estudo, o cálculo foi realizado utilizando a variável índice de corrupção. Os resultados apontam que existe a formação de clusters espaciais de corrupção nos dois períodos avaliados (2015 e 2016).

4.6.2 INDICADOR LOCAL DE ASSOCIAÇÃO ESPACIAL (LISA)

A partir da decomposição do Índice de Moran Global, pode ser construído um mapa de dependência espacial local, o chamado *Local Indicators of Spatial Association* que permite verificar as relações espaciais mais detalhadamente. Os resultados desta estimação podem ser verificados na figura 3 a seguir.

Figura 3 – Índice de Associação Espacial Local (LISA)



Fonte: Elaborado pelos autores com o software GeoDa.

A partir do cálculo do *LISA*, é possível observar a formação de clusters espaciais do tipo *High-High* (*alto-alto*) composto por 35 e 40 microrregiões ao norte do Brasil para os anos de 2015 e 2016, respectivamente. Observa-se que há um aparente transbordamento da prática de corrupção entre essas localidades. Já em relação às regiões ao sul observou-se que a formação de um cluster do tipo *Low-Low* (*baixo-baixo*), o que indica haver um transbordamento de boas práticas e menor incidência de corrupção.

Nas demais regiões (126) a correlação espacial não foi significativa. Os resultados indicam a necessidade de ações efetivas de combate à corrupção e ao fomento a melhoria da gestão dos recursos públicos nas regiões do norte do país. Pode-se inferir ainda a necessidade de que sejam realizadas ações coordenadas entre as regiões, uma vez que há um aparente transbordamento das práticas corruptas.

Alguns possíveis canais para o efeito da corrupção espalhar-se entre regiões vizinhas podem ser: (i) facilidade de transações econômicas com custos de transação reduzidas, pela proximidade; (ii) questões socioculturais, tais como história, cultura, linguagem, contexto étnico entre vizinhos; (iii) aproximação dos governos para transações políticas por proximidade geográfica, assim como pela similaridade de desafios e questões similares que enfrentam, conforme apontado por Kelejian et al. (2013).

4.7 CONSIDERAÇÕES FINAIS

O trabalho teve por objetivo propor um Índice Microrregional de Corrupção que compreende todo o território brasileiro, esta abrangência ao nível de 220 microrregiões é

inédita na literatura brasileira. Uma das vantagens em estudar a microrregião se deve a possibilidade de verificar as disparidades locais já que pode haver sensíveis diferenças dentro de uma mesma região. Diferente de quando se analisa ao nível dos estados.

Além da variável do Cadirreg, utilizada na literatura, acrescenta-se à estimação do IMRC as variáveis referentes a 27.975 processos judiciais de corrupção, em todas as Procuradorias Regionais do Ministério Público Federal, com jurisdição nos 5.570 municípios brasileiros, além da análise dos dados da avaliação da transparência em todos esses municípios. Dados importantes para a identificação microrregional da corrupção.

Ademais, a partir do índice calculado, avalia-se a existência de dependência espacial, na qual os resultados apontam a formação de clusters espaciais da corrupção do tipo (*high-high*) alta incidência para os estados do Norte e (*low-low*) baixa incidência para os estados do Sul. Observou-se que os estados com maior incidência de corrupção em 2015 foi o Tocantins, e em 2016, Roraima. Em relação a menor incidência de corrupção, os resultados apontam o Rio Grande do Sul com menor IMRC em 2015 e Santa Catarina em 2016.

O índice geral de corrupção para o Brasil que foi calculado neste trabalho apresentou redução entre os anos de 2015 e 2016, passando de $IMRC=0,23$ em 2015, para $IMRC=0,21$ em 2016. Esse resultado apresenta consonância com os dados do Índice de Percepção da Corrupção, calculado pela Transparência Internacional - TI, no qual o Brasil melhorou em torno de 5%, variação também percebida pelo índice proposto neste artigo que ficou na casa dos 8%. Os resultados encontrados neste trabalho apresentam consonância com o Índice de Corrupção Geral criado por Boll (2010). Pesquisas adicionais devem ser realizadas para identificar as causas da redução do IMRC entre os anos de 2015 e 2016.

O IMRC apresenta um indicador de corrupção por regiões de atuação das procuradorias federais, podendo assim ser estratégico no combate a corrupção em nível microrregional, cooperando para a atuação do poder público em políticas de combate à corrupção. Ademais, foi elaborado um arquivo de mapa georeferenciado (*Shape file*) para as áreas de atuação das procuradorias federais, ou seja, um arquivo com a área de jurisdição das 220 Procuradorias do Ministério Público Federal, que abrange todos os municípios.

Duas principais limitações do trabalho podem ser destacadas, a primeira refere-se à dificuldade de estudar práticas criminosas, pois como a corrupção ocorre às escondidas, fatores objetivos, tais como o Cadirreg e o número e processos judiciais poderiam representar a eficácia do Estado em combater esses crimes. Em segundo lugar, devido aos dados disponíveis de transparência, a análise somente foi possível de ser realizada para os anos de

2015 e 2016, um painel com um período maior poderia fornecer uma percepção mais ampla da distribuição da corrupção no Brasil.

Por fim, este artigo visa cooperar com a focalização dos recursos de controle e fiscalização no combate à corrupção, oferecendo uma visão sobre o comportamento da corrupção espacial no Brasil, além do indicador microrregional para 220 regiões. A partir do índice é possível a especialização no processo de combate a esse fenômeno complexo. Cita-se ainda, a disponibilização à literatura de um índice atualizado para realização de pesquisas diversas.

REFERÊNCIAS

- ABRAMO, C.W. Percepções pantanosas: a dificuldade de medir a corrupção. *Novos Estudos-CEBRAP, SciELO Brasil*, n. 73, p. 33–37, 2005.
- ALMEIDA, Eduardo. *Econometria Espacial Aplicada*. Editora alínea. 2012.
- ANSELIN, Luc. Local indicators of spatial association—LISA. *Geographical analysis*, v. 27, n. 2, p. 93-115, 1995.
- ATTILA, Joseph. Is corruption contagious? An econometric analysis. *An Econometric Analysis*. Norwegian Institute of International Affairs (NUPI) Working Paper, n. 742, 2008.
- AVRITZER, Leonardo et al (Org.). *Corrupção: Ensaio e Críticas*. Belo Horizonte: Editora UFMG, 2012.
- AZEVEDO, A. e ALVES, M. . A relação entre emendas à lei do orçamento federal e corrupção municipal no Brasil: evidências de relatórios do órgão brasileiro de auditoria interna . *ANPAD* , 3: 414 - 33. 2010.
- BAKLOUTI, Nedra; BOUJELBENE, Younes. Shadow economy, corruption, and economic growth: An empirical analysis. *The Review of Black Political Economy*, v. 47, n. 3, p. 276-294, 2020.
- BARDHAN, Pranab. Corruption and Development: A Review of Issues. *Journal of Economic Literature*. 35. 1320-1346 1997.
- BOLL, José L. S. A corrupção governamental no Brasil: construção de indicadores e análise da sua incidência relativa nos estados brasileiros. *Concurso de Monografias da CGU*. 2010.
- BORSKY, S; KALKSCHMIED, K. Corruption in space: A closer look at the world's subnations. *European Journal of Political Economy*, vol. 59, p.400-422, 2019.
- BECKER, Gary S. Crime and punishment: An economic approach. *Journal of political economy*, v. 76, n. 2, p. 169-217, 1968.
- BECKER, Sascha O.; EGGER, Peter H.; SEIDEL, Tobias. Common political culture: Evidence on regional corruption contagion. *European Journal of Political Economy*, v. 25, n. 3, p. 300-310, 2009.
- BROLLO, Fernanda et al. The political resource curse. *American Economic Review*, v. 103, n. 5, p. 1759-1796, 2013.
- CAMPOS, Francisco de Assis Oliveira; CASTELAR, Ivan; SOARES, Ricardo Brito. Fatores associados à corrupção municipal em transferências de recursos da União. *Nova Economia*, v. 28, p. 879-911, 2018.
- CARDENAS, J., Y MIJANGOS, M. “Acerca del marco teórico de la corrupción”, en N. González Martín (coord.), *Estudios Jurídicos en homenaje a Marta Morineau*. Tomo ii: sistemas jurídicos contemporáneos. *Derecho comparado. Temas diversos* (pp. 167-226). México: UNAM. 2005.

CARDENAS, Roberto et al. MRAS observer for sensorless control of standalone doubly fed induction generators. *IEEE Transactions on Energy conversion*, v. 20, n. 4, p. 710-718, 2005.

CARRARO, A. et al. Proposta para estimação da corrupção regional no Brasil. *Política e Sociedade*, Florianópolis, v. 14, n. 31, p. 326-352, 2015.

COSTA, A. R. F. et al. Corrupção e Transparência: Análise de sua incidência a partir da relação entre as irregularidades das contas públicas e o indicador de transparência dos estados brasileiros. *Anais do 10º Congresso UFSC de Controladoria e Finanças*. Florianópolis, p. 1-18, 2020.

CRUZ RODRIGUEZ, Ignacio Javier. Comportamiento espacial de la corrupción en México a nivel entidad federativa 2001-2010. *Espiral (Guadalaj.)*, Guadalajara, v. 24, n. 70, p. 53-84, dic. 2017.

DINCER, Oguzhan. Does corruption slow down innovation? Evidence from a cointegrated panel of US states. *European Journal of Political Economy*, v. 56, p. 1-10, 2019.

DONFOUET, H. P. P., Jeanty, P. W., & Malin, E. Analysing spatial spillovers in corruption: A dynamic spatial panel data approach. *Papers in Regional Science*, <https://doi.org/10.1111/pirs.12231> Dong, G., Harris, R., Jones, K., & Yu, J. 2015.

FORTINI, Cristiana; MOTTA, Fabrício. Corrupção nas licitações e contratações públicas: sinais de alerta segundo a Transparência Internacional. *A&C-Revista de Direito Administrativo & Constitucional*, v. 16, n. 64, p. 93-113, 2016.

FERRAZ, Claudio; FINAN, Frederico. Exposing corrupt politicians: the effects of Brazil's publicly released audits on electoral outcomes. *The Quarterly journal of economics*, v. 123, n. 2, p. 703-745, 2008.

GARCIA, Gilson Piqueras. Proposta de um indicador de corrupção no Brasil baseado em dados do Tribunal de Contas da União. *Revista da CGU*, v. 14, n. 25, p. 129-143, 2022.

GRÜNDLER, Klaus; POTRAFKE, Niklas. Corruption and economic growth: New empirical evidence. *European Journal of Political Economy*, v. 60, p. 101810, 2019.

HUNTINGTON, S. *Political order in changing societies*. New Haven, 1968.

KAUFMANN, Daniel; KRAAY, Aart; MASTRUZZI, Massimo. *Governance matters III: Governance indicators for 1996-2002*. World bank policy research working paper, n. 3106, 2003.

KELEJIAN, Harry H.; MURRELL, Peter; SHEPOTYLO, Oleksandr. Spatial spillovers in the development of institutions. *Journal of Development Economics*, v. 101, p. 297-315, 2013.

LAMBSDORFF, Johann Graf. How corruption affects productivity. *Kyklos*, v. 56, n. 4, p. 457-474, 2003.

LEFF, N. H. *Economic development through bureaucratic corruption*. American Behavioral Scientist, Sage Publications, v. 8, n. 3, p. 8-14, 1964.

- LOPES, Yuri Chagas. Descentralização da análise de corrupção no Brasil: efeitos de spillover e ciclos políticos corruptos. 2021.
- LV, Zhike; GAO, Zhenya. The effect of corruption on environmental performance: does spatial dependence play a role?. *Economic Systems*, v. 45, n. 2, p. 100773, 2021.
- MATOS, Paulo. Uma nota sobre o impacto da corrupção no endividamento dos Estados brasileiros. *Revista Brasileira de Economia*, v. 72, p. 186-195, 2018.
- MAURO, P. Corruption and the composition of government expenditure. *Journal of Public Economics*, Elsevier, v. 69, n. 2, p. 263–279, 1998.
- MELO, Felipe Luiz Neves Bezerra de; SAMPAIO, Luciano Menezes Bezerra; OLIVEIRA, Renato Lima de. Corrupção burocrática e empreendedorismo: uma análise empírica dos estados brasileiros. *Revista de Administração Contemporânea*, v. 19, p. 374-397, 2015.
- MINISTÉRIO PÚBLICO FEDERAL. Entenda o caso da Operação Lava Jato. <http://www.mpf.mp.br/grandes-casos/lava-jato/linha-do-tempo>. 2020.
- MORAN, P.A.P., The interpretation of statistical maps. *Journal of the Royal Statistical Society. Series B (Methodological)*,10(2): 243-251, 1948.
- RAUSCH, Rita Buzzi; SOARES, Maurélio. Controle social na administração pública: a importância da transparência das contas públicas para inibir a corrupção. *Revista de Educação e Pesquisa em Contabilidade (REPeC)*, v. 4, n. 3, p. 23-43, 2010.
- ROSE-ACKERMAN, S. The Economics of Corruption. *Journal of Public Economics*, v. 4, n. 2, p. 187-203, 1975.
- SOBRAL, Eryka Fernanda Miranda. Corrupção e os seus efeitos sobre a dinâmica do crescimento econômico regional: uma análise do caso brasileiro. Dissertação de Mestrado. Universidade Federal de Pernambuco. 2014.
- SUI, Bo; FENG, Gen-Fu; CHANG, Chun-Ping. The pioneer evidence of contagious corruption. *Quality & Quantity*, v. 52, p. 945-968, 2018.
- SOUZA, A. I.; ARAÚJO, M. P. Uma análise econômica da corrupção no Brasil a partir de indicadores objetivos e subjetivos 1998-2013. *Revista Espacios*, v. 38, n. 9, p. 4, 2017.
- SOTO, R. La corrupción desde una perspectiva económica. *Estudios Públicos*, (89), 23-62. 2003.
- SOUSA, M. C. A corrupção governamental nos municípios do Ceará nos anos de 2002 a 2011: construção de indicadores e análise. Dissertação (Mestrado em Economia) – Universidade Federal do Ceará, Faculdade de Economia, Administração, Atuária e Contabilidade, Fortaleza, 2018.
- ZAKHAROV, Nikita. Does corruption hinder investment? Evidence from Russian regions. *European Journal of Political Economy*, v. 56, p. 39-61, 2019.

Apêndice A3

Apêndice A3 Figura 1 - Componentes da Avaliação Transparência Municipal

ESPELHO DA AVALIAÇÃO TRANSPARÊNCIA MUNICIPAL

Site do ente avaliado:	XXXXXXXXXXXXXXXXXXXX
Site do e-SIC:	XXXXXXXXXXXXXXXXXXXX
Inquérito Civil Nº:	X.XX.XXX.XXXXXX/XXXX-XX
Data da Avaliação:	XX/XX/XXXX

O ente possui informações sobre Transparência na internet?

2. O site contém ferramenta de pesquisa de conteúdo que permita o acesso à informação?

3. Há informações sobre a receita nos últimos 6 meses, incluindo natureza, valor de previsão e valor arrecadado?

4. As despesas apresentam dados dos últimos 6 meses contendo:

a. Valor do empenho

b. Valor da liquidação

c. Favorecido

d. Valor do pagamento

5. O site apresenta dados nos últimos 6 meses contendo:

a. Íntegra dos editais de licitação

b. Resultado dos editais de licitação (vencedor é suficiente)

c. Contratos na íntegra

6. O ente divulga as seguintes informações concernentes a procedimentos licitatórios com dados dos últimos 6 meses?

a. Modalidade

b. Data

c. Valor

Ente Avaliado: XXXXXXXXXXXXXXXX

d. Número/Ano de edital

e. Objeto

7. O site apresenta:

a. A prestação de contas (Relatório de Gestão) do ano anterior

b. Relatório Resumido da Execução Orçamentária (RRO) dos últimos 6 meses

c. Relatório de Gestão Fiscal (RGF) dos últimos 6 meses

d. Relatório estatístico contendo a quantidade de pedidos de informação recebidos, atendidos e indeferidos, bem como informações genéricas sobre os solicitantes

8. O site possibilita a gravação de relatórios em diversos formatos eletrônicos, abertos não proprietários, tais como planilhas e texto (CSV), de modo a facilitar a análise das informações?

9. Possibilidade de entrega de um pedido de acesso de forma presencial:

a. Existe indicação precisa no site de funcionamento de um SIC físico

b. Há indicação do órgão

c. Há indicação de endereço

d. Há indicação de telefone

e. Há indicação dos horários de funcionamento

10. Há possibilidade de envio de pedidos de informação de forma eletrônica (E-SIC)?

11. Apresenta possibilidade de acompanhamento posterior da solicitação?

12. A solicitação por meio do e-SIC é feita de forma fácil e simples sem a exigência de pelo menos um dos seguintes itens de identificação do requerente que dificultem ou impossibilitem o acesso à informação, tais como: envio de documentos, assinatura reconhecida, declaração de responsabilidade, maioridade?

13. No site está disponibilizado o registro das competências e estrutura organizacional do ente?

14. O Portal disponibiliza endereços e telefones das respectivas unidades e horários de atendimento ao público?

15. Há divulgação de remuneração individualizada por nome do agente público?

16. Há divulgação de diárias e passagens por nome de favorecido e constando, data, destino, cargo e motivo de viagem?

Considerações do Avaliador.

Apêndice A3 Figura 2 - Fatores utilizados pelo MPF para a estimação do índice de transparência municipal.

FATORES ASSOCIADOS AS VARIÁVEIS DO QUESTIONÁRIO DE TRANSPARÊNCIA		FATOR 1	FATOR 2	FATOR 3	FATOR 4	FATOR 5
VARIÁVEL/NOME DO FATOR		Licitações	Rendas/Despesas	Acessibilidade	e-sic	Gestão
1. O ente possui informações sobre Transparência na internet?			0.3864			
2. O site contém ferramenta de pesquisa de conteúdo que permita o acesso à informação?			0.7512			
3. Há informações sobre a receita nos últimos 6 meses, incluindo natureza, valor de previsão e valor arrecadado?			0.9315			
4. As despesas apresentam dados dos últimos 6 meses contendo:			0.9367			
	a. Valor do empenho		0.9068			
	b. Valor da liquidação		0.9385			
	c. Favorecido					
	d. Valor do pagamento					
5. O site apresenta dados nos últimos 6 meses contendo:		0.5824				
	a. Íntegra dos editais de licitação	0.5760				
	b. Resultado dos editais de licitação (vencedor é suficiente)	0.3368				
	c. Contratos na íntegra	0.9334				
	a. Modalidade	0.9312				
	b. Data	0.9136				
	c. Valor	0.9136				
	d. Número/Ano de edital	0.9136				
	e. Objeto	0.9342				
6. O ente divulga as seguintes informações concernentes a procedimentos licitatórios com dados dos últimos 6 meses?						
	a. A prestação de contas (Relatório de Gestão) do ano anterior					0.5323
	b. Relatório Resumido da Execução Orçamentária (RRO) dos últimos 6 meses					0.7987
	c. Relatório de Gestão Fiscal (RGF) dos últimos 6 meses					0.8028
	d. Relatório estatístico contendo a quantidade de pedidos de informação					
7. O site apresenta:						
	a. Existência de contas (Relatório de Gestão) do ano anterior					
	b. Relatório Resumido da Execução Orçamentária (RRO) dos últimos 6 meses					
	c. Relatório de Gestão Fiscal (RGF) dos últimos 6 meses					
	d. Relatório estatístico contendo a quantidade de pedidos de informação					
8. O site possibilita a gravação de relatórios em diversos formatos eletrônicos, abertos não proprietários, tais como planilhas e texto (CSV), de modo a facilitar a análise das informações?			0.3519			
	a. Existe indicação precisa no site de funcionamento de um SIC físico			0.8083		
	b. Há indicação do órgão			0.8389		
	c. Há indicação de endereço			0.9176		
	d. Há indicação de telefone			0.8799		
	e. Há indicação dos horários de funcionamento			0.7446		
9. Possibilidade de entrega de um pedido de acesso de forma presencial:						
	a. Existe indicação precisa no site de funcionamento de um SIC físico					
	b. Há indicação do órgão					
	c. Há indicação de endereço					
	d. Há indicação de telefone					
	e. Há indicação dos horários de funcionamento					
10. Há possibilidade de envio de pedidos de informação de forma eletrônica (E-SIC)?						0.8722
11. Apresenta possibilidade de acompanhamento posterior da solicitação?						0.5245
12. A solicitação por meio do e-SIC é feita de forma fácil e simples sem a exigência de pelo menos um dos seguintes itens de identificação do requerente que dificultem ou impossibilitem o acesso à informação, tais como: envio de						0.8736
13. No site está disponibilizado o registro das competências e estrutura organizacional do ente?						
14. O Portal disponibiliza endereços e telefones das respectivas unidades e horários de atendimento ao público?						
15. Há divulgação de remuneração individualizada por nome do agente público?						
16. Há divulgação de diárias e passagens por nome de favorecido e constando, data, destino, cargo e motivo de viagem?						

Factor	Variance	Proportion	Cumulative
Factor1	591.237	0.2832	0.2832
Factor2	508.641	0.2436	0.5268
Factor3	427.652	0.2048	0.7316
Factor4	230.277	0.1103	0.8419
Factor5	194.307	0.0931	0.9350
Factor6	0.60838	0.0291	0.9641

Apêndice A3 Tabela 1 - Índice Microrregional de Corrupção

Estado	Procuradoria M.P.	Pop. PRM	Índice Corrupção 2015	Índice Corrupção 2016
ACRE	PR-AC	571116	0,326685	0,342379
ACRE	PRM-C. DO SUL	218985	0,293989	0,305215
ALAGOAS	PR-AL	2224244	0,301787	0,294397
ALAGOAS	PRM-ARAPIRACA	1097486	0,408626	0,414500
AMAPA	PR-AP	750912	0,438819	0,505854
AMAZONAS	PR-AM	3374993	0,412957	0,331318
AMAZONAS	PRM-TABATINGA	221580	0,472969	0,322864
AMAZONAS	PRM-TEFÉ	277170	0,593451	0,778377
BAHIA	PR-BA	4342686	0,185683	0,108719
BAHIA	PRM-ALAGOINHAS	848765	0,348338	0,352563
BAHIA	PRM-B. J. LAPA	549098	0,286910	0,307510
BAHIA	PRM-BARREIRAS	469107	0,348310	0,361133
BAHIA	PRM-C. FORMOSO	871812	0,345044	0,366305
BAHIA	PRM-EUNAPOLIS	388862	0,325485	0,266217
BAHIA	PRM-FEIRA	1753857	0,250981	0,264038
BAHIA	PRM-GUANAMBI	612603	0,321385	0,385797
BAHIA	PRM-ILHEUS	1165921	0,345367	0,309574
BAHIA	PRM-IRECÊ	714705	0,247207	0,404382
BAHIA	PRM-JEQUIE	815321	0,305279	0,347723
BAHIA	PRM-JUAZEIRO	520045	0,303592	0,367036
BAHIA	PRM-PAULO AFONS	610122	0,315188	0,318645
BAHIA	PRM-T. FREITAS	444445	0,203084	0,302971
BAHIA	PRM-VIT. CONQUI	1019022	0,290929	0,275555
CEARA	PR-CE	1281024	0,404296	0,319736
CEARA	PRM-CRATEUS	285515	0,482851	0,269917
CEARA	PRM-ITAPIPOCA	3650119	0,162000	0,120124
CEARA	PRM-J. NORTE	1423397	0,412606	0,283850
CEARA	PRM-LIMOEIRO	993365	0,320483	0,263143
CEARA	PRM-SOBRAL	1209371	0,302486	0,308616
DISTRITO FEDERAL	PR-DF	2570160	0,305377	0,317108
ESPÍRITO SANTO	PR-ES	2157174	0,145491	0,127168
ESPÍRITO SANTO	PRM-CACHOEIRO	671549	0,169747	0,251888
ESPÍRITO SANTO	PRM-COLATINA-ES	385486	0,267467	0,221366
ESPÍRITO SANTO	PRM-LINHARES	329725	0,219531	0,102191
ESPÍRITO SANTO	PRM-S.MATEUS	341115	0,242624	0,211892
GOIAS	PR-GO	3472865	0,262642	0,254543
GOIAS	PRM-ANAPOLIS	1167741	0,271589	0,184450
GOIAS	PRM-ITUMBIARA	248811	0,245904	0,231606
GOIAS	PRM-LUZIANIA	989189	0,305654	0,218350
GOIAS	PRM-RIO VERDE	644616	0,213004	0,245925
MARANHAO	PR-MA	4381355	0,466440	0,452394
MARANHAO	PRM-BACABAL	517330	0,601398	0,514559
MARANHAO	PRM-BALSAS	490400	0,421557	0,260149
MARANHAO	PRM-CAXIAS	773099	0,427059	0,294551
MARANHAO	PRM-IMPERATRIZ	688700	0,468540	0,514985
MATO GROSSO	PR-MT	1106651	0,198179	0,210932
MATO GROSSO	PRM-B.DO GARÇAS	316932	0,295311	0,339646
MATO GROSSO	PRM-CACERES	296186	0,498919	0,381482
MATO GROSSO	PRM-JUÍNA	562624	0,205909	0,162262
MATO GROSSO	PRM-RONDONOPOLI	363918	0,238667	0,181803
MATO GROSSO	PRM-SINOP	578046	0,142654	0,129751
MATO GROSSO DO SUL	PR-MS	1144789	0,344473	0,353719
MATO GROSSO DO SUL	PRM-CORUMBA	129498	0,329928	0,295974
MATO GROSSO DO SUL	PRM-COXIM	130921	0,204217	0,291848
MATO GROSSO DO SUL	PRM-DOURADOS	520461	0,323367	0,268620
MATO GROSSO DO SUL	PRM-NAVIRAÍ	156553	0,250039	0,250430
MATO GROSSO DO SUL	PRM-PONTA PORA	242484	0,337094	0,461366
MATO GROSSO DO SUL	PRM-TRES LAGOAS	289904	0,308722	0,292342
MINAS GERAIS	PR-MG	5925153	0,163777	0,139939
MINAS GERAIS	PRM-DIVINÓPOLIS	1126833	0,231396	0,199895
MINAS GERAIS	PRM-GOV VALADAR	737934	0,348704	0,308658
MINAS GERAIS	PRM-IPATINGA	910630	0,270095	0,237821
MINAS GERAIS	PRM-ITUIUTABA	279828	0,320001	0,346436
MINAS GERAIS	PRM-JANAÚBA	484564	0,373986	0,307181
MINAS GERAIS	PRM-JUIZ FORA	866378	0,161229	0,174554
MINAS GERAIS	PRM-M. CLAROS	1330845	0,383650	0,378630
MINAS GERAIS	PRM-MANHUAÇU	960100	0,258329	0,312468

MINAS GERAIS	PRM-P. ALEGRE	1287329	0,248634	0,283968
MINAS GERAIS	PRM-P. MINAS/MG	446825	0,280226	0,270823
MINAS GERAIS	PRM-PARACATU	382869	0,333407	0,382055
MINAS GERAIS	PRM-PASSOS	604662	0,244494	0,265776
MINAS GERAIS	PRM-S.J.DEL REI	829822	0,196156	0,202253
MINAS GERAIS	PRM-SETE LAGOAS	600275	0,214194	0,233727
MINAS GERAIS	PRM-TEOF. OTONI	956774	0,366590	0,474116
MINAS GERAIS	PRM-UBERABA	693919	0,193921	0,225982
MINAS GERAIS	PRM-UBERLANDIA	916251	0,226509	0,293902
MINAS GERAIS	PRM-VARGINHA	642095	0,232002	0,112931
MINAS GERAIS	PRM-VIÇOSA	751011	0,231087	0,217566
PARA	PR-PA	4493079	0,285734	0,291822
PARA	PRM-ALTAMIRA-PA	239862	0,411499	0,628836
PARA	PRM-ITAITUBA	198897	0,355257	0,396208
PARA	PRM-MARABA	753095	0,328522	0,328366
PARA	PRM-PARAGOMINAS	579461	0,351705	0,347330
PARA	PRM-REDENÇÃO	521027	0,290827	0,410146
PARA	PRM-SANTAREM	839480	0,323476	0,442013
PARA	PRM-TUCURUI	464638	0,279480	0,191641
PARAIBA	PR-PB	1565615	0,236443	0,206539
PARAIBA	PRM-C.GRANDE	993718	0,350910	0,252401
PARAIBA	PRM-GUARABIRA	305341	0,332589	0,115366
PARAIBA	PRM-MONTEIRO	214338	0,408595	0,612520
PARAIBA	PRM-PATOS/PB	251826	0,488527	0,506869
PARAIBA	PRM-SOUSA	613047	0,407163	0,506095
PARANA	PR-PR	3471736	0,267827	0,173070
PARANA	PRM-APUCARANA	292365	0,225822	0,191840
PARANA	PRM-C. MOURAO	327794	0,165850	0,156768
PARANA	PRM-CASCADEL	792409	0,193142	0,092657
PARANA	PRM-F.BELTRAO	354034	0,203230	0,236882
PARANA	PRM-FOZ	445677	0,289233	0,283549
PARANA	PRM-GUARAPUAVA	545222	0,236110	0,119206
PARANA	PRM-GUAÍRA	55212	0,274790	0,229244
PARANA	PRM-JACAREZINHO	336374	0,231897	0,147348
PARANA	PRM-LONDRINA	1172257	0,166699	0,127484
PARANA	PRM-MARINGA	946483	0,200466	0,154863
PARANA	PRM-P.GROSSA	938555	0,190330	0,123030
PARANA	PRM-PARANAGUA	283981	0,190552	0,192695
PARANA	PRM-PARANAVAI	236155	0,211045	0,179916
PARANA	PRM-PATO BCO	276294	0,209695	0,100322
PARANA	PRM-UMUARAMA	375312	0,327646	0,354360
PARANA	PRM-UNIÃO VIT	231832	0,208134	0,126461
PERNAMBUCO	PR-PE	4200271	0,256440	0,228635
PERNAMBUCO	PRM-C.S.AGOSTIN	471230	0,397606	0,236058
PERNAMBUCO	PRM-CARUARU	1366979	0,262730	0,271507
PERNAMBUCO	PRM-GARANHUNS	1015241	0,366833	0,375527
PERNAMBUCO	PRM-GOIANA	343901	0,289024	0,258115
PERNAMBUCO	PRM-PALMARES	379979	0,424049	0,305795
PERNAMBUCO	PRM-PETROLINA	428552	0,210092	0,186246
PERNAMBUCO	PRM-S. TALHADA	518712	0,327785	0,359628
PERNAMBUCO	PRM-SALGUEIRO	549978	0,247997	0,260901
PIAUI	PR-PI	1752957	0,423370	0,370907
PIAUI	PRM-CORRENTE	176082	0,692023	0,359242
PIAUI	PRM-FLORIANO	280812	0,427386	0,535979
PIAUI	PRM-PARNAIBA	326424	0,305439	0,201438
PIAUI	PRM-PICOS-PI	374729	0,481120	0,357851
PIAUI	PRM-S.R.NONATO	275384	0,633757	0,450564
RIO DE JANEIRO	PR-RJ	6693154	0,122045	0,153713
RIO DE JANEIRO	PRM-ANGRA REIS	224905	0,256405	0,279949
RIO DE JANEIRO	PRM-CAMPOS	666486	0,175960	0,224567
RIO DE JANEIRO	PRM-ITAPERUNA	285435	0,238205	0,242249
RIO DE JANEIRO	PRM-MACAE	432928	0,086220	0,101830
RIO DE JANEIRO	PRM-N.FRIBURGO	502108	0,247347	0,170940
RIO DE JANEIRO	PRM-NITEROI	638581	0,067406	0,061020
RIO DE JANEIRO	PRM-PETROPOLIS	504476	0,264030	0,216543
RIO DE JANEIRO	PRM-RESENDE-RJ	185697	0,151197	0,123930
RIO DE JANEIRO	PRM-S.GONÇALO	1715058	0,188466	0,133626
RIO DE JANEIRO	PRM-S.J. MERITI	3300899	0,092401	0,081479
RIO DE JANEIRO	PRM-S.P.ALDEIA	586326	0,260701	0,237320
RIO DE JANEIRO	PRM-V.REDONDA	725120	0,159402	0,147820
RIO GRANDE DO NORTE	PR-RN	1884190	0,348879	0,249331
RIO GRANDE DO NORTE	PRM-ASSU	251723	0,394483	0,314936

RIO GRANDE DO NORTE	PRM-CAICÓ-RN	294094	0,334166	0,164907
RIO GRANDE DO NORTE	PRM-CEARÁ-MIRIM	256998	0,300624	0,255620
RIO GRANDE DO NORTE	PRM-MOSSORO	464548	0,345664	0,309285
RIO GRANDE DO NORTE	PRM-PAU FERROS	256957	0,528412	0,514520
RIO GRANDE DO SUL	PR-RS	2990970	0,129139	0,068631
RIO GRANDE DO SUL	PRM-B.GONCALVES	325052	0,184343	0,091850
RIO GRANDE DO SUL	PRM-BAGE	173383	0,160330	0,172230
RIO GRANDE DO SUL	PRM-C. DA CANOA	291746	0,146941	0,211126
RIO GRANDE DO SUL	PRM-CANOAS	586868	0,062241	0,029103
RIO GRANDE DO SUL	PRM-CAXIAS SUL	865218	0,113731	0,083655
RIO GRANDE DO SUL	PRM-CRUZ ALTA	191500	0,151386	0,135029
RIO GRANDE DO SUL	PRM-ERECHIM	278347	0,180639	0,127908
RIO GRANDE DO SUL	PRM-LAJEADO	309430	0,148359	0,127909
RIO GRANDE DO SUL	PRM-N.HAMBURGO	1035672	0,103651	0,064518
RIO GRANDE DO SUL	PRM-P. MISSÕES	253605	0,131335	0,283449
RIO GRANDE DO SUL	PRM-P.FUNDO	713973	0,122122	0,188438
RIO GRANDE DO SUL	PRM-PELOTAS-RS	584466	0,117290	0,126137
RIO GRANDE DO SUL	PRM-R.GRANDE	271733	0,272392	0,206518
RIO GRANDE DO SUL	PRM-S.ANGELO	395829	0,121989	0,058892
RIO GRANDE DO SUL	PRM-S.LIVRAMENT	264095	0,169832	0,215190
RIO GRANDE DO SUL	PRM-S.MARIA	611612	0,152488	0,086749
RIO GRANDE DO SUL	PRM-SANTA ROSA	247278	0,159950	0,069262
RIO GRANDE DO SUL	PRM-STA CRUZ SU	554684	0,069417	0,079568
RIO GRANDE DO SUL	PRM-URUGUAIANA	259013	0,204383	0,153680
RONDONIA	PR-RO	871782	0,212840	0,132932
RONDONIA	PRM-GUAJARÁ-MIR	73128	0,425064	0,563254
RONDONIA	PRM-JI PARANÁ	579852	0,349860	0,139310
RONDONIA	PRM-VILHENA	223769	0,355091	0,430647
RORAIMA	PR-RR	496936	0,516215	0,595049
SANTA CATARINA	PR-SC	993517	0,197397	0,104685
SANTA CATARINA	PRM-BLUMENAU	611078	0,148487	0,067132
SANTA CATARINA	PRM-CAÇADOR	327751	0,069321	0,099207
SANTA CATARINA	PRM-CHAPECO	490084	0,180325	0,061699
SANTA CATARINA	PRM-CONCORDIA	176245	0,161967	0,126241
SANTA CATARINA	PRM-CRICIUMA	596930	0,239168	0,134631
SANTA CATARINA	PRM-ITAJAI	848188	0,088744	0,105678
SANTA CATARINA	PRM-JARAGUAS	370286	0,057630	0,031619
SANTA CATARINA	PRM-JOÃO ABA	187823	0,103829	0,044865
SANTA CATARINA	PRM-JOINVILLE	717964	0,092772	0,046250
SANTA CATARINA	PRM-LAGES	290298	0,114504	0,127729
SANTA CATARINA	PRM-MAFRA	190290	0,087336	0,105632
SANTA CATARINA	PRM-RIO DO SUL	288333	0,150287	0,099025
SANTA CATARINA	PRM-S.MIGUEL	252003	0,134571	0,267633
SANTA CATARINA	PRM-TUBARAO	364773	0,161404	0,148656
SAO PAULO	PR-SP	12681315	0,072440	0,069408
SAO PAULO	PRM-ANDRADINA	280319	0,245599	0,208245
SAO PAULO	PRM-ARACATUBA	590581	0,216086	0,192556
SAO PAULO	PRM-ARARAQUARA	607577	0,205115	0,194121
SAO PAULO	PRM-ASSIS	252436	0,272403	0,284875
SAO PAULO	PRM-BARRETOS	210934	0,140955	0,328530
SAO PAULO	PRM-BAURU	1072224	0,180241	0,139173
SAO PAULO	PRM-BRAGANÇA	732168	0,202769	0,238906
SAO PAULO	PRM-CAMPINAS	2408149	0,161225	0,125468
SAO PAULO	PRM-CARAGUATA	310240	0,223604	0,248112
SAO PAULO	PRM-FRANCA	527897	0,109955	0,099690
SAO PAULO	PRM-GUARATIN	474355	0,275458	0,252718
SAO PAULO	PRM-GUARULHOS	2961740	0,169192	0,168360
SAO PAULO	PRM-ITAPEVA	413461	0,210735	0,264613
SAO PAULO	PRM-JALES	302094	0,355714	0,220814
SAO PAULO	PRM-JAU	313347	0,191017	0,098482
SAO PAULO	PRM-JUNDIAI	805313	0,126063	0,167862
SAO PAULO	PRM-MARÍLIA	801060	0,213237	0,136356
SAO PAULO	PRM-OSASCO	2624506	0,184028	0,261172
SAO PAULO	PRM-OURINHOS	323874	0,216502	0,083172
SAO PAULO	PRM-P.PRUDENTE	622681	0,231775	0,274640
SAO PAULO	PRM-PIRACICABA	2018589	0,173770	0,122268
SAO PAULO	PRM-R.PRETO	1708745	0,207312	0,132044
SAO PAULO	PRM-REGISTRO	283588	0,205693	0,223145
SAO PAULO	PRM-S.BERNARDO	2702071	0,163639	0,136216
SAO PAULO	PRM-S.CARLOS	565700	0,150147	0,211932
SAO PAULO	PRM-S.J.B.VISTA	759205	0,189972	0,114967
SAO PAULO	PRM-S.J.CAMP	1042561	0,150876	0,097048

SAO PAULO	PRM-S.J.R.PRETO	1358965	0,250405	0,199449
SAO PAULO	PRM-SANTOS	1781620	0,159034	0,160709
SAO PAULO	PRM-SOROCABA	1894753	0,137191	0,147475
SAO PAULO	PRM-TAUBATE	603236	0,169142	0,170681
SERGIPE	PR-SE	2219574	0,355403	0,338143
TOCANTINS	PR-TO	642940	0,563557	0,713813
TOCANTINS	PRM-ARAGUAINA	561228	0,532632	0,447595
TOCANTINS	PRM-GURUPI	292712	0,469281	0,491152

5. CONSIDERAÇÕES FINAIS

Esta tese teve por objetivo a análise empírica de três problemas de sociais. No primeiro ensaio, constatou-se a relevância de estudar a evasão escolar sob a perspectiva de gênero, com evidências de que meninos e meninas evadem da escola por motivos distintos. Nossos resultados refletem a forma como a sociedade tem se organizado com a alocação de tempo entre escola, trabalho e tarefas de cuidado familiar. Identificamos uma possível sobrecarga das meninas, que estão expostas a um alto risco de evasão tanto pelo motivo de casamento e filhos, quanto por necessidade de trabalho. Em relação à etapa de ensino, as razões de riscos de evadir em virtude do trabalho aumentam tanto da primeira para a segunda etapa do ensino fundamental, quanto do ensino fundamental para o ensino médio.

Em termos de políticas públicas, ressaltamos a importância de políticas que visem prevenir a gravidez na adolescência e o casamento precoce. Ademais, identifica-se a importância da oferta de vagas em creches públicas que visem manter as mães na escola, assim como no mercado de trabalho, após o nascimento dos filhos.

No segundo ensaio, avaliamos a política de Educação de Jovens e Adultos, que visa reparar a evasão escolar e a baixa escolaridade da população. A partir das estimativas realizadas, foi possível observar menor probabilidade dos indivíduos que cursam EJA obterem certificação via Enceja, tanto para aqueles que buscam certificação em nível fundamental, quanto para o nível médio. Tal fato indica um pior desempenho do grupo participante da política de EJA e possivelmente a baixa qualidade dos cursos de Educação de Jovens e Adultos. Na certificação para o ensino médio identificou-se menor probabilidade de aprovação no Enceja para os indivíduos que cursam EJA, em todas as disciplinas analisadas, sendo o efeito negativo de maior magnitude para Linguagens e Redação, e Matemática. Na certificação em Nível Fundamental verificou-se desempenho inferior dos estudantes da EJA nas disciplinas de Linguagens e Redação, e Ciências da Natureza.

Sobre possíveis políticas públicas identifica-se a necessidade de adotar meios para qualificar o ensino da EJA e levar em conta sua interação com o Enceja, verificando questões de competição e rivalidade entre as duas políticas. Deve-se ainda serem consideradas questões já apontadas pela literatura, tais como metodologia de ensino da EJA, pedagogia, qualificação dos professores, tendo em vista as especificidades de jovens e adultos que retornam ao ambiente escolar.

Por fim, no terceiro ensaio propusemos um Índice Microrregional de Corrupção com dados objetivos de abrangência em 220 microrregiões do Brasil. Além da variável do Cadirreg,

largamente utilizada na literatura, acrescenta-se à estimação do IMRC as variáveis referentes a 27.975 processos judiciais de corrupção, em todas as Procuradorias Regionais do Ministério Público Federal, com jurisdição nos 5.570 municípios brasileiros, além da análise dos dados da avaliação da transparência em todos esses municípios. Dados importantes para a identificação microrregional da corrupção. Adicionalmente, a partir do índice calculado, avalia-se a existência de dependência espacial, na qual os resultados apontam a formação de clusters espaciais da corrupção do tipo (*high-high*) alta incidência para os estados do Norte e (*low-low*) baixa incidência para os estados do Sul.

O IMRC apresenta um indicador de corrupção podendo assim ser estratégico no combate a corrupção em nível microrregional, cooperando para a atuação do poder público com a focalização dos recursos de controle e fiscalização no combate à corrupção, oferecendo uma visão sobre o comportamento espacial da corrupção no Brasil, além do indicador para 220 regiões.

Pelas pesquisas realizadas nesta tese entende-se que o objetivo de analisar empiricamente três problemas sociais em áreas que compreendem questões de interesse público foi atingido. Com os resultados aqui aventados, espera-se contribuir com a literatura e com gestores públicos, tendo por finalidade qualificar programas de educação, além apresentar indícios sobre a localização e o comportamento espacial da corrupção no Brasil.



Pontifícia Universidade Católica do Rio Grande do Sul
Pró-Reitoria de Pesquisa e Pós-Graduação
Av. Ipiranga, 6681 – Prédio 1 – Térreo
Porto Alegre – RS – Brasil
Fone: (51) 3320-3513
E-mail: propesq@pucrs.br
Site: www.pucrs.br