

Documentos para Disseminação

Memória Institucional

28



POPULAÇÃO, TERRITÓRIO E ESTATÍSTICAS PÚBLICAS



25 anos da Pós-Graduação
stricto sensu da ENCE

Presidente da República
Luiz Inácio Lula da Silva

Ministra do Planejamento e Orçamento
Simone Nassar Tebet

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE

Presidente
Marcio Pochmann

Diretora-Executiva
Flávia Vinhaes Santos

ÓRGÃOS ESPECÍFICOS SINGULARES

Diretoria de Pesquisas
Gustavo Junger da Silva

Diretoria de Geociências
Maria do Carmo Dias Bueno

Diretoria de Tecnologia da Informação
Marcos Vinícius Ferreira Mazoni

Centro de Documentação e Disseminação de Informações
José Daniel Castro da Silva

Escola Nacional de Ciências Estatísticas
Jorge Abrahão de Castro

UNIDADE RESPONSÁVEL

Escola Nacional de Ciências Estatísticas

Ministério do Planejamento e Orçamento
Instituto Brasileiro de Geografia e Estatística - IBGE
Escola Nacional de Ciências Estatísticas

Documentos para Disseminação

Memória Institucional 28

População, Território e Estatísticas Públicas

25 anos da Pós-Graduação
stricto sensu da ENCE



Rio de Janeiro
2026

Instituto Brasileiro de Geografia e Estatística - IBGE

Av. Franklin Roosevelt, 166 – Centro – 20021-120 – Rio de Janeiro – Brasil

ISSN 0103-6335 **Documentos para Disseminação**

ISSN 0103-6459 **Memória institucional**

Divulga textos sobre aspectos históricos do IBGE e/ou de seus estudos e pesquisas bem como sobre personalidades que contribuíram para a história do Instituto.

ISBN 978-85-240-4700-8

© IBGE. 2026

As opiniões emitidas nesta publicação são de exclusiva e inteira responsabilidade do(s) autor(es), não exprimindo, necessariamente, o ponto de vista do IBGE.

Em virtude do prazo disponível para o cumprimento do cronograma editorial, os originais desta publicação não foram submetidos aos protocolos completos de normalização e editoração.

Capa

Marcos Balster Fiore - Centro de Documentação e
Disseminação de Informações - CDDI

Ilustração

Fabio Muniz de Moura -
Gerência de Editoração/Centro de Documentação e
Disseminação de Informações - CDDI

**DADOS INTERNACIONAIS DE CATALOGAÇÃO NA
PUBLICAÇÃO (CIP)**

Gerência de Biblioteca, Informação e Memória do IBGE

População, Território e Estatísticas Públicas: 25 anos da Pós-
Graduação stricto sensu da ENCE / Escola Nacional de
Ciências Estatísticas. - Rio de Janeiro: IBGE, 2026.

312 p. - (Documentos para disseminação. Memória institucional,
ISSN 0103-6459; n. 28).

ISBN 978-85-240-4700-8

1. População - Brasil. 2. Território - Brasil. 3. Estatística. I. Escola
Nacional de Ciências Estatísticas (Brasil). II. Série.

CDU: 314.1
SOC

Sumário

- 7 **Apresentação**
- 9 **Introdução**
Angelita Alves de Carvalho
- Seção 1
Reflexões e impactos dos 25 anos do Programa de Pós-Graduação em População, Território e Estatísticas Públicas (PTEP)
- 15 **As recentes conquistas (2018-2023) do Programa de Pós-Graduação em População, Território e Estatísticas Públicas (PTEP)**
Angelita Alves de Carvalho
- 19 **Potencialidades e Limitações na Formação em População, Território e Estatísticas Públicas: Reflexões a partir da Pesquisa de Egressos 2023**
Angelita Alves de Carvalho, Alinne Magalhães Gomes
- 33 **Inserção de mestres e doutores egressos da ENCE nos registros de emprego formal**
Marcos Dantas Hecksher
- Seção 2
Desafios e Limitações na Produção e Uso de Estatísticas Públicas
- 45 **Produção estatística: velhos e novos desafios**
Natália da Silva Castro
- 59 **Discussão sobre a natureza do conceito de população de difícil acesso**
João Gabriel Malaguti, Letícia de Carvalho Giannella, Alinne de Carvalho Veiga

- 73 A Plataforma Nilo Peçanha e a pesquisa no campo da educação profissional: limites e possibilidades**
Eilson Castro Soares de Oliveira, Barbara Cobo Soares
- Seção 3
Estimação de Indicadores e Monitoramento de Políticas Públicas e Objetivos de Desenvolvimento Sustentável
- 93 Access to clean drinking water: complementary indicators of Monitoring Goal 6.1 of the 2030 Agenda at the local level in Brazil**
Ana Clara Cancelo Cruz, Andrea Diniz da Silva
- 105 Alternativas com base em dados abertos para o cálculo do Indicador ODS 11.7.1: Proporção da área construída das cidades que é espaço público aberto para uso de todos**
Isis Gonçalves Peixoto
- 125 Expansão e interiorização da educação superior: Programas e políticas públicas de acesso da população brasileira**
Lidiane Braga Sales de Souza, Barbara Cobo Soares, Denise Britz do Nascimento Silva
- 145 Indicadores de (des)equilíbrio da oferta e demanda no ensino superior brasileiro**
Luiz André Ribeiro Zardo, Pedro Luis do Nascimento Silva, Joana Simões de Melo Costa
- Seção 4
Análises Setoriais e de Gênero sobre Mercado de Trabalho
- 169 Estimativas de desocupação trimestrais para estratos geográficos de Minas Gerais: estimação em pequenos domínios com séries temporais**
Caio Gonçalves, Luna Hidalgo, Denise Britz do Nascimento Silva, Jan van den Brakel
- 189 Uma perspectiva interseccional dos marcadores sociais da diferença de gênero e raça no mercado de trabalho brasileiro (2014-2018)**
Rodrigo Santos Martins, César Augusto Marques da Silva, Paula Alves de Almeida
- 211 A importância das características do trabalho feminino para a fecundidade: reflexões para o Brasil em 2015**
Angelita Alves de Carvalho, Ana Elisa Mello Britto
- 227 Características demográficas dos NEETs no Brasil relacionadas à (não) procura por ocupação**
Marcos Filgueiras de Sousa, Miguel Antonio Pinho Bruno, Angelita Alves de Carvalho

Seção 5

Outras análises acerca de População, Território e Estatísticas Públicas**257 Fatores de risco associados à prática de maus-tratos contra pessoas idosas no município de São Paulo**

Ana Carolina Soares Bertho, Bruna Matter dos Santos, Marília Miranda Forte Gomes, Jair Lício Ferreira dos Santos, Yeda Aparecida de Oliveira Duarte

271 Cidades inteligentes no Brasil: sistema de mensuração com foco em bem-estar social

Renata Curi de Moura Estevão Nagatomi, Daniel Takata Gomes, Julia Célia Mercedes Strauch

290 Estimating conditioning factors to proficiency via a Bivariate Multilevel Model using MCMC

Natália Raquel de Souza Pires, Pedro Luis do Nascimento Silva, James Brown

Convenções

-	Dado numérico igual a zero não resultante de arredondamento;
..	Não se aplica dado numérico;
...	Dado numérico não disponível;
x	Dado numérico omitido a fim de evitar a individualização da informação;
0; 0,0; 0,00	Dado numérico igual a zero resultante de arredondamento de um dado numérico originalmente positivo; e
-0; -0,0; -0,00	Dado numérico igual a zero resultante de arredondamento de um dado numérico originalmente negativo.

Apresentação

A tríade ensino, pesquisa científica e extensão é uma marca constitucional de todas as instituições de ensino superior no Brasil. A prática dessas atividades, contudo, está longe da estabilidade: é mutável e reflete as dinâmicas socioeconômicas, com suas questões históricas e tendências emergentes. Assim, no País, permanece a busca pela superação das diversas desigualdades estruturais, tal qual emergem desafios a práticas acadêmicas que considerem quadros teóricos e métodos inovadores, replicáveis e inclusivos. No IBGE, a produção de conhecimento não foge a essa dinâmica, mas ganha contornos específicos, relacionados a estudos e pesquisas nas áreas de Estatística e Geociências que fomentem a realização de sua missão institucional: “Retratar o Brasil com informações necessárias ao conhecimento de sua realidade e ao exercício da cidadania”.

É nesse quadro geral que se situa a Escola Nacional de Ciências Estatísticas - ENCE, identificada com uma dupla singularidade: é vinculada ao IBGE desde 1953, ano de fundação de sua atuação acadêmica, e é a instituição mais longeva a manter uma Graduação em Estatística em toda a América Latina. A ENCE, em suas dimensões material e imaterial, constitui um dos patrimônios do IBGE construídos ao longo dos quase 90 anos de atuação do Instituto. Para isso, o ensino ministrado atualmente na Escola é estruturado em três pilares: graduação, pós-graduação e capacitação/treinamento.

Ao longo de sua trajetória, a ENCE foi constantemente repensada, seja na construção dos níveis de ensino ofertados, seja no eixo central de sua ação, ou, ainda, no próprio conteúdo de cada modalidade de ensino. Nesse contexto, em 1998, foi criado o seu Programa de Pós-Graduação *stricto sensu*, com o nome de Pós-Graduação em Estudos Populacionais e Pesquisas Sociais, que, atuando com um curso de Mestrado desde a sua origem, mudou de designação, em 2014, para Pós-Graduação em População, Território e Estatísticas Públicas, com a criação do curso de Doutorado. Associando a prudência necessária à produção de estatísticas públicas que primem pela confiabilidade e pela robustez metodológica, o Programa de Pós-Graduação *stricto sensu* da ENCE foi paulatinamente consolidado nos cenários nacional e internacional, chegando, em 2023, aos 25 anos de existência com uma trajetória acadêmica notável.

O presente livro, *25 anos da Pós-Graduação stricto sensu da ENCE*, constitui um marco desse evento. Ele espelha o esforço contínuo de reflexão sobre o desenvolvimento do Programa de Pós-Graduação *stricto sensu* da ENCE, ao mesmo tempo em que exemplifica resultados de pesquisas realizadas pelo corpo docente e pelo alunado egresso de seus Cursos de Mestrado e Doutorado. Se, em obras anteriores, há o registro de questões da primeira e segunda décadas do Programa, esta registra os primeiros cinco anos da terceira década, em mais um capítulo da história da instituição

Considerando especialmente o período recente, há uma série de avanços no atual Programa de Pós-Graduação *stricto sensu* da ENCE que valem ser destacados, pois foram criadas e consolidadas, no decorrer dos últimos anos, iniciativas típicas de uma instituição com Doutorado, entre as quais: maior integração com o Curso de Graduação, principalmente por meio de estágios docentes nos quais atuam alunos do Programa; estabelecimento de convênios internacionais de cooperação para execução de projetos de pesquisa; criação de um núcleo de pesquisas na ENCE; continuidade da realização de pesquisas científicas relacionadas à produção institucional do IBGE no âmbito das estatísticas públicas e de informações geocientíficas; adoção de ações afirmativas nos processos seletivos de ingresso; e aumento da produção qualificada em periódicos, entre outras iniciativas. Esse cenário propiciou um importante reconhecimento do Programa, que recebeu em 2023, pela primeira vez, a nota 5 na avaliação da pós-graduação realizada pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES.

Tal movimento resultou de uma evolução coletiva, bem como da acumulação de conhecimento e de práticas que levaram à maturação do Programa de Pós-Graduação *stricto sensu* da ENCE em face das várias mudanças institucionais, sociais e tecnológicas ocorridas desde a sua fundação. Sem a atuação dos diversos docentes que por ele passaram ao longo de sua existência, essa conquista, por certo, não seria possível. Assim, como registro histórico, cabe destacar que o Programa teve nove Coordenadores nos seus primeiros 25 anos. Em ordem cronológica, foram: Victor Hugo de Carvalho Gouvêa, Jane Souto de Oliveira, Cesar Ajara, Aída Verdugo Lazo, José Eustáquio Diniz Alves, Maysa Sacramento de Magalhães, Suzana Marta Cavenaghi, César Augusto Marques da Silva e Angelita Alves de Carvalho. Em 2025, a décima Coordenadora assumiu: Ana Carolina Soares Bertho.

Com um total de 17 capítulos, esta obra traz reflexões sobre a trajetória da Pós-Graduação em População, Território e Estatísticas Públicas e apresenta resultados de alguns estudos realizados, mais recentemente, no âmbito das suas 2 linhas de pesquisa: População, território e condições de vida; e Produção de Estatísticas Públicas. Estas demonstram o compromisso da ENCE com as questões contemporâneas no avanço do saber, contribuindo, assim, para a missão institucional do IBGE. Ao mesmo tempo em que é um ponto de chegada que registra, no presente, o histórico recente, este livro também é um ponto de partida para o futuro, que está em aberto e é múltiplo, idealmente não colonizado pelo presente.

Desejamos a todos uma boa leitura!

Jorge Abrahão de Castro

Coordenador-Geral da Escola Nacional de Ciências Estatísticas

César Augusto Marques da Silva

Coordenador-Geral Adjunto da Escola Nacional de Ciências Estatísticas

Introdução

Em comemoração aos 25 anos do Programa de Pós-Graduação em População, Território e Estatísticas Públicas da Escola Nacional de Ciências Estatísticas - ENCE, ocorrido em 2023, o presente livro, estruturado em cinco seções, reúne uma coletânea de artigos que discorrem sobre a jornada exitosa desse Programa, bem como sobre os seus reflexos nas trajetórias acadêmica e profissional do alunado egresso, e apresenta estudos inovadores e relevantes que abordam temas centrais para a compreensão das dinâmicas sociais, econômicas e educacionais no Brasil, especialmente no que tange à produção estatística, às políticas públicas e ao mercado de trabalho.

A seção de abertura oferece um olhar interno acerca do Programa, com três artigos: o primeiro contempla suas importantes conquistas, entre os anos de 2018 e 2023, com destaque para o alcance da nota 5; outros dois se dedicam aos impactos do Programa, a partir da análise e da discussão de dados relacionados ao alunado egresso. Um deles apresenta os principais resultados e reflexões decorrentes de uma pesquisa de campo realizada, em 2023, como parte das iniciativas voltadas à autoavaliação institucional do Programa. O outro artigo analisa a inserção do alunado egresso dos Cursos de Mestrado e Doutorado do Programa, no período de 2010 a 2022, e sua inserção no mercado de trabalho formal, com base nos dados da Relação Anual de Informações Sociais - RAIS 2021, de responsabilidade do Ministério do Trabalho e Emprego. Esta seção, portanto, fornece subsídios importantes para repensar estratégias de formação contínua, ao mesmo tempo em que traz evidências concretas sobre a relevância do Programa de Pós-Graduação em População, Território e Estatísticas Públicas da ENCE, tanto na formação de profissionais com alta qualificação quanto em sua inserção no mercado de trabalho.

A segunda seção é composta de três artigos. O primeiro discute os desafios enfrentados no processo de produção de estatísticas públicas e as diversas limitações que o condicionam, entre as quais se destacam as relacionadas à dinâmica de acesso aos dados, ao financiamento dessa produção, aos recursos humanos necessários para tal e à legislação vigente, apontando, ainda, a necessidade de aprimoramento das metodologias e a importância de uma abordagem crítica frente às velhas e novas demandas por dados confiáveis. O segundo arti-

go discute a natureza do conceito de “população de difícil acesso”, tendo como base a teoria de relações de poder nas Ciências Sociais, e propõe uma extensão de tal conceito com vistas à melhor compreensão das dificuldades enfrentadas na coleta de informações em contextos complexos. O terceiro artigo mostra os limites e as possibilidades de utilização dos dados da Plataforma Nilo Peçanha - PNP para pesquisa sobre a atuação da Rede Federal de Educação Profissional, Científica e Tecnológica - RFEPCT, cuja origem remonta a 1909 com a criação, à época, de 19 escolas de aprendizes artífices pelo então Presidente da República Nilo Peçanha.

Temas essenciais para o desenvolvimento sustentável e a formulação de políticas públicas são discutidos na terceira seção, que inclui quatro artigos. Os dois primeiros contribuem para a estimação de indicadores e o monitoramento de metas globais, como os Objetivos de Desenvolvimento Sustentável - ODS, com foco no ODS 6.1, que visa, até 2030, alcançar o acesso universal e equitativo à água potável e segura para todos; e no ODS 11.7.1, que, embora ainda não seja calculado no Brasil devido à ausência de dados, objetiva medir a proporção de áreas urbanas destinadas a espaços públicos abertos para uso universal, conforme disposto na Agenda 2030 para o Desenvolvimento Sustentável, adotada, em 2015, pela Organização das Nações Unidas - ONU. Os dois artigos seguintes versam sobre a expansão e a interiorização da educação superior no Brasil, com base na análise de programas e políticas orientados ao acesso à educação, bem como sobre os desequilíbrios entre a oferta e a demanda desse nível de ensino no País.

A seção 4 dedica-se a análises setoriais e de gênero no mercado de trabalho brasileiro, por meio de quatro artigos. O primeiro apresenta estimativas trimestrais do número de pessoas de 14 anos ou mais de idade desocupadas, no Estado de Minas Gerais, utilizando para esse fim séries temporais de 10 estratos geográficos, com base nos dados da Pesquisa Nacional por Amostra de Domicílios Contínua - PNAD Contínua, do primeiro trimestre de 2012 ao segundo trimestre de 2022. O segundo artigo, pautado na PNAD Contínua relativa ao período de 2014 a 2018, marcado pela crise no mercado de trabalho brasileiro, constata um efeito interativo entre os marcadores sociais “cor ou raça” e “sexo”. O terceiro artigo analisa a relação entre os diferentes tipos de ocupação feminina e as suas jornadas de trabalho com o número de filhos tidos, com base na PNAD 2015, a última realizada anualmente e cujos dados são os únicos disponíveis para esse tipo de análise no Brasil. O quarto artigo analisa as características sociodemográficas e econômicas da população residente de 15 a 29 anos de idade que se encontra na condição NEET (*not in employment, education or training*), a partir dos dados da PNAD anual (3º trimestre de 2014) e da PNAD Contínua (3º trimestre de 2016).

Encerrando a publicação, a última seção apresenta três artigos que efetuam outras análises abarcadas pelo Programa População, Território e Estatísticas Públicas. O primeiro trata da ocorrência de maus-tratos contra pessoas idosas, cujas informações foram obtidas da Pesquisa Saúde, Bem-Estar e Envelhecimento - SABE 2015, coordenada pela Organização Pan-Americana da Saúde - OPAS, que entrevistou 1 224 pessoas de 60 anos ou mais de idade residentes no Município de São Paulo. O segundo artigo apresenta metodologia para construção de um sistema de indicadores que perpassam variados aspectos econômicos, sociais, de meio ambiente, e de governança para mensurar a aplicação do conceito de cidades inteligentes aliado ao bem-estar social, considerando-se as municipalidades brasileiras com mais de 100 000 habitantes, com base nos dados do Censo Demográfico 2010. O terceiro artigo, por fim, propõe uma abordagem metodológica inovadora ao analisar conjuntamente o desempenho em Português e Matemática por meio de modelos multiníveis bivariados aplicados ao SAEB. Ao evidenciar como fatores socioeconômicos e escolares atuam de forma integrada na aprendizagem, o trabalho reforça a importância de análises sensíveis às realidades locais.

Este livro é resultado do esforço conjunto do corpo docente do Programa de Pós-Graduação em População, Território e Estatísticas Públicas da ENCE, bem como do alunado egresso dos Cursos de Mestrado e Doutorado desse Programa no período de 2018 a 2025, cujas expertises enriquecem os debates acadêmico e técnico empreendidos sobre temas essenciais para o desenvolvimento social do Brasil. Esperamos que essas contribuições inspirem novas pesquisas, políticas públicas mais eficazes e uma valorização ainda maior das estatísticas como ferramenta fundamental para a transformação social.

Angelita Alves de Carvalho

Coordenadora da pós-graduação da ENCE 2021-2025

Seção 1

Reflexões e impactos dos 25 anos do Programa de Pós- Graduação em População, Território e Estatísticas Públicas (PTEP)

As recentes conquistas (2018-2023) do Programa de Pós-Graduação em População, Território e Estatísticas Públicas (PTEP)

Angelita Alves de Carvalho*

O Programa de Pós-Graduação em População, Território e Estatísticas Públicas (PTEP) completou 25 anos de atuação no ensino *stricto sensu* em 2023 e, nesse mesmo ano, alcançou a nota 5 na avaliação da CAPES, representando um marco importante em sua trajetória institucional. Nestes 25 anos o programa alcançou a marca de 402 mestres formados e 21 doutores e passou por diversas modificações que foram implementadas a fim de adequar os objetivos dos cursos, linhas de pesquisa e disciplinas às novas demandas e realidade de pesquisa da área. Uma das mudanças mais importantes desses primeiros 20 anos diz respeito à criação do doutorado em 2015, o que possibilitou o crescimento da pesquisa bem como consolidação da trajetória formativa ofertada pelo programa. Detalhes acerca desse histórico de criação dos cursos e reestruturações implementadas estão muito bem descritas em Silva (2020), capítulo introdutório da obra comemorativa aos 20 anos do programa, na qual se delineia o percurso do PTEP até o ano de 2018. Assim, nessa breve introdução, busca-se dar continuidade à história do programa, apresentando, de forma sintética, os principais avanços e destaques ocorridos no quinquênio 2018–2023.

Desde sua criação, o PTEP tem se destacado por oferecer uma formação acadêmica diferenciada em sua área de avaliação na CAPES, marcada pela articulação entre diferentes campos disciplinares e abordagens metodológicas voltadas à pesquisa demográfica, econômica e social. Tal diferencial se intensifica pela vinculação institucional ao IBGE, que proporciona uma formação teórica aliada à prática na produção e análise de estatísticas públicas. Essa relação favorece o acesso dos discentes aos pesquisadores do Instituto, promovendo um ambiente de formação conectado às etapas de planejamento, execução e análise das pesquisas que compõem as estatísticas demográficas, sociais e territoriais do país.

Como reflexo dessa proposta formativa, o programa consolidou-se ao longo de seus 25 anos, com progressivo crescimento e busca contínua por excelência acadêmica. Os avanços alcançados até aqui vão além dos critérios avaliativos da CAPES, expressando um amadurecimento institucional que se reflete tanto na formação discente quanto na produção científica e no impacto social do programa. Serão destacados abaixo alguns desses pontos que exemplificam essa consolidação do programa em diferentes frentes.

Um dos marcos importantes do período refere-se à redução do tempo médio para conclusão dos cursos. Historicamente, esse aspecto constituía um ponto

* Professora permanente do programa de pós-graduação da ENCE/IBGE. Coordenadora do programa entre 2021-2025.

crítico, sobretudo no mestrado, em razão da conciliação entre estudos e atividades profissionais por parte dos discentes. A criação da figura do “leitor” — docente convidado a acompanhar e oferecer pareceres ao longo do desenvolvimento das dissertações — foi fundamental nesse processo. Introduzido no segundo ano do mestrado, o leitor colabora na análise de viabilidade do projeto e no acompanhamento do seminário de pesquisa, contribuindo para a qualificação dos trabalhos e para o cumprimento dos prazos. Essa prática contribuiu significativamente para a diminuição dos tempos de defesa, ainda que a pandemia tenha provocado, pontualmente, a necessidade de prorrogações.

A internacionalização também se intensificou no período, ampliando as oportunidades para discentes e docentes. Desde a implantação do doutorado, em 2015, o Programa passou a participar do Programa de Doutorado Sanduíche no Exterior (PDSE), o que fortaleceu as experiências formativas internacionais e estimulou a construção de redes colaborativas. Além da participação contínua de estudantes estrangeiros, principalmente oriundos da América Latina e da África, em 2023 teve início um projeto de destaque no âmbito da internacionalização: “Aspectos da Dinâmica Demográfica Brasileira pós COVID-19 a partir dos dados do Censo Demográfico 2022”, financiado pelo Programa PROBRAL/CAPES, em parceria com o Max Planck Institute for Demographic Research (MPIDR), da Alemanha. O projeto envolveu docentes da ENCE, pesquisadores do IBGE e da UERJ, e possibilitou missões de intercâmbio, realização de minicursos e estágios de doutorado sanduíche e pós-doutorado no MPIDR. Essa iniciativa evidencia a inserção internacional do programa e sua capacidade de cooperação científica de alto nível.

O fortalecimento das conexões entre a pós-graduação e a graduação também foi valorizado na avaliação da CAPES. A participação dos docentes do PTEP em atividades da graduação em Estatística, como a oferta de disciplinas, orientação de monografias e supervisão de projetos de iniciação científica, tem ampliado a interlocução entre os dois níveis formativos. Essa aproximação tem, inclusive, favorecido a transição de alunos da graduação para o mestrado, fortalecendo o ciclo formativo da instituição. Adicionalmente, a extensão universitária tem se expandido, com destaque para projetos como o LabData e o Rio em Dados, que evidenciam o impacto social do programa.

A aplicabilidade das pesquisas desenvolvidas no âmbito do programa sempre foi uma de suas marcas. A partir de 2019, essa vocação se consolidou ainda mais com a celebração de convênios com órgãos públicos, como o Ministério do Turismo, que estimularam a realização de dissertações e teses com base em dados da PNAD Contínua. A participação dos alunos em projetos vinculados ao IBGE, como o Observa Censo e testes cognitivos de questionários, também tem fortalecido a formação prática e aplicado diretamente os conhecimentos adquiridos ao longo do curso.

O período também foi marcado pelo crescimento da produção acadêmica, alavancado por estratégias como a consolidação do doutorado, a criação de grupos de pesquisa, o aumento do número de pós-doutorandos, o estímulo à publicação de artigos em periódicos e anais de congressos, e o crescimento da captação de recursos. O projeto “Censo Demográfico 2022: agregando valor com análises da qualidade, dos métodos, das inovações e dos resultados”, aprovado no edital PDPG Pós-Doutorado Estratégico da CAPES, resultou na concessão de bolsas de pós-doutorado, reforçando as atividades de pesquisa e qualificação do corpo acadêmico. Em 2023, além dessas bolsas, outros três pesquisadores realizaram estágio pós-doutoral no Programa.

Ainda nesse contexto, consolidaram-se dois novos grupos de pesquisa registrados no CNPq: o grupo Big Data e Estatísticas Públicas e o GENERIS – Grupo de Pesquisas sobre Gênero, Sexualidades, Reprodução e suas Interseccionalidades. O GENERIS participou ativamente da realização do Seminário FEMINA 2023, atividade vinculada ao Festival Internacional de Cinema Feminino, promovendo debates acadêmicos sobre questões de gênero, representatividade

e direitos humanos, o que evidencia a ampliação da agenda temática do programa e seu comprometimento com a diversidade.

O aumento de projetos financiados, a elevação do número de bolsas de produtividade concedidas a docentes e o fortalecimento da qualificação profissional – com estágios pós-doutorais e participação em eventos científicos nacionais e internacionais – refletem o amadurecimento acadêmico do corpo docente. O incremento dos recursos financeiros do Programa, especialmente por meio do PROAP/CAPES, permitiu expandir o apoio a publicações, organização e participação em eventos, consolidando a visibilidade institucional. Entre os eventos co-organizados destacam-se o Congresso Internacional de Estatística (ISI), o Seminário NIEM/UFRJ, e a Semana NIC.br de Metodologias de Pesquisa. Discentes e docentes participaram de diversas conferências científicas como ISI, ENAPUR, ENANPEGE, SINAPE, CONGEO, ESAMP, MGEST, CINABEH, ABET, ABEP, ALAP e Desfazendo Gênero, entre outras, reforçando o compromisso com a interdisciplinaridade e a cooperação interinstitucional.

Outro avanço relevante foi a adoção de política de ações afirmativas no processo seletivo a partir de 2023. No mestrado, das 20 vagas ofertadas, 6 foram reservadas para ações afirmativas; no doutorado, das 10 vagas, 3 foram destinadas ao mesmo fim, totalizando 8 vagas com reserva. Essas políticas vêm acompanhadas de ações institucionais, como o planejamento de implantação de uma Gerência de Acompanhamento Psicopedagógico, prevista no PDI 2021–2025, com foco no suporte individualizado a discentes, especialmente àqueles em situação de vulnerabilidade ou que necessitem de apoio psicopedagógico.

Entre os desafios enfrentados no período, cabe destaca-se o impacto da pandemia de COVID-19, que exigiu uma reestruturação emergencial das atividades. Em resposta às restrições sanitárias, o programa organizou suas ações em três frentes principais: oferta de aulas remotas, readequação das atividades de pesquisa e reformulação dos processos seletivos. Essas medidas permitiram a manutenção da qualidade da formação, o fortalecimento do vínculo com os discentes e a produção de conhecimento relevante no contexto da crise sanitária. A rápida adaptação às novas condições demonstrou a resiliência e a capacidade institucional do programa, assegurando sua continuidade acadêmica e científica em um cenário adverso.

Em síntese, entre os 20 e 25 anos o programa foi caracterizado por uma série de reformulações e aprimoramentos que permitiram ao PTEP incorporar de maneira robusta o curso de doutorado, ampliar sua produção científica, fortalecer a internacionalização, consolidar vínculos com a graduação e com a gestão pública, e intensificar sua atuação social. Esses elementos foram determinantes para a elevação da nota do Programa de 4 para 5 na avaliação da CAPES, conquista celebrada como reconhecimento de um esforço coletivo, mas compreendida como um novo ponto de partida para alcançar patamares ainda mais elevados de excelência.

Para o próximo quadriênio e no marco dos 30 anos do programa, espera-se a consolidação de novas parcerias internacionais que viabilizem a vinda de docentes estrangeiros e a mobilidade de discentes e pós-doutorandos, bem como o fortalecimento das atividades de extensão e o aprofundamento das estratégias de impacto social. A reestruturação curricular em curso e a revisão dos regulamentos internos poderão incorporar novos objetivos e metas, contribuindo para ampliar o alcance e o papel do programa no campo de Planejamento Urbano e Regional/Demografia.

Potencialidades e Limitações na Formação em População, Território e Estatísticas Públicas: Reflexões a partir da Pesquisa de Egressos 2023

Angelita Alves de Carvalho*

Alinne Magalhães Gomes**

Resumo

No ano de 2023, o Programa de Pós-Graduação (PPG) em População, Território e Estatísticas Públicas da ENCE/IBGE completou 25 anos. Como parte das iniciativas de autoavaliação, o programa realizou uma pesquisa de campo com os seus egressos para analisar suas experiências durante o processo de formação e avaliar os impactos do programa em suas trajetórias pessoais e profissionais. Este artigo sistematiza os principais achados e reflexões decorrentes dessa iniciativa. A pesquisa foi realizada por meio de um formulário on-line (Google Forms), entre os meses de maio e agosto de 2023, com respostas voluntárias e anônimas. Foram obtidas 157 respostas válidas, o que representa cerca de 40% dos egressos até o ano de 2021. A maior parte dos respondentes concluiu o programa no último quinquênio (2017-2022). Entre os principais achados, destaca-se a excelente avaliação do programa pela grande maioria dos egressos, demonstrando a sua relevância para a trajetória acadêmica e profissional dos respondentes. A pesquisa ainda contribuiu para a identificação de aspectos essenciais para a melhoria contínua do programa, destacando tanto seus pontos fortes, que devem ser mantidos e potencializados, quanto os desafios enfrentados durante a formação, que demandam atenção para serem minimizados.

* Professora do programa de pós-graduação da ENCE.

** Professora do programa de pós-graduação da ENCE.

Introdução

A autoavaliação nos Programas de Pós-Graduação (PPGs) é um tema recente e tem sido cada vez mais sugerido pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) como uma importante atividade, tendo sido, inclusive, incorporada como um dos aspectos considerados em seu processo avaliativo. Isso porque a autoavaliação é uma prática estratégica, na medida em que fornece subsídios para o conhecimento da realidade institucional, e a partir da análise dessas informações, possibilita a elaboração de ações para o aprimoramento contínuo dos indicadores. No contexto da pós-graduação, esse aprimoramento contínuo contribui diretamente no processo de ensino-aprendizagem e no desenvolvimento de pesquisas que tragam visibilidade e respostas para os problemas enfrentados pela sociedade (Brasil, 2019; Leite *et al.*, 2020).

O processo de autoavaliação, além de favorecer essa (re)construção da identidade, contribui para o reconhecimento e fortalecimento das particularidades do programa e promove um envolvimento que ultrapassa o momento da avaliação externa (Brasil, 2019; Galdino, 2011; Leite *et al.*, 2020). Nessa perspectiva, em 04 de julho de 2018, a CAPES estabeleceu uma Comissão com o propósito de “implantar uma sistemática de autoavaliação no âmbito dos programas de pós-graduação” (Brasil, 2019, p.4).

Em relação às mudanças geradas pela implementação da autoavaliação, Leite *et al.* (2020, p. 347) destacam:

Se, até o momento, a educação superior do país esteve pautada por estratégias de avaliação influenciadas por políticas de ressonância internacional que repercutiram nas decisões sobre investimentos e organização institucional, neste momento a CAPES abre uma janela para que a diversidade de culturas e a riqueza dos contextos emergentes reflua para dentro da pós-graduação.

Trata-se, portanto, do estabelecimento de um processo contínuo de reflexão crítica sobre a prática (Brasil, 2019). Esse processo dialoga com as ideias de Freire (2005, p. 38), ao destacar que: “É pensando criticamente a prática de hoje ou de ontem que se pode aprimorar a próxima prática.”

Destaca-se que a autoavaliação não se configura como algo pontual ou um produto isolado, ela possui um caráter sistêmico e contínuo. Se constitui como um processo formativo, que busca compreender os significados do conjunto das atividades realizadas hoje, para a melhoria da qualidade educativa e maior relevância social amanhã (Brasil, 2019; Galdino, 2011; Leite *et al.*, 2020).

Para além das complexidades e pluralidades que o processo de autoavaliação assume em cada programa de pós-graduação, é de fundamental importância o envolvimento de todos os membros da instituição, em uma perspectiva que articule o passado, o presente e a visão de futuro. Ao convocar a comunidade institucional para esse processo, torna-se essencial incluir, além de docentes, discentes e técnicos administrativos, também os egressos, cujas vivências ao longo de sua trajetória no programa constituem um valioso repertório para a reflexão coletiva. É preciso realizar a inserção dos egressos de forma mais profícua no sistema de avaliação, pois isso favorece a melhoria das condições de ensino e representa mais um passo em direção à ampla participação nos processos avaliativos da pós-graduação brasileira (Brasil, 2019; Leite *et al.*, 2020).

O acompanhamento sistemático e as pesquisas com egressos, embora representem apenas uma das etapas do processo de autoavaliação, configuram-se como práticas estratégicas. Distanciados do cotidiano da pós-graduação, os egressos são capazes de elaborar uma visão mais crítica e reflexiva sobre os processos vivenciados, compreendendo, com o tempo, os impactos da formação em sua atuação profissional e cidadã. Ao compartilharem suas percepções, contribuem de forma valiosa para o aprimoramento contínuo dos Programas de Pós-Graduação (PPGs). As informações coletadas a partir desses relatos subsidiam decisões mais conscientes e bem fundamentadas, orientando ajustes nas práticas institucionais e promovendo melhorias na

experiência de aprendizagem daqueles que atualmente se encontram em processo de formação (Espartel, 2009; Maccari; Teixeira, 2014).

Espartel (2009, p.104) destaca a importância de captar as percepções dos egressos por meio de dois enfoques complementares: a “avaliação direta”, voltada à análise de aspectos “operacionais e estruturais”; e a “avaliação indireta”, em uma análise temporal ampliada, que permite ao egresso refletir sobre os impactos da formação em sua trajetória profissional ao longo do tempo. A integração desses dois tipos de avaliação possibilita uma compreensão mais abrangente da experiência formativa, permitindo identificar fragilidades e potencialidades do programa, bem como delinear “estratégias de manutenção de relacionamento com os egressos”.

Estabelecer e cultivar uma relação duradoura com os egressos, acompanhando sua evolução profissional ao longo dos anos e compreendendo como percebem sua formação à medida que vivenciam diferentes experiências, é de grande importância para as Instituições de Ensino Superior (IES). Conforme destacam Chi *et al.* (2012), os egressos ocupam uma posição estratégica, pois ao atuarem em diferentes contextos sociais e profissionais, carregam consigo a marca da instituição. Nessas vivências, reproduzem na prática os conhecimentos e valores adquiridos durante a formação, projetando a instituição para a sociedade. Ciente da relevância dos egressos e do valor de suas perspectivas, este artigo tem como objetivo sistematizar as informações obtidas e refletir acerca dos dados levantados na pesquisa de egressos realizada no Programa de Pós-Graduação em População, Território e Estatísticas Públicas (PTEP) da Escola Nacional de Ciências Estatísticas (ENCE), vinculada ao Instituto Brasileiro de Geografia e Estatística (IBGE).

Aspectos Metodológicos

O PTEP teve início em 1998, com a oferta do curso de mestrado. Posteriormente, em 2015, passou a ofertar também o curso de doutorado. No ano de 2023, ao comemorar 25 anos de existência, o programa promoveu uma pesquisa com os seus egressos em celebração à data e como parte das ações de construção e fortalecimento do processo de autoavaliação. A iniciativa teve como objetivo identificar o perfil dos alunos já formados, considerando a diversidade do público atendido bem como mapear as contribuições da formação para suas trajetórias e reunir sugestões voltadas ao aprimoramento contínuo dos cursos.

A pesquisa de egressos em 2023 foi conduzida em duas etapas. A primeira consistiu na extração e análise do banco de dados do sistema acadêmico da ENCE, o qual permitiu a elaboração de uma listagem de todos os trabalhos de conclusão de curso do programa e, conseqüentemente, os egressos desde sua criação. Nesse banco de dados, o qual teve a data final em dezembro de 2021, constavam um total de 410 defesas, sendo 395 do curso de mestrado e 15 de doutorado. Contudo, o número de egressos identificados foi de 399, dado que 11 discentes obtiveram ambos os títulos no programa. Esse banco de dados possibilitou também inferir que:

- 51% dos estudantes eram do sexo feminino e 49% do sexo masculino.
- A média de idade ao ingressar no programa era de 27 anos.
- 74% dos discentes residiam no município do Rio de Janeiro ao iniciarem os cursos.
- Apenas 14 discentes, representando 3,5% do total, eram estrangeiros, refletindo a predominância de alunos brasileiros.

A partir desse breve panorama, buscou-se atualizar o banco de dados, uma vez que muitos endereços de e-mail eram inválidos. Para isso, foi realizada uma busca pelos nomes dos egressos nas plataformas sociais digitais tais como Facebook, Instagram e LinkedIn para atualização dos e-mails. Como resultado dessa ação, grande parte dos contatos foram resgatados e, em seguida, foi enviado um e-mail convidando para a participação na pesquisa. A mensagem trazia um *link* para preenchimento de um questionário no Google Forms, cuja participação era voluntária e anônima.

A pesquisa foi realizada entre os meses de maio e agosto de 2023, resultando em 157 respostas válidas, o que corresponde a aproximadamente 40% do total de egressos do programa. Dentre os respondentes, a maioria (48%) concluiu o curso no quinquênio 2017-2022.

Resultados e Discussões

Perfil dos respondentes

A maior parte dos respondentes pertence ao sexo masculino (54%), enquanto 45% pertencem ao sexo feminino. Embora o programa costume receber alunos de outras nacionalidades, 99% dos participantes da pesquisa eram brasileiros. A baixa taxa de respostas dos egressos estrangeiros sinaliza a necessidade de estratégias de acompanhamento mais específicas para esse grupo, os quais tem demonstrado perder o contato com a instituição após a conclusão do curso.

Com relação à raça/cor, a maioria dos respondentes se identificou como branca (67%), seguida por pardos (21%), pretos (9%) e amarelos (3%). A fim de ampliar o acesso de pessoas negras à pós-graduação, a ENCE estabeleceu, em 2023, as normas para a política de ações afirmativas em seus processos seletivos. Essa política, regulamentada pelas Instruções Normativas IN-ENCE nº 16/2023 e nº 17/2023, definem que entre 25% e 30% das vagas dos cursos de pós-graduação deverão ser reservadas para pessoas autodeclaradas pretas e pardas e comprovadamente pertencentes às populações indígenas ou membros de outros povos e comunidades tradicionais. A partir de 2023, os editais dos processos seletivos da pós-graduação passaram a obedecer às instruções normativas, apresentando a distribuição detalhada das vagas por modalidade. As ações afirmativas foram implementadas e já passaram a valer para as turmas com ingresso em 2024. Espera-se que, nos próximos anos, os efeitos dessa política possam ser observados com a maior diversidade entre os ingressantes, e consequentemente, no aumento de egressos negros em estatísticas futuras do programa.

Em relação à formação acadêmica, 65% possuíam o título de mestre, 29% possuíam o título de doutor e 6% concluíram o pós-doutorado. Observa-se que grande parte deles cursou a graduação em instituições públicas de ensino superior. A predominância de mestres entre os egressos é um dado esperado, uma vez o número de vagas destinadas ao curso mestrado representa o dobro das vagas de doutorado. Adicionalmente, considerando a recente aprovação do doutorado (primeira defesa ocorreu em 2018), é esperado que o programa possua um quantitativo maior de mestres em relação ao quantitativo de doutores.

A principal formação universitária dos egressos foi em estatística (38%), seguida por economia (18%). Por um lado, esse dado evidencia a continuidade da formação dos egressos do curso de graduação em estatística da ENCE, que optam por continuar os seus estudos na instituição. Por outro lado, reflete o terreno fértil que o programa da ENCE representa para esses profissionais. Por ser a primeira escola de estatística da América Latina e estar inserida no IBGE, a instituição possui um corpo docente e uma grade de disciplinas voltados para os estudos de técnicas e metodologias estatísticas.

O intervalo entre a conclusão da graduação e o ingresso no mestrado foi, predominantemente, de 1 a 5 anos, seguido pelos discentes que ingressaram com menos de 1 ano após a conclusão da graduação. Com relação ao ingresso no doutorado, a maioria iniciou o curso com menos de 1 ano após a conclusão do mestrado. Esses dados indicam que nossos egressos possuem uma preferência pela continuidade de sua formação, sem interrupções ou com baixo intervalo entre elas. Essa característica pode indicar que esses profissionais buscaram por mais qualificação diante das dificuldades de uma alocação adequada no mercado de trabalho após a conclusão da graduação ou mesmo do mestrado. Esse comportamento pode ter relação também com o fato de que o ingresso imediato ao mestrado ou ao doutorado pode parecer mais

fácil para aqueles que já estão inseridos no meio acadêmico, uma vez que já conhecem o ritmo de estudos e entendem que a saída e um futuro retorno para o meio acadêmico necessitaria de um novo processo de adaptação. Vale destacar que, no caso dos egressos de mestrado da ENCE, é possível integralizar as disciplinas já cursadas caso o ingresso no doutorado ocorra em até 6 anos após a conclusão do mestrado. Essa possibilidade pode representar um ganho para o egresso interessado em prosseguir no programa, pois, com a redução da carga de disciplinas, ele dispõe de mais tempo para se dedicar exclusivamente à pesquisa ou até acelerar a conclusão do curso.

Vivências durante os cursos

Sobre a trajetória dos alunos nos cursos, observou-se que 66% dos respondentes foram bolsistas, o que indica um elevado apoio e um esforço contínuo do programa em proporcionar condições para que os alunos se mantenham vinculados ao curso e o concluam. As bolsas recebidas pelos alunos incluem aquelas concedida pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), as de monitoria do IBGE, bem como as provenientes do convênio entre o IBGE e o Ministério do Turismo. Destaca-se que esta última, além de oferecer valores superiores às bolsas CAPES e IBGE, proporciona aos bolsistas uma aproximação com a produção das estatísticas públicas, uma vez que eles atuam diretamente com a manipulação das bases de dados do referido convênio, no caso do Ministério do Turismo, a PNAD Contínua, especificamente o suplemento de turismo. Atualmente, um convênio similar foi estabelecido com o Ministério da Saúde, junto a Pesquisa Nacional de Demografia e Saúde.

Embora o recebimento de bolsas possa amenizar dificuldades financeiras, a trajetória dos discentes ainda é marcada por outros desafios. Sobre as principais dificuldades para a conclusão dos cursos (Gráfico 1), a conciliação das atividades acadêmicas com a vida profissional foi a mais citada pelos respondentes. Isso porque grande parte dos respondentes já exerciam atividades remuneradas ao iniciar o programa ou começaram a trabalhar durante sua trajetória nos cursos.

Conciliar os compromissos de um mestrado ou doutorado com uma rotina profissional não é uma tarefa trivial, exigindo do aluno um elevado grau de planejamento, disciplina e resiliência. A falta de tempo imposta pelo acúmulo de responsabilidades é desafiadora e pode influenciar diretamente na permanência ou no desligamento do discente em um PPG (Colombo, 2023).

A segunda maior dificuldade apontada, pelos respondentes foi a elaboração da dissertação ou tese. Segundo Freitas (2002), trata-se de uma fase desafiadora da vida acadêmica, exigindo muita concentração, disciplina, organização e ainda uma boa interação entre orientando e orientador.

No mesmo nível da dificuldade na elaboração da dissertação ou tese encontram-se as dificuldades financeiras. Embora 66% dos respondentes tenham sido bolsistas e o programa seja gratuito, essa questão está relacionada ao fato de que as bolsas oferecidas pelos PPGs são, na maioria das vezes, insuficientes para permitir que os estudantes se dediquem exclusivamente à pós-graduação, dado seus valores reduzidos, especialmente em uma cidade com o custo de vida elevado, como o Rio de Janeiro. Esse cenário torna-se ainda mais crítico para os estudantes que precisaram se mudar para cursar o programa.

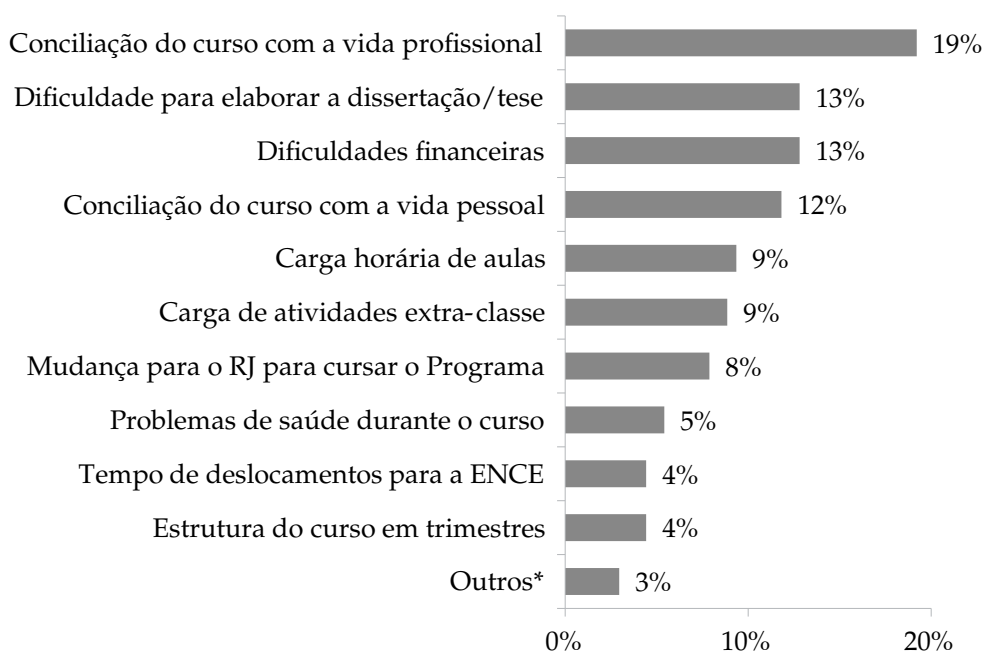
Cabe mencionar que, durante muitos anos, os valores das bolsas permaneceram congelados e somente em 2023 houve um reajuste de 40%. O valor da bolsa de mestrado passou de R\$ 1.500 para R\$ 2.100,00, e a de doutorado de 2.200 para R\$ 3.100 (Brasil, 2023a). Além do reajuste, em 2023 a CAPES regulamentou o acúmulo da bolsa com atividade remunerada ou outros rendimentos, permitindo conciliar o recebimento de bolsa de pós-graduação (mestrado, doutorado e pós-doutorado) e emprego com carteira assinada. Anteriormente, só era permitido aos bolsistas da CAPES ter vínculo empregatício remunerado e acumular bolsas

em atividades profissionais vinculadas à pesquisa que estava sendo realizada (Brasil, 2023b). Espera-se que, nos próximos levantamentos, as dificuldades financeiras não configurem entre as principais dificuldades dos egressos, evidenciando que as modificações no valor das bolsas impactaram positivamente a trajetórias dos estudantes e auxiliaram a reduzir os problemas financeiros enfrentados durante a realização dos cursos

Três outras dificuldades apontadas foram a conciliação do curso com a vida pessoal e as cargas de leitura das aulas e das atividades extraclases. Essas questões estão alinhadas com os estudos de Miranda et al. (2023) e Rezende et al. (2017), que destacam que o elevado volume de leitura e as diversas demandas durante o curso exigem grande esforço e dedicação de tempo dos discentes. Diante da grande quantidade de atividades e da limitação temporal, os estudantes encontram dificuldades em conciliar as exigências do curso com os compromissos pessoais, familiares, sociais e profissionais. Esse contexto contribui para níveis elevados de estresse, tornando ainda mais desafiador alcançar o desempenho acadêmico desejado.

As informações reveladas durante a pesquisa com os egressos evidenciam a necessidade de uma abordagem coletiva envolvendo docentes e coordenação na elaboração de estratégias para reduzir essas dificuldades, por exemplo, repensando a carga de leitura e atividades das disciplinas. Além dessas estratégias, também é necessário promover ações que auxiliem os discentes na organização pessoal e no enfrentamento do estresse.

Gráfico 1 – Dificuldades enfrentadas para a conclusão dos cursos



Fonte: Pesquisa com Egressos PTEP, 2023.

Nota: Outros inclui problemas com a coordenação, com professores, com alunos e com infraestrutura física

Outro dado importante revelado pela pesquisa é que grande parte dos egressos relatou não ter tido conhecimento sobre os projetos de pesquisa disponíveis, levando-os a não se envolverem com as iniciativas existentes. Considerando os impactos positivos desse envolvimento, tanto na graduação quanto na pós-graduação, foi consolidado o Núcleo de Pesquisa da ENCE. Com o objetivo de dar visibilidade aos projetos existentes, desde 2022 o Núcleo promove a divulgação anual do Catálogo de Projetos da ENCE. O catálogo é disponibilizado no site¹ e contém um resumo de todas as iniciativas que estão sendo desenvolvidas pelos docentes e discentes, sejam elas de pesquisa ou extensão. Em paralelo, recentemente, foram criados dois grupos de pesquisa da ENCE, os quais têm divulgado suas atividades e pesquisas nas redes sociais da Escola. Espera-se que, tanto o catálogo quanto a divulgação das pesquisas nas re-

¹ <https://ence.ibge.gov.br/index.php/descricao/projetos-nupesq>

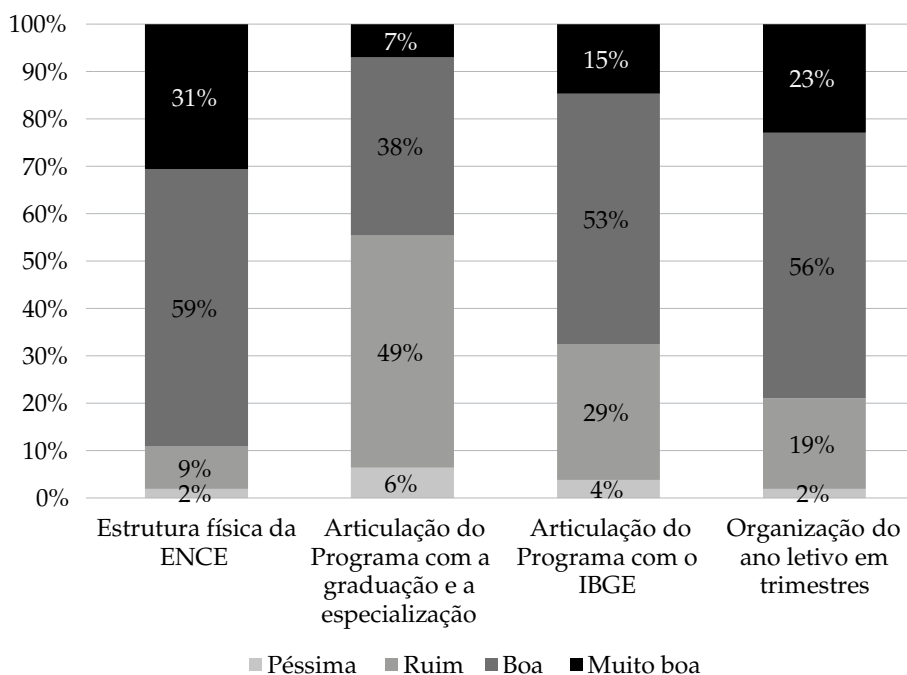
des sociais, permitam maior visibilidade dos projetos de pesquisa existentes, despertando o interesse e o envolvimento dos discentes. Conforme sugerido por Freitas e Souza (2018), a participação estruturada e contínua em projetos e grupos de pesquisa tem sido percebida como uma atividade de grande contribuição na formação dos mestres e doutores. Essas iniciativas enriquecem o processo formativo, possibilitando a criação de redes, a publicação conjunta de artigos científicos, o encontro com diferentes temas e metodologias de estudo e outras atividades não contempladas diretamente nas disciplinas.

Com relação às disciplinas obrigatórias e optativas, os egressos avaliaram como adequadas, indicando ainda que cursaram as disciplinas de seu interesse. As atividades de monitoria e estágio docente também foram avaliadas positivamente, sendo reconhecidas pela sua contribuição na formação acadêmica. Por outro lado, os egressos consideraram que o Seminário de Dissertação contribuiu pouco para a sua formação, diferentemente da Qualificação de Tese, que foi considerada como uma atividade de muita contribuição. A partir dessas informações, no processo de reestruturação do programa, está prevista a criação de um exame de qualificação de mestrado, nos moldes do exame de qualificação de doutorado, buscando melhorar o impacto dessa atividade na trajetória acadêmica dos mestrandos.

Avaliação do programa

A avaliação feita pelos egressos a respeito da estrutura física da ENCE, da articulação entre o PTEP com os outros níveis de ensino da escola (graduação e especialização) e com o IBGE, bem como sua organização em trimestres está exposto no Gráfico 2.

Gráfico 2 - Avaliação da estrutura física, articulação entre o PPG e os demais cursos da ENCE, articulação entre o PPG e o IBGE e a modalidade em trimestres



Fonte: Pesquisa com egressos PTEP, 2023.

Destaca-se que a estrutura física da ENCE foi o item mais bem avaliado, com 31% dos respondentes a considerando como muito boa e 59% boa. Embora a distribuição em trimestres não tenha recebido uma avaliação negativa, com apenas 21% dos egressos avaliando como péssimo ou ruim, a desaprovação pelo formato trimestral do programa tem sido apontada por discentes e docentes em outros instrumentos de avaliação do programa. Dessa forma, a coordenação atual do programa iniciou um processo de diálogo interno para a reestruturação do programa e sua organização em semestres, além de outras adequações do regulamento. Para atender a essa de-

manda, foi instaurada uma comissão para que a nova estrutura do programa seja apresentada aos estudantes em 2025 e sua efetivação a partir da turma de 2026.

O Gráfico 2 revela também que a articulação entre o PTEP com o IBGE tem sido considerada boa para 68% dos egressos e ruim para 32%. Apesar da ENCE estar inserida no IBGE, a articulação entre as áreas que lidam diretamente com a produção da estatística pública e a área acadêmica ainda é desafiadora. Acredita-se que para alguns estudantes essa articulação seja mais evidente por abordarem em suas pesquisas temas vinculados à linha de pesquisa Produção de Estatísticas Públicas e por buscarem participar de seminários, divulgações ou mesmo eventos do instituto. Para outros, a conexão pode ser ainda mais direta por atuarem como bolsistas dos convênios citados, cujas atividades necessitam da interação com as áreas do IBGE. Contudo, para outros alunos, que abordam em suas pesquisas temas que não dialogam fortemente com o IBGE e que não são bolsistas, essa articulação pode não ser evidente.

A partir dessa percepção, algumas iniciativas têm sido implementadas buscando aproximar os discentes das áreas que lidam diretamente com as pesquisas e conseqüentemente aumentar a articulação com o IBGE. Em 2022, quatro alunos participaram do programa Observa Censo, programa que tem como objetivo o compartilhamento de conhecimentos sobre a coleta dos dados do Censo Demográfico 2022. Nessa iniciativa, os discentes vivenciaram a aplicação dos questionários do Censo 2022 em quatro capitais brasileiras: Belém, Aracaju, Belo Horizonte e Curitiba.

Em 2023, os alunos tiveram a oportunidade de participar do Teste Cognitivo de um questionário para Estatísticas sobre Pessoas com Deficiência. Esse teste tem o objetivo de desenvolver e aprimorar as perguntas sobre educação inclusiva.

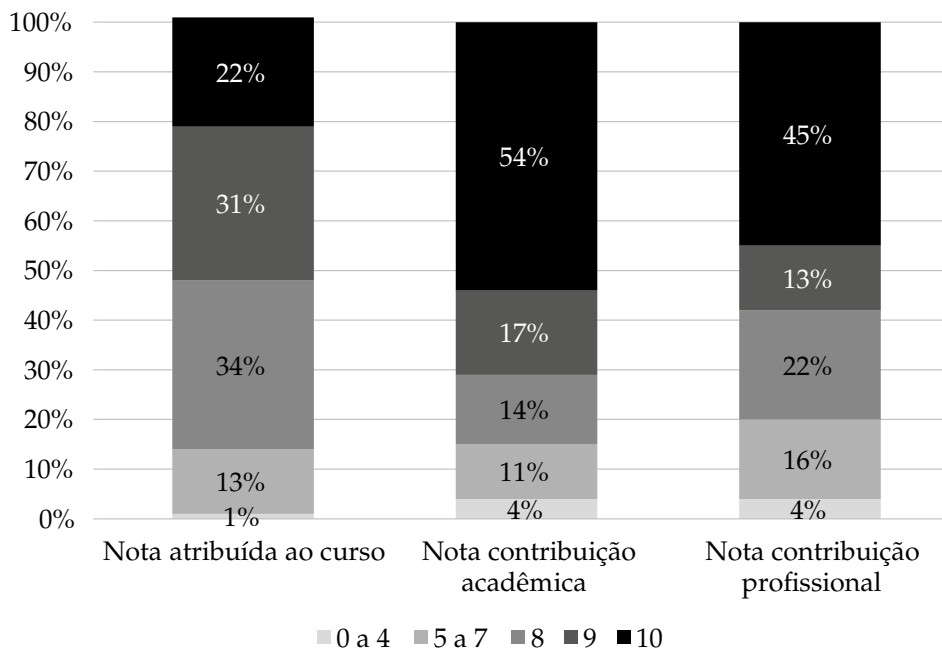
Cabe destacar que, em 2022, os alunos participaram do teste cognitivo do questionário da PNDS sobre as perguntas de orientação sexual e identidade de gênero. Essa iniciativa ocorreu nas dependências da ENCE e os discentes atuaram no processo de recrutamento de possíveis respondentes nos arredores da escola, na aplicação dos questionários e na escrita do relatório de percepção das respostas.

Destaca-se também a participação de discentes no teste piloto do questionário do Censo de População de Rua realizado em Niterói, em 2023. Os alunos passaram três dias imersos em atividades de treinamento de entrevistadores, treinamento de observadores, coleta e reunião de avaliação. Essa e as demais experiências citadas aproximam os discentes das diversas etapas da produção estatística. São iniciativas que promovem experiências práticas, proporcionando atividades que demandam dos discentes uma postura ativa e a aplicação dos conhecimentos adquiridos por meio das leituras e disciplinas, enriquecendo a formação dos futuros pesquisadores.

O Gráfico 2 também apresenta que a articulação entre o PTEP e os demais cursos da ENCE foi considerada boa para 45% dos egressos e ruim para 55%. Embora os cursos ocorram no mesmo espaço físico, promover a interação entre os diversos públicos ainda é uma tarefa a ser aprimorada. Uma das formas de articulação é por meio do estágio docente, no qual os doutorandos do programa atuam em parceria com os docentes das disciplinas do curso de graduação em estatística e interagem com os graduandos. Além do estágio docente, a escola promove diversos eventos como Tardes Culturais, Mostra de Talentos, Ence de Portas Abertas, aulas inaugurais e seminários que podem contribuir para promover a interação dos alunos dos diferentes níveis.

Durante a pesquisa, os respondentes puderam ainda atribuir uma nota geral ao programa e à contribuição do curso em suas trajetórias acadêmica e profissional (Gráfico 3).

Gráfico 3 - Notas atribuídas ao curso e à sua contribuição para a trajetória acadêmica e profissional

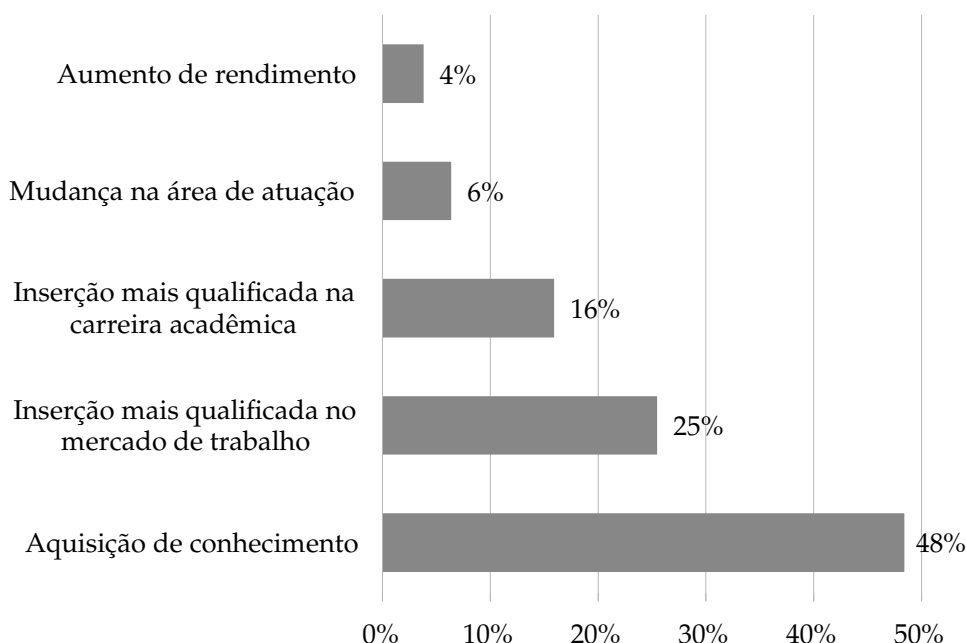


Fonte: Pesquisa com egressos PTEP, 2023.

Mais de 80% dos respondentes atribuíram notas entre 8 e 10 para todos os três quesitos. Percebe-se que a nota relativa à contribuição acadêmica foi a maior, em que 54% atribuíram nota 10.

Apesar de baixo, o maior percentual de notas baixas (0 a 7) foi atribuído para a contribuição profissional. Como muitos egressos já eram servidores concursados quando ingressaram nos cursos, acredita-se que pela sua condição estável no mercado de trabalho, o título impactou pouco para a sua trajetória em seus locais de trabalho.

Além da atribuição das notas, foi solicitado que os egressos indicassem as principais contribuições do programa em sua trajetória. O Gráfico 4 mostra as principais contribuições do programa e revela que a grande parte dos egressos considera a aquisição de conhecimento como a principal contribuição. Em seguida, a inserção mais qualificada tanto na vida profissional quanto acadêmica foram indicadas como contribuições importantes. Mudança da área de atuação e aumento da remuneração também foram citados. Esses e outros aspectos têm sido encontrados em estudos que analisaram o impacto dos programas de pós-graduação na vida dos egressos. O estudo de Pontes (2023) revelou que os egressos indicaram impactos na qualificação, empregabilidade e renda, além de aumento da eficiência e maior segurança para participarem de processos seletivos.

Gráfico 4 – Principais contribuições do programa na trajetória dos egressos

Fonte: Pesquisa com egressos PTEP, 2023.

Por fim, ao se analisar a situação dos egressos no mercado de trabalho antes e após a conclusão dos cursos, observa-se uma inserção mais favorável após a passagem pelo programa. Houve uma redução no número de respondentes que se encontravam sem trabalho, procurando emprego, atuando por conta própria ou como bolsistas. Em contrapartida, verificou-se um aumento no percentual de egressos empregados no setor público e no setor privado.

Considerações Finais

A pesquisa com os egressos explicitou importantes aspectos do programa, demonstrando seus pontos fortes e as oportunidades de melhoria.

Como pontos fortes, destaca-se a estrutura física e a relação da ENCE com o IBGE. O instituto conta com um corpo técnico especializado, com conhecimentos e habilidades que podem enriquecer o percurso formativo proposto pelo programa. Nesse sentido, apesar da articulação ter sido bem avaliada, é um tópico que deve ser aprimorado, seja por meio de pesquisadores da área técnica atuando nas disciplinas, minicursos e seminários do programa, seja por meio da inserção de alunos em atividades práticas da produção das estatísticas públicas, como coleta de dados, testes de campo e outras aproximações. A realização de eventos conjuntos também se apresenta como uma boa prática. A ENCE já realizou eventos em parceria com a Diretoria de Geociências (DGC) e há previsão de replicar essa ação com a Diretoria de Pesquisas (DPE).

A avaliação com notas elevadas por grande parte dos respondentes também foi um ponto forte evidenciado pela pesquisa, indicando a percepção de impactos positivos diretos na vida acadêmica e profissional.

Além dos pontos positivos, a avaliação evidenciou fraquezas e limitações importantes. Sobre essas dificuldades, destaca-se a conciliação da vida acadêmica com a profissional e as limitações financeiras. Essas questões apontam para a insuficiência da política científica do país, a qual não disponibiliza um número suficiente de bolsas, bem como bolsas com valores que possam garantir a dedicação integral aos cursos de pós-graduação.

Outro aspecto que merece destaque foi o baixo envolvimento dos alunos nos projetos e grupos de pesquisa já instituídos no programa. Considerando as potencialidades dessas atividades, evidenciou-se a necessidade de ações visando a criação, o fortalecimento e a divulgação dos grupos de investigação existentes na escola.

Além de identificar forças e fragilidades, a pesquisa evidenciou os impactos nas trajetórias dos egressos. Observou-se que, além da aquisição de conhecimentos, a passagem pelo programa contribuiu para uma melhor inserção no mercado de trabalho.

Espera-se que os resultados obtidos sejam considerados na formulação de ações de aprimoramento no planejamento estratégico, visando à ampliação do impacto do programa nas dimensões acadêmica, profissional e pessoal dos estudantes.

Referências

- BRASIL. Ministério da Educação. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES. *Autoavaliação de Programas de Pós-Graduação - Grupo de Trabalho*. Brasília, 2019. Disponível em: <https://www.gov.br/capes/pt-br/centrais-de-conteudo/10062019-autoavaliacao-de-programas-de-pos-graduacao-pdf> . Acesso em: 20 junho 2024.
- BRASIL. Ministério da Educação. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES. *Portaria nº 33, de 16 de fevereiro de 2023a. Dispõe sobre o reajuste dos valores das bolsas da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES, no país*. Disponível em: <https://cad.capes.gov.br/ato-administrativo-detalhar?idAtoAdmElastic=10902> . Acesso em: 12 set. 2025.
- BRASIL. Ministério da Educação. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES. *Portaria nº 133, de 10 de julho de 2023b. Regulamenta o acúmulo de bolsas de mestrado, doutorado e pós-doutorado concedidas pela CAPES no País com atividade remunerada ou outros rendimentos*. Disponível em: <https://cad.capes.gov.br/ato-administrativo-detalhar?idAtoAdmElastic=12302>. Acesso em: 12 set. 2025.
- CHI, Hongmei; JONES, Edward L.; GRANDHAM, Lakshmi P. *Enhancing mentoring between alumni and students via smart alumni system*. *Procedia Computer Science*, v. 9, p. 1390–1399, 2012. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1877050912002748> . Acesso em: 17 abr. 2025.
- COLOMBO, Daniel Gama e. *Os desafios do emprego durante o doutorado e o acúmulo com bolsas de pesquisa*. Brasília, 2023. Disponível em: <https://www.ipea.gov.br/cts/pt/central-de-conteudo/artigos/artigos/378-os-desafios-do-emprego-durante-o-doutorado-e-o-acumulo-com-bolsas-de-pesquisa> . Acesso em: 29 abr. 2025.
- ESCOLA NACIONAL DE CIÊNCIAS ESTATÍSTICAS (ENCE). *Instrução Normativa ENCE/IBGE nº 16, de 28 de julho de 2023. Estabelece normas para as ações afirmativas nos cursos de graduação e pós-graduação da ENCE*. Disponível em: <https://drive.google.com/file/d/1hejB8L31cCIH2eOk7yD9evXKvc8u7Qew/view> . Acesso em: 12 set. 2025.
- ESCOLA NACIONAL DE CIÊNCIAS ESTATÍSTICAS (ENCE). *Instrução Normativa ENCE/IBGE nº 17, de 28 de julho de 2023. Estabelece normas e procedimentos para a realização de heteroidentificação racial complementar à autodeclaração racial de candidato(a) preto(a) e pardo(a) para preenchimento de vagas reservadas nos cursos de graduação e pós-graduação da ENCE*. Disponível em: https://drive.google.com/file/d/1vXuVS_aLKSv37NPrjki1NI9iT7K0FH/view . Acesso em: 12 set. 2025.
- ESPARTEL, Lélis Balestrin. *O uso da opinião dos egressos como ferramenta de avaliação de cursos: o caso de uma instituição de ensino superior catarinense*. *Revista Alcance*, v. 16, n. 1, p. 102-114, 2009.
- FREITAS, Maria Ester de. *Viver a tese é preciso!: Reflexões sobre as aventuras e desventuras da vida acadêmica*. *Revista de administração de empresas*, v. 42, p. 1-6, 2002.
- FREITAS, Maria de Fatima Quintal de; SOUZA, Jusamara. *Pensar a formação e a pesquisa na pós-graduação stricto sensu*. *Educar em Revista*, v. 34, n. 71, p. 09-18, 2018.
- FREIRE, Paulo. *Pedagogia da Autonomia: saberes necessários à prática educativa*. 31. ed. São Paulo: Paz e Terra (Coleção Leitura), 2005.
- GALDINO, Mary Neuza Dias. *A autoavaliação institucional no ensino superior como instrumento de gestão*. Fundação CESGRANRIO/Universidade do Grande Rio, 2011.
- LEITE, Denise et al. *A autoavaliação na Pós-Graduação (PG) como componente do processo avaliativo CAPES*. *Avaliação: Revista da Avaliação da Educação Superior (Campinas)*, v. 25, n. 02, p. 339-353, 2020.

MACCARI, Emerson Antonio; TEIXEIRA, Gislaine Cristina dos Santos. *Estratégia e planejamento de projeto para acompanhamento de alunos egressos de programas de pós-graduação stricto-sensu*. Revista de Administração da Universidade Federal de Santa Maria, v. 7, n. 1, p. 101-116, 2014.

MIRANDA, Gilberto José *et al.* *Dificuldades, Preocupações e Estresse na Pós-Graduação*. Revista Gestão Universitária na América Latina-GUAL, p. 24-43, 2022.

PONTES, Alinildo Ferreira de, *A contribuição da pós-graduação para a formação acadêmica e profissional dos egressos*. 2023. 118 f. Dissertação (Programa de Pós-Graduação em Políticas Públicas, Gestão e Avaliação da Educação Superior - Mestrado Profissional do Centro de Educação) - Universidade Federal da Paraíba, João Pessoa, 2023.

REZENDE, Marise Santana de; MIRANDA, Gilberto José; PEREIRA, Janser Moura; CORNACCHIONE JUNIOR, Edgard Bruno. *Estresse e desempenho acadêmico na pós-graduação stricto sensu em ciências contábeis no Brasil*. Arquivos Analíticos de Políticas Educativas, 25 (96), 2017.

Inserção de mestres e doutores egressos da ENCE nos registros de emprego formal

Marcos Dantas Hecksher*

Resumo

Dos 180 mestres e doutores formados pela Escola Nacional de Ciências Estatísticas (ENCE) de 2010 a 2022, 70% estavam formalmente empregados por pessoas jurídicas no Brasil ao fim de 2021 segundo os microdados identificados da Relação Anual de Informações Sociais (RAIS), do Ministério do Trabalho e Emprego (MTE). Essa proporção é levemente superior aos 64% observados no conjunto dos mestres e doutores com até 64 anos de idade segundo a Pesquisa Nacional por Amostra de Domicílios (PNAD) Contínua do último trimestre de 2021. Sem o objetivo de inferir relações causais, mas sim de descrever a inserção em empregos formais dos egressos 2010-2022 da pós-graduação *stricto sensu* em População, Território e Estatísticas Públicas da ENCE, este capítulo detalha sua distribuição na RAIS 2021, a última disponível à época do levantamento, por vários critérios: período de admissão no emprego, tipo de vínculo, natureza jurídica do empregador, ocupação, setor de atividade, tamanho e local do estabelecimento, jornada semanal e outros. Essa descrição busca identificar a maior parte das atividades laborais em que o conhecimento produzido e disseminado no curso tende a fluir e ser aplicado. A remuneração é desagregada por quintos da distribuição, curso concluído, sexo, cor/raça e faixa etária dos egressos.

* Professor do programa de pós-graduação da ENCE.

Introdução

Até onde pode chegar o conhecimento produzido dentro de uma escola? A incontáveis lugares, por muitos meios existentes e a serem inventados. No caso de um programa de mestrado e doutorado, um dos canais mais evidentes pelos quais o conhecimento se dissemina é a inserção profissional das pessoas que ele forma. Não por acaso, “destino, atuação e avaliação dos egressos do programa em relação à formação recebida” é um dos itens previstos pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), do Ministério da Educação, a serem examinados quadrienalmente em cada curso de pós-graduação *stricto sensu* mantido no Brasil (Portaria Nº 122/2021, art. 24, II, c).

A Pós-Graduação em População, Território e Estatísticas Públicas da Escola Nacional de Ciências Estatísticas (ENCE) buscou reunir informações sobre a inserção profissional dos mestres e doutores que formou em duas bases de dados: 1) uma pesquisa com questionário elaborado por equipe do próprio programa, enviado aos egressos cujo contato foi possível encontrar para preenchimento e devolução via internet; e 2) a base de microdados identificados da Relação Anual de Informações Sociais (RAIS), do Ministério do Trabalho e Emprego (MTE). Este capítulo relata os resultados encontrados nos registros administrativos da RAIS de 2021, a última disponível quando a análise foi realizada. Já os resultados da pesquisa via internet são apresentados em outro capítulo deste livro.

Buscas nos registros da RAIS dos egressos de vários cursos já foram reportadas pela Capes (2018) e por muitos outros pesquisadores e instituições de ensino superior¹, mas esta é a primeira análise concentrada em mestres e doutores formados pela ENCE.

Ressalte-se que este capítulo não apresenta uma avaliação do impacto do curso oferecido pela ENCE – ou do efeito causado por ele – sobre a inserção dos concluintes no mercado de trabalho formal, mas sim estatísticas descritivas sobre a parcela deles encontrada nos registros da RAIS. Assim, os resultados encontrados não devem ser atribuídos diretamente ao curso, nem trazem hipóteses sobre o que ocorreria na sua ausência, mas ajudam a compreender as maneiras mais frequentes de inserção dos seus concluintes em empregos formais em instituições privadas e públicas do Brasil. Informam uma parte importante dos vínculos profissionais pelos quais o conhecimento desenvolvido na ENCE deve fluir e encontrar aplicações.

Método da Análise

A análise descritiva apresentada neste capítulo resulta da integração entre duas bases de dados:

- a) a base de alunos formados no mestrado ou doutorado da ENCE até 2022, mantida nos registros da coordenação do programa; e
- b) a Relação Anual de Informações Sociais (RAIS/MTE) 2021 identificada.

Note-se que são utilizados dados de alunos que concluíram o curso e foram aprovados em suas defesas de dissertação ou tese até 2022, mas a RAIS utilizada se refere a 2021 e, como veremos, inclui postos de trabalho em que parte dos alunos ingressaram antes de entrar no programa ou enquanto passavam por ele. Isso reflete uma opção por utilizar o conjunto mais atualizado de dados disponíveis em cada uma das bases no momento em que o estudo foi realizado, em fevereiro de 2023. Mais uma vez, ressalte-se que o objetivo é descrever quantos concluintes tinham empregos formais em certa data (31/12/2021) e suas características, não inferir quantas ou quais dessas colocações foram obtidas por causa do curso.

O pareamento entre as duas bases de microdados foi feito pelo número do Certificado de Pessoa Física (CPF) de cada concluinte de mestrado ou doutorado na ENCE e confirmado por data de nascimento e nome completo, considerando a possibilidade de erros de digitação e

¹ Ver, por exemplo: Coutinho, 2022; FACE/UFG, 2021; Ferreira e Almeida, 2021; Griboski, Bedritichuk e Ferreira, 2017; Silveira, 2018; UnB, 2023; Vaz, 2020; Vieira, Honorato e Rodrigues, 2022; Vieira, Paul e Barbosa, 2023.

trocas de sobrenome. Para concentrar o centro da análise no último período avaliativo de interesse da Capes, a base integrada foi filtrada para restringir-se aos alunos que completaram seus cursos nos anos mais recentes, de 2010 a 2022.

As análises descritivas apresentam a distribuição dos concluintes, devidamente desidentificados, entre categorias e classes de valores de diversas variáveis disponíveis na RAIS, seja na base sobre os vínculos (com informações de cada emprego formal de cada pessoa) ou na base sobre os estabelecimentos (com detalhes sobre cada pessoa jurídica empregadora).

Resultados

Dos 180 egressos da ENCE que completaram seus cursos de mestrado ou doutorado de 2010 a 2022, 70% (126 pessoas) estavam formalmente empregados por pessoas jurídicas no Brasil em 31/12/2021 segundo os registros administrativos da RAIS mantidos pelo MTE (tabela 1). Essa proporção é levemente superior aos 64% observados no conjunto dos mestres e doutores em qualquer área de formação com até 64 anos de idade (idade máxima, ao fim de 2021, entre os egressos da pós-graduação *stricto sensu* da ENCE) residentes em todo o Brasil segundo a Pesquisa Nacional por Amostra de Domicílios (PNAD) Contínua, do Instituto Brasileiro de Geografia e Estatística (IBGE), referente ao último trimestre de 2021.

Tabela 1 - Presença na RAIS 2021 de mestres e doutores formados na ENCE de 2010 a 2022

	Pessoas	Encontradas em 2021 na RAIS	Empregadas em dez/2021	% encontrado em 2021	% empregado em 31/dez
Mestrado (exclui os 8 que concluíram ambos)	162	112	110	69	68
Doutorado (inclui os 8 que concluíram ambos)	18	16	16	89	89
Todos os egressos 2010-2022 (concluintes)	180	128	126	71	70

Fontes: RAIS/MTE 2021 e Cadastro de mestres e doutores formados pela ENCE/IBGE de 2010 a 2022.

O percentual de 64% da PNAD Contínua citado como base de comparação inclui somente empregados com carteira assinada do setor privado e do setor público, além de servidores estatutários e militares. Afinal, a RAIS só registra vínculos dos empregados formais e, portanto, não esgota o universo de egressos da ENCE com algum trabalho, certamente maior do que 70%. Na PNAD Contínua, por exemplo, o percentual de mestres e doutores com algum trabalho e qualquer posição na ocupação – incluindo também aquelas não registradas na RAIS: empregadores, empregados sem carteira nos setores público e privado, trabalhadores por conta própria, empregados domésticos e trabalhadores familiares auxiliares – chega a 89% no último trimestre de 2021².

A tabela 1 também mostra que o percentual com registro de emprego formal em 31/12/2021 era maior no caso das 18 pessoas que completaram o doutorado (89%) do que entre as 162 que completaram o mestrado (68%) entre 2010 e 2022 na ENCE. Nessa e nas demais análises daqui por diante, as 8 pessoas que concluíram ambos os cursos na ENCE foram computadas no grupo de doutores e não no de mestres formados pelo programa.

Entre as computadas como mestres, duas tiveram vínculos de emprego formal ao longo do ano de 2021 segundo a RAIS, mas não os mantinham em 31/12/2021. Nas demais estatísticas a serem apresentadas, serão consideradas todas as 128 pessoas registradas na RAIS 2021 por

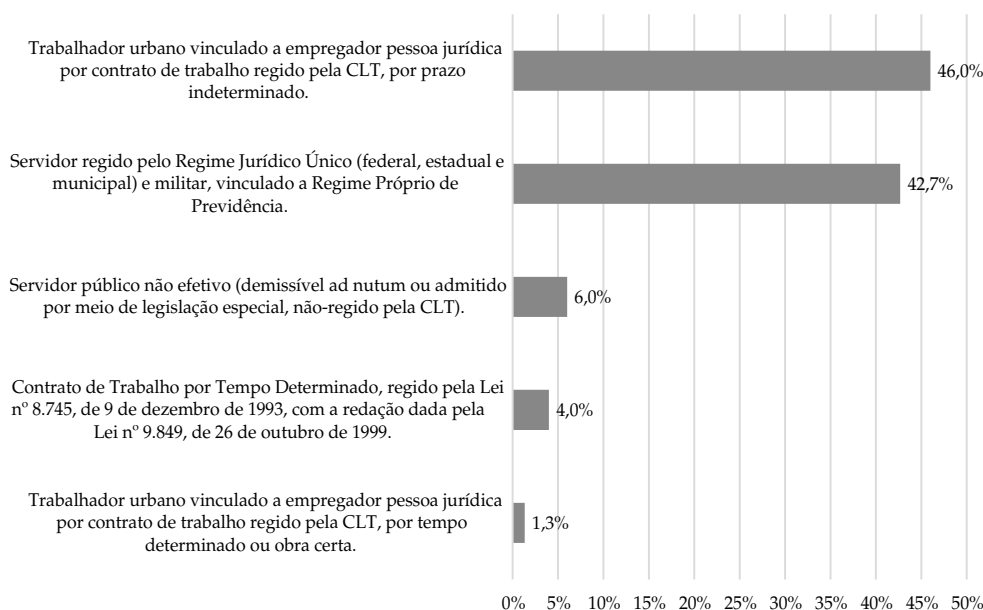
² A mencionada pesquisa com os egressos do programa, descrita em outro capítulo deste livro, encontrou, entre os que foram localizados e responderam o questionário, 66,2% em ocupações formais e 94,6% em alguma ocupação no mercado de trabalho em 31/12/2022, um ano após a data de referência da RAIS analisada no presente capítulo.

vínculos em qualquer período daquele ano. A maioria (83,6%) dessas 128 pessoas teve somente um vínculo registrado em 2021, enquanto as demais tiveram dois (15,6%) ou três (0,8%) vínculos, concomitantes ou subsequentes. Assim, a maior parte das análises a seguir refere-se aos 150 vínculos empregatícios de 128 mestres e doutores da ENCE encontrados na RAIS 2021.

Os anos de admissão nesses 150 vínculos variam de 1991 a 2021, com a maioria (51,3%) concentrada nos cinco últimos anos (2017-2021). A maioria absoluta (59,3%) dos vínculos de 2021 teve data de admissão posterior à conclusão do último curso completo na ENCE. Esse resultado é o que se encontra mesmo incluindo na conta pessoas formadas somente em 2021 ou 2022, casos em que nenhuma admissão foi posterior à formatura – o que seria impossível entre os formandos de 2022 e pouco provável entre os de 2021. Outros 30,0% dos vínculos empregatícios mantidos em 2021 já haviam sido iniciados antes do ingresso da pessoa em seu último curso completo na ENCE e os demais 10,7% começaram durante sua passagem pelo curso.

O gráfico 1 mostra os tipos de vínculo encontrados, sendo de tempo indeterminado em 88,7% dos casos. Entre os 150 vínculos observados em 2021, 46,0% eram contratos de tempo indeterminado regidos pela Consolidação das Leis do Trabalho (CLT), 42,7% eram vínculos de servidores públicos com cargo efetivo ou militares, 6,0% eram de servidores públicos não efetivos em funções comissionadas, 4,0% eram temporários da administração pública e 1,3% eram temporários pela CLT.

Gráfico 1 - Tipos de vínculo na RAIS 2021 de mestres e doutores formados na ENCE de 2010 a 2022 (%)



Fonte: RAIS/MTE 2021. Base: 150 vínculos de 128 mestres e doutores formados na ENCE em 2010-2022.

A maioria (58,7%) dos empregadores eram instituições públicas, mistas ou paraestatais, contra 41,3% privadas. Esses dois percentuais são agregações a partir de 18 naturezas jurídicas detalhadas na RAIS, aqui listadas em ordem decrescente de participação no total: fundação federal (21,3%), sociedade empresária limitada (12,0%), autarquia federal (9,3%), associação privada (9,3%), órgão público do poder executivo estadual ou do Distrito Federal (6,7%), sociedade autônoma fechada (6,7%), órgão público do poder executivo federal (6,0%), órgão público do poder executivo municipal (6,0%), fundação privada (5,3%), sociedade anônima aberta (4,0%), fundação

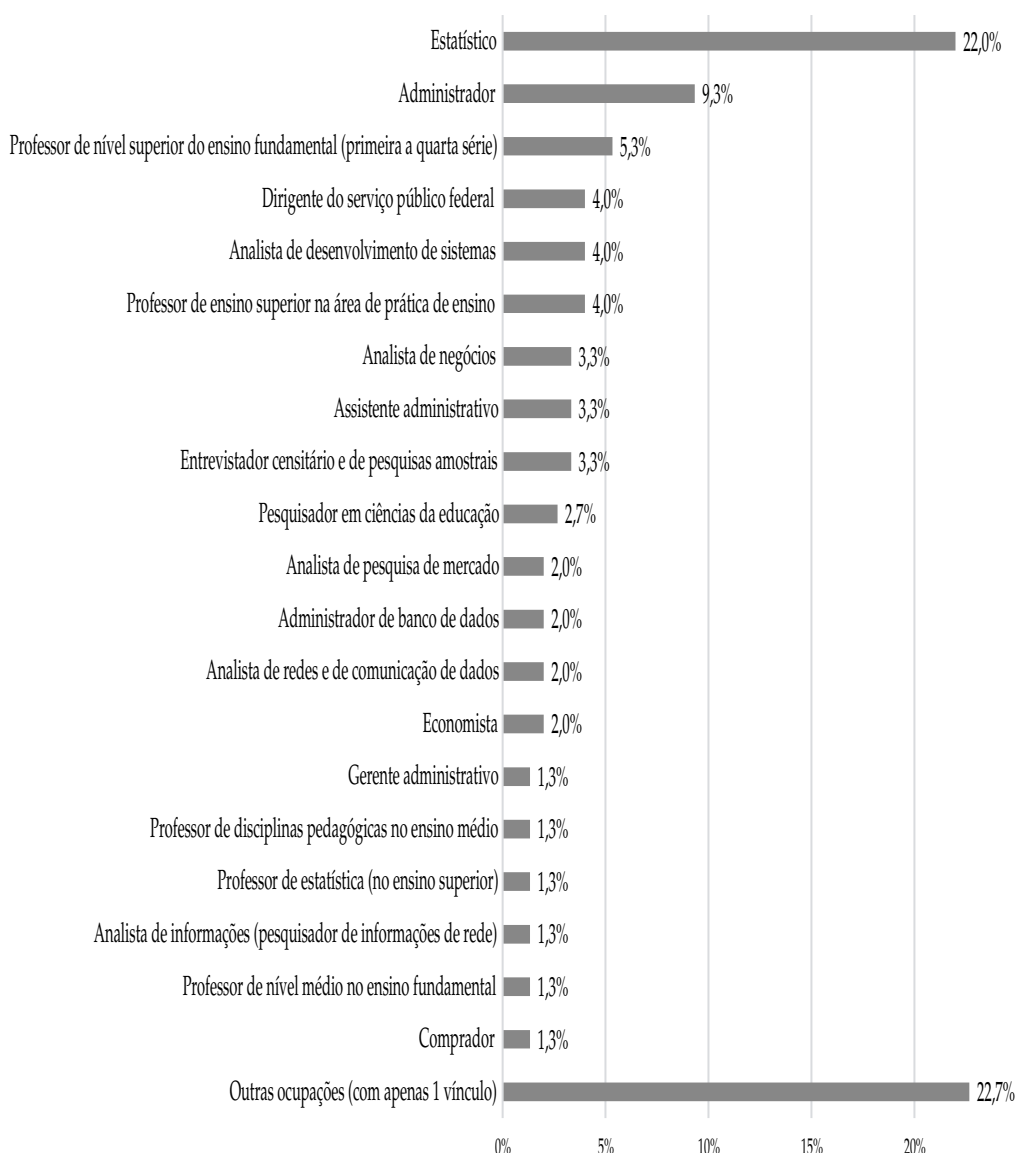
estadual ou do Distrito Federal (2,7%), serviço social autônomo (2,7%), sociedade de economia mista (2,0%), sociedade simples limitada (2,0%), empresa pública (1,3%), entidade sindical (1,3%), autarquia estadual ou do Distrito Federal (0,7%) e cooperativa (0,7%).

Quando se consideram os setores dessas pessoas jurídicas empregadoras no nível mais detalhado da Classificação Nacional de Atividades Econômicas (CNAE 2.3), a sete dígitos, a categoria mais frequente é Administração pública em geral, que responde por 41,3% dos vínculos dos egressos. Em seguida, em ordem decrescente, os maiores percentuais encontrados são: Educação superior – graduação e pós-graduação (6,7%); Educação superior – graduação (3,3%); Atividades associativas não especificadas anteriormente (3,3%); Outras atividades auxiliares dos serviços financeiros não especificadas anteriormente (2,7%); Atividades de apoio à educação, exceto caixas escolares (2,7%); e Atividades de associações de defesa de direitos sociais (2,0%). Os demais setores incluem empresas de energia, telefonia, desenvolvimento de programas de computador, previdência, planos de saúde, consultoria, pesquisas de mercado e de opinião pública, defesa e outras.

A maioria dos vínculos (54,0%) se concentra em estabelecimentos grandes com pelo menos mil vínculos empregatícios. Outros 34,0% ocorrem em estabelecimentos com 100 a 999 vínculos, 10,7% naqueles com 5 a 99 vínculos e apenas 1,3% nos de 1 a 4 vínculos.

O município de localização da ENCE, o Rio de Janeiro, é também onde se registra na RAIS o endereço da maioria (54,7%) dos estabelecimentos que empregam seus mestres e doutores egressos. Ressalte-se que esse endereço-sede de registro nem sempre corresponde ao local de efetiva atuação dos empregados, mas vale reportar os outros mais frequentes: Brasília - DF (12,0%), São Paulo - SP (10,7%), Niterói - RJ (5,3%), Belo Horizonte - MG (2,7%), Cuiabá - MT (2,7%), Duque de Caxias - RJ (1,3%), São José dos Campos - SP (1,3%) e Florianópolis - SC (1,3%). O Estado do Rio de Janeiro (RJ) é a unidade da federação de 64,0% dos estabelecimentos. Entre os estabelecimentos registrados no município do Rio de Janeiro, os bairros mais frequentes dos endereços de registro são: Castelo (30,5%), Centro (15,9%), Botafogo (9,8%), Santo Cristo (8,5%), Cidade Nova (6,1%), Barra da Tijuca (4,9%), Rio Comprido (3,7%), Cidade Universitária (2,4%), Copacabana (2,4%), Laranjeiras (2,4%), Maracanã (2,4%) e Saúde (2,4%).

O gráfico 2 mostra a distribuição dos egressos de 2010-2022 em ocupações registradas na RAIS 2021 segundo a Classificação Brasileira de Ocupações (CBO 2002). Os mais frequentes, que somam 58,7% dos vínculos, são: Estatístico (22,0%); Administrador (9,3%); Professor de nível superior do ensino fundamental – primeira a quarta série (5,3%); Dirigente do serviço público federal (4,0%); Analista de desenvolvimento de sistemas (4,0%); Professor de ensino superior na área de prática de ensino (4,0%); Analista de negócios (3,3%); Assistente administrativo (3,3%); e Entrevistador censitário e de pesquisas amostrais (3,3%). A variedade é bem grande, com os 150 vínculos distribuídos em 54 categorias, das quais 34 registram um único vínculo. O grupo de 34 ocupações com apenas um vínculo responde junto por 22,7% de todos os vínculos, mais do que qualquer ocupação específica isoladamente. Esse resultado é compatível com a versatilidade dos profissionais formados no curso, que tem caráter transdisciplinar e atrai alunos graduados em áreas diversas.

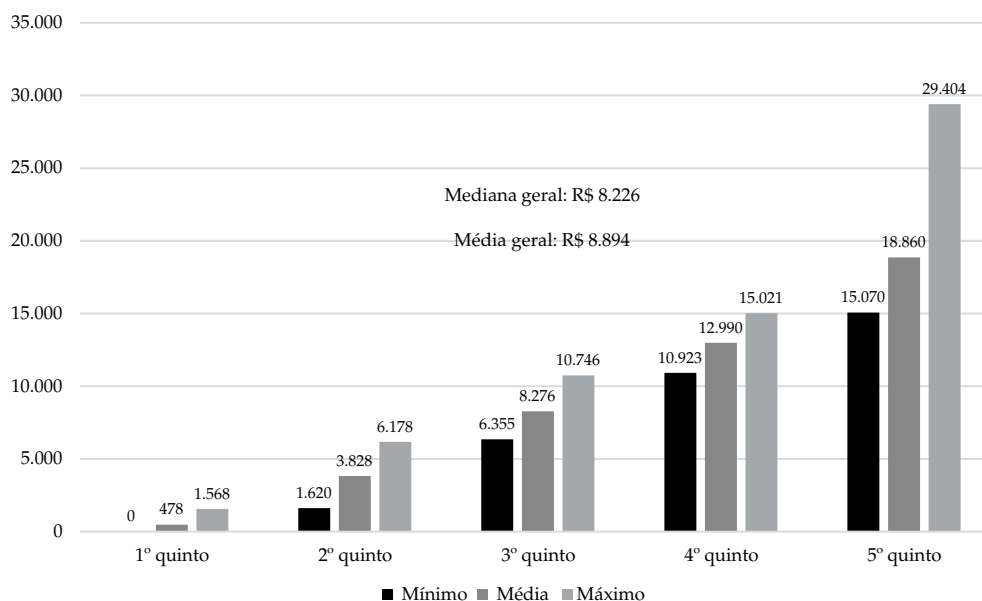
Gráfico 2 - Ocupação na RAIS 2021 de mestres e doutores formados na ENCE de 2010 a 2022 (%)

Fonte: RAIS/MTE 2021. Base: 150 vínculos de 128 mestres e doutores formados na ENCE em 2010-2022.

Aproximadamente dois terços (67,3%) dos vínculos dos egressos na RAIS 2021 têm jornadas contratuais de 40 horas semanais de trabalho. Outros 19,3% têm jornadas de 42 a 44 horas semanais e os 13,3% restantes têm jornadas de até 36 horas semanais.

Considerando a remuneração mensal bruta nominal registrada em todos os vínculos formais de cada pessoa ao longo dos 12 meses de 2021 e tomando como zeros os meses sem vínculo ou remuneração na RAIS, observa-se um valor médio de R\$ 8.894 e uma mediana de R\$ 8.226. A distribuição desse indicador de remuneração entre os 128 mestres e doutores formados de 2010 a 2022 na ENCE é detalhada por quintos no gráfico 3. No quinto de menor remuneração registrada na RAIS, que inclui pessoas com vínculos de emprego formal por períodos curtos em 2021, o duodécimo do valor acumulado no ano foi de até R\$ 1.568, com média de R\$ 478. Na outra ponta da distribuição, os 20% mais bem remunerados ganharam de R\$ 15.070 a R\$ 29.404 por mês, com média de R\$ 18.860.

Gráfico 3 - Remuneração mensal mínima, média e máxima por quinto da distribuição na RAIS 2021 de mestres e doutores formados na ENCE de 2010 a 2022 (%)



Fonte: RAIS/MTE 2021. Base: 150 vínculos de 128 mestres e doutores formados na ENCE em 2010-2022.

A tabela 2 mostra o número de pessoas e os valores mediano e médio da remuneração em vários subgrupos dos egressos, definidos por curso, sexo, cor/raça e faixa etária. A remuneração média é mais alta entre os 16 que fizeram o curso de doutorado na ENCE (R\$ 11.346) do que entre os 112 que concluíram somente mestrado (R\$ 8.544) na instituição.

Tabela 2 - Número de pessoas e remunerações mensal mediana e média na RAIS 2021 por curso, sexo, cor/raça e faixa etária de mestres e doutores formados na ENCE de 2010 a 2022

	Número de pessoas	Remuneração Mediana (R\$)	Remuneração Média (R\$)
Mestrado	112	7,825	8,544
Doutorado	16	13,856	11,346
Homens	66	11,138	10,336
Mulheres	62	6,045	7,359
Preta	3	1,841	1,289
Parda	15	1,361	4,278
Branca	37	6,178	6,765
Não ident.	73	12,133	11,234
25 a 29 anos	12	1,704	3,492
30 a 39 anos	73	6,510	7,402
40 a 49 anos	34	14,126	13,168
50 a 59 anos	7	15,363	12,165
60 a 64 anos	2	11,666	11,666
Total	128	8,226	8,894

Fontes: RAIS/MTE 2021 e Cadastro de mestres e doutores formados pela ENCE/IBGE de 2010 a 2022.

A remuneração média também é mais alta entre os 66 homens (R\$ 10.336) do que entre as 62 mulheres (R\$ 7.536), reproduzindo o padrão da desigualdade observada no conjunto do mercado de trabalho. É verdade que o grupo tem mais homens (10) do que mulheres (6) com doutorado e é paritário entre concluintes de mestrado (56 homens e 56 mulheres), mas as remunerações médias dos homens se mantêm mais altas mesmo dentro de cada nível de formação.

A variável cor/raça só estava disponível na base da RAIS, onde é preenchida pelos empregadores e frequentemente deixa de ser informada. No serviço público do poder executivo federal, por exemplo, o padrão era não registrar essa informação até 2023, quando a autodeclaração passou a ser requerida a todos os servidores. Com isso, a maior parte dos egressos analisados têm cor/raça não identificada na RAIS 2021 e este é o grupo com a maior remuneração média (R\$ 11.234). Entre as pessoas que tiveram informação de cor/raça registrada, as três identificadas como pretas tiveram a menor remuneração média (R\$ 1.289). Outras 15 foram identificadas como pardas, com remuneração média de R\$ 4.278 e, novamente em linha com a desigualdade observada no conjunto do mercado de trabalho, as 37 identificadas como brancas tiveram a maior média (R\$ 6.765) entre aquelas com cor/raça reportada. Mesmo a remuneração desse último grupo ainda fica abaixo da média geral dos egressos, acima da qual figura apenas a média dos que tiveram cor/raça não identificada.

As faixas etárias mostradas na tabela 2 foram definidas pela idade de cada egresso em 31/12/2021. Mais da metade do grupo (73 pessoas) tinha então de 30 a 39 anos de idade. Essa faixa etária modal era seguida em frequência pela de 40 a 49 anos (34 pessoas) e pela de 25 a 29 anos (12 pessoas). Sete pessoas tinham de 50 a 59 anos e duas tinham de 60 a 64 anos. A remuneração média aumenta fortemente da faixa etária mais jovem até a de 40 a 49 anos e passa a decrescer levemente nas duas seguintes. Já a mediana atinge seu pico na faixa de 50 a 59 anos e recua somente na de 60 a 64 anos.

Considerações Finais

Sete em cada dez egressos da pós-graduação *stricto sensu* da ENCE em População, Território e Estatísticas Públicas tinha pelo menos um emprego formal ao fim de 2021 segundo os registros administrativos da RAIS mantidos pelo MTE. A maioria dos vínculos formais do grupo eram de tempo indeterminado, com jornada de 40 horas semanais e se concentravam em instituições públicas, mistas ou paraestatais. As atividades que mais os empregavam eram a administração pública e a educação. A maioria atuava em grandes estabelecimentos localizados na mesma cidade do curso, o Rio de Janeiro, em ocupações variadas, mas com destaque para estatístico, administrador e professor. As remunerações mensais variavam ao redor de uma média de R\$ 8.544, com desigualdades que reproduziam em menor escala as encontradas no conjunto do mercado de trabalho.

Os resultados relatados neste capítulo mostram que a maioria dos mestres e doutores formados na ENCE tem conseguido se inserir formalmente no mercado de trabalho, seja em empregos nos quais ingressaram antes, durante ou depois de passar pelo curso – e, portanto, independentemente de haver impacto do curso na empregabilidade, questão não avaliada aqui. Isso reforça a expectativa de que os conhecimentos desenvolvidos no programa possam e devam ser aplicados em diversas atividades produtivas.

O conhecimento produzido dentro de uma escola pode chegar muito longe, cruzando o tempo e o espaço de forma ilimitada, por mais caminhos do que se pode imaginar. Oportunidades de trabalho constituem um caminho especial para o conhecimento fluir, gerar resultados concretos e espalhar benefícios, para muito além das pessoas que passaram pela escola. Por isso, a inserção dos mestres e doutores formados pela ENCE em empregos formais no Brasil pode ser considerada mais um importante sinalizador, entre vários outros, da capacidade do programa para beneficiar o país.

Referências

CAPES. *Relatório Técnico da DAV. Egressos da Pós-Graduação: Áreas Estratégicas*. Relatório Técnico DAV 1/2017. CAPES/MEC, 2018. Disponível em: <https://www.gov.br/capes/pt-br/centrais-de-conteudo/19122018-cartilha-dav-egressos-pdf>.

COUTINHO, J. P. P. *O efeito da qualidade do ensino superior na inserção e salários futuros no mercado de trabalho formal do Brasil*. Dissertação de Mestrado em Economia, UnB, 2022.

FACE/UFG. *Acompanhamento de Alunos e Egressos*. FACE - Curso de Ciências Econômicas. Disponível em: <https://eco.face.ufg.br/p/38142-acompanhamento-de-alunos-e-egressos>. Atualizada em 12/10/2021.

FERREIRA, F. D. S.; ALMEIDA, A. T. C. *Ação Afirmativa e Ganhos Iniciais Após a Graduação: Uma Análise Para Egressos da UFPB*. XXVI Encontro Regional de Economia, Anpec, 2021.

GRIBOSKI, C. M.; BEDRITICHUK, A. G. A.; FERREIRA, G. V. *Autoavaliação institucional: uma análise da formação e inserção profissional dos egressos da UNB*. Florianópolis, 3o Simpósio Avaliação da Educação Superior (Avalies), 2017.

SILVEIRA, P. B. V. *The economic benefits of higher education: from high school to labor market*. Tese de Doutorado em Economia, Fundação Getulio Vargas, Escola de Economia de São Paulo, 2018.

UnB. *Avaliação UnB - Pesquisa de Egressos*. Disponível em: <https://avaliacao.unb.br/pesquisa-de-egressos>. Acesso em: 14/10/2023.

VAZ, K. C. S. *Trajetória ocupacional e área de formação: uma análise para os concluintes do Enade 2011*. 44º Encontro Anual da Anpocs, 2020.

VIEIRA, A. H. P.; HONORATO, G.; RODRIGUES, L. *Educação superior e resultados no mercado de trabalho no Brasil: uma revisão da literatura e dos dados disponíveis*. Revista Brasileira de Sociologia, vol. 10, núm. 25, pp. 193-218, 2022.

VIEIRA, A. H. P.; PAUL, J-J.; BARBOSA, M. L. O. *The entry of graduates of higher education degrees into teaching*. Revista da Avaliação da Educação Superior, Campinas; Sorocaba, v. 28, e023006, 2023.

Seção 2

Desafios e Limitações na Produção e Uso de Estatísticas Públicas

Produção estatística: velhos e novos desafios

Natália da Silva Castro*

Resumo

Para enfrentar velhos e novos desafios, produtores de estatística vêm buscando formas de adaptar-se às mudanças sociais surgidas desde o fim da segunda guerra mundial. Um exemplo dessa adaptação é uma transição do uso de fontes únicas de dados para o uso de múltiplas fontes, o uso de alguma combinação de dados provenientes de registros administrativos e de pesquisas, assim como de big data. O presente trabalho identifica alguns desses desafios a serem enfrentados no processo de produção de estatísticas públicas, considerando diversas limitações com destaque para o acesso aos dados, financiamento, recursos humanos e legislação. As limitações levantadas constituem algumas reflexões que se configuram, de certo modo, em inquietações, indagações e percepções que têm influência de práticas profissionais relatadas na literatura e observadas nos institutos oficiais. As limitações abordadas estão na pauta de produtores de estatística, em especial institutos de estatística, e constituem verdadeiros desafios na produção de estatísticas públicas. Mostra-se com isso que, seja nos países mais desenvolvidos ou naqueles países em desenvolvimento, o empenho deve ser constante para que estejam continuamente preparados para melhorar seus sistemas estatísticos, desenvolvendo métodos e fontes que respondam mais adequadamente às demandas. Destaca-se que houve evolução dos institutos de estatística para a produção estatística, seja quanto a adequação de métodos, seja na ampliação das fontes utilizadas para produzir suas estatísticas. No entanto, há que se considerar que os institutos vêm enfrentando dificuldades, quer seja por insuficiência nos recursos tecnológicos, orçamentários, humanos ou por legislação nacional que não se adaptaram às necessidades.

Palavras-chave: Informação estatística; Estatísticas públicas; Desafios para produção estatística.

* Doutora pelo programa de pós-graduação da ENCE.

Introdução

Ao término da segunda guerra mundial, os países estavam destruídos. Nos países vencedores e nos países derrotados, era premente fazer algo no intuito de promover um renascimento. A alocação de recursos financeiros, voltados para o planejamento da reconstrução, dependia de estatísticas confiáveis e de boa qualidade, que se tornou uma matéria-prima importante para a tomada de decisões. Por esse motivo, a segunda metade do século XX pode ser considerada a era de ouro das agências produtoras de estatística (Senra, 2001).

Os produtores de estatísticas recebiam, à época, um estímulo nunca visto. As demandas foram muitas e houve condições para a oferta ser à mesma altura. Considerando que a produção de estatísticas dependia de pesquisas, apontadas como caras num contexto de recursos limitados, sejam eles recursos financeiros ou recursos humanos, uma alteração nesse cenário de prosperidade poderia facilmente extenuar os processos de produção estatística (ibidem).

Ao longo dos anos, esse cenário foi se alterando e a produção de estatísticas públicas buscou formas de adaptar-se às mudanças. Um exemplo dessa adaptação é a transição do uso de fonte única de dados para o uso de múltiplas fontes. Alguns Países têm feito combinação de dados provenientes de registros administrativos e de pesquisas com levantamentos tradicionais periódicos. Nos Estados Unidos, por exemplo, registros como os da Receita Federal e da Estadual, referentes a impostos, bem como o cadastro das autorizações para a construção civil, são utilizadas como fontes importantes de dados para estimar a população de áreas pequenas. Mesmo em países em desenvolvimento cujos registros administrativos não tenham tanta qualidade quanto nos países desenvolvidos, pode-se apontar a existência de informações que podem ser aproveitadas para produzir boas estatísticas públicas (Hakkert, 2019).

Nesse sentido é notório que ajustes tenham sido feitos no intuito de superar desafios, porém existem algumas limitações que merecem um lugar de debate e reflexão, como uma maior competição pelos recursos financeiros, além de aumento nas exigências de recursos humanos capazes de atender ao aumento da produção estatística. Também pode-se indicar questões normativas com limitações a serem superadas, visto que, dada a crescente relevância da produção estatística para a sociedade, existe uma demanda específica por leis que visem proteger e fortalecer a produção estatística, bem como das formas de estabelecer melhorias no acesso aos dados.

Dada a complexidade de tratar das limitações que se colocam frente a produção estatística, é importante indicar que este trabalho aborda o assunto de forma que a ordenação dos problemas enfrentados não se refere a uma escala de nível de desafios a serem superados nem de prioridade. Acrescenta-se também que não há pretensão de esgotar tais desafios. Assim, este trabalho foca naqueles desafios que mais de perto falam aos produtores de estatísticas considerando as diversas limitações com destaque para o acesso aos dados, financiamento, recursos humanos e legislação.

A dinâmica de obtenção de dados

O acesso aos dados se configura um desafio que está presente. Para pesquisas tradicionais, geralmente no modo presencial, tem a questão de chegar ao respondente. Fatores como violência, medo e desconfiança dificultam o acesso a esses. Há também a dificuldade de acesso aos povos e comunidades tradicionais, tais como indígenas, quilombolas e ribeirinhos. Isso porque, embora a premência em ter estatísticas sobre esses grupos populacionais, existe todo um protocolo a ser seguido para coletar seus dados em localizações geográficas nem sempre de fácil acesso. No caso do uso do telefone, tem-se barreiras devido ao uso do celular, que com seus identificadores de chamada fazem com que as pessoas não o atendam que a chamada vem de um número desconhecido. Para pesquisas feitas pela internet, embora a tecnologia esteja disponível, ainda não está disponível a toda população. Tratando-se do uso de fontes de dados alternativos para a produção estatística, como big data e registros administrativos, existe também uma série de

questões em aberto, como a cobertura da população de interesse, a qualidade dos dados, sua confidencialidade e a privacidade dos indivíduos. Tudo isso pode ser posto como dificuldade em acessar dados.

A começar por questões de localização e tempo de pesquisa, Czaja e Blair (2003) indicam uma relação inversa entre o tamanho das cidades, já que quanto menor a cidade, maiores as taxas de resposta e vice e versa. Com relação ao tempo de conclusão da pesquisa, indicam que, em caso de pesquisas de abrangência nacional, cobrir todos os fusos horários e diferenças regionais pode apresentar limitações. Assim, a relação espaço e tempo pode afetar a taxa de resposta, que é um indicador de qualidade das pesquisas, levando a imprecisão dos dados.

As pesquisas por telefone também têm problemas que se relacionam à taxa de resposta. Isso porque telefones celulares, que assumiram o lugar dos telefones fixos, têm tecnologia que facilitam a recusa de chamadas pelos usuários, como identificadores e bloqueadores. Esses aspectos tornam mais difícil e demorado o contato com os respondentes. Também não se deve ignorar que o uso do telefone para telemarketing tornou-se tão incômodo que o público em geral ficou mais avesso a ligações telefônicas não solicitadas. Conseqüentemente, o número médio de tentativas de contato telefônicos necessários para concluir uma entrevista aumentou (Czaja e Blair, 2003).

Para os casos de coleta de dados por meio do modo presencial, tem-se como vantagem melhores taxas de respostas quando comparado aos demais, já que é mais difícil recusar o contato e a solicitação de colaboração pessoalmente do que pelo telefone, por exemplo. No entanto, há limitações relacionadas com mudanças estruturais na sociedade. Como, por exemplo, ao tratar de pesquisas domiciliares, onde houve mudanças na composição e no perfil dos moradores que habitam os domicílios, visto que as famílias estão menores e com seus moradores cada vez mais fora de casa. Isso ilustra como é cada vez mais desafiador obter respostas em pesquisa no modo de coleta em questão (ibidem).

Além disso, se constituem como desafio para coleta de dados no modo presencial, questões relacionadas com a segurança pública, pois as pessoas estão cada vez mais receosas de dar acesso a desconhecidos não só aos seus dados, mas principalmente, a seus domicílios e empresas. A tentativa de entrevistas em apartamentos, tipos de residências comuns atualmente, também pode ter limitações devido à dificuldade de acesso aos condomínios. Há indícios também de que quanto mais pessoal é o modo de coleta de dados, os entrevistados tendem a dar respostas não tão precisas a perguntas sensíveis, isso quando não se negam a respondê-las (Czaja e Blair, 2003).

Fellegi (2004) aponta outras questões enfrentadas para produção estatística que também podem afetar o acesso aos dados, as quais se relacionam com confidencialidade dos dados e privacidade dos pesquisados. Assim, o autor indica que há um declínio da confiança social de que os dados sejam mantidos confidenciais, principalmente, ao que tange interferências políticas nas estatísticas divulgadas para o público. E essa confiança parece estar em queda desde os anos imediatos ao pós-guerra. Considera-se que a confiança sempre foi uma característica a qual os sistemas estatísticos são dependentes, principalmente quando se trata de pesquisas bem elaboradas e se espera uma boa garantia de qualidade. Logo, baixa ou nenhuma confiança pode resultar em taxas de erro elevadas.

Quanto à privacidade é indicado que a evolução da tecnologia da informação e suas características estimulou que as preocupações com a privacidade fossem intensificadas. Até mesmo a exploração intensiva dos registros administrativos e sua ligação com as informações da pesquisa são elementos que podem influenciar nas questões de privacidade e, portanto, no acesso aos dados. Assim, as preocupações com a privacidade são um desafio estratégico fundamental para a produção estatística e a atenção para o tema existe desde a chegada dos meios de comunicação eletrônica, isto é, há pelo menos 30 anos (Fellegi, 2004).

No entanto, para o autor, os medos já perceptíveis foram impulsionados nos últimos anos com o avanço tecnológico. Tanto que, nos países desenvolvidos, foram instituídos grupos para tratar sobre proteção de privacidade, pensado em estratégias de melhorar a capacidade de mobi-

lizar a opinião pública a favor dos produtores de estatística. Isso porque a grande preocupação do público com a privacidade pode minar a confiança tanto em participar das pesquisas, quanto em seus resultados. É relevante acrescentar ainda que a forma de enfrentamento desse desafio será decisiva para o fortalecimento e continuidade da produção estatística nos próximos anos

Ainda no que tange os contratempos relacionados à confidencialidade dos dados e que refletem no seu acesso, Braaksma *et al* (2014) acrescentam que a sociedade passou a se preocupar com o fato de seus dados estarem sendo reutilizados sem o seu consentimento. Isso aponta para um desafio central sobre a utilização do big data na produção estatística e pode ter reflexo no acesso à informação de modo geral.

Primeiramente, porque são propriedade privadas. Mesmo os dados abertos, disponíveis para acesso gratuito por todas as pessoas, são de propriedade de alguma instituição que pode, a qualquer tempo, limitar o seu acesso. Logo, o seu compartilhamento pode requerer contratos e envolver custos. Também, têm-se o fato de que, para produzir esse tipo de dado, nem sempre são seguidos os mesmos padrões comumente usados por fontes tradicionais. Como efeito, os diferentes fatores poderão afetar a qualidade da estatística produzida.

Como se pode ver, as limitações de acesso aos dados são decorrentes de diferentes fatores e têm em comum o fato de a sociedade adaptar seu modo de vida ao longo do tempo. Isso pode ser visto, inicialmente, por meio da queda de cooperação dos respondentes e dos desafios enfrentados para se obter acesso a dados de registros administrativos ou big data.

(Des)financiamento de bens públicos

As relações sociais, econômicas e culturais passaram por mudanças em uma escala mundial. Pessoas, mercadorias e informações começaram a circular de forma intensa, onde o aprofundamento internacional no que se refere às diversas dimensões de uma sociedade, especialmente à econômica, teve reflexo também na produção de estatísticas públicas em todo o mundo. Logo, é relevante apontar como essa internacionalização provocou uma remodelação do próprio capitalismo e como as relações de Estado foram se adaptando a essas mudanças.

Chamado de globalização o fenômeno cuja característica são as grandes transformações nos domínios social, econômico e também nas relações entre Estado e capital, se relaciona fortemente com o encurtamento dos recursos enfrentado pelos produtores estatísticos em todo o mundo (Menezes, 2010).

A começar pelo papel do Estado, que se amplia ou se estreita, à medida em que se modificam as condições de reprodução do capital e, conforme o estabelecimento de uma hegemonia do mundo financeiro, provocou a adoção de políticas econômicas de cunho mais liberal. Juntando-se a esse fato, a crise fiscal e financeira que muitos países passaram, chega-se a frequente escassez de recursos necessários à implementação do processo de financiamento nos principais centros de produção de estatística (Carvalho Júnior, 2016).

Nessa perspectiva se aponta a atuação do Estado como principal provedor de bens públicos, sendo esses protagonistas do fortalecimento econômico e social das nações. Bens públicos são aqueles bens ou serviços que podem ser consumidos por diferentes pessoas, sem que esse consumo resulte na falta desse mesmo bem para outro indivíduo. Essa característica é chamada de não-rivalidade. Outra característica que define se o bem é público ou não é a sua não exclusividade, isto é, para ser bem público o acesso ao bem ou serviço não se dá somente para aqueles que pagam ou contribuem de alguma forma para poder acessá-lo. Os benefícios são quase universais em termos de Estado. No entanto, existe um problema relacionado à provisão desses bens, que são os custos e quem seriam os responsáveis por seu financiamento (Menezes, 2010).

É sob tais condições que se gestam o financiamento da produção de estatística pública. Portanto, não é incomum, que os produtores enfrentem limitações no que tange à obtenção de recursos, visto se tratar de um bem que será usado por todos. Esse é um desafio que se amplifica pelas crises financeiras, que se tornaram constantes em um mundo globalizado. Assim, a ques-

tão de quem ofertará os recursos financeiros, frente às políticas de investimento e de alocação de recursos cada vez mais competitivos entre diversos demandantes desses, são necessidades apresentadas aos produtores de estatística.

Senra (2001) reforça o argumento posto ao descrever que houve transição da era de ouro, para uma era de escassez de recursos, sendo esse gatilho para diferentes desafios, principalmente, quando se trata da temática do financiamento da produção estatística. Para o autor, essa diminuição dos investimentos, em razão de crises financeiras, acompanhado pelo aumento na demanda pela produção estatística, é um desafio que se mostra crescente. Segundo seus argumentos, esse desfinanciamento resulta no enfraquecimento do processo de construção da produção estatística em suas várias etapas.

Ressalta-se ainda que, quanto maiores os custos dessa produção, mais difícil se torna seu financiamento. Portanto, é importante reavaliar essas etapas do processo de produção estatística, visto que existem formas de dar continuidade a essa produção, sem incorrer nesses grandes custos. Uma dessas etapas é a coleta de dados, utilizada principalmente na realização de pesquisa com entrevista. Essa prática é a mais frequente ao redor do mundo e também a prática mais dispendiosa, dificultando ainda mais a adoção em cenários que incluem cortes no orçamento.

Outras etapas que cooperam para elevar os custos da coleta de dados, são as campanhas publicitárias e algumas ferramentas, como cartas informativas, folhetos de divulgação e outros materiais com informações de pesquisa, fornecido pelos produtores de estatística para ser usado pela mídia a fim de despertar o interesse do público e incentivar a participação dos entrevistados. Ou, ainda, para que a desconfiança seja menor, sendo, portanto, fatores que, além de contribuir para o aumento da taxa e da precisão das respostas (Fellegi, 2003), ajudam a diminuir custo pela redução do número de visitas ao respondente.

Uma ilustração que mostra a dimensão do encurtamento do financiamento para fins de produção estatística nos últimos anos, são os cortes orçamentários nos institutos de estatística, caracterizando um grande desafio, representado por diversos adiamentos de importantes pesquisas no Brasil. A exemplo disso, tem-se o Censo Demográfico de 1990 que foi adiado em um ano devido à falta de recursos, e a contagem populacional de 2016 que não ocorreu pela mesma razão. Em seguida, o Censo Agropecuário que deveria ter ocorrido em 2015, foi autorizado somente para 2017 (IBGE, 2016). Em 2020, o Censo Demográfico foi adiado pela pandemia de COVID-19, mas já enfrentava incertezas quanto à manutenção do orçamento (IBGE, 2021a), que inicialmente era de 3,4 bilhões de reais e foi readequado para 2,3 bilhões (Cano *et al*, 2021). Em 2021, um novo corte de 90% no orçamento inviabilizou o Censo, até que, em 2022, após intervenção junto ao Congresso e uma emenda ao Plano de Lei Orçamentária Anual, o financiamento foi finalmente assegurado (IBGE, 2021b).

Embora o corte de orçamento tenha criado dificuldades iniciais, o uso de TICs surge como uma solução para otimizar o processo, compensando parcialmente a limitação orçamentária. No entanto, o uso dessas novas tecnologias implica, por sua vez, uma formação contínua de competências específicas na área, o que exige esforços na formação do capital humano, além de investimentos em equipamentos. Assim, para os produtores de estatística serem capazes de satisfazer as limitações dos avanços tecnológicos, seja desde o treinamento de pessoal capacitado para melhor aproveitamento, na renovação do quadro técnico, na criação de um plano de carreira, e até na remuneração adequada, se faz necessário mais investimentos. Os impactos desses investimentos podem, portanto, ser percebidos na capacidade de produção, na interface com os usuários e na produção científica (Carvalho e Figueirôa, 2010).

Tratando-se ainda de infraestrutura tecnológica, essa pode se tornar obsoleta com a falta de investimentos, comprometendo a capacidade de produzir informações na qualidade e quantidade necessárias ou ainda na que deverá produzir no futuro. Questões de recursos financeiros estão intimamente ligadas a várias contingências relacionadas ao processo de produção estatística, como no caso do quadro de pessoal que, conforme é abordado neste trabalho, demanda redimensionamentos diversos, dependendo de recursos financeiros para isso. Dado que a pro-

dução das principais estatísticas públicas se encontra em um patamar de bem público, os usuários desses produtos não têm um benefício econômico direto ou tangível, assim, ainda que o principal financiador seja o Governo, há uma limitação ao fazer o financiamento de tal produção (Leonardo *et al*, 2017).

A conjuntura é ainda menos favorável em relação ao mundo em desenvolvimento, onde os recursos, não são suficientes para atender às suas necessidades. Em países em desenvolvimento, se faz necessária pesquisas de observações ao longo de um certo tempo, as chamadas pesquisas longitudinais, bem como da existência de um sistema de informação maior, no qual os dados de pesquisas domiciliares podem ser pareados com dados de registros administrativos. Tudo isso torna a produção da informação tipicamente mais cara do que para países que já possuem sistemas mais consolidados de produção. Garantir financiamento para esses sistemas é claramente um problema a ser enfrentado. (Fellegi, 2004).

Pode-se concluir que a diminuição ou falta de financiamento da produção de estatísticas, vista como um bem público e capaz de beneficiar a sociedade, tem consequências para produção estatística e representa um desafio a ser enfrentado. Indubitavelmente, as pressões orçamentárias desempenham um papel importante no que tange o desenvolvimento no uso de novas fontes de dados, bem como os registros administrativos e, principalmente, big data. Adicionalmente, nota-se a importância de ter uma perspectiva econômica dessa situação, já que o desfinanciamento da produção estatística corresponde a um risco importante, que não pode ser desprezado, principalmente pela conjuntura atual. A manutenção de capacidade de investimentos que, com as crises fiscais, ficou seriamente prejudicada, pode ocasionar, se algo não for feito com premência, implicações negativas no futuro. Especialmente no que se refere às oportunidades apresentadas pelas TICs.

Capital humano: um recurso valioso

Quando se fala em produção estatística, deve-se ter em mente que são inúmeros os profissionais envolvidos nas mais diferentes atividades do processo como um todo. Uma gama de saberes e técnicas, adquiridos na prática e academicamente, que fazem esses profissionais envolvidos terem capacidades únicas e de extremo valor. A passos lentos, esses profissionais vão se aperfeiçoando, principalmente nos institutos nacionais de estatística (INEs), e se tornaram um importante recurso que pode, direta ou indiretamente, afetar a produção estatística. Ao exercer suas atividades, sejam administrativas ou técnicas, a discussão sobre os recursos humanos merece atenção, sendo esse tema um desafio, já que no fim de todo um processo produtivo, tudo depende deles.

Posto que a democratização do acesso à informação é um dos principais elementos presentes nas sociedades modernas, a produção estatística passa a ter um papel importante na formação do conhecimento e da cidadania. Portanto, faz-se necessário um sistema integrado de instituições orientadas para essa produção estatística, de forma que esses produtos sejam confiáveis e bem estruturados, oferecendo credibilidade para a tomada de decisão dos usuários. Nesse sentido, é fundamental que se tenha quadro de pessoas com formação apropriada para esse tipo de trabalho, de forma que possam contribuir com suas capacidades para a produção da estatística pública (Carvalho e Figueirôa, 2010).

São imprescindíveis quadros de profissionais com formação adequada, amplo conhecimento da realidade e com acesso aos níveis mais micro da informação, para conhecer, em detalhes, as suas potencialidades, capacidades explicativas e limitações. A contratação de bons entrevistadores, por exemplo, é apontada como crucial para o sucesso da coleta de dados (Fellegi, 2003). Acrescenta-se ainda que para o caso em que a coleta de dados ocorre de forma menos tradicional, essa demanda por recursos humanos não é diferente.

Questões referentes aos recursos humanos são apontadas como um ponto a ser melhorado para a produção estatística. O quadro técnico precisa ser renovado de forma a considerar uma convivência entre as gerações que participam dos processos, para que aqueles que estão

começando a trabalhar na área possam absorver a cultura dos órgãos e instituições produtores de estatística. É preciso ir além de um olhar voltado para a modernização dos processos e de conhecimento técnico (Cunha, 2010).

É necessário ter um conhecimento tácito que não se transmite via manual dada a natureza complexa em que está envolta a produção estatística (Cunha, 2010). Tal conhecimento de que trata o autor está relacionado a conceitos, conhecimento dos processos e também dos produtos. Um exemplo da importância do que foi apontado são aquelas estatísticas que exigem continuidade, como são os relacionados às séries temporais, de importância vital para tomada de decisões econômicas e outras mais.

Dificuldades diferentes podem resultar na instabilidade dos produtores de estatística, limitando avanços no que tange essas questões de demanda de quadro de pessoal técnico e especializado. Fatores como falta de capacitação, demora na renovação do quadro, planos de carreira pouco atraentes, remuneração não competitiva, entre outros, portanto, têm impactado a capacidade de produção. Tais problemas acarretam em baixa retenção desses profissionais e, conseqüentemente, uma descontinuidade da produção técnica e científica. Podendo até mesmo reduzir a interface com os usuários, afetando a confiança do público nas produções estatísticas (Cunha, 2010).

Tratando-se de capacitação do quadro técnico de produtores estatísticos, é categórico que o treinamento de todos deve ser cuidadosamente planejado para garantir o desempenho uniforme e que a compreensão dos conceitos da pesquisa seja totalmente transmitida (Fellegi, 2003). Porém, dado um tempo cada vez mais imperativo e orçamentos cada vez mais apertados, se torna um desafio alcançar a eficiência almejada. Adicionalmente, deve-se ter sob perspectiva que o processo de produção estatística vai além do contato com os entrevistados para obtenção de dados. Implica em ter profissionais comprometidos, com disponibilidade para atuar, muitas vezes, em locais de difícil acesso e fazer perguntas sensíveis, o que reforça a importância desse recurso humano.

Ainda a respeito de profissionais que atuam nas pesquisas tradicionais, é importante saber que fica a cargo do entrevistador tarefas que têm grande impacto na qualidade da pesquisa, como a responsabilidade de preparar e agendar entrevistas; garantir que as informações sejam coletadas da unidade correta; obter a cooperação do entrevistado para minimizar a não resposta; fazer as perguntas e registrar as respostas com precisão para evitar erros; e, às vezes, a edição de respostas. Tudo isso respeitando procedimentos de segurança que garantam a confidencialidade dos dados. Muitas vezes essas tarefas são realizadas por pessoas que estão acumulando outras tarefas, tendo sob sua responsabilidade ainda o acompanhamento de falhas de edição e não resposta, a codificação dos dados, o controle de documentação usada e outras tarefas mais, sobrecarregando os já sobrecarregados programas de trabalho (Senra, 2001).

O desenvolvimento das novas TICs são, sem dúvidas, uma vantagem para tornar os recursos humanos mais eficientes em uma tentativa de atender à crescente demanda da produção estatística. Mas a utilização das novas tecnologias implica, por sua vez, uma formação contínua para desenvolvimento de competências tecnológicas que exige, em contrapartida, esforços específicos na formação desse capital humano (Cervera e Martin-Guzman, 2002). Portanto, a pluralidade de atribuições, redesenhadas e reforçadas pelo avanço na demanda por estatísticas públicas, exige novos saberes e novas habilidades por parte do quadro de pessoal, tornando a necessidade pelos recursos humanos um crescente desafio. Deve-se dar uma atenção especial aquelas demandas direcionadas para as necessidades de trabalhos inovadores (Cunha, 2010).

No que se refere à formação acadêmica para atuar em instituições que estejam envolvidas com produção estatística, Pfeffermann (2015) propõe um inevitável questionamento, que diz respeito sobre a oferta de programas de ensino voltados para área. Ele assume que as universidades não estão preparando os alunos, sendo os ocupantes das vagas nas instituições produtoras, profissionais formados nos mais diferentes campos do saber. Logo, esse é um desafio a ser enfrentado, visto que a necessidade de mais e uma melhor formação foi levantada em vários fóruns de discussão sobre o assunto. Os problemas seriam, especificamente, a escassez de cursos sobre

temas relacionado ao trabalho de produção estatística e a falta de pesquisadores especialistas para ensinar.

Os alunos não são suficientemente expostos aos processos e problemas importantes subjacentes ao trabalho de produção estatística. Somado a isso, tem-se o fato de que as estatísticas mudaram muito drasticamente nas últimas décadas, com o que é conhecido como estatística clássica, dando lugar a métodos novos, mais aprimorados e intensivos em tecnologia, como de análise de big data (Pfeffermann, 2015). Algumas universidades têm buscado oferecer programas que desenvolvam habilidades de ensino e pesquisa capazes de ajudar na produção estatística. No entanto, são raros os programas de pós-graduação que sejam voltados à produção de estatísticas oficiais. Vale destacar o programa de pós-graduação População, Território e Estatísticas Públicas da Escola Nacional de Ciências Estatísticas (ENCE) do Instituto Brasileiro de Geografia e Estatística (IBGE), os programas de pós-graduação Economia, Ciência de Dados, Estatística, Finanças e Sociologia da École Nationale de la Statistique et de l'administration économique Paris (ENSAE) e os programas de graduação e pós-graduação voltados para órgãos do setor público do Data Science Campus da Office for National Statistics (ONS) do Reino Unido.

Apesar das opções de capacitação acadêmicas atuais estarem evoluindo, ainda possuem algumas falhas importantes. Primeiro, porque se deve ter muito cuidado ao alinhar as áreas de habilidade com as necessidades práticas da produção estatística. Em segundo lugar, porque é necessário estar preparados para fazer investimentos sustentados e focados em áreas de treinamento prioritários, como os que compreendem o uso de novos métodos e fontes para produção estatística. Deve-se incluir incentivos adequados e estruturas institucionais para todos aqueles envolvidos nos processos de formação (Pfeffermann, 2015).

A omissão no que se refere aos cuidados com os recursos humanos tende a desestimular aqueles que possuem interesse na área, e isso pode vir a afetar todos os processos da produção estatística, principalmente no que se refere à sua modernização. No longo prazo, pode afetar a legitimidade e o reconhecimento, entre outros atores significativos, dos institutos e outros produtores, em seus próprios países e também fora deles (ibidem).

Outro desafio importante, quanto aos riscos associados a pessoas, e que se relacionam com a legitimidade e reconhecimento da produção estatística, se refere à fraude, ou seja, quando, por diferentes meios concebidos pela engenhosidade humana, um indivíduo obtém vantagem sobre outro, por meio de sugestões falsas ou mesmo a omitindo verdades. Os riscos de fraudes dentro de uma organização que produz, direta ou indiretamente, estatísticas é um importante ponto de preocupação, e não somente para produtores de estatística, mas também para os usuários de seus produtos. Questões relacionadas à fraude não devem ser ignoradas, pois podem apresentar impactos negativos para as organizações, devendo ser abordadas, especialmente, sob a ótica da prevenção. Assim, recomenda-se que instituições, públicas e privadas, promovam ações para conscientização sobre os prejuízos das fraudes e treinamento para seu reconhecimento, divulgando seus mecanismos de detecção e fortalecendo um dos seus recursos mais valiosos (Silva *et al.*, 2020).

É evidente que a produção estatística necessita cada vez mais de equipes multidisciplinares, com conhecimentos e habilidades diversas. Além disso, está faltando uma formação continuada desses profissionais que devem ter, principalmente, o domínio de técnicas estatísticas com uso intensivo em TICs, entre outros temas. Vale ressaltar, que existem falhas quanto ao exercício competente no uso dessas técnicas estatísticas mais recentes e que isso põe em risco a continuidade mais eficiente da produção. Essa lacuna indica uma carência de implementação nos processos de aprendizado continuado e na formação de uma rede de parceiros, com a participação de centros de pesquisas e universidades, no âmbito nacional e internacional, que se inter-relacionarão em trabalhos em rede, seminários, cursos e outros mais (Carvalho e Figueiró, 2010).

Pelo exposto identifica-se uma deficiência no fomento ao desenvolvimento das competências profissionais, no sentido de que esses tenham uma formação apropriada, podendo incorporar as novas tecnologias e manter a informação atualizada e de forma contínua, a fim de atender

às expectativas dos usuários no que se refere a produção estatística. São necessidades reais, também apresentadas aos produtores de estatística, a contratação de profissionais que tenham as qualificações técnicas apropriadas. Conclui-se também que o repasse do conhecimento é uma limitação, quando se trata dos recursos humanos para a produção estatística, visto que faltam não somente pesquisadores para atuar na produção, mas também para ensinar e, até mesmo, faltam programas acadêmicos. Faltam planos de carreira atraentes, com remunerações adequadas a fim de não descontinuar a produção estatística, aumentando a quantidade dos desafios existentes.

Legislação para produção estatística

É estratégico que governos e produtores de estatística atuem em conjunto, com objetivo de modernizar a legislação que rege a produção estatística, viabilizando atualizações normativas que se adaptem à nova realidade. São urgentes maiores atuações no campo legislativo, em frentes como cooperação compulsória do respondente, garantia de recursos financeiros e de continuidade da produção estatística. As leis devem atender às necessidades contemporâneas nesse campo, a fim de possibilitar avanços na produção. Na atual conjuntura, onde cada vez mais dados são gerados e são cada vez mais necessários para tomada de decisão, torna-se imperativo discutir sobre como a legislação pode reduzir limitações para a produção estatística.

Deve-se refletir sobre o fato de que uma vez que estamos na era da globalização, em geral, há necessidade de produzir estatísticas comparáveis internacionalmente e isso se torna cada dia mais evidente. Tal necessidade pode ser vista em diversos países que fazem parte de blocos ou mecanismos com cooperação internacional, como a União Europeia, os BRICS e o Mercado Comum do Sul - MERCOSUL. A harmonização de normas administrativas e legais, também deve acompanhar essa tendência de padronização, cada vez mais necessária, frente ao fato de que os registros administrativos estão sendo mais usados como fontes para produção estatística (Cervera e Martiin-Gúzman, 2002). Embora os registros administrativos gerem dados relevantes, se faz necessário ainda uma regulação abrangendo um sistema de informação mais integrado e que englobe de forma equilibrada os diferentes dados de forma mais articulada (MacFeely, 2016).

Existe também, atualmente, uma grande preocupação com a perda de privacidade, o que provocou um movimento na União Europeia no sentido de promover alterações na legislação sobre proteção de dados. Isso ocorreu para reforçar os direitos de proteção de dados dos cidadãos, incluindo, entre outras coisas, o direito de ser esquecido (MacFeely, 2016), sugerindo que existe um desafio para produção estatística, que se relaciona com questões de privacidade. Para contornar essa situação, com a maior eficiência e eficácia possível, a legislação deve ter a possibilidade de promover melhorias adicionais eficazes, com novas disposições legais plenamente estabelecidas, no sentido de garantir a crescente demanda pela produção estatística.

No Brasil, apesar de existir um Plano Nacional de Estatísticas Básicas e um Plano Nacional de Informações Estatísticas, essas não acompanharam a dinâmica da realidade, refletindo em uma falta de dinamismo entre os institutos de estatística com outras entidades e instituições (Schwartzman, 1996). Esse é um ponto importante visto que contornar os desafios depende de um trabalho permanente de busca de consenso entre os produtores de estatística, os detentores dos dados, públicos e privados, e o legislador.

Há grande vulnerabilidade, seja pela ineficiência das políticas de Estado, que acabam não garantindo estabilidade de longo prazo, seja em relação às consistências e firmezas jurídicas que poderiam dar uma base melhor para os produtores de estatística. Esses fatores colocam tais produtores de estatística sujeitos a mudanças políticas e a crises econômicas. Os autores também tecem críticas aos aparatos jurídicos existentes que, segundo eles, não asseguram a obtenção dos diferentes tipos de dados, de forma autônoma para a escolha, a definição e a divulgação da produção estatística. (Carvalho e Figueirôa, 2010).

A regulamentação deve ter como um dos focos principais a privacidade e a proteção do sigilo da informação, pois a disponibilização de dados individuais é uma constante demanda, especialmente de pesquisadores (Pfeffermann, 2015). Logo, há claramente uma necessidade de proteger a privacidade, garantindo a confidencialidade, o que se relaciona com dois aspectos importantes diferentes. Primeiramente, os relativos à proteção de dados digitais, que carece de debate de domínio jurídico. Depois, aquele relativo a garantias de como os dados divulgados podem ser usados. Há um conflito presente entre a demanda dos pesquisadores por mais dados detalhados e a responsabilidade dos produtores de estatística de proteger a confidencialidade dos indivíduos, exercendo pressão sob a produção estatística (Carvalho e Figueirôa, 2010).

Mostrando-se preocupados com tais questões, os produtores de estatística tentam lidar com o desafio de atender a demanda por acesso a dados, ao mesmo tempo em que protege a confidencialidade das informações individuais. Acompanhando essa tendência, o IBGE disponibiliza uma sala para acesso a dados restritos, onde os pesquisadores podem acessar arquivos de microdados, sem colocar em risco a confidencialidade, ou seja, garantindo a privacidade dos informantes. Dessa forma, o IBGE assegura o sigilo das informações individuais, conforme estabelecido na Lei Nº 5.534, de 14.11.68 (IBGE, 2022).

No que compete os registros administrativos, o desafio é garantir seu acesso por parte dos produtores de estatística, bem como garantir a sua qualidade. Em alguns países, como por exemplo Portugal, há previsão legal de que todo registro administrativo deve se submeter ao órgão coordenador do sistema estatístico, a fim de avaliar sua compatibilidade técnica para produção de estatísticas. No entanto, essa prática não está bem estabelecida em todo o mundo. Outro fenômeno que merece um olhar mais atento, também se referindo a fontes de dados, é o aumento da disponibilidade de big data. Isso porque, em geral, esses tipos de dados são de propriedade de organizações privadas que, além de nem sempre observarem o princípio fundamental 6 – Confidencialidade, podem não facilitar o acesso aos dados. No entanto é inevitável que seu uso fará parte da produção estatística, exigindo cooperação entre os detentores desse conjunto de dados e os produtores de estatística, o que deve ser previsto em lei (IBGE, 2018).

A manipulação de dados pessoais, que compõem o grande conjunto de dados disponíveis em fontes tradicionais e alternativas, como as de big data, ganhou espaço com a tecnologia, principalmente da amplificação do uso da internet. Assim, a preocupação com a proteção dos direitos fundamentais de privacidade das pessoas, levou a um movimento que motivou a criação da recente Lei Geral de Proteção de dados - LGPD no Brasil, inspirada pela General Data Protection Regulation - GDPR, em vigor nos países da União Europeia (Sá, 2019).

Embora tenha sido um avanço no âmbito jurídico, a LGPD possui suas limitações. O Artigo 7, por exemplo, que determina que o consentimento dado pelo titular para o tratamento dos dados deve referir-se a finalidades determinadas. Assim, uma empresa só pode tratar os dados dos seus usuários para desenvolver e comercializar especificamente seus produtos, o que não ajuda quando se trata de fomentar a produção estatística (BRASIL, 2018). Existem outras leis brasileiras que têm implicações na produção estatística como o Decreto 10.046, de 9 de outubro de 2019 que dispõe sobre a governança no compartilhamento de dados no âmbito da administração pública federal e institui o Cadastro Base do Cidadão e o Comitê Central de Governança de Dados (BRASIL, 2019).

Ao discorrer sobre os desafios no âmbito das leis fica clara a necessidade de uma revisão da legislação já existente, visando a modernização da produção estatística. Para tanto, são indispensáveis os dispositivos que facilitem o trabalho em conjunto com empresas privadas possuidoras de dados. Ressalta-se também, que um desafio para a legislação é a harmonização de leis que oportunizem o uso de fontes mais disponíveis e acessíveis, como as de big data, que sejam de fácil interpretação e aplicação e se baseiem em contextos sociais, econômicos e políticos de seus próprios países. Enfatiza-se que a lei deve possibilitar a modernização da produção estatística, diminuindo os obstáculos contra a transformação.

Conclusão

Este trabalho identificou desafios enfrentados pelos produtores de estatísticas públicas, considerando a escalada do uso de produtos estatísticos para tomada de decisões, colocando-os no papel de protagonista em um mundo onde tais decisões se dão cada vez mais sob a ótica de dados, que traduzidos em informações, definem as diferentes relações sociais.

Com o passar dos anos, ocorreram mudanças na sociedade, promovendo ainda mais a busca por estatísticas públicas e, também, ampliando os desafios para este tipo de produção. Desse modo, este trabalho permitiu conhecer melhor os desafios que se mostram presentes em diferentes escalas e domínios, como de acesso aos dados, financiamento, recursos humanos e legislação, e que tem sido tema de debate para os produtores de estatística em todo o mundo, principalmente nos institutos de estatística.

No que se refere aos dados, insumos utilizados na produção estatística, existe um desafio duplo que é garantir ao máximo a confidencialidade, ao mesmo tempo em que se preserva a utilidade destes dados divulgados aos usuários. Tal desafio está associado a dois fatos: a premência de obter cada vez mais informações e ao surgimento das novas tecnologias. Assim, frente à intensificação no surgimento das novas fontes de dados, em consequência dos avanços tecnológicos, ampliaram-se as possibilidades de revolucionar a produção estatística, podendo desencadear um movimento no sentido de modernizar essa produção. É relevante apontar que a progressão se dá sobremaneira nos aspectos tecnológicos, com potencial para ser uma boa estratégia de enfrentamentos nas diferentes limitações.

No entanto entende-se que em decorrência do desfinanciamento a modernização dos sistemas estatísticos pode se dar de forma lenta, ficando muitas vezes defasada em relação à tecnologia mais recente, bem como pode fomentar limitações relacionadas aos recursos humanos. Isso porque embora esteja posta a necessidade de envolvimento de profissionais com uma formação capaz de atender às múltiplas demandas para a produção estatística, não há programas acadêmicos suficientes para formar tais profissionais e, principalmente, não há atratividade para aqueles que queiram fazer carreira na área.

Adicionalmente, destaca-se a importância de uma revisão nas legislações estatísticas nacionais no intuito de contornar tais desafios, sendo primordial a ampliação dessas leis para além dos institutos de estatística. Principalmente aquelas que versam sobre a proteção do sigilo das informações, no intuito de oferecer maior segurança quanto à confidencialidade. Dito isso, é importante indicar que esses institutos possuem capacidade para influenciar outros produtores, pois têm produzido boas experiências, como a transição no uso de fontes únicas de dados para o uso de múltiplas fontes, especialmente no que tange ao uso de dados de levantamentos não tradicionais, como big data. No entanto, há que se considerar que os institutos vêm enfrentando dificuldades quer seja por insuficiência nos recursos tecnológicos, orçamentários, humanos ou por leis que não se adaptaram às necessidades.

Entende-se que para superar os desafios, os produtores de estatística precisarão definir meios de trabalhar com diferentes parceiros - públicos, privados, nacionais, internacionais - a fim de aproveitar os benefícios oportunizados pelas TICs, bem como no uso de novas fontes de dados. Os desafios que acompanharão a exploração dessas oportunidades, não são totalmente novos e se inserem nos domínios já discutidos ao longo deste trabalho que se constitui uma contribuição de caráter exploratório, uma vez que foi realizado um estudo acerca dos desafios enfrentados para a produção estatística. Por fim, indica-se que o principal objetivo presente é a geração de novos conhecimentos que sejam úteis para o avanço do debate e se faz relevante face a falta de trabalhos científicos sobre o assunto no Brasil.

Agradecimentos

A autora agradece a Escola Nacional de Ciências Estatísticas – ENCE por ter possibilitado o desenvolvimento do trabalho de dissertação de mestrado que resultou no estudo apresentado, assim como a professora Andréa Diniz da Silva que me orientou de forma respeitosa e compreensiva, dando valiosos conselhos sempre que necessário. Adicionalmente agradece às professoras Denise Britz do Nascimento Silva e Carmem Aparecida do Valle Costa Feijó pelas sugestões que resultaram na melhoria do trabalho feito. Por fim, ao Instituto Brasileiro de Geografia e estatística – IBGE, que proporcionou inúmeras experiências profissionais que estimularam o avanço no conhecimento na área de produção estatística.

Referências

- BRAAKSMA, B.; STRUIJS, P.; DAAS, PJ. *Official statistics and Big Data*. Big Data & Society. 2014. doi:10.1177/2053951714538417. Acesso em: 24 maio 2022.
- BRASIL. *Decreto de Lei nº 10.046, de 9 de outubro de 2019. Dispõe sobre a governança no compartilhamento de dados no âmbito da administração pública federal e institui o Cadastro Base do Cidadão e o Comitê Central de Governança de Dados*. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2019-2022/2019/Decreto/D10046.htm - art34>. Acesso em: 24 maio de 2022.
- CANO, E. C.; Uchôas, A. A.; Gallina, F. *Os desafios do IBGE e a realização do censo demográfico 2022*. Disponível em: <<https://periodicosonline.uems.br/index.php/GEOF/article/view/6553>>. Acesso em: 13 de jun 2022.
- CARVALHO JUNIOR, Cesar Vaz de; FIGUEIRÔA, Edmundo Sá Barreto; GUIMARÃES, José Ribeiro Soares. *Informação, planejamento, cidadania e desafios contemporâneos do sistema estatístico nacional*. Anais, p. 1-17, 2016.
- CARVALHO, César Vaz de.; FIGUEIRÔA, Edmundo Sá Barreto. *As instituições públicas estaduais brasileiros voltadas para a produção e disseminação de estatísticas públicas, estudos e pesquisas e planejamento e os desafios da contemporaneidade*. Boletim de estatísticas públicas ANIPES (Associação Nacional das Instituições Públicas de Planejamento, Pesquisa e estatística). 2010.
- CERVERA, Jose Luis; MARTÍN-GUZMÁN, Pilar. *The Training Of Official Statisticians: Challenges And Experiences*. Instituto Valenciano de Estadística Spain. 2002.
- CUNHA, Adrião. *Análise comparada da legislação estatística do Brasil com as de Alemanha, Argentina, Áustria, Canadá, México e Portugal*. Boletim de estatísticas públicas ANIPES (Associação Nacional das Instituições Públicas de Planejamento, Pesquisa e estatística). 2010.
- CZAJA, Ronald; BLAIR, Johnny J. *Designing Surveys: A Guide To Decisions And Procedures*, 2nd ed. London: Pine Forge Press. 2003.
- FELLEGI, Ivan. P. *Official Statistics—Pressures and Challenges* ISI President’s Invited Lecture, *International Statistical Review*, 72, 1, 139–155, Printed in The Netherlands International Statistical Institute. 2004.
- HAKKERT, Ralph. *Fontes e dados demográficos*. Textos didáticos, p. 1-71, 2019
- IBGE. *Acesso à informação – Sala de acesso a dados restritos*. Disponível em: <<https://www.ibge.gov.br/aceso-informacao/sala-de-aceso-a-dados-restritos.html>>. Acesso em: 19 de maio de 2022.
- IBGE. *Estatísticas (Cada Vez) Melhores Para O Brasil: A Importância De Um Sistema Estatístico Nacional E Por Que É Necessário Revisar A Legislação*. Carta da Presidência IBGE. Proposta preliminar para discussão. 2018.
- IBGE. *Nota de esclarecimento: A Complementação Do Orçamento Do Censo 2022*. Comunicados da Direção. 15 de outubro de 2021b.
- IBGE. *Nota de esclarecimento: Censo Demográfico e PLOA-2022*. Comunicados da Direção. 31 de agosto de 2021a.
- IBGE. *Nota de esclarecimento: Corte no orçamento inviabiliza realização do Censo Agropecuário em 2017*. Comunicados da Direção. 18 de abril de 2016.
- LEONARDO, Fabrizio Clares; CALAIS, Gilson de Oliveira Silva; COPPEDE, Junior Wagner. *Desafios e Oportunidades para a Fundação Seade: Sua Transformação e Adaptação ao Complexo e Dinâmico Ambiente das Estáticas Oficiais*. Dissertação apresentada à Escola de Administração de Empresas de São Paulo, da Fundação Getúlio Vargas, como requisito para obtenção do título de Mestre em Gestão e Políticas Públicas. 2017.

MACFEELY, S. *The continuing evolution of official statistics: some challenges and opportunities*, Journal of Official Statistics, 32(4), pp. 789–810. doi: 10.1515/jos-2016-0041. 2016.

MENEZES, Clarice Cristine Ferreira. *Cooperação internacional e patrimônio mundial*. Revista Brasileira de História & Ciências Sociais Volume 2 - Número 3 - Julho de 2010

PFEFFERMANN, Danny. *Methodological issues and challenges in the production of official statistics*. 24TH ANNUAL MORRIS HANSEN LECTURE. Journal of Survey Statistics and Methodology (2015) 3, 425–483. 2015.

SÁ, Marcelo Dias. *Análise do Impacto da Nova Lei de Proteção de Dados Pessoais nas aplicações de Internet das coisas: Aplicações mobile do governo*. Trabalho de conclusão de curso de especialização em informática gestão de tecnologia da informação. UFMG. 2019.

SCHWARTZMAN, S. O. *O Sistema Estatístico Nacional*. Documento preparado para uma reunião no Ministério do Planejamento. 1996. Disponível em: <http://www.schwartzman.org.br/simon/sistema_est.htm>. Acesso em: 15 mar. 2022.

SENRA, N. C. *Informação estatística e política pública: desafios futuros no presente*. Transinformação, v. 13, n. 1, p. 57-65, 2001. DOI: 10.1590/S0103-37862001000100005 Acesso em: 04 mar. 2022.

SILVA, C. A.; SANTOS, E. A.; RIBEIRO, F.; ROSA, F. S. *Consciência, Prevenção e Detecção contra Fraude: Estudo com Auditores Internos e Contadores do Setor Público Brasileiro*. XX USP Internacional Conference in Accounting. São Paulo, 2020.

Discussão sobre a natureza do conceito de população de difícil acesso

João Gabriel Malaguti*
Letícia de Carvalho Giannella**
Alinne de Carvalho Veiga***

Resumo

A produção de estatísticas públicas enquanto campo de pesquisa não se remete apenas aos métodos, aqui entendidos como o ferramental matemático e estatístico, mas também se remete às discussões e análises relacionadas a questões como o que se deve mensurar, como o fazer e o definir e, também, em quais grupos essa medida dever ser realizada.

Este presente artigo inicia uma discussão sobre o conceito de “população de difícil acesso”, tendo como base a teoria de relações de poder das ciências sociais, principalmente na expressão dessas relações como identidade e território, dois aspectos que vêm a caracterizar diversas populações de interesse para diversos campos, como profissionais do sexo e povos ribeirinhos. Com base neste diálogo entre diferentes ciências, propõe-se uma extensão do conceito visando melhorar sua aplicação.

* Mestre e doutorando pelo programa de pós-graduação da ENCE.

** Professora do programa de pós-graduação da ENCE.

*** Professora do programa de pós-graduação da ENCE .

Introdução

Ao se debruçar sobre a história do campo da Estatística, pode-se considerar que seu *motif*, sua origem, é a população humana. Dedução esta se dá por meio dos censos populacionais que atuaram como primeira afirmação desta ciência. Enquanto que os estudos sobre jogos de azar que inspiraram a probabilidade datam do ano de 1565 (HALD, 2005), e as referências mais antigas à inferência foram feitas pelo árabe al-Kindi no século IX da Era Comum (BROEMELING, 2011), existem referências a censos populacionais nos papiros de Kahun, que datam da época do Império Médio do Egito, aproximadamente entre os anos 2055 e 1650 Antes da Era Comum (SHAW & BLOXAM, 2020; VALBELLE, 1987).

Enquanto ciência do Estado, a Estatística tem interesse em produzir informações sobre a população para que a máquina pública saiba, por exemplo, quem e quanto taxar e quais as necessidades desta sociedade (MARTIN, 2001). Mas o aparato estatístico também é parte de um aparelho ideológico maior, que auxilia o Estado a criar realidades (ALTHUSSER, 1970). Enquanto ciência da metodologia, aqui tendo o sentido de estatística enquanto sua aplicação por outros que não o Estado, tem interesse em entender ou explicar fenômenos sociais. Como posto por Ian Hacking,

A Estatística tem ajudado a determinar a forma das leis sobre a sociedade e o caráter dos fatos sociais. Tem engendrado conceitos e classificações dentro das ciências humanas. Além disso, a coleta de estatísticas tem criado, no mínimo, uma grande máquina burocrática. A Estatística pode pensar nela mesma como apenas fonte de informação, mas é em si parte da tecnologia de poder em um estado moderno. (HACKING, 1991, p. 181)¹

Temos então que os censos populacionais, e a Estatística em si, são instrumentos de controle e de entendimento (EYRE, 2013; FOUCAULT, 2008). Até mesmo a ausência da estatística também é instrumento de poder (FRIED, 2014), visto que a escolha por parte do poder de quais estatísticas calcular e quais delas disseminar moldam tanto o Estado quanto a sociedade, focando ou desfocando diferentes características (SENRA, 2005).

Conceitos básicos para a discussão

A identificação e classificação das populações é parte inerente da estatística, independente de sua aplicação ser voltada para o controle destas ou não. Ao definir conceitos, ao classificar, ao definir população de um modo geral, os pesquisadores realizam uma construção da realidade que merece atenção e não simplesmente uma aceitação. É de interesse deste trabalho, portanto, entender melhor o que é um conceito antes de se propor a analisar o conceito de população, como um todo, e de população de difícil acesso.

O papel do analista é interpretar, mas sempre tendo em mente que ele traz uma carga semântica que pode alterar as respostas (SANTOS, 1982). Um conceito, longe de ser visto como simples abstração, representação, generalização, é também, em si, uma realidade. Ao pretender traduzir, reproduzir ou rerepresentar o real, produzem-se realidades. Toda construção da realidade é um instrumento de poder por si só (RAFFESTIN, 1993). Então, também é papel da ciência devolver conceitos ao conjunto específico que lhes deu o sentido, não apenas usar o “senso comum”. Nas palavras de Milton Santos (2000, p. 104), “É indispensável afinar os conceitos que tornem operacional o nosso enfoque”.

“Tornar operacional” é parte inerente à teoria científica: definições são limitantes por natureza, mas possibilitam um diálogo, criam uma linguagem comum. É natural, no processo de tentar entender o desconhecido, que sejam feitas tentativas de nomenclaturas e taxonomias

¹ Traduzido pelo autor a partir do original: “Statistics has helped determine the form of laws about society and the character of social facts. It has engendered concepts and classifications within the human sciences. Moreover the collection of statistics has created, at the least, a great bureaucratic machinery. It may think of itself as providing only information, but it is itself part of the technology of power in a modern state.”

para a instrumentalização de termos (SENRA, 2005). Ao escolher um foco de pesquisa, estamos realizando um recorte do real e definindo um campo específico de leitura, sob um determinado olhar ao mesmo tempo (HAESBAERT, 2010). No entanto, devemos ser claros que o foco tem seu preço. Algo se ganha e algo se perde, um foco em algo implica em uma falta de foco em outro. Isto não significa que o foco é em si uma falha ou algo ruim, mas sim que o foco é uma escolha do pesquisador e, portanto, deve ser considerado e justificado ou explicado. Cabe a cada pesquisador, diante da problemática que pretende abordar, optar pelo melhor sistema teórico-conceitual, consciente das implicações deste na inter-relação dos conceitos.

Conceitos são instrumentos de pesquisa, mediações (conexões) entre o racional e o real, e, ao mesmo tempo, intervenções capazes de ter efeitos sobre o real (sejam esses efeitos a favor ou contra mudanças) (GALLO, 2008). Existem ao menos cinco questões que precisam ser abordadas para se revelar as propriedades dos conceitos. Estas são, como postuladas por Valter Cruz (2013):

- Relação entre conceito e problema/campo de problematização
- Relação entre conceito e plano de imanência/solo epistemológico
- Relação entre conceito e complexidade/multiplicidade
- Relação entre conceito e lócus de enunciação
- Relação entre conceito e historicidade

Essencialmente, é necessário entender (i) o motivo pelo qual o conceito foi criado, sua *raison d'être*; (ii) com quais teorias/epistemologias o conceito se baseia/relaciona; (iii) quais os múltiplos elementos que formam e definem o conceito e suas relações entre si, além das relações entre conceitos distintos dentro de uma mesma constelação; (iv) a condição social, a situação e a localização geo-histórica dos autores; e (v) a história do próprio conceito, como seu uso e significados se alteraram (ou não) ao longo do tempo (CRUZ, 2013). Tendo então feito esta pequena discussão sobre conceitos em geral, pode-se abordar a discussão sobre a população.

De um ponto de vista puramente estatístico, população (ou universo, em alguns textos) é o conjunto de elementos cujas propriedades se investigam (BOLFARINE & BUSSAB, 2005; COCHRAN, 1977). A população é o agregado de elementos, sendo caracterizada pelos mesmos e definidos em conjunto, tanto os elementos quanto a população (KISH, 1965). Isto significa que não é possível definir estatisticamente a população a não ser a partir da identidade dos elementos que a compõem e, ao definir a identidade dos elementos, também se define uma população.

Essa identidade estatística é composta pelas características de interesse (também chamadas de conteúdo), unidades amostrais (sejam as pessoas diretamente ou conjuntos de pessoas, como turmas de uma escola), localização (ou "extensão geográfica") e temporalidade (KISH, 1965). Um exemplo de população com identidade completa seria "pessoas (unidade) vítimas de violência doméstica (conteúdo) residentes na cidade do Rio de Janeiro (extensão) no ano de 2021 (temporalidade).

Existem vários tipos de população cujas análises e definições interessam para a amostragem, como população alvo, população referida e população amostrada. No entanto, para este trabalho, o principal foco é a população de difícil acesso. A população de difícil acesso pode ser definida como aquela população cuja participação em pesquisas é considerada difícil, seja devido a características dos indivíduos ou pelo desenho da pesquisa em si (STOOP, 2014).

Apesar de alguns autores, como Tourangeau (2014) definirem populações de difícil acesso como um subtipo de *hard-to-survey populations* ("populações de difícil pesquisa", em tradução livre) outros, inclusive na mesma publicação (SMITH, 2014) continuam a utilizar "populações de difícil acesso" como o termo mais abrangente, assim como será feito aqui.

Tourangeau (2014) inclui como outros subgrupos: *hard-to-sample* (de difícil amostragem, para os quais inexitem listagens próprias e atualizadas), *hard-to-identify* (de difícil identificação, para os quais os indivíduos, motivados por estigma ou marginalização, não identificam a si ou a outros como membros da população), *hard-to-reach* (de difícil alcance, para os quais os indivíduos são difíceis de localizar ou contactar), *hard-to-persuade* (de difícil persuasão, aqueles que, quando contactados se recusam a participar) e *hard-to-interview* (de difícil entrevista, para

os quais existem obstáculos físicos, cognitivos ou linguísticos para a realização da entrevista). Estes subgrupos não são necessariamente conjuntos disjuntos. Por exemplo, pessoas surdas são ao mesmo tempo de difícil amostragem e de difícil entrevista, se considerarmos uma entrevista oral.

Populações de difícil acesso são, em grande parte, populações cujo tamanho é relativamente pequeno quando comparada a população geral (chamadas também de populações raras) e “escondidas socialmente”, para usar o termo de Smith (2014) (chamadas populações ocultas). Entre essas populações ocultas, Smith destaca aquelas cujos membros praticam alguma atividade ilegal, como profissionais do sexo, imigrantes não documentados e usuários de drogas; aqueles cujos membros preferem não declarar seu status, como vítimas de violência sexual ou doméstica; populações marginalizadas, como homossexuais e alcoólatras; e os ricos.

Importante notar que os exemplos apresentados por Smith são dependentes da sociedade (estadunidense) na qual ele se encontrava. Prostituição no Brasil, por exemplo, não é criminalizada e até se encontra na Classificação Brasileira de Ocupação publicada pelo Ministério do Trabalho (RODRIGUES, 2009). Mesmo assim, profissionais do sexo ainda são populações de difícil acesso, pela marginalização que sofrem. Outra nota de interesse é que Stoop (2014) considera o grupo dos ricos como exceção e caso particular.

Em sua revisão, Smith (2014) então enumera doze grupos de difícil acesso (Quadro 1), os quais considera sujeitos a fatores distintos e apresentados de maneira independente, embora Smith aquiesca, ao fim, ao fato de que existe considerável sobreposição entre os grupos.

Quadro 1 – Grupos de difícil acesso de acordo com Smith (2014)

#	Populações
01	Socialmente ou geograficamente isoladas
02	Em locais isolados ou remotos
03	Móveis
04	Pobres
05	Ricas
06	Criminosas
07	Não respondentes
08	Ocupadas
09	Estigmatizadas e marginalizadas
10	Extremistas políticos
11	De difícil entrevista
12	Minorias linguísticas

Fonte: Elaborado pelo autor, com base em Smith, 2014.

Mesmo sem entrar em questões mais sociais, é possível notar, por exemplo, que existe uma sobreposição entre as populações socialmente ou geograficamente isoladas e aquelas em locais isolados ou remotos, visto que grupos em isolamento geográfico muitas vezes se encontram em locais remotos.

Se, no entanto, focarmos em questões de marginalização e estigma, pode-se perceber que estas não estão limitadas à nona linha do Quadro 1, mas permeiam o mesmo. As ligações das populações marginalizadas com populações socialmente ou geograficamente isoladas, populações pobres, populações criminosas e minorias linguísticas são bastante evidentes. São afetadas também as populações em locais isolados ou remotos, que, por muitas vezes, lidam com o estigma geográfico, e as populações móveis, várias das quais também são marginalizadas. Alguns exemplos são os *Romani* da Europa continental, os *Travellers* da Irlanda e Reino Unido, imigrantes não documentados, refugiados, deslocados e pessoas em situação de rua. Ademais, pode-se considerar as intersecções destas com as outras populações, como por exemplo a

sobreposição das populações pobres com as populações trabalhadoras ou ocupadas, o que expandiria ainda mais a influência de tais questões.

Com algumas das ferramentas de Valter Cruz apresentadas anteriormente, pode-se analisar com mais cuidado essas construções que procuram definir e detalhar o conceito de população de difícil acesso.

A problemática que gerou o conceito pode ser vista como puramente metodológica, uma maneira de rotular grupos de indivíduos que as pesquisas têm problemas em alcançar. Uma visão não contemplada por esta lente é que, se tratando de populações humanas, questões sociais são prováveis (ou até certas) de aparecerem, trazendo consigo todas as suas complicações.

Como muitos conceitos existentes na estatística — nas ciências como um todo, para ser mais preciso — o conceito de população de difícil acesso foi construído a partir do Norte Global e não necessariamente será aderente às realidades dos países do Sul Global.² Pode-se notar isso a partir do exemplo dado por Smith (2014) mencionado anteriormente, de profissionais do sexo como populações criminosas, o que pode ser verdade em certos países, como os Estados Unidos, mas não é parte da realidade brasileira.

Pode-se também analisar o lócus de enunciação em duas escalas diferentes: um nível geo-histórico, e outro (menor) da pesquisa em si. O geo-histórico porque diferentes locais e momentos terão diferentes relações com suas respectivas populações. Por exemplo, a população de Palestinos quando observada no mesmo território antes e depois da Guerra da Palestina e da criação do estado de Israel é bastante diferente, indo de cidadãos a refugiados no que era sua própria nação (ARURI, 2001); sem tocar no como as concepções sobre a população palestina mudam de acordo com quem enuncia o conceito. Quando é enunciado por israelenses, a definição da população será profundamente distinta de como quando é enunciado por palestinos. Analogamente, fixando-se o período e observando territórios vizinhos, temos o exemplo da população de estudantes negros nos mesmos anos de 1958-1959 nos estados vizinhos da Virgínia e Virgínia Ocidental, onde temos grandes diferenças. Enquanto que a Virgínia Ocidental acatou os resultados do processo *Brown v. Board of Education of Topeka* e dessegregou suas escolas, a Virgínia preferiu fechar as escolas, causando um êxodo no qual as famílias que podiam, enviavam seus filhos para morar com parentes ou amigos em outros estados (SHETTERLY, 2016). Neste caso, temos uma mesma população, em um mesmo momento histórico tendo relações diferentes para com seu território.

Analisar o lócus ao nível da pesquisa em si, aqui, diz respeito ao desenho da pesquisa e como este se relaciona com a população na qual será aplicado. Se o desenho da pesquisa contempla um questionário escrito e um falado, tanto pessoas cegas quanto pessoas surdas são “de fácil acesso”, por assim dizer, mas se o desenho possui apenas um tipo de questionário, um dos grupos seria de difícil acesso, por exemplo. Uma pesquisa com orçamento limitado terá mais dificuldade em alcançar populações mais distantes geograficamente, se depender de entrevistas face-a-face; no entanto, se utilizar de tecnologias como telefones ou internet a distância tem menor importância do que o acesso a tais tecnologias.

O termo “população de difícil acesso”, então, é incapaz de diferenciar entre uma população que é de difícil acesso por causa do desenho (e financiamento) da pesquisa e uma população que é de difícil acesso por sua própria natureza. O mesmo conceito pode ser enunciado em contextos diferentes e não se referir ao mesmo tipo de população.

Como dito anteriormente, o foco, longe de ser uma falha, é uma escolha do pesquisador e, portanto, precisa ser explicitado. O pesquisador precisa estar atento ao que esta lente lhe permite e o que passa a ficar fora de foco, por assim dizer.

Uma questão no cerne das populações de difícil acesso (e, também, das populações humanas de modo geral) é a identidade. Não só o processo de identificação é inerente a pesquisas tanto na definição da população de interesse quanto na determinação de indivíduos

² Ver, por exemplo, o fato do livro-texto sobre populações de difícil acesso, *Hard-to-Survey Populations*, ter apenas 2 autores contribuindo do sul global, de um total de 76 autores (TOURANGEAU ET AL., 2014).

que pertencem a tal população, o processo também permite a criação de relações entre pessoas que compartilham alguma faceta de identidade (BERGER & LUCKMANN, 1966).

A partir da identidade, outros tópicos - outros focos - podem ser derivados. Os mesmos rótulos que permitem a identificação são aqueles que fazem parte de questões sobre estigma e marginalização (BECKER, 2018). Estas lentes obscurecem algumas das classes apresentadas por Smith, como as populações ricas e as de não respondentes, mas ao mesmo tempo permite outras análises, focando nas populações mais vulneráveis. Acredita-se que entender estes aspectos da sociedade e seus efeitos sobre as pesquisas é crucial para que se possa discutir como melhorar a metodologia aplicada nestas. Toda investigação da sociedade será afetada pela mesma, desde quais perguntas são o motivo da pesquisa até o processo de ir a campo.

Esquecer ou ignorar a vulnerabilidade social das populações marginalizadas e estigmatizadas mesmo quando se discute metodologia pura é perigoso. Precisa-se tomar extremo cuidado quando se discute, por exemplo, como identificar e localizar tais grupos dado que essa informação pode ser utilizada contra estes.

Em 1933, uma meta central da Alemanha Nazista era a identificação e subsequente destruição de sua comunidade de judeus. Graças aos esforços da empresa de tabulação de censos IBM, por meio de sua subsidiária alemã *Deutsche Hollerith Maschinen Gesellschaft* ou *Dehomag*, os nazistas obtiveram as listas necessárias para o Holocausto a partir do censo de 1933 (BLACK, 2012; SELTZER, 1998).

Menos de uma década depois, em 19 de fevereiro de 1942, presidente americano Franklin Delano Roosevelt assina a Ordem Executiva 9066, que permite ao secretário de guerra designar “áreas militares”, preparando para o internamento de cerca de 100.000 japoneses (e seus descendentes americanos) em campos de concentração (ROOSEVELT, 1942). Enquanto que a agência responsável pelo censo americano defendeu por anos que apenas fornecera aos militares as tabulações do censo de 1940 à nível de setores censitários e quarteirões (US CENSUS BUREAU, 2005; SELTZER & ANDERSON, 2000), estudos mais recentes por Seltzer e Anderson (2007) revelaram que microdados também foram fornecidos.

Relações entre os conceitos de identidade e de território

O segundo volume do livro *O Segundo Sexo* de Simone de Beauvoir (*Le Deuxième Sexe*, no original) abre com a frase “Ninguém nasce mulher: torna-se mulher.” (BEAUVOIR, 1949, p. 11). A uma primeira vista é uma frase simples, mas guarda consigo a teoria de construção da identidade. A identidade é o resultado de inúmeras relações de poder (COLLINS & BILGE, 2021) que agregam significados às características dos indivíduos, ou, em outras palavras, construindo socialmente a identidade (BERGER & LUCKMANN, 1966).

É claro, isto não significa que por ser um construto social a identidade não tem efeitos reais sobre os indivíduos e a sociedade em geral. Por exemplo, a cor de pele negra é uma característica biológica, mas a ideia (errada) que os negros têm maior resistência a dor (física ou emocional) é um estereótipo social sem base na realidade biológica e com claros efeitos tanto históricos (o uso de negros em cirurgias experimentais sem anestesia) quanto atuais (nos EUA negros são menos prováveis de serem receitados analgésicos do que brancos sofrendo das mesmas enfermidades) (WASHINGTON, 2006).

A identificação de pessoas como membros de uma certa população se dá em parte por compartilhar características, mas também por sua ausência (FREYRE, 2009). A população *queer*, por exemplo, inclui todos aqueles que não são cis e heterossexuais, independente de suas sexualidade e gêneros: uma pessoa trans e heterossexual é considerada *queer*, assim como uma pessoa cis e assexual.

Os rótulos de identidade podem ser aplicados a um indivíduo por ele mesmo, o chamado autorreconhecimento, ou por outros, como é o caso da identificação por parte de pesquisadores. Estes rótulos podem ou não ter a mesma definição para ambos os lados da pesquisa

(pesquisadores e pesquisados) e precisa-se tomar cuidado para que a população pesquisada seja realmente a população de interesse. No início da crise de AIDS dos anos 80 nos Estados Unidos, a população de homens gays era considerada uma população de risco mas esta tanto incluía homens gays sexualmente inativos quanto excluía outros homens que faziam sexo com homens mas não se consideravam gays. A metodologia e a terminologia tiveram de considerar estes fatores para melhorar as pesquisas sobre a epidemia de AIDS em seus variados aspectos. Compare, por exemplo, o uso de “homem homossexual” (*male homosexual*) em Gupta et al. (1989) com o uso de “homens que fazem sexo com homens” (*men who have sex with men*) em Rispel e Metcalf (2009) e Cloete, Kalichman e Simbayi (2013).

Os rótulos em si carregam traços de status principais e auxiliares (HUGHES, 1945), com o principal sendo, de maneira simples, a definição dos mesmos; e os traços auxiliares sendo características informalmente esperadas pela sociedade na qual o indivíduo rotulado se encontra. O exemplo apresentado por Hughes e ecoado por Becker (2018) é que, na sociedade norte-americana, presume-se que um indivíduo rotulado como “médico” seja de classe média alta, branco, do sexo masculino e protestante, quando seu traço principal define apenas que a pessoa cumpriu certos requisitos e se encontra licenciada a praticar medicina. A partir destes traços auxiliares que indivíduos são vistos como desviantes (*deviant*) e consequentemente estigmatizados e marginalizados (BECKER, 2018).

O estigma é um dos principais motivos pelo qual certos grupos são relutantes em se identificar para aqueles que não fazem parte do grupo (no caso, os pesquisadores) (MCCORMICK & ZHENG, 2014). Existem riscos que indivíduos correm caso sejam identificados publicamente como membros de uma população estigmatizada (MAY-CHAHAL & CAWSON, 2005), como por exemplo:

- Serem sujeitados a discriminação, desprezo, vergonha ou ridículo em suas interações com outros (AMSTADTER *et al.*, 2008);
- Sofrerem danos financeiros ou danos às reputações (KANOUSE *et al.*, 1991);
- Serem sujeitados à sanções legais, como ações administrativas ou mesmo penas criminais (ROSS *et al.*, 2007);
- Serem sujeitados à perseguição, ameaças ou represálias como resultado da participação na pesquisa (BERRY & GUNN, 2014).

A construção e, de maneira análoga, a desconstrução dos significados atrelados aos rótulos ocorre por meio das relações de poder dentro de uma sociedade e pode ser de maneira passiva (mudanças na cultura e tradições sem um ator específico) ou ativa (muitas vezes, propaganda). Isto se dá tanto com a rotulação de pessoas quanto com a rotulação de lugares (SLATER, 2017; O'DONNELL, 2013).

O território, que será discutido em mais detalhes ainda nesta seção, merece um maior destaque por ser, posto de forma simples, as relações de poder expressas no espaço (SOUZA, 1995). O território é ao mesmo tempo subordinado e condicionante (CORRÊA, 1992), isto é, ao mesmo tempo que as pessoas definem os seus vários territórios, os territórios também ajudam a definir as pessoas. De fato, o espaço é tão interligado com essas discussões que na França se utiliza o termo *etiquetage* (etiquetar, rotular) para indicar a estigmatização das pessoas por seus locais de moradia (SECCHI, 2013).

Importante notar que para um rótulo ter efeitos de estigma ele não precisa ser verdadeiro, apenas percebido como tal (BECKER, 2018). Um exemplo extremo, embora emblemático, é o caso das bruxas de Salem. Entre 1692 e 1693, mais de 200 pessoas foram acusadas de bruxaria e rotuladas como bruxos e bruxas na Colônia da Baía de Massachusetts. Destas, 30 foram julgadas culpadas e 19 foram enforcadas (SCHIFF, 2015).

A perseguição àqueles rotulados permaneceu no imaginário norte-americano por séculos e permitiu que o dramaturgo Arthur Miller reconhecesse algo similar na sociedade de sua época (MILLER & WELLAND, 1992): pessoas taxadas por aqueles no poder como inimigos da moral, poderosos o bastante para infiltrar uma sociedade correta com o intuito de corromper

as “pessoas de bem” e compactuando com o mal. No século XVII eram chamadas de bruxas, mas durante a Guerra Fria o rótulo tinha outro nome: Comunistas.

A peça de Miller, *As Bruxas de Salem (The Crucible, no original)*, é uma alegoria ao Macarthismo, quando os EUA perseguia aqueles que rotulava como comunistas. Atores, escritores, músicos, diretores e outros artistas foram barrados de exercer suas atividades em Hollywood e outros, como o engenheiro Hsue-Shen Tsien, foram deportados, sendo sua filiação aos comunistas real ou não (CHANG, 1995).

As várias identidades de uma pessoa não são estáticas e interagem entre si, alterando as relações entre essa pessoa e a sociedade na qual está incluída. A abordagem metodológica que contempla a ideia que os efeitos das identidades não são independentes é a chamada interseccionalidade (COLLINS & BILGE, 2021).

Evitando as chamadas abordagens aditivas, os estudos interseccionais buscam entender como as várias desigualdades se moldam mutuamente, como os efeitos (multiplicativos) das opressões interagem com a sociedade (KING, 1988). É prudente, então, não assumir a independência de fatores sociais quando existe um complexo arcabouço teórico que argumenta contra esta visão.

Uma grande parte dos estudos interseccionais foca no estudo do entrelace entre raça, gênero e classe social, mas alguns estudos também incorporam outros aspectos, como o geográfico (COLLINS & BILGE, 2021). Os povos tradicionais, por exemplo, são definidos, entre outros fatores, por suas relações com seus respectivos territórios (HAESBAERT, 2020).

Mais do que um solo inerte sobre o qual as pessoas vivem, o espaço geográfico afeta e é afetado por aqueles que o ocupam, nas mais variadas escalas. O país em que uma pessoa cresce, por exemplo, tende a influenciar quais línguas a pessoa aprende, e estas por sua vez, ditam sua exposição a certos textos e ideias. Em uma escala mais reduzida, o bairro em que uma pessoa mora pode influenciar a maneira que outras pessoas a tratam. É de interesse, então, também discutir questões sócio-geográficas e, para isso, precisa-se antes definir alguns conceitos.

Muitos ainda consideram região, território e espaço como sinônimos, mas longe de serem determinações monolíticas e apolíticas, o espaço apropriado e (geo)grafado historicamente deve ser questionado, ao invés de ser ignorado (PORTO-GONÇALVES, 2006). Impressas nas concepções de espaço utilizado estão as relações sociais que tanto interessam à estatística e às ciências humanas.

Pode-se iniciar uma abordagem aos conceitos geográficos, dos quais o principal para este trabalho é o conceito de território. Este encontra-se obrigatoriamente mergulhado na relação com outros conceitos (como espaço ou região) frente ou com os quais ele pode ser mais bem definido, na chamada “constelação de conceitos” (HAESBAERT, 2010).

A categoria-chave da constelação, da qual os outros dois conceitos a serem tratados surgem, é o espaço. Seguindo a visão de Henri Lefebvre, o espaço e sua produção (material ou cultural, por exemplo) podem ser estudados a partir da tríade concebido/percebido/vivido (ALVES, 2019).

Espaço concebido é a dimensão espacial ligada às relações de produção, ao conhecimento, ao planejamento, à ordem instituída. Trata-se do espaço planejado, das normas técnicas, que se apresenta como neutro e apolítico apesar de não o ser. Por exemplo, o Estado, enquanto sujeito de dominação política, intervém no espaço através de suas instituições (BOVOLenta, 2018). A dimensão do vivido ou espaço de representação está ligada ao lado subterrâneo e sutil da vida social e também à arte, vivido por meio das imagens e dos símbolos que o acompanham. O espaço percebido é relativo à prática social e expressa a relação entre a realidade cotidiana e a urbana. O espaço percebido também pode ser entendido como resultado do embate entre o vivido e o concebido. Esses elementos são indissociáveis, mas muitas vezes uma das dimensões se projeta acima das outras, ocultando-as.

Para a academia e os estudos em geral, o espaço concebido se sobrepõe aos outros, com o espaço muitas vezes sendo tratado como palco sobre o qual fenômenos acontecem sem influenciar o espaço ou serem influenciados por este; e outras vezes tratado apenas como covariável em um

estudo maior. No entanto, ao ir a campo, o espaço encontrado pelos pesquisadores é o percebido, mediando entre o espaço vivido da população de interesse e o espaço concebido do desenho da pesquisa.

A região não é o foco deste trabalho, mas será definida com o intuito de evitar confusão entre esta e o conceito de território. Comumente, região é tratada como porção do espaço delimitada por algum critério ou dotada de alguma característica própria (HAESBAERT, 2019). Regiões podem ser constituídas por um único ou apenas um grupo de propriedades (regiões simples ou elementares) ou reunindo num mesmo conjunto todas as propriedades ou dimensões (região complexa ou síntese) (HAESBAERT, 2010).

Do espaço, nasce o território; este é fundamentalmente um espaço definido e delimitado por e a partir de relações de poder (SOUZA, 1995), mas não necessariamente o poder político estatal (CATAIA, 2011; MOYSÉS, 2011). O território se constrói assim, entre a dominação hegemônica, isto é, por aqueles no poder, e a apropriação e resistência subalterna, por aqueles excluídos ou marginalizados. Essa ideia de que outras fontes de poder além do Estado constroem territórios difere do conceito anglo-saxão de território, sendo específico da perspectiva latino-americana (HAESBAERT, 2020).

A informação e a energia são aspectos constituintes do poder e cada sociedade organiza seu espaço combinando esses elementos e, desta maneira, modela seu território (SAQUET, 2007). Essa modelagem manifesta e condiciona o exercício do poder (RIBEIRO, 2005). Apesar de o principal ser o poder, isto não quer dizer que a cultura e mesmo a economia não sejam relevantes ou não estejam “contemplados” ao se lidar com esse foco (SOUZA, 2009).

Territórios existem e são construídos e desconstruídos nas mais variadas escalas espaciais e temporais, desde ruas até países, desde dias até séculos. Podem ter um caráter permanente ou uma existência periódica, cíclica (BARROS et al., 2017), como é o caso dos territórios de profissionais do sexo, que se refazem às noites e são desfeitos com a chegada do dia (SOUZA, 1995).

A discussão sobre territórios é de particular interesse para a conversa sobre populações de difícil acesso porque o território é um campo de forças, uma rede de relações sociais, definindo o limite não só geográfico, mas também entre o grupo e os outros (SOUZA, 1995).

As populações criam seus próprios territórios, podendo estes variar em extensão, estabilidade, contiguidade e até temporalidade. Os territórios de favela são contíguos e estáveis temporalmente, podendo durar décadas em um mesmo espaço. Já os territórios de profissionais do sexo, por exemplo, são fluidos, com seus limites instáveis e em constante mudança, e sua temporalidade periódica, em geral existindo apenas às noites (SOUZA, 1995). O território do tráfico de drogas, por sua vez, além de fluido é descontínuo, sendo chamado de território-rede, na qual os seus pontos (ou nós) se intercalam com nós de outras redes, sobrepostas em um mesmo espaço, com as quais disputam o mesmo mercado consumidor (BARROS et al., 2017).

A identidade e o território informam, afetam e moldam um ao outro, sujeitos às relações de poder e seus rótulos. Como coloca Doreen Massey (2017), para possibilitar o “pensar geográfico” devemos considerar o poder, e assim, reconhecer as geografias através das quais o mundo é construído. Uma reflexão sobre a “geografia mental” - as imagens, os rótulos - que as pessoas carregam consigo sobre o mundo é necessária. Mais do que apenas uma covariável a ser inserida em um modelo, o território é algo que pesquisadores deveriam considerar em todas as etapas da pesquisa, assim como deveriam considerar a identidade.

Retornando ao conceito de população de difícil acesso

Tendo abordado estas discussões geo-sociológicas, pode-se voltar ao ponto focal deste artigo: o conceito de população de difícil acesso. Tal como existe agora, o conceito é certamente vago. O termo pode ser enunciado em contextos diferentes e não se referir ao mesmo tipo de população. Decorre então que este conceito precisa ser expandido, se não reformulado.

Primeiramente, a incapacidade de distinguir entre uma população que é de difícil acesso por causa do desenho (e financiamento) da pesquisa e uma população que é de difícil acesso por sua própria natureza deve ser resolvida. Pesquisadores precisam saber o que faz uma população ser considerada de difícil acesso para que possam começar a planejar como melhor alcançar tais populações. Enquanto que as populações de difícil acesso por motivo de desenho (ou financiamento) da pesquisa podem utilizar de estudos na área de metodologia para fundamentar suas possíveis soluções, as populações de difícil acesso por motivos de natureza, identidade ou território pedem um diálogo maior entre os pesquisadores e membros dessas populações para que as soluções sejam planejadas em conjunto. Para ambos os casos, a difusão das soluções também é importante.

O outro ponto que precisa ser destacado aqui é a influência das relações de poder na sociedade, nas mais variadas escalas. Se a identidade é indissociável da concepção de população, argumenta-se aqui que o território em sua concepção latino-americana também o é. A identidade e o território informam, afetam e moldam um ao outro e ambos são sujeitos às relações de poder não só em um contexto geral, mas também no contexto dos estudos. Os pesquisadores, talvez mais do que qualquer outro grupo, precisam refletir sobre suas “geografias mentais” – assim como suas “sociologias mentais”, por assim dizer – e os efeitos, mesmo inconscientes, destas sobre as pesquisas realizadas.

Se o objetivo da ciência é entender, então é um desserviço agir como se certas áreas da ciência fossem isoladas umas das outras. Uma pintura não é apenas um conjunto de cores sobre uma superfície, nem apenas os materiais utilizados em sua composição. Um poema não é redutível à sua métrica, nem a prosa à sua estrutura. Uma aproximação - um diálogo - das diversas áreas do conhecimento é necessária para que todos possam não só ampliar as maneiras de ver e entender o mundo, mas também para que se possa gerar novas ideias.

Propõe-se então, timidamente, a expansão do conceito em duas subclasses:

“populações de difícil acesso baseadas no desenho” e “populações de difícil acesso baseadas em sua natureza”. O primeiro grupo diz respeito àquelas populações que poderiam ser alcançadas com algum ajuste no desenho ou financiamento da pesquisa como, por exemplo, a tradução de questionários em outras línguas para se alcançar minorias linguísticas. Já o segundo grupo continuaria a ser de difícil acesso, mesmo com uma pesquisa “perfeita”.

Obviamente, pela natureza interseccional da identidade humana, certa população pode pertencer aos dois grupos ao mesmo tempo. Ainda assim, a classificação é interessante para que se possa melhor entender como o conceito está sendo enunciado (e o que ele se pretende a significar) e para pautar estudos sobre as populações de difícil acesso baseadas em sua natureza com menos variáveis de confusão.

Referências

- ALTHUSSER, L. *Ideologia e Aparelhos Ideológicos do Estado*. Lisboa: Editorial Presença, 1970.
- ALVES, G. D. A. *A produção do espaço a partir da tríade lefebvriana concebido/percebido/vivido*. GEOUSP Espaço e Tempo (Online), v. 23, n. 3, p. 551–563, out. 2019. ISSN 2179-0892, 1414-7416. Disponível em: <<https://www.revistas.usp.br/geousp/article/view/163307>>.
- AMSTADTER, A. B. et al. *Service utilization and help seeking in a national sample of female rape victims*. Psychiatric services, Am Psychiatric Assoc, v. 59, n. 12, p. 1450–1457, 2008.
- ARURI, N. H. (Ed.). *Palestinian refugees: the right of return*. Sterling, Va: Pluto Press, 2001. (Pluto Middle East studies). ISBN 978-0-7453-1777-9 978-0-74531776-2.
- BARROS, A. P. S.; VIEIRA, D. C. D. M.; CHAGAS, C. A. N. *Território, rede e tráfico de drogas: uma aproximação teórico-conceitual da atividade de comercialização de substâncias ilícitas na Cidade Universitária Prof. José da Silveira Netto (UFPA), em Belém/PA*. InterEspaço: Revista de Geografia e Interdisciplinaridade, v. 3, n. 10, p. 43, dez. 2017. ISSN 2446-6549.
- BEAUVOIR, S. de. *Le deuxième sexe*. Paris: Gallimard, 1949.
- BECKER, H. S. *Outsiders: Estudos de sociologia de desvio*. Rio de Janeiro: Zahar, 2018.
- BERGER, P. L.; LUCKMANN, T. *The social construction of reality: A treatise in the sociology of knowledge*. Garden City: Doubleday, 1966.
- BERRY, S.; GUNN, P. *Conducting research on vulnerable and stigmatized populations*. In: Hard-to-survey populations. Cambridge: Cambridge University Press, 2014. p. 368–378.
- BLACK, E. *IBM and the Holocaust: The Strategic Alliance Between Nazi Germany and America's Most Powerful Corporation-Expanded Edition*. [S.l.]: Dialog press, 2012.
- BOLFARINE, H.; BUSSAB, W. d. O. *Elementos de amostragem*. São Paulo: Editora Blucher, 2005.
- BOVOLENTA, D. *Dominação e apropriação na produção do espaço urbano: o caso do bairro Maré, Rio de Janeiro (RJ)*. In: Dominação e apropriação na luta por espaço urbano. Rio de Janeiro: Consequência, 2018. p. 223–253
- BROEMELING, L. D. *An account of early statistical inference in arab cryptology*. The American Statistician, Taylor & Francis, v. 65, n. 4, p. 255–257, 2011.
- CATAIA, M. A. *Território político: fundamento e fundação do estado*. Sociedade & Natureza, v. 23, n. 1, p. 115–125, 2011. ISSN 1982-4513.
- CHANG, I. *Thread of the silkworm*. New York: Basic Books, 1995. ISBN 978-0465-00678-6 978-0-465-08716-7.
- CLOETE, A.; KALICHMAN, S. C.; SIMBAYI, L. C. *Layered stigma and HIV/AIDS: experiences of men who have sex with men (msm) in South Africa*. In: Stigma, discrimination and living with HIV / AIDS. New York: Springer, 2013. p. 259–269.
- COCHRAN, W. G. *Sampling techniques*. Hoboken, New Jersey: John Wiley & Sons, 1977.
- COLLINS, P. H.; BILGE, S. *Interseccionalidade*. [S.l.]: Boitempo Editorial, 2021.
- CORRÊA, R. L. *Corporação, práticas espaciais e gestão do território*. Anuário do Instituto de Geociências, v. 15, p. 35–41, jan. 1992. ISSN 1982-3908, 01019759.
- CRUZ, V. D. C. *A teoria como caixa de ferramentas: reflexões sobre o uso dos conceitos na pesquisa em geografia*, 2013.
- EYRE, C. *The use of documents in Pharaonic Egypt*. Oxford: Oxford University Press, 2013. (Oxford studies in ancient documents). OCLC: ocn868390820. ISBN 978-0-19-967389-6

- FOUCAULT, M. *Segurança, território, população*. São Paulo (SP): Martins Fontes, 2008. OCLC: 817044535. ISBN 978-85-336-2377-4. FREYRE, G. *Sociologia: Introdução ao estudo de seus princípios*. Recife: É Realizações, 2009.
- FRIED, S. J. *Quantify this: Statistics, the state, and governmentality*. Washington, D.C.: Georgetown University, 2014.
- GALLO, S. *Filosofia e o exercício do pensamento conceitual na educação básica*. *Educação e Filosofia*, v. 22, n. 44, p. 55–78, 2008.
- GUPTA, S.; ANDERSON, R. M.; MAY, R. M. *Networks of sexual contacts: implications for the pattern of spread of HIV*. *AIDS (London, England)*, v. 3, n. 12, p. 807–817, 1989.
- HACKING, I. *How should we do the history of statistics? The Foucault effect: Studies in governmentality*, Harvester Wheatsheaf, 1991.
- HAESBAERT, R. *Regional-global: dilemas da região e da regionalização na geografia contemporânea*. Rio de Janeiro: Bertrand Brasil, 2010.
- HAESBAERT, R. *Conceitos fundamentais da Geografia - região*. *GEOgraphia*, v. 21, n. 45, p. 117–120, 2019.
- HAESBAERT, R. *Território(s) numa perspectiva latino-americana*. *Journal of Latin American Geography*, v. 19, n. 1, p. 141–151, 2020. ISSN 1548-5811.
- HALD, A. *A history of probability and statistics and their applications before 1750*. Hoboken, New Jersey: John Wiley & Sons, 2005.
- HUGHES, E. C. *Dilemmas and contradictions of status*. *American journal of Sociology*, University of Chicago Press, v. 50, n. 5, p. 353–359, 1945.
- KANOUSE, D. E. *et al. Response to the AIDS Epidemic. A Survey of Homosexual and Bisexual Men in Los Angeles County*. [S.l.], 1991.
- KING, D. K. *Multiple Jeopardy, Multiple Consciousness: The Context of a Black Feminist Ideology*. *Signs*, v. 14, n. 1, p. 42–72, 1988.
- KISH, L. *Survey sampling*. Hoboken, New Jersey: John Wiley & Sons, 1965.
- MARTIN, O. *Da estatística política à sociologia estatística. desenvolvimento e transformações da análise estatística da sociedade (séculos XVII-XIX)*. *Revista Brasileira de História, SciELO Brasil*, v. 21, p. 13–34, 2001.
- MASSEY, D. *A mente geográfica*. *GEOgraphia*, v. 19, n. 40, p. 36–40, 2017.
- MAY-CHAHAL, C.; CAWSON, P. *Measuring child maltreatment in the united kingdom: a study of the prevalence of child abuse and neglect*. *Child abuse & neglect*, Elsevier, v. 29, n. 9, p. 969–984, 2005.
- MCCORMICK, T. H.; ZHENG, T. *Network-based methods for accessing hard-to-survey populations using standard surveys*. In: *Hard-to-survey populations*. Cambridge: Cambridge University Press, 2014. p. 485–502.
- MILLER, A.; WELLAND, D. *The Crucible: Arthur Miller*. [S.l.]: Longman, 1992.
- MOYSÉS, Y. d. S. *As múltiplas dimensões do processo de des-reterritorialização vivenciado pela comunidade de Itambi (Itaboraí – Rio de Janeiro)*. In: *Anais do XIV Encontro Nacional da ANPUR*. [S.l.]: ANPUR, 2011.
- O'DONNELL, J. *A invenção de Copacabana: culturas urbanas e estilos de vida no Rio de Janeiro (1890-1940)*. [S.l.]: Editora Schwarcz-Companhia das Letras, 2013.
- PORTO-GONÇALVES, C. W. *De saberes e de territórios: diversidade e emancipação a partir da experiência latino-americana*. *GEOgraphia*, v. 16, p. 37–52, 2006.

- RAFFESTIN, C. *Por Uma Geografia do poder*. Brasília: Atica, 1993. OCLC: 683716279. ISBN 978-85-08-04290-6.
- RIBEIRO, A. C. T. *Território usado e humanismo concreto: o mercado socialmente necessário*. In: Anais do X Encontro de Geógrafos da América Latina. São Paulo: USP, 2005.
- RISPEL, L. C.; METCALF, C. A. *Breaking the silence: South African HIV policies and the needs of men who have sex with men*. *Reproductive Health Matters*, v. 17, n. 33, p. 133–142, jan. 2009. ISSN 0968-8080, 1460-9576.
- RODRIGUES, M. T. *A prostituição no brasil contemporâneo: um trabalho como outro qualquer?* *Revista Katálysis, SciELO Brasil*, v. 12, p. 68–76, 2009.
- ROOSEVELT, F. D. *Executive order 9066*. US National Archives & Records Administration, 1942.
- ROSS, M. W. *et al.* *Stigma consciousness concerns related to drug use and sexuality in a sample of street-based male sex workers*. *International Journal of Sexual Health, Taylor & Francis*, v. 19, n. 2, p. 57–67, 2007.
- SANTOS, C. N. F. d. *A desordem é só uma ordem que exige uma leitura mais atenta*. *Revista de administração municipal*, n. 165, p. 6–17, 1982.
- SANTOS, M. *O papel ativo da geografia: um manifesto*. *Território*, v. 5, p. 103–109, 2000.
- SAQUET, M. A. *Abordagens e concepções sobre território*. São Paulo: Expressão Popular, 2007.
- SCHIFF, S. *The Witches: Salem, 1692*. [S.l.]: Hachette UK, 2015.
- SECCHI, B. *La città dei ricchi e la città dei poveri*. [S.l.]: Gius. Laterza & Figli Spa, 2013.
- SELTZER, W. *Population Statistics, the Holocaust, and the Nuremberg Trials*. *Population and Development Review*, v. 24, n. 3, p. 511, set. 1998. ISSN 00987921. Disponível em: <https://www.jstor.org/stable/2808153?origin=crossref>.
- SELTZER, W.; ANDERSON, M. *After Pearl Harbor: The proper role of population data systems in time of war*. In: Annual meeting of the Population Association of America. [S.l.: s.n.], 2000.
- SELTZER, W.; ANDERSON, M. *Census confidentiality under the second war powers act (1942-1947)*. In: Population Association of America annual meeting, New York. [s.n.], 2007. Disponível em: .
- SENRA, N. *O Saber e o poder das estatísticas: Uma história das relações dos estatísticos com os estados nacionais e com as ciências*. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística, 2005. ISBN 8524037563.
- SHAW, I.; BLOXAM, E. G. (Ed.). *The Oxford handbook of Egyptology*. First edition. Oxford, United Kingdom ; New York, NY: Oxford University Press, 2020. (Oxford handbooks). ISBN 978-0-19-927187-0.
- SHETTERLY, M. L. *Hidden figures: The untold story of the African American women who helped win the space race*. [S.l.]: HarperCollins UK, 2016.
- SLATER, T. *Territorial Stigmatization: Symbolic Defamation and the Contemporary Metropolis*. In: *The SAGE Handbook of New Urban Studies*. 1 Oliver's Yard, 55 City Road London EC1Y 1SP: SAGE Publications Ltd, 2017. p. 111–125. ISBN 978-1-4129-1265-5 978-1-4739-8260-4. Disponível em: <https://sk.sagepub.com/reference/the-sage-handbook-of-new-urbanstudies/i951.xml>.
- SMITH, T. W. *Hard-to-survey populations in comparative perspective*. In: *Hardto-survey populations*. Cambridge: Cambridge University Press, 2014. p. 21–36.
- SOUZA, M. L. *O território: sobre espaço e poder, autonomia e desenvolvimento*. In: *Geografia: conceitos e temas*. Rio de Janeiro: Bertrand Brasil, 1995. p. 77–116.

SOUZA, M. L. “Território” da divergência (e da confusão): em torno das imprecisas fronteiras de um conceito fundamental. In: Territórios e territorialidades: teorias, processos e conflitos. São Paulo: Expressão Popular, 2009. p. 57–72.

STOOP, I. *Representing the populations: what general social surveys can learn from surveys among specific groups*. In: *Hard-to-survey populations*. Cambridge: Cambridge University Press, 2014. p. 225–244.

TOURANGEAU, R. *Defining hard-to-survey populations*. In: *Hard-to-survey populations*. Cambridge: Cambridge University Press, 2014. p. 3–20.

TOURANGEAU, R, et al. (eds.). *Hard-to-survey populations*. Cambridge: Cambridge University Press, 2014.

US CENSUS BUREAU, *Census confidentiality and privacy, 1790-2002*. 2004. Disponível em: <https://www.census.gov/prod/2003pubs/conmono2.pdf>.

VALBELLE, D. *Les recensements dans l’Egypte pharaonique des troisième et deuxième millénaires in sociétés urbaines en Egypte et au Soudan*. Cahier de recherches de l’Institut de papyrologie et d’égyptologie de Lille, n. 9, p. 33–49, 1987.

WASHINGTON, H. A. *Medical apartheid: The dark history of medical experimentation on Black Americans from colonial times to the present*. New York: Doubleday Books, 2006.

A Plataforma Nilo Peçanha e a pesquisa no campo da educação profissional: limites e possibilidades

Eilson Castro Soares de Oliveira*

Barbara Cobo Soares**

Resumo

O presente artigo mostra os limites e possibilidades de utilização dos dados da PNP – Plataforma Nilo Peçanha para pesquisa sobre a atuação da RFEPCT– Rede Federal de Educação Profissional, Científica e Tecnológica. A plataforma dissemina estatísticas oficiais, que retratam uma rede centenária com variados formatos de organização de suas unidades educacionais. Foram analisadas a agenda de produção de estatísticas, bem como os metadados da PNP, colocando em evidência as fontes de informações, periodicidade de divulgação e documentação. São apresentadas limitações de cobertura dos registros e dificuldades para o usuário dos microdados. Apesar dos problemas, aponta-se a riqueza de informações disponíveis para retratar várias dimensões de atuação da RFEPCT, que oferta cursos de diversos tipos e níveis de ensino, situando a importância do usuário para o desenvolvimento da plataforma.

* Doutorando do programa de pós-graduação da ENCE.

** Professora do programa de pós-graduação da ENCE.

Introdução

A educação profissional compreende uma multiplicidade de redes de ensino, como o Sistema S¹, as redes públicas municipais, estaduais e federal, bem como escolas e universidades privadas. Estas instituições atuam em vários níveis e modalidades de ofertas de cursos, que visam suprir diversas demandas da formação para o mundo do trabalho, desde cursos rápidos que buscam formações específicas, como os cursos FIC – Formação Inicial e Continuada, até grades curriculares que necessitam de maior complexidade e tempo de duração, como os cursos técnicos de nível médio e tecnológicos de graduação e pós-graduação.

Para ofertar essa variedade de cursos, as unidades de ensino podem se especializar na oferta de apenas um nível de ensino ou adotar modelos organizacionais que visem ter uma atuação pluricurricular. A oferta de cursos pode ir desde o ensino fundamental até a pós-graduação, sendo possível diversos formatos institucionais que podem ter especificidades quanto a atividade docente, organização pedagógica ou a articulação com a pesquisa e extensão.

Um exemplo pode ser visto na diferença de estruturação dos cargos docentes pelas redes de ensino. Na RFEPT – Rede Federal de Educação Profissional, Científica e Tecnológica, o cargo é denominado de Professor EBTT – Ensino Básico, Técnico e Tecnológico, pois este pode atuar ao mesmo tempo em diversos níveis de ensino, uma vez que as unidades educacionais são pluricurriculares. Na rede pública estadual de São Paulo os cursos técnicos são ofertados nas ETCs – Escolas Técnicas e os de nível superior nas FATECs – Faculdades de Tecnologia, onde existe a separação da carreira de professor do ensino básico e do ensino superior.

Assim, a educação profissional possui uma série de especificidades, sendo possível haver diversos formatos institucionais na organização das unidades de ensino. Analisar esta realidade demanda a produção de informações, que possibilitem acompanhar as políticas educacionais, por meio de indicadores sobre as diversas dimensões da atuação das unidades de ensino. Tal análise, pode contribuir para o aperfeiçoamento das ações e programas, bem como para a construção de uma sociedade democrática, que vise a inclusão dos diversos segmentos populacionais ao acesso da formação para o mundo do trabalho.

Atualmente existem ações que buscam acompanhar a educação profissional, havendo levantamentos efetuados pelos órgãos oficiais de estatística. O INEP – Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, publicou um Anuário Estatístico específico em 2019 (INEP, 2021) e possui quesitos da educação profissional no Censo Escolar e no Censo da Educação Superior. O IBGE investiga informações no suplemento anual de educação da PNAD Contínua – Pesquisa Nacional por Amostra de Domicílios Contínua desde 2016.

As próprias redes de ensino coletam informações para guiar suas ações. Schwartzman (2016) destaca que existem levantamento em várias instituições, como o Centro Paula Souza² e o SENAI - Serviço Nacional de Aprendizagem Industrial. Em 2018 a SETEC – Secretaria de Educação Profissional e Tecnológica implementou a PNP como ferramenta de coleta, validação e disseminação de estatísticas oficiais da RFEPT (MORAES *et al.*, 2020).

O objetivo da PNP foi formar um todo informacional coerente para a realidade vivenciada nas diversas unidades de ensino, por meio do fornecimento de indicadores planejados para a gestão da rede (PRADO; BRITO; NUNES, 2022). Assim, a plataforma é o resultado de um grande esforço, mas apresenta limitações quanto a cobertura, a documentação, a qualidade das informações e a falta de explicação de conceitos, como será visto na discussão do artigo. Apesar das limitações e de ser uma iniciativa recente, já é utilizada como ferramenta de apoio a gestão e na pesquisa acadêmica da educação profissional.

¹ O Sistema S é formado por entidades corporativas patronais, subsidiadas por impostos e que atuam prestação de vários serviços, como a educação profissional. Atualmente é composto por SENAI - Serviço Nacional de Aprendizagem Industrial, SESC – Serviço Social do Comércio, SESI – Serviço Social da Indústria, SENAC – Serviço Nacional de Aprendizagem do Comércio, SEBRAE – Serviço Brasileiro de Apoio às Micro e Pequenas Empresas, SENAR – Serviço Nacional de Aprendizagem Rural, SEST – Serviço Social de Transporte, SENAT – Serviço Nacional de Aprendizagem do Transporte e SESCOOP – Serviço Nacional de Aprendizagem do Cooperativismo.

² O Centro Paula Souza é a autarquia do Governo de São Paulo que administra a rede da estadual de educação profissional.

O CONIF – Conselho Nacional das Instituições da Rede Federal de Educação Profissional, Científica e Tecnológica³ utiliza as estatísticas sistematizadas pela PNP como o principal insumo para construção dos indicadores, que estabelecem os critérios de distribuição de recursos entre seus membros (FORPLAN, 2020).

A Plataforma é utilizada como fonte de informação em pesquisas acadêmicas, servindo como insumo de aferição da oferta de ensino. É possível encontrar diversas referências de estudos com uso dos dados da PNP, como Alves e Melo (2021) que discutem a capilaridade da oferta do EMI – Ensino Médio Integrado no IFAL – Instituto Federal de Alagoas, enquanto Zacaria, Rebeque e Lima (2020) estudaram a oferta de licenciaturas nos IFs – Institutos Federais de Educação, Ciência e Tecnologia. Por sua vez, a evasão e o abandono de estudantes foi tema de investigação em Oliveira et al. (2022), Barbosa (2021), Faustino-Ferber e Martins (2022) e Da Silva e Borges (2019). A preocupação em mensurar aspectos ligados a gestão dos IFs usando dados da PNP aparece nas publicações de Silva, Sá e Freitas (2022) e Panosso *et al.* (2021). Barbosa, Lopes e Araújo (2020) também possuem interesse ligado a aspectos da gestão, buscando verificar o desempenho de alguns Campi no processo de expansão dos IFs.

O papel social da RFEFCT visando a inclusão dos diversos segmentos populacionais é uma preocupação reportada na literatura, com investigações que operam sobre atributos sociais, como as questões de gênero, racial e educação especial. Paiva e Silva (2021) realizaram uma análise de gênero nos cursos técnicos de um *campus* no IFMT, verificando a participação feminina ligada as dimensões do acesso, da evasão e da conclusão. Giarola *et al.* (2021) investigaram a questão racial de estudantes de cursos de licenciatura, enquanto Oliveira e Carvalho Delou (2021) buscaram investigar a educação especial, apontando a inadequação das informações provenientes do Censo Escolar e da PNP sobre deficiências e condições dos estudantes matriculados neste segmento.

Este último artigo sobre lacunas de informações da educação especial na PNP levanta a importância de haver estudos que reflitam sobre a produção e utilização das estatísticas oficiais da Plataforma. Nesse sentido, Rodrigues e Pereira da Silva (2022) e Prado, Brito e Nunes (2022) tornaram a PNP um objeto de estudo em si, onde os primeiros avaliaram, em conjunto com outras bases de informação, o potencial da Plataforma enquanto uma fonte de dados sobre evasão escolar. Prado, Brito e Nunes (2022), por sua vez, fizeram uma análise qualitativa do impacto da PNP nas políticas de avaliação, monitoramento e controle, enquanto mecanismo de gestão da RFEFCT. Ambas as análises são importantes na medida em que uma maior utilização da PNP por pesquisadores que investigam a educação profissional pode possibilitar a familiarização com o método utilizada no levantamento, bem como contribuir com a melhoria dos aspectos investigados e da qualidade dos dados, ao apontar lacunas nas informações disponibilizadas pela plataforma.

Nota-se que a PNP tem sido utilizada na literatura recente tanto como fonte de dados para análise da educação profissional, como objeto de investigação em si, em termos de potencial para subsidiar novos estudos. Como forma de contribuir com os esforços que visam investigar a realidade da educação profissional, o presente artigo foca no segundo campo de estudo, ou seja, sobre as potencialidades e limitações da PNP.

Considera-se que esta é uma base de dados oficiais recente, com possibilidades de utilização dos seus dados tanto na gestão da educação profissional quanto na pesquisa acadêmica sobre estrutura e desempenho deste segmento educacional. Por isso, na segunda seção será discutida a expansão da RFEFCT e a agenda de produção de estatísticas oficiais. Na terceira seção é discutida a relação da política educacional com a produção de estatísticas oficiais, observando aspectos como a cobertura da RFEFCT na população-alvo. Na quarta seção são apresentados alguns aspectos referentes aos metadados da Plataforma, como as fontes de informações, periodicidade de divulgação, documentação, entre outros aspectos. Na quinta seção são apresentadas algumas limitações para o usuário dos microdados da PNP. Por fim, são apresentadas as considerações finais.

³ Instituição que congrega os 38 (trinta e oito) Institutos Federais, o CEFET-RJ, o CEFET-MG e o Colégio Pedro II.

A expansão da Rede Federal de Educação Profissional, Científica e Tecnológica e a agenda de produção de estatísticas oficiais

A RFEFCT – Rede Federal de Educação Profissional, Científica e Tecnológica tem origem em 1909 com a criação de 19 escolas de aprendizes e artífices, pelo então presidente Nilo Peçanha. A própria nomeação da PNP – Plataforma Nilo Peçanha homenageia o criador das primeiras unidades e coloca em evidência a história de uma rede educacional centenária (MORAES, 2020b).

Esta rede passou por várias reformas educacionais ao longo das décadas, tendo sido estruturada em variados modelos de organização institucional no seu percurso histórico. As primeiras escolas de aprendizes e artífices somadas a instalação dos ensinamentos agrícolas, foram sendo transformadas em escolas técnicas, escolas agrotécnicas e CEFETs – Centros Federais de Educação Tecnológica (RAMOS, 2014).

Nas duas primeiras décadas deste século, foram implementadas políticas públicas visando a expansão da RFEFCT, juntamente com a criação e transformação de novos modelos institucionais. A rede passou por uma forte expansão, saindo de 144 unidades em 2006 (MEC, 2023) para chegar a 656 unidade em 2022 (SETEC, 2022).

Em 2005, por meio da Lei nº 11.184, foi criada a UTFPR – Universidade Tecnológica Federal do Paraná, a partir da transformação do CEFET do Paraná e tornando-se uma Universidade com atuação na oferta de graduações, cursos técnicos e formação continuada (BRASIL, 2005), ou seja, uma instituição da educação profissional que atua desde a educação básica, mas que possui foco de atuação no ensino superior.

Em 2008 foram criados 38 IFs – Institutos Federais de Educação, Ciência e Tecnologia a partir da instalação de novas unidades e transformação de outras que já existiam, como CEFETs, ETVs – Escolas Técnicas Vinculadas a Universidades e Escolas Agrotécnicas, com o foco de atuação na educação profissional com formato multicampi distribuídos em todo o território nacional (BRASIL, 2008).

Os IFs foram criados pela Lei nº 11.892 de 2008 com vocação pluricurricular e status de universidade com oferta do nível superior, visando articular ensino, pesquisa e extensão, mas com o foco de atuação no nível médio, em especial com a oferta de cursos técnicos na modalidade EMI - Ensino Médio Integrado.

No processo de expansão da rede e formatação da nova política pública, com foco no modelo dos IFs, foi possibilitado a todas as unidades da RFEFCT aderir ao novo formato institucional. Porém, algumas unidades preferiram por não modificar sua estrutura e mantiveram seu antigo modelo. Este foi o caso dos CEFETs de Minas Gerais e do Rio de Janeiro, que optaram por não se transformar em IFs, possuindo o anseio de um dia virem a ser Universidades Tecnológicas, bem como de algumas ETVs (VIDOR *et al.*, 2011).

Em 2012, o CPEI – Colégio Pedro II passou a ter mesma estrutura institucional dos IFs (BRASIL, 2012). Cabe lembrar que o CPEI é uma instituição com quase dois séculos de existência, tendo sido criada como colégio em 1837 (BRASIL, 1837) e que além da atuação na educação profissional, ainda oferta vagas no ensino básico regular, possuindo matrículas na educação infantil, ensino fundamental e ensino médio generalista.

Portanto, a RFEFCT é formada por uma variedade de instituições, que conjugam novos formatos institucionais, com unidades de ensino de longa tradição de atuação. Possui formatos distintos de organização acadêmica, sendo composta por 2 CEFETs – Centros Federais de Educação Tecnológica (CEFET-MG e CEFET-RJ); a UTFPR; as 23 ETVs; 38 IFs; e o Colégio Pedro II. Cabe destacar que, apesar de coexistirem vários formatos institucionais, a maior parte das unidades, e consequentemente, das matrículas, estão nos IFs.

Em paralelo a esse processo de expansão e criação de novos formatos de organização institucional, a RFEFCT passou a ser cobrada por melhorias no monitoramento estatístico de seus resultados pelos órgãos de controle, devido a “falhas e incompletudes numéricas dos levanta-

mentos” (MORAES, 2019, p. 41). O TCU – Tribunal de Contas da União, por meio dos acórdãos 480/2005 e 2.267/2005, chamou a atenção para a necessidade de informações sobre a atuação da RFEPCT, cobrando a produção e sistematização de estatísticas e listando 11 indicadores para integrarem o acompanhamento das instituições da rede.

Além da cobrança pelos órgãos de controle, os IFs, por serem instituições com formato pluricurricular, já nasceram com a necessidade de cumprir percentuais legais na distribuição de suas matrículas pelos cursos ofertados. De acordo com a sua lei de criação⁴, os IFs devem ofertar pelo menos 50% de vagas para os cursos técnicos EMI e 20% para as licenciaturas. Além destes, o Decreto 5.840 de 2006 estabeleceu que pelo menos 10% das vagas sejam destinadas ao ingresso para o PROEJA - Programa Nacional de Integração da Educação Profissional com a Educação Básica na Modalidade de Educação de Jovens e Adultos.

Há, de um lado, a cobrança dos órgãos de controle sobre informações estatísticas da RFEPCT, e, de outro, a própria legislação estabelece critérios quantitativos para distribuição de vagas dentro das unidades de ensino. Porém, apesar desses apontamentos, as informações das estatísticas educacionais disponíveis não se adequam a realidade da rede.

Para Moraes (2019), no processo de expansão das unidades educacionais havia uma “síndrome do botão mágico”, onde acreditava-se que a simples implementação de novos sistemas informáticos iria garantir o acesso a estatísticas adequadas. O MEC – Ministério da Educação possui um órgão próprio para a produção de estatísticas oficiais educacionais: o INEP – Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Este conta com corpo técnico e funções específicas de produção de informações e avaliações educacionais. O INEP busca retratar a educação básica e superior de forma separada, por meio de duas pesquisas principais: o Censo Escolar e o Censo da Educação Superior. Essas pesquisas são organizadas com coletas e aspectos metodológicos próprios que levam em consideração o nível de ensino, sendo investigados alguns aspectos da educação profissional.

Porém, de acordo com o “guia de referência metodológica” da PNP (MORAES *et al.*, 2020) a soma das duas pesquisas não é adequada para retratar a totalidade da educação profissional em especial a atuação da RFEPCT, pois:

Os dois Censos não utilizam a mesma metodologia de contagem, não sendo correto somar os seus resultados; [...]
O Censo Escolar não contabiliza, no ano de referência, os alunos que ingressam nos cursos técnicos no segundo semestre; [...]
As contagens não contemplam a oferta dos Cursos de Qualificação Profissional e de Pós-Graduação; [...] (MORAES *et al.*, 2020, p. 07).

As principais pesquisas do INEP não conseguem automaticamente reconstruir o todo de atuação de uma unidade de ensino desta rede, não respondendo as necessidades de informações requeridas pelos órgãos de controle, bem como sobre os percentuais legais de oferta das matrículas.

De uma forma geral, a produção de estatísticas oficiais demanda vários tipos de esforços, como a adequação conceitual e metodológica, marco legal de referência e a atuação de uma complexa rede de atores composta pelas várias instituições da rede. Estes aspectos foram levados em consideração na construção da PNP.

Recentemente foi aprovada a Lei nº 14.645/2023, que alterou a Lei de Diretrizes e Bases da Educação Nacional estabelecendo a avaliação nacional de cursos da educação profissional e tecnológica, onde deverão ser consideradas:

[...] estatísticas de oferta, fluxo e rendimento, a aprendizagem dos saberes do trabalho, a aderência da oferta ao contexto social, econômico e produtivo local e nacional, a inserção dos egressos no mundo do trabalho e as condições institucionais de oferta. (BRASIL, 2023)

⁴ Lei 11.892 de 2008.

Sancionada em agosto do corrente ano, esta Lei tem potencial de provocar efeitos na produção de estatísticas oficiais das diversas redes de ensino que atuam na educação profissional e, em particular, na PNP.

As políticas públicas e as estatísticas oficiais

A produção das estatísticas oficiais possui relação direta com a forma de estruturação das políticas públicas, que demandam continuamente informações para sua implementação, monitoramento e avaliação. Sob a perspectiva do objeto de estudo do presente artigo, é importante, portanto, compreender as lacunas e limitações da PNP – Plataforma Nilo Peçanha à luz do modo como as instituições da RFEPC – Rede Federal de Educação Profissional, Científica e Tecnológica estão atualmente organizadas.

A PNP possui sua execução centrada na SETEC – Secretaria de Educação Profissional e Tecnológica, o órgão que executa a política pública de educação profissional, mantendo colaboração com o órgão responsável pela produção de estatísticas oficiais dentro do MEC – Ministério da Educação, tendo em vista que a própria idealização e articulação de aspectos metodológicos foram operacionalizados em colaboração com técnicos do INEP – Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (MORAES, 2020b).

A PNP foi instituída pela portaria nº 1/2018/MEC/SETEC, sendo que no mesmo instrumento normativo foi definido a REVALIDE – Rede de Coleta, Validação e Disseminação das Estatísticas. Esta rede operacionaliza a coleta e validação dos dados da PNP, de acordo com atribuições em três níveis, do governo federal, da instituição e da unidade de ensino, sendo constituída pelos seguintes atores:

- i. os responsáveis pelo registro acadêmico local (RA) de cada unidade de ensino da Rede Federal;
- ii. os diretores de cada unidade de ensino;
- iii. os Pesquisadores Institucionais (PIs), ou cargo equivalente que responda pela produção da estatística educacional, de cada instituição;
- iv. os dirigentes máximos de cada instituição da Rede Federal;
- v. a Diretoria de Desenvolvimento da Rede Federal de Educação Profissional, Científica e Tecnológica [...] (BRASIL, 2018a).

A PNP é operacionalizada por uma rede, que vai, portanto, desde o registro acadêmico local da unidade de ensino até a SETEC, no MEC, onde fica a Diretoria de Desenvolvimento da Rede Federal. A produção e sistematização de informações é feita de forma colaborativa com vários atores, centrado no mesmo órgão que executa e gerencia a política pública da educação profissional, a SETEC, com a colaboração de profissionais de diversos níveis hierárquicos dentro de várias instituições espalhadas por todo o país.

Sobre a estruturação da REVALIDE, cabe ressaltar a centralidade da SETEC na articulação da sistematização das informações em âmbito federal. Porém, uma parte significativa da RFEPC não está diretamente ligada em termos de gestão dos recursos à SETEC. Este é um aspecto que necessita maior reflexão, pois a organização da estrutura da política pública e estruturação de redes de ensino pode impactar na própria produção das informações estatísticas.

Apesar da PNP buscar espelhar a RFEPC por meio de suas estatísticas oficiais, existem limitações de cobertura devido as diversas formas de organização institucional das unidades de ensino. As informações consolidadas na PNP não cobrem, por exemplo, a UTFPR – Universidade Tecnológica Federal do Paraná que, por ser uma organização universitária, não está sob a gestão direta da SETEC, mas na SESU – Secretaria de Educação Superior do MEC. Este é um problema de cobertura, onde uma instituição da educação profissional, que também oferta cursos técnicos, fica fora das estatísticas oficiais sistematizadas no âmbito da SETEC.

As ETVs – Escolas Técnicas Vinculadas a Universidades, por sua vez, também possuem problemas de cobertura. Estas são geridas dentro das estruturas de Universidades Federais, possuindo indicadores consolidados na PNP, mas com lacunas de informações na Plataforma como o “Gasto Corrente por Matrícula – GCM”. Segundo a ficha técnica, esse indicador não é calculado devido “impossibilidade de segregar os dados financeiros” das ETVs (MORAES et al., 2020, p. 34). Ou seja, a PNP não consegue dimensionar os gastos financeiros com as escolas técnicas que estão dentro das estruturas das universidades que não estão sob sua gestão direta.

A Portaria nº 50/2018 instituiu o grupo de especialistas, selecionando três grupos de atores com a finalidade de definir o calendário de 2019 e aprimorar a coleta, validação, capacitação e disseminação da PNP. O primeiro é a SETEC, que coordena todo o processo de gestão da PNP. O segundo é o CONIF – Conselho Nacional das Instituições da Rede Federal de Educação Profissional, Científica e Tecnológica, instituição que congrega todos os IFs – Institutos Federais de Educação, Ciência e Tecnologia, os CEFETs – Centros Federais de Educação Tecnológica e o Colégio Pedro II. Por último, o CONDETUF - Conselho Nacional de Dirigentes das Escolas Técnicas Vinculadas às Universidade Federais. Ou seja, o próprio grupo de especialistas pressupõe a existência de dois segmentos dentro da RFEPCT. Assim, para o usuário das estatísticas públicas da PNP – Plataforma Nilo Peçanha, é necessário ter o cuidado de analisar o recorte institucional em que trabalhará os dados, tendo em vista que apesar da plataforma buscar retratar a RFEPCT – Rede Federal de Educação Profissional, Científica e Tecnológica, há importantes lacunas de cobertura referentes à atuação da UTFPR – Universidade Tecnológica Federal do Paraná e das ETVs – Escolas Técnicas Vinculadas a Universidades.

Os dados da Plataforma Nilo Peçanha

A PNP – Plataforma Nilo Peçanha é uma base que disponibiliza uma grande quantidade de informações para a pesquisa em educação profissional, possibilitando a investigação de vários aspectos da realidade vivenciada pela RFEPCT – Rede Federal de Educação Profissional, Científica e Tecnológica. Anualmente são divulgadas informações sobre diversos aspectos, como os servidores que compõem a força de trabalho da rede, reportando os totais de professores e técnicos em assuntos educacionais por nível de ensino. São divulgadas estatísticas sobre os recursos financeiros destinados ao pagamento de pessoal, no investimento e no custeio da rede. Do mesmo modo são disponibilizados diversos indicadores que mensuram dimensões importantes para o acompanhamento das políticas educacionais, como evasão, eficiência acadêmica, verticalização, entre outros aspectos. Apenas para ter a magnitude da riqueza de informações, na Tabela 1, são apresentados os totais de unidades, cursos e matrículas por tipo de organização acadêmica da RFEPCT na PNP para o ano de 2022, onde havia informações de mais de um milhão e meio de estudantes matriculados em mais de 11 mil cursos, sendo estes distribuídos em 656 unidades de ensino.

Tabela 1 - Total de unidades, cursos e matrículas por organização acadêmica na RFEPCT – Brasil – 2022

Organização Acadêmica	Unidades	Cursos	Matrículas
Total	656	11.814	1.513.075
CEFET	17	319	37.319
Colégio Pedro II	14	111	18.183
ETVs	23	288	20.178
Instituto Federal	602	11.096	1.437.395

Fonte: SETEC. PNP – Plataforma Nilo Peçanha 2022.

Essa grande quantidade de informações pode ser investigada em vários recortes de pesquisa, como em diferentes perfis específicos de cursos, desde as formações iniciais e continuadas, os cursos técnicos de nível médio, subsequentes, concomitantes e integrados, até o nível superior, em graduações de bacharelados, licenciaturas, tecnólogos e pós-graduações. Para além dos diferentes níveis de ensino, os pesquisadores podem buscar recortes com dados sobre as ofertas de cursos no território, bem como dos vários formatos de organização institucional da rede. Compreender os procedimentos metodológicos da coleta de dados de uma base com informações estatísticas é importante para o usuário poder utilizar adequadamente os dados, viabilizando a reflexão sobre suas potencialidades, bem como verificar as limitações.

Os dados da PNP são extraídos de três sistemas de registros administrativos do Governo Federal: o SISTEC – Sistema Nacional de Informações da Educação Profissional e Tecnológica, o SIAPE – Sistema Integrado de Administração de Recursos Humanos e o SIAFI – Sistema Integrado de Administração Financeira do Governo Federal (BRASIL, 2018; MORAES *et al.*, 2020).

Os dados são coletados por um módulo de coleta e validação, desenvolvido pelo NEXT – Núcleo de Pesquisa e Desenvolvimento para Excelência e Transformação do Setor Público da UnB – Universidade de Brasília (MORAES *et al.*, 2020). Após a extração dos dados dos sistemas, estes são processados através de regras de consistência e passam pela validação através da rede REVALIDE, sendo posteriormente divulgados para a utilização dos usuários através de microdados e do módulo de disseminação da Plataforma.

O processo de revalidação da PNP, onde são detectadas inconsistências e feitas as devidas correções, possui uma descrição breve na documentação, não ficando clara a metodologia completa destas correções, bem como quais são as variáveis que passam por esse procedimento e o motivo pelo qual, ao final do processo, ainda permaneça na base de dados uma grande quantidade de inconsistências em algumas variáveis, como categorias “NA” ou “S/I”, que serão analisadas na próxima seção.

Conforme mencionado anteriormente, a PNP divulga suas informações desde 2018, com periodicidade anual, tendo como referência os dados do ano anterior. Mas, justamente o conceito de ano de referência pode ser um fator de dúvidas para o usuário da Plataforma. A PNP estabelece duas referências temporais, sendo o ano de edição, que indica o período de publicação dos dados, e o ano base, que indica o período de coleta dos dados. Assim, caso o pesquisador queira trabalhar com os dados que retratem o ano de 2017, este deve utilizar a edição 2018 que tem por ano base 2017. Essa é uma situação simples, mas que pode gerar confusão na escrita de resultados em pesquisas acadêmicas, por colocar a referência de dois anos numa mesma publicação. Para evitar confusões, neste artigo foi utilizado como referência apenas o ano base.

A PNP conta atualmente com 6 edições, tendo resultados publicados de 2017 a 2022 em tabelações disponibilizadas no módulo de disseminação em dois formatos: Tableau⁵ e Power BI⁶. O usuário também pode acessar diretamente os microdados⁷ de 5 edições, de 2017 a 2021⁸, em quatro bases: matrículas, eficiência acadêmica, servidores e financeiro (MORAES *et al.*, 2020), permitindo análises em diversas escalas e temáticas.

Como forma de orientar o usuário na utilização dos dados, é mantida uma ampla documentação que define uma série de conceitos utilizados pela plataforma. Esse é um aspecto positivo, pois houve um grande esforço para definir conceitos e adequar as informações da base de dados a realidade da educação profissional. Para um maior entendimento dos conceitos utilizados, é necessário consultar um conjunto de instrumentos legais que fornecem o arcabouço institucional da política educacional e da Plataforma, estabelecendo definições importantes na coleta e operacionalização das informações. Fundamentais neste aspecto são as portarias SETEC nº 01/2018,

⁵ Disponível em: <<https://www.gov.br/mec/pt-br/pnp/versao-tableau>>. Acesso em: 04 nov. 2023.

⁶ Disponível em: <<https://app.powerbi.com/view?r=eyJrjoiZDhkNGNiYzgtMjQ0My00OGVILWJjNzYtZWQwYjI2OTlhYWY1IiwidCI6IjllNjgyMzU5LWQxMjgtNGVhYyU4LTgyYjJhMTUzNDZmZi99>>. Acesso em: 04 nov. 2023.

⁷ Disponível em: <<https://dadosabertos.mec.gov.br/pnp>>. Acesso em: 04 nov. 2023.

⁸ Até a submissão desse manuscrito não haviam sido publicados microdados dos anos base de 2022.

que instituiu a PNP; SETEC 146/2021, que definiu os conceitos da PNP; e MEC 1.162/2018, que regulamentou o conceito de aluno-equivalente⁹.

Dois documentos são igualmente importantes para o usuário compreender melhor a sistematização de informações na Plataforma: o dicionário de variáveis e o guia de referência metodológica. Aqui cabe destacar que embora tenha havido grande esforço para a descrição detalhada das informações coletadas, o último guia de referência metodológica é de 2020 e o dicionário de variáveis é o mesmo desde a edição 2018. A não atualização da documentação dificulta a compreensão do usuário, tendo em vista que a PNP possui edição anual e apresentou um aumento significativo das variáveis disponibilizadas. Na base de matrículas dos microdados, o quantitativo de variáveis saiu de 25, em 2017, para 35, em 2018, passando para 54 variáveis nos anos de 2019 e 2020. Em 2021 a PNP possuía 56 variáveis.

Um aspecto que causa estranheza na documentação é o fato do “guia de referência metodológica”, que é o principal instrumento de informações do usuário, ter sido publicado com autores, sendo aqui referenciado como Moraes *et al.* (2020). As ações da PNP são centradas na SETEC, que coordena e implementa todo o processo de coleta, validação e disseminação de estatísticas oficiais. Por isso, seria mais adequado a documentação apresentar esse órgão público como a referência na publicação do guia.

No seu curto período de existência, o usuário que utiliza os microdados pode acessar uma grande variedade de informações que cobrem várias temáticas relativas aos estudantes da rede, apesar de nem todas suficientemente documentadas nas referências metodológicas da Plataforma, constituindo-se, portanto, como uma importante fonte de dados com amplas possibilidades de estudos sobre a RFEPCT, não obstante a atenção que precisa ser dada aos aspectos metodológicos aqui apontados e que serão melhor especificados na seção seguinte.

Limitações na utilização dos microdados da PNP

Nesta seção, como forma de mostrar algumas limitações da PNP – Plataforma Nilo Peçanha, são discutidos resultados obtidos com o uso dos microdados da base matrículas, referentes aos anos-base de 2017 a 2021, dos cursos técnicos na modalidade EMI – Ensino Médio Integrado dos IFs – Institutos Federais de Educação, Ciência e Tecnologia. Para o processamento das informações foi utilizado o *software* livre *R Core Team* (2021), sendo que, em todo o processo, o módulo de disseminação da PNP serviu como parâmetro de validação dos processamentos realizados.

A primeira dificuldade encontrada, conforme ressaltado na seção anterior, foi compreender alguns conceitos utilizados nas divulgações dos dados e reproduzir resultados do módulo disseminação utilizando os microdados, como, por exemplo, nos resultados utilizando as variáveis sobre unidades de ensino.

Na documentação disponibilizada, o termo unidade de ensino é conceituado como “todas as unidades organizacionais que possuam matrículas vinculadas no ano de referência” (MORAES *et al.*, 2020, p. 118). Essa definição não explicita completamente o que se entende por unidade de ensino, levando em conta que existem diversos formatos organizacionais na RFEPCT, conforme já também apresentado anteriormente. Existem nos microdados três variáveis referentes as unidades de ensino - Código da Unidade de Ensino – SISTEC, Cód. Unidade e Unidade de Ensino – sem muita clareza da distinção entre elas. Analisando os totais de unidades de ensino por variável (Tabela 2), observa-se que as três variáveis são divergentes em quantitativo e não permitem ao usuário reproduzir o número de unidades de ensino publicados no módulo de disseminação, reproduzido na última linha da referida tabela.

⁹ O conceito de aluno equivalente corresponde a uma ponderação de uma determinada matrícula pela carga horária e fator de esforço do curso. Este é importante para análises que necessitem considerar cursos de diversos níveis e cargas horárias, como na distribuição de recursos ou na relação professor-aluno.

Tabela 2 - Total de unidades, por ano, segundo as categorias de unidades de ensino dos microdados e do módulo disseminação da PNP – Brasil – 2017/2021

Categoria	Ano				
	2017	2018	2019	2020	2021
Código da Unidade de Ensino - SISTEC	-	647	653	654	652
Cod Unidade	-	-	653	654	656
Unidade de Ensino	637	640	646	647	648
Unidade de ensino	643	647	653	654	656

Fonte: SETEC. PNP – Plataforma Nilo Peçanha 2017/2021.

A variável “Código da Unidade de Ensino – SISTEC” não está presente no ano de 2017 e possui valor divergente em 2021, tendo valores iguais ao módulo disseminação em 2018, 2019 e 2020. A variável “Cod Unidade” possui os valores iguais ao do módulo disseminação, mas esta não existia nas divulgações de 2017 e 2018. A variável “Unidade de Ensino” possui totais divergentes em todos os anos. Assim, falta para o usuário, informações ano a ano que possibilitem fazer a reconstrução exata das informações por unidade de ensino.

Outra limitação para o usuário dos microdados da PNP é a ausência de informações na documentação, que pode gerar confusões, como ocorre nas variáveis idade, sexo, cor ou raça e renda familiar. A idade não consta como variável na edição de 2017 e nas outras 4 edições, como o dicionário de variáveis é o mesmo, consta como “idade do estudante no dia 31/12/2018”. Assim, o usuário encontra um problema devido a não atualização do dicionário de variáveis desde 2018. Como forma de contornar o problema, aqui a informação foi entendida como a idade do estudante no dia 31 de dezembro do ano base.

Nos microdados da PNP existem pessoas com idades que não são adequadas ao curso que frequenta. Nas matrículas dos IFs no EMI, que corresponde a cursos de nível médio, foram encontradas pessoas com idades de 0 (zero) até 13 (treze) anos, ou seja, idades que não são compatíveis com o nível de ensino médio. Também há informações ausentes, ou seja, sem o preenchimento da idade. Essa situação indica erros nos microdados da PNP quanto à variável idade. Apesar destes, deve se levar em consideração que as matrículas com idade de 14 anos ou mais representam a grande maioria das observações chegando a 99,7% do total em 2021.

Na variável sexo, aparecem nos microdados as seguintes categorias: “F”, “M”, “Feminino”, “Masculino” e “S/I”. Na falta de informações precisas sobre as categorias “F” e “M” pode-se tabular estas como Feminino e Masculino, respectivamente, e chegar aos mesmos totais do módulo de disseminação. A categoria “S/I” pode representar matrículas que não contêm a informação, mas, isso não fica claro na documentação. A categoria S/I aparece apenas nos anos de 2000 e 2021, representando um baixo percentual, menor que 1%. Apesar da grande quantidade de informações, em algumas variáveis a PNP não apresenta detalhamento suficiente para entender importantes aspectos da coleta e metodologia, como ocorre na análise dos quesitos de cor ou raça e renda familiar.

No quesito de cor ou raça a forma como cada pessoa se define pode mudar e não há o registro se as informações provêm de um sistema de auto-atribuição, onde o próprio indivíduo informa a categoria a que pertence, ou heteroatribuição, onde outro indivíduo informa a categoria, como os pais/responsáveis, o servidor do registro escolar da unidade de ensino, ou de outra forma. Tampouco há descrição de quando é coletada essa categoria, se no processo seletivo ou em outra etapa pelo registro acadêmico. A informação de pertencimento racial é sensível a percepção do indivíduo, bem como é parte integrante de políticas de acesso, como as cotas, sendo importante para o pesquisador que irá utilizar os dados saber como é o processo de coleta, bem como os conceitos utilizados.

De forma análoga, não há informações para o usuário da Plataforma como foi atribuída a faixa de renda familiar de cada matrícula, se pela escola, pelo indivíduo matriculado ou por país/responsáveis em algum instrumento de declaração, ou de outra forma. Não só há dúvidas

quanto à forma de coleta, como as categorias utilizadas da forma como estão disponibilizadas nos microdados não permitem fazer uma avaliação adequada sobre esse importante aspecto. No dicionário de variáveis aparece a seguinte descrição para renda familiar: “Faixa de renda per capita familiar do aluno, declarada na PNP pela instituição (Opções: $0 < RFP \leq 0,5$; $0,5 < RFP \leq 1$; $1 < RFP \leq 1,5$; $1,5 < RFP \leq 2,5$; $2,5 < RFP \leq 3,5$; $RFP > 3,5$; Não declarada)”. Essa descrição indica que é considerada a renda per capita da família do estudante, porém, não fornece os critérios e decisões metodológicas para a agregação em faixas de rendimento, pois não são explicados os motivos para o estabelecimento das categorias que foram agregadas as faixas de renda. Também não fica delimitado o conceito de faixa de renda, se estas são medidas do salário-mínimo ou outra forma de agregação.

Nas Tabelas 3 e 4 são apresentadas as distribuições percentuais das matrículas por cor ou raça e renda familiar de 2017 à 2021. Apesar destas serem informações importantes para verificar os impactos da atuação do EMI para diversos segmentos populacionais, bem como refletir e analisar as políticas voltadas à inclusão como a assistência estudantil e as cotas, não é possível fazer inferências devido a grande quantidade de matrículas na categoria “não declarada”.

Tabela 3 - Percentual de matrículas por cor ou raça no EMI dos Institutos Federais – Brasil – 2017/2021

Cor ou raça	Ano				
	2017	2018	2019	2020	2021
Amarela	0,9	1,8	1,0	1,0	0,9
Branca	27,6	26,2	28,6	32,2	33,1
Indígena	0,5	0,6	0,5	0,5	0,5
Parda	33,8	36,6	40,6	43,6	43,7
Preta	5,7	6,4	7,1	7,8	7,7
Não declarada	31,5	28,4	22,1	15,0	14,1

Fonte: SETEC. PNP – Plataforma Nilo Peçanha 2017/2021.

Tabela 4 - Percentual de matrículas por Renda Familiar no EMI dos Institutos Federais – Brasil – 2017/2021

Renda Familiar	Ano				
	2017	2018	2019	2020	2021
$0 < RFP \leq 0,5$	18,8	22,0	23,8	26,8	26,8
$0,5 < RFP \leq 1,0$	14,1	16,2	16,1	17,7	18,3
$1,0 < RFP \leq 1,5$	8,8	9,7	10,0	11,1	11,3
$1,5 < RFP \leq 2,5$	6,5	6,9	6,4	7,3	7,0
$2,5 < RFP \leq 3,5$	3,0	2,9	2,8	3,2	3,1
$RFP > 3,5$	4,0	3,5	3,6	4,1	4,1
Não declarada	44,7	38,9	37,3	29,8	29,3

Fonte: SETEC. PNP – Plataforma Nilo Peçanha 2017/2021.

A categoria “Não declarada” apresentou significativa redução nas cinco edições da PNP para as variáveis cor ou raça e renda familiar, o que indica melhoria da qualidade dos dados. Cabe lembrar que as estatísticas da plataforma dependem de uma ampla rede de atores fornecerem dados. A melhoria da captação de informações pode ocorrer devido a melhor compreensão dos diversos atores envolvidos na apuração e registro das informações.

Apesar da redução da não declaração ao longo das edições, há ainda 14% de registros na variável cor ou raça e 29% na variável renda familiar nesta condição em 2021. Isto inviabiliza a

utilização desses quesitos para a pesquisa, pois não há como saber a real situação da distribuição das matrículas pelas categorias de interesse.

Causa estranheza permanecer uma grande quantidade de informações não declaradas após um processo de validação, porém, a documentação não informa o motivo desta permanência, se por problemas operacionais, não sendo possível a coleta ou correção, ou se o processo metodológico de validação não atua sobre as duas variáveis.

No Quadro 1 foi sintetizado as principais dificuldades e possíveis ajustes que podem facilitar a utilização dos microdados pelo usuário PNP. Alguns ajustes são de difícil execução, como fortalecer os procedimentos para correção dos dados nas variáveis cor ou raça, renda familiar e cor ou raça. Outras ações mais simples podem trazer impactos positivos, como a padronização de rótulos das variáveis e categorias, bem como ajustes na documentação que forneçam mais informações sobre as categorias, conceitos e procedimentos na coleta e processamento dos dados.

Quadro 1: Dificuldades encontradas para a utilização dos microdados da Plataforma Nilo Peçanha

Variável	Dificuldades	Ajustes
Unidade de ensino	Possui três variáveis que não permitem reproduzir o total de unidades do módulo disseminação	Padronizar uma variável que possa reproduzir o módulo disseminação
Idade	Matrículas abaixo da idade escolar.	Fortalecer os procedimentos de correção
Sexo	Matrículas rotuladas como "Feminino", "Masculino", "F", "M" e "S/I"	Padronizar a nomenclatura das categorias e explicar a categoria "S/I"
Cor ou raça	Grande quantidade da categoria "Não declarada"	Instituir procedimentos de correção
	Falta de informação sobre a coleta	Informar o formato de coleta da variável, se por instrumento de autoatribuição ou heteroatribuição.
Renda Familiar	Grande quantidade da categoria "Não declarada"	Fortalecer os procedimentos de correção
	Falta de informação sobre os procedimentos de coleta, categorias e rótulos	Informar na documentação a forma de coleta, se por instrumento de autoatribuição ou heteroatribuição e explicar de forma exaustiva o conceito e categorias utilizadas.
Tipo de Oferta	Possui duas nomenclaturas "Tipo Oferta" e "Tipo de Oferta", estando duplicada em duas na edição 2020	Padronizar a variável

Fonte: SETEC. PNP – Plataforma Nilo Peçanha 2017/2021.

Portanto, apesar da grande quantidade de informações, a PNP ainda possui limites nos dados fornecidos, que vão desde a documentação, até a qualidade de algumas informações, o que pode dificultar ou até mesmo inviabilizar análises, como o perfil econômico ou racial dos estudantes matriculados, diretriz importante para políticas de inclusão educacional. Há espaço para ações que visem o ajuste e aprimoramento da base de modo a possibilitar maior utilização pelos usuários.

Considerações finais

A PNP é uma base de dados que oferta estatísticas oficiais da educação profissional, uma iniciativa importante que busca adequar metodologicamente as informações coletadas e disponibilizadas à realidade vivenciada pela RFEPCT.

Deve-se considerar que a rede é composta por instituições pluricurriculares que possuem uma grande variedade de formações. Refletir essa diversidade é uma potencialidade da PNP ao permitir a investigação de perfis específicos de cursos. Também há a possibilidade de, a partir dos microdados, buscar informações sobre as ofertas de cursos espalhados no território, podendo o pesquisador operar em vários recortes institucionais, de abrangência geográfica, níveis de ensino e tipos de ofertas curriculares.

A maior utilização da PNP por usuários com o objetivo de pesquisa pode ajudar justamente ao maior entendimento da atuação da RFEPCT. Sendo esta atividade importante para reflexões sobre o desenvolvimento da educação profissional. Do mesmo modo, uma maior comunidade de usuários demandando informações da SETEC, pode contribuir no desenvolvimento das estatísticas oficiais, visando um maior entendimento metodológico da base de dados, bem como indicando os aperfeiçoamentos necessários.

Em suas seis edições a plataforma vem apresentando crescimento da quantidade de informações divulgadas, sendo uma rica fonte de dados, que atualmente já contribui com a gestão de políticas educacionais, bem como é uma fonte buscada por pesquisadores para investigar a realidade da RFEPCT.

Devido à importância da PNP, foram analisados os seus limites e possibilidades na pesquisa educacional, como uma fonte de dados para o usuário de estatísticas públicas. Como resultado foi possível verificar que os dados possibilitam várias análises para a pesquisa acadêmica em educação profissional, porém, alguns aspectos, tais como a origem dos dados, as limitações da PNP e a agenda de produção de estatísticas, merecem maior atenção da comunidade de usuários dessas estatísticas públicas, sendo importante para análises futuras.

Os dados da PNP são extraídos de registros administrativos de três fontes: o SISTEC, o SIAPE e o SIAFI. São necessárias investigações que analisem as qualidades e limitações dessas três fontes, bem como seus impactos nos resultados da PNP. Um exemplo são os dados financeiros, que não permitem o cálculo de alguns indicadores para as ETVs, bem como as informações não declaradas em variáveis importantes como cor ou raça ou rendimento familiar dos estudantes.

A PNP apresenta limitações quanto à cobertura, documentação, a qualidade das informações e a falta de explicação de conceitos. Deve-se destacar que há um rico conjunto de informações, mas que podem ter maior potencial para a pesquisa, com a implementação de ações que visem a combater as limitações da plataforma. É necessário que sejam tomadas ações para a atualização contínua da documentação, como o dicionário de variáveis e do guia de referência metodológica, de modo que o usuário possa reproduzir através dos microdados as informações do módulo de disseminação. Do mesmo modo, é importante buscar explicitar melhor conceitos e colocar informações e decisões metodológicas, como visto nas variáveis idade, sexo, cor ou raça e renda familiar.

Por último, deve-se aprofundar reflexões que tematizem sobre a agenda de produção de informações, analisando os aspectos positivos, bem como as limitações da produção de informações da PNP ser feita dentro do órgão que executa e gerencia a política pública da educação profissional, bem como os futuros efeitos na plataforma em decorrência da recente aprovação da avaliação nacional dos cursos da educação profissional e tecnológica, estabelecendo a obrigação da produção de estatísticas.

Referências

- ALVES, D. DOS S.; MELO, B. M. DE. *Entre tensões e contradições: a capilaridade do Ensino Médio Integrado na Educação Profissional e Tecnológica*. Retratos da Escola, v. 15, n. 32, p. 581–596, 2021.
- BARBOSA, L. *Ensino Médio Integrado e Abandono Escolar: Uma análise do Instituto Federal do Paraná (2017-2019)*. Reflexão e Ação, v. 29, n. 3, 2021.
- BARBOSA LOPES, A. D. O.; ARAÚJO, R. M. DE. *Avaliação de desempenho dos Institutos Federais: um estudo de caso na terceira expansão da Rede Federal de Educação Profissional em Pernambuco*. Pensamento & Realidade, v. 35, n. 1, p. 113–130, 2020.
- BRASIL. *Decreto de 2 de dezembro de 1837. Convertendo o Seminário de S. Joaquim em collegio de instrução secundaria, com a denominação de Collegio de Pedro II, e outras disposições.*, [1837]. Disponível em: https://www2.camara.leg.br/legin/fed/decret_sn/1824-1899/decreto-36979-2-dezembro-1837-562344-publicacaooriginal-86295-pe.html. Acesso em: 24 out. 2023.
- BRASIL. *Lei nº 9.394, de 20 de dezembro de 1996. Estabelece as diretrizes e bases da educação nacional*. Brasília: Congresso Nacional, [1996]. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/l9394.htm. Acesso em: 25 abr. 2023.
- BRASIL. *Tribunal de Contas da União. Acórdão nº 480/2005. Plenário. Relator: Ministro Benjamin Zymler*. Sessão de 27 de abril de 2005a. Disponível em: < <https://pesquisa.apps.tcu.gov.br/redireciona/acordao-completo/%22ACORDAO-COMPLETO-22466%22>>. Acesso em: 05 nov. 2023.
- BRASIL. *Lei nº 11.184, de 7 de outubro de 2005. Dispõe sobre a transformação do Centro Federal de Educação Tecnológica do Paraná em Universidade Tecnológica Federal do Paraná e dá outras providências*, [2005]. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2004-2006/2005/lei/111184.htm. Acesso em: 13 jun. 2023.
- BRASIL. *Tribunal de Contas da União. Acórdão nº 2.267/2005. Plenário. Relator: Ministro Benjamin Zymler*. Sessão de 13 de dezembro de 2005b. Disponível em: < <https://pesquisa.apps.tcu.gov.br/redireciona/acordao-completo/%22ACORDAO-COMPLETO-26325%22>>. Acesso em: 05 nov. 2023.
- BRASIL. *Lei nº 11.892, de 29 de dezembro de 2008. Institui a Rede Federal de Educação Profissional, Científica e Tecnológica, cria os Institutos Federais de Educação, Ciência e Tecnologia, e dá outras providências*. Brasília: Congresso Nacional, [2008]. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2007-2010/2008/lei/l11892.htm. Acesso em: 07 jun. 2023.
- BRASIL. *Lei nº 12.677, de 25 de junho de 2012. Dispõe sobre a criação de cargos efetivos, cargos de direção e funções gratificadas no âmbito do Ministério da Educação, destinados às instituições federais de ensino; altera as Leis nºs 8.168, de 16 de janeiro de 1991, 11.892, de 29 de dezembro de 2008, e 11.526, de 4 de outubro de 2007; revoga as Leis nºs 5.490, de 3 de setembro de 1968, e 5.758, de 3 de dezembro de 1971, e os Decretos-Leis nºs 245, de 28 de fevereiro de 1967, 419, de 10 de janeiro de 1969, e 530, de 15 de abril de 1969; e dá outras providências*. Brasília: Congresso Nacional, [2012]. Disponível em: https://www.planalto.gov.br/ccivil_03/_Ato2011-2014/2012/Lei/L12677.htm#art5. Acesso em: 24 out. 2023.
- BRASIL. *Lei nº 12.796, de 4 de abril de 2013. Altera a Lei nº 9.394, de 20 de dezembro de 1996, que estabelece as diretrizes e bases da educação nacional, para dispor sobre a formação dos profissionais da educação e dar outras providências*. Brasília: Congresso Nacional, [2013]. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2013/lei/l12796.htm. Acesso em: 25 abr. 2023.
- BRASIL. Ministério da Educação (MEC). Secretaria de Educação Profissional e Tecnológica (SETEC). *Portaria nº 1, de 3 de janeiro de 2018. Institui a Plataforma Nilo Peçanha - PNP, a Rede de Coleta, Validação e Disseminação das Estatísticas da Rede Federal de Educação Profissional, Científica e Tecnológica - Revalide*. Diário Oficial da União, Brasília, DF, n. 3, 4 jan. 2018a. Seção 1, p. 10. Disponível em:

<https://www.in.gov.br/web/dou/-/portaria-n-1-de-3-de-janeiro-de-2018-1590408>. Acesso em: 25 abr. 2023.

BRASIL. Ministério da Educação (MEC). Secretaria de Educação Profissional e Tecnológica (SETEC). *Portaria nº 1.162, de 9 de novembro de 2018. Regulamenta o conceito de Aluno-Equivalente e de Relação Aluno por Professor, no âmbito da Rede Federal de Educação Profissional, Científica e Tecnológica*. Diário Oficial da União, Brasília, DF, n. 217, 9 nov. 2018b. Seção 1, p. 35. Disponível em: <https://www.in.gov.br/web/dou/-/portaria-no-1-162-de-9-de-novembro-de-2018-49743560>. Acesso em: 25 abr. 2023.

BRASIL. Ministério da Educação (MEC). Secretaria de Educação Profissional e Tecnológica (SETEC). *Portaria nº 50, de 12 de novembro de 2018. Institui o grupo de especialistas para o desenvolvimento dos trabalhos da Plataforma Nilo Peçanha 2019 (Ano Base 2018)*. Diário Oficial da União, Brasília, DF, n. 220, 16 nov. 2018c. Seção 2, p. 23. Disponível em: <https://www.in.gov.br/web/dou/-/portaria-n-50-de-12-de-novembro-de-2018-50490973>. Acesso em: 25 abr. 2023.

BRASIL. Ministério da Educação (MEC). Secretaria de Educação Profissional e Tecnológica (SETEC). *Portaria nº 146, de 25 de março de 2021. Define conceitos e estabelece fatores para uso na Plataforma Nilo Peçanha - PNP e para cálculo dos indicadores de gestão das Instituições da Rede Federal de Educação Profissional, Científica e Tecnológica*. Diário Oficial da União, Brasília, DF, n. 58, 26 mar. 2021. Seção 1, p. 48. Disponível em: <https://www.in.gov.br/web/dou/-/portaria-n-146-de-25-de-marco-de-2021-310597431>. Acesso em: 25 abr. 2023.

DA SILVA, J. R.; BORGES, L. F. F. *Políticas de acesso nos processos seletivos da Rede Federal de Educação Profissional no Brasil: Uma análise de desempenho*. Revista Ciências Humanas, v. 12, n. 3, 2019.

ENCE, Escola Nacional de Ciências Estatísticas. *Apresentação da Pós-Graduação em População, Território e Estatísticas Públicas*. Disponível em: <https://ence.ibge.gov.br/index.php/mestrado-e-doutorado/apresentacao>. Acesso em: 04 nov. 2023.

FORPLAN – Fórum de Planejamento e Administração do CONIF. *Metodologia da Matriz Orçamentária da Rede de Ensino Profissional e Tecnológica 2020*. Disponível em: <<https://www.ifms.edu.br/gestao/planejamento/orcamento/metodologia-da-matriz-conif-2020.pdf>>. Acesso em: 11 abr. 2023.

FAUSTINO-FERBER, A. P.; MARTINS, A. M. *Formação inicial de professores: análise da evasão em cursos de Pedagogia no Instituto Federal de São Paulo*. Revista Principia, v. 59, n. 4, p. 1137–1157, 2022.

GIAROLA, A. M. et al. *Investigação das características de cor, investimento por aluno e relação de concluintes das licenciaturas da Rede Federal de Educação Profissional, Científica e Tecnológica*. La Salle (Canoas, Rio Grande do Sul, Brazil), v. 26, n. 1, p. 1, 2021.

INEP, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Anuário Estatístico da Educação Profissional e Tecnológica: Ano base 2019*. Brasília: INEP, 2021.

MARTINS, R. *Considerações para a promoção de respostas integrais de enfrentamento da Covid-19 no Brasil*. In: ABEP, Associação Brasileira de Estudos Populacionais. *População e Desenvolvimento: impactos multidimensionais da pandemia da Covid-19 no Brasil*. Campinas, SP: Traço Publicações e Design/ABEP, 2021. p. 60-66. Disponível em : <https://brazil.unfpa.org/sites/default/files/pub-pdf/e-book_populacao_e_desenvolvimento_em_debate_unfpa_e_abep_2021.pdf>. Acesso em: 09 jun. 2023.

MEC, Ministério da Educação. *Expansão da Rede Federal*. Disponível em: <http://portal.mec.gov.br/setec-programas-e-aco-es/expansao-da-rede-federal>. Acesso em: 04 nov. 2023.

MORAES, G. H.; ALBUQUERQUE, A. E. *As Estatísticas da Educação Profissional e Tecnológica: silêncios entre os números da formação de trabalhadores*. Textos para Discussão, v. 45. Brasília: INEP, 2019. Disponível em: <http://td.inep.gov.br/ojs3/index.php/td/article/view/3884/3560>. Acesso em: 30 nov. 2021.

MORAES, G. H. *et al.* *Plataforma Nilo Peçanha: guia de referência metodológica*. Brasília: Editora Evobiz, 2020.

MORAES, G. H. *A Plataforma Nilo Peçanha: desafios à construção de estatísticas e indicadores de avaliação para a Rede Federal*. In: MORAES, G. H. *et al.* (Org.). *Avaliação da Educação Profissional e Tecnológica: um campo em construção*. 1ed. Brasília: INEP, 2020b. p. 423-476.

OLIVEIRA, W. M. DE; CARVALHO DE LOU, C. M. *Estudantes público-alvo da educação especial nos institutos federais: quem são?* Tear (Canoas), v. 10, n. 1, 2021.

OLIVEIRA, I. D. S. *et al.* *Construção de um data warehouse para uma análise multidimensional da evasão escolar - estudo de caso da Plataforma Nilo Peçanha*. *Revista Principia - Divulgação Científica e Tecnológica do IFPB*, 2022.

PAIVA, T. S. Z. N.; SILVA, J. S. *A Participação Feminina nos Cursos Técnicos Integrados ao Ensino Médio da Educação Profissional e Tecnológica*. *Revista brasileira de informática na educação*, v. 29, p. 993-1006, 2021.

PANOSSO, O. *et al.* *Eficiência dos gastos em educação técnica no Brasil: análise dos IFs - Institutos Federais*. *Revista Fatec Zona Sul*, v. 7, n. 6, p. 13-30, 31 ago. 2021.

PRADO, D. P. F.; BRITO, V. L. F. DE; NUNES, C. P. *Concepções e perspectivas da Plataforma Nilo Peçanha: regulação e emancipação*. *Estudos em Avaliação Educacional*, v. 33, p. e07594, 16 dez. 2022.

R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing, 2021.

RAMOS, M. N. *História e política da educação profissional*. Curitiba: Instituto Federal do Paraná, 2014.

RODRIGUES, H. K.; PEREIRA DA SILVA, L. *A utilização de plataformas digitais na transparência de dados da evasão na EAD: Análise de um curso técnico a distância no IFB*. *Revista Nova Paideia*, v. 4, n. 3, p. 657-671, 2022.

SETEC – Secretaria de Educação Profissional e Tecnológica. *Plataforma Nilo Peçanha: módulo disseminação*. Disponível em: <https://app.powerbi.com/view?r=eyJrIjoiZDhkNGNiYzgtMjQ0My00OGVILWJjNzYtZWQwYjI2OThhYWY1IiwidCI6IjllNjgyMzU5LWQxMjgtNGVhYyYjU4LTgyYjJhMTUzNDZjZj9>. Acesso em: 04 nov. 2023

SCHWARTZMAN, S. *Legitimidade, controvérsias e traduções em estatísticas públicas*. *Teoria & Sociedade*. Belo Horizonte, n.1, 1997, p.9-38.

SCHWARTZMAN, S. *Educação média profissional no Brasil: situação e caminhos*. São Paulo: Fundação Santillana, 2016.

SILVA, T. H. O.; SÁ, E. M. DE; FREITAS, B. A. *Gestão educacional: proposta de uma abordagem objetiva para mensuração da eficiência da rede federal de educação*. *Revista Gestão universitária na América Latina*, v. 15, n. 1, p. 226-252, 2022.

ZACARIA, E. G. D.; REBEQUE, P. V.; LIMA, F. B. G. DE. *Dez anos de criação dos Institutos Federais de Educação, Ciência e Tecnologia: sobre a oferta de licenciaturas nas áreas de ciências e matemática*. Tear (Canoas), v. 9, n. 1, 2020.

VIDOR, A. REZENDE, C. PACHECO, E. CALDAS, L. *Institutos federais: Lei no 11.892 de 29/12/2008 – Comentários e reflexões*. In: PACHECO, E (Org.). *Institutos Federais: uma revolução na educação profissional e tecnológica*. São Paulo, Brasília: Fundação Santillana - Editora Moderna, 2011.

Seção 3

Estimação de Indicadores e Monitoramento de Políticas Públicas e Objetivos de Desenvolvimento Sustentável

Access to clean drinking water: complementary indicators of Monitoring Goal 6.1 of the 2030 Agenda at the local level in Brazil

Ana Clara Cancelo Cruz*

Andrea Diniz da Silva**

Abstract

Among the global efforts related to public health and the fight against social inequalities, the 2030 Agenda presents the issue of drinking water in SDG 6 - Drinking Water and Sanitation. To monitor achievement regarding access to drinking water, the United Nations (UN) presented target 6.1 - By 2030, achieve universal and equitable access to safe and affordable drinking water for and indicator 6.1.1 - Proportion of the population that uses safely managed drinking water services. Taking into account Brazilian particularities and specific challenges regarding data availability, we present an analysis of indicator 6.1.1 in Brazil and propose 8 complementary indicators to assist monitoring target 6.1, with evidence from the city of Rio de Janeiro, which allowed analyzing particularities involving access to safely managed drinking water not covered by indicator 6.1.1 at the local level. The indicators and this study can improve discussions regarding incentive policies for the implementation and monitoring of the SDG 6 at the local level.

Keywords: Drinking water; SDG 6; SDG indicators; 2030 Agenda.

* Mestre pelo programa de pós-graduação da ENCE.

** Professora do programa de pós-graduação da ENCE.

Introduction

The year 2020 was marked by the emergence of a new human coronavirus, SARS-CoV-2, which can cause Coronavirus Disease 2019 (or COVID-19). Due to the ease of propagation of this virus and the growing number of cases of the disease around the world, a global pandemic was announced on January 30th, thus classifying the outbreak as a public health emergency of international importance – the highest alert level of the World Health Organization (WHO). According to monitoring carried out by Johns Hopkins University, until January 2021, there had been more than 92 million cases globally, causing almost 2 million deaths from COVID-19. Brazil was among the countries with the highest numbers of confirmed cases in the world, in June 2021 having occupied third position in the ranking, behind only the United States and India.

To prevent the spread of the virus, it was (and still is) recommended that people take a number of precautions, such as using personal protective equipment (PPE, chiefly masks) and properly washing hands with soap and running water whenever there is contact with surfaces possibly contaminated by the virus (OPA, 2020). Additionally, it is recommended to wash all parts of the body, clothing, packages of purchased products and other objects that have come into contact with surfaces subject to contamination, to reduce the chances of contamination through these vectors. To enable such precautions, the availability of water in sufficient quantity and quality for the population to carry out hygiene procedures as recommended by the WHO is extremely important. Access to clean water has also long been included in international agendas to fight social inequalities and improve basic human rights, but it has become even more important in the present context.

The 2030 Agenda for Sustainable Development, instituted by the UN in 2015, addresses the issue through Sustainable Development Goal 6 - Drinking Water and Sanitation (SDG 6), that proposes to ensure availability and sustainable management of water and sanitation for all. As water is an essential resource for life, the social aspect involving this issue is also regarded, in addition to the use from agriculture, industry, and energy sectors. By 2024, the SDG 6 presented eight targets and eleven indicators.

The focus of this work is the target 6.1: By 2030, achieve universal and equitable access to safe and affordable drinking water for all. This target is monitored by the indicator 6.1.1: Proportion of the population that uses safely managed drinking water services (UN BRAZIL, 2020). However, since each country has specific characteristics and challenges related to each objective, the UN encourages adapting targets and indicators considering national realities and priorities as well as producing complementary indicators so that they are successful in monitoring the goals.

Complementary indicators must fulfill the objective of evaluating the best possible situation, including the complexities, considering local policies and government programs, in this particular case involving access to clean drinking water, which can provide greater adherence to its environmental and social reality. In addition, the availability of accurate data with continuous production must be taken into account. In Brazil, the main source of data used for monitoring target 6.1, the Continuous National Household Survey (PNAD Contínua) provides estimates only at national and state levels. In this context, monitoring target 6.1 at the local level can only be carried out through complementary indicators calculated based on alternative data sources.

To help improve the monitoring of target 6.1 at all levels, we propose eight complementary indicators that can be calculated at the national and local levels in Brazil, using data from the National Sanitation Information System (SNIS), the Brazilian Institute of Geography and Statistics (IBGE) and the Database of the Unified Health System (DATASUS).

The present paper is an extension of the work done Cruz (2021). It is divided into three sections in addition to this introduction. The second section describes the situation of target 6.1 in Brazil, according to indicator 6.1.1 presented by the UN and calculated by the IBGE with data from the Continuous PNAD. The third section introduces the proposed complementary indica-

tors by means of an illustration at the local level of the city of Rio de Janeiro, for the period from 2016 to 2018. Discussion is presented at the fourth and last section.

Drinking water in Brazil: Indicator 6.1.1

In Brazil, the monitoring of universal access to water is done through target 6.1, which is expressed as “by 2030, achieving universal and equitable access to safe drinking water for all” (UN BRAZIL, 2020). This target refers to the supply of safe drinking water to households, i.e., water free from contamination, available, when necessary, in sufficient quantity and quality for the consumption needs of the population, in an equitable manner (ANA, 2019). The UN proposes that the monitoring of this target be done through indicator 6.1.1: Proportion of the population using safely managed drinking water services. In Brazil, this indicator has been calculated since 2016 using data from the Continuous National Household Survey (PNAD-C), carried out by the Brazilian Institute of Geography and Statistics (IBGE), which is the national statistical office.

The indicator is composed of variables that indicate whether the household has safe water services, which is characterized by simultaneously having conditions such as an adequate source of water, which can be a general network, artesian well, protected shallow well, protected spring or stored rainwater, and availability of water when needed, free from fecal, chemical and/or pipework contamination.

The indicator is calculated as follows:

$$\text{Indicator 6.1.1} = \frac{[(A + B + C) \times 100] + (D \times 50)}{E}$$

Where:

A = People living in households supplied by the general water network, with water pipes in the house or on the land/property, with a supply frequency of at least 4 days a week or with a reservoir, water tank, cistern, to store water.

B = People living in households supplied by artesian wells, with water piped into the home or on the land/property.

C = People living in households supplied in other ways, with water piped inside the home.

D = People living in households supplied in other ways, with water piped only on the land or property”.

E = People living in permanent private households.

The proportion of people using safely managed water services in Brazil, when measured by 6.1.1, remained unchanged between 2016 and 2017, at 97.9%. In 2018, this proportion increased to 98.2% (ODS Brazil, 2020). In absolute numbers, during the period analyzed, approximately 3.9 million people attained access to securely managed water services. However, despite this progress, it is important to note there was still a significant number of people who were left out. In 2018, this number exceeded 3.8 million. This shows that even though the indicator presents a high proportion of inclusion, around 98% of the Brazilian population, there are still many people without access to such services. For this goal to be met, i.e., for all Brazilians to have access to safely managed water services by 2030, the government needs to do its part through public programs and policies.

Despite indicating the proximity of the universalization of the drinking water service in the country, indicator 6.1.1 does not reveal issues involving the topic in its entirety and full complexity, since it cannot be calculated at the local level. Brazil is a country with a huge social inequality, where socioeconomic contrasts are observed not only between regions and states, but

also within them. Therefore, a situation at national level does not always depicts what happens at the local level. Thus, it is important to analyze the topic through complementary indicators, which in addition to allowing new dimensions to be discussed, allow an understanding of what happens at local levels. That is the case of the selected indicators proposed here.

Proposed complementary indicators

The complementary indicators were calculated using data from three sources: National Sanitation Information System (SNIS), a unit linked to the National Sanitation Secretariat of the Ministry of Regional Development; the Department of Information Technology of the Unified Health System (DATASUS), linked to the Ministry of Health and the Brazilian Institute of Geography and Statistics (IBGE). SNIS provides to the users an open microdata base with information on the provision of water and sewage services, urban solid waste management and drainage and management of urban rainwater, collected from municipal sanitation, while DATASUS provides mortality and survival data, demographic and socioeconomic information, as well as information on healthcare for the population, hospital and outpatient network records, among others. As national statistical office, IBGE provide data form surveys, census and registers, among others.

From SNIS, five indicators were selected for direct use and others were used to calculate new indicators. From DATASUS, data on people hospitalized due to waterborne diseases were taken and from IBGE data on the resident population were taken. All data used are available for all Brazilian municipalities, so it is possible to replicate the study presented here for other municipalities.

To illustrate the proposed indicators' contributions in monitoring target 6.1, we chose the city of Rio de Janeiro, given its importance in social and economic terms. It is the second most populous municipality in Brazil, with demographic density of 5,265.82 inhabitants/km² (IBGE Cidades, 2021). These indicators help to explain the challenges and pressure on the water supply service. The high demographic density reveals the particular form of occupation of urban space, which ultimately influences the effectiveness of the provision of basic sanitation services, due to the lack of physical space for the installation of public facilities. These data help to locate the study areas and understand the context in which the problem of access to water is inserted, as is typical of large metropolises.

The reference period of the study, 2016 to 2018, coincides with the year of implementation of the 2030 Agenda and the most recent year for which there were data released by all data sources as of this writing.

Complementary indicators

Eight complementary indicators are proposed, five of which are already calculated regularly by the SNIS, and three new ones are presented by us. They are, respectively: Total water service index (%) – IAA; Urban water service index (%) – IAU; Incidence of excessive residual chlorine (%) – ICR; Incidence of excessive turbidity (%) – IT; and Incidence of excessive total coliforms (%) – IC; in addition: Effective per capita consumption ratio for recommended consumption – RER; Urban-rural water supply ratio – RUR; and Proportion of people hospitalized due to waterborne diseases (%) – PDVH. Their composition is as follows:

Total water service index – IAA

This indicator refers to the population served by water supply services in the reference year, expressed as a percentage (%). Since target 6.1 refers to the population's access to drinking

water, the indicator in question is important for monitoring it, especially regarding the scope of water supply services.

$$IAA = \frac{PSW}{PRES} \times 100$$

Where:

PSW= Total population served by water supply.

PRES= Total resident population with water supply, according to IBGE.

Urban Water Service Index - IAU

This indicator shows the proportion of the population living in the urban area served by water supply services in the reference year in relation to the total urban population, expressed as a percentage (%). Like the "Water Service Index", this indicator defines the scope of water supply services, but it concerns more specifically the urban area. Therefore, this indicator can provide important information on possible inequalities in water supply services between rural and urban areas.

$$IAU = \frac{PUSW}{PURES} \times 100$$

Where:

PUSW = Urban population served by water supply.

PURES = Population living in urban area with water supply.

Quality of Water: ICR - Incidence of analyses with non-standard residual chlorine, IT - Incidence of excessive turbidity and IC – Incidence of excessive total coliforms

When it comes to safely managed drinking water, it is essential to value the good quality of the physical and chemical parameters of the water that reaches the population through the treatment and supply network. This makes it possible to investigate this issue through a set of indicators provided by the SNIS, namely: incidence of non-standard residual chlorine analyses, incidence of non-standard turbidity analyses and incidence of non-standard total coliform analyses, all expressed as a percentage (%). These indicators characterize the percentage of samples with a given physical and chemical parameter of the water outside the standard recommended by the Ministry of Health, through Ordinance 2,914/2011.

Incidence of analyses with non-standard residual chlorine – ICR

$$ICR = \frac{SWC}{STT} \times 100$$

Where:

SWC = Number of samples with non-standard residual chlorine

STC = Number of samples to test for residual chlorine.

Incidence of non-standard turbidity analyses – IT

$$IT = \frac{SWT}{STT} \times 100$$

Where:

SWT = Number of samples with non-standard turbidity.

STT = Number of samples to test turbidity.

Incidence of non-standard total coliform analyses – IC

$$IC = \frac{SWC}{STC} \times 100$$

Where:

SWC = Number of samples with non-standard total coliforms.

STC = Number of samples to test total coliforms.

Urban-rural water supply ratio – RUR

This indicator aims to point out possible differences in the provision of services to urban and rural populations. To this end, it shows the relationship between the proportion of people served by water supply services in urban centers and in locations outside these centers, that is, in rural areas. This indicator can help monitor actions that aim to ensure equitable access to drinking water for the general population, regardless of their area of residence. The indicator serves as a thermometer, indicating whether there is an imbalance in the provision of services according to the area, urban or rural. This indicator is derived from two other important indicators, one that measures rural provision and the other that measures urban provision, used here as numerator and denominator.

$$RUR = \frac{PUSW/PUR}{PRSW/PRR}$$

Where:

PUSW = Urban population served with water supply.

PUR = Population residing in urban areas.

PRSW = Urban population served with water supply.

PRR = Population residing in rural areas.

Effective consumption-recommended consumption ratio – RER

As previously stated, the average water consumption recommended by the UN per day is 110 liters per inhabitant. Thus, this indicator aims to measure how close the population of a given location is to the daily water consumption target and is another important indicator for monitoring the issue of accessible water for all, as dictated by target 6.1. The closer to 1, the closer the population's daily water consumption is to the recommended amount. However, if this number is much higher than 1, it is also necessary to evaluate the issue of water waste, since residents would be using much more water than recommended. The conscious use of water is one of the most important issues regarding the environment today, and it is important to monitor water consumption to avoid waste and other more serious problems related to the misuse of this resource, such as lack of water, pollution, among others.

$$\text{RER} = \frac{(\text{VWC} - \text{VWE})}{\text{PWS}} \times \frac{1.000.000}{365}$$

Where:

VWC = Volume of water consumed, expressed in 1000 m³/year.

VWE = Volume of treated water exported, expressed in 1000 m³/year.

PWS = Population living in an area with water supply,

Proportion of people hospitalized due to waterborne diseases (per 100.000 inhabitants) - PDVH

This indicator makes it possible to monitor the proportion of the population affected by the main waterborne diseases per 100,000 inhabitants. According to the National Health Foundation (2010), the main diseases related to direct transmission through contaminated water, through ingestion or contact are: cholera, typhoid and paratyphoid fevers, shigellosis, amoebiasis, diarrhea and gastroenteritis of presumed infectious origin, other infectious intestinal diseases, schistosomiasis, leptospirosis and hepatitis A. In addition, the indicator can also assist in another SDG of the 2030 Agenda, such as SDG 3 – Good Health and Well-Being -, in particular target 3.3 – By 2030, end the epidemics of AIDS, tuberculosis, malaria and neglected tropical diseases, and combat hepatitis, waterborne diseases, and other communicable diseases.

$$\text{PDVH} = \frac{\text{PHD}}{\text{PWS}} \times 100.000$$

Where:

PHD = Population hospitalized due to waterborne diseases.

PWS = Population living in an area with water supply.

The IAA and the IAU have equal proportions (Table 1). This is because the city of Rio de Janeiro has particularities that classify it as predominantly urban, with more than 50.000 inhabitants in an area of dense occupation and a high degree of urbanization (IBGE, 2017). In fact, when data from official surveys are analyzed, the rural population in the city of Rio de Janeiro is considered zero. The value of the IAA corresponds to the ratio between the total population served with water supply and the total population of the city. Thus, the drop in the proportion between 2016 and 2018 indicates that the municipality is moving away from reaching target 6.1. The impact of the increase in relative scarcity is worrying, since the lack of treated water can cause numerous problems for the population, including the increase in waterborne diseases.

Among the indicators that relate to the quality of water supplied to the population, only the ICR showed a decrease in its proportion (Table 1). The other two indicators, IT and IC, showed an increase. It should be noted that the concern should not be centered only on the proportion of the population with water supply service, but also on the quality and safety of the product. These indicators can be of great value for the assessment of the quality of the water supplied and, consequently, for the safety issue present in the target text.

Table 1: Proposed SNIS complementary indicators by year, Rio de Janeiro City

Indicators (%)	Reference Years		
	2016	2017	2018
IAA - Total water service index	99.0	99.2	97.4
IAU - Urban water service index	99.0	99.2	97.4
ICR - Incidence of excessive residual chlorine	5.8	0.7	0.5
IT - Incidence of excessive turbidity	8.0	6.9	10.9
IC – Incidence of excessive total coliforms	0.8	5.9	8.0

Source: National Sanitation Information System (SNIS).

The RER indicator shows how close the per capita consumption of the population is to the value recommended by the UN, i.e., 110 liters per day per inhabitant. The value considered adequate for this indicator is 1 or slightly above 1, which would indicate that the population is consuming the recommended amount of water or a bit more. However, the values of this indicator showed a different reality. During the three years analyzed, the RER presented values close to 3 (Table 2), which means that the per capita consumption was close to three times the recommended amount. This is a sign of excessive consumption and/or waste. It should be recalled that drinking water is a resource undergoing an accelerated process of depletion as a result of climate change, pollution and contamination of drinking water and soil degradation. Furthermore, this pattern of consumption can jeopardize equitable access to drinking water if the resource is scarce in the region. In this case, discussing the parsimony in the use of the resource is important. It is not acceptable for part of the population to have too much water supply, while another suffers from a lack of water. This adversity highlights the importance of the debate on how access to drinking water is associated with social inequality, occurring throughout the country. Due to the importance of this issue, more studies should be conducted to investigate the imbalance of per capita consumption comparing different regions, states or municipalities in Brazil. For this purpose, the RER is a useful indicator.

To add a new dimension to the issue, the RUR indicator can be used. It reveals possible differentials between urban and rural areas, by showing the relationship between the proportion of people with water supply services in urban areas and in rural areas. Rio de Janeiro does not have any rural areas, so the RUR indicator does not apply for our purposes here. However, this indicator can be calculated for other locations in Brazil with both urban and rural areas. Since there is a generally a smaller supply of piped water in rural areas, RER may point to consumption below the recommended level in such areas, while in urban areas the values can be above this threshold. To have a better picture, the RUR and RER should be examined jointly

The PDVH contributes to the assessment of the safety of the water consumed by the population. For its calculation, the main diseases related to direct transmission by contaminated water are considered, through ingestion or contact, according to the National Health Foundation (2010). These include cholera, typhoid and paratyphoid fevers, shigellosis, amoebiasis, diarrhea, infectious gastroenteritis, other infectious intestinal diseases, schistosomiasis, leptospirosis and hepatitis A. The indicator expresses the number of people hospitalized due to waterborne diseases, as recorded in DATASUS, per 100.000 inhabitants.

The PDVH remained close to zero and on a downward path during the period studied. However, even with the PDVH close to zero, it is noteworthy that there are still people in Rio de Janeiro who suffer from waterborne diseases. This issue must be analyzed especially from the perspective of the most vulnerable people, who have difficulty in accessing health services, basic sanitation and adequate food. In addition, it should be noted that the PDVH is an indicator based on the hospitalized population and not on the total number of people with waterborne diseases, which is likely higher due to underreporting. It is noteworthy that access to good-quality water is also present in the SDG 3: Health and Well-Being, target 3.3: “By 2030, end the

epidemics of AIDS, tuberculosis, malaria and neglected tropical diseases, and fight hepatitis, waterborne diseases, and other communicable diseases". This can also be measured using the PDVH indicator.

Table 2: Proposed new complementary indicators by year, Rio de Janeiro City

Indicators	Reference Years		
	2016	2017	2018
RUR - Urban-rural water supply ratio	NA	NA	NA
RER - Effective consumption-recommended consumption ratio	2,974	2,974	2,972
PDVH - Proportion of people hospitalized due to waterborne diseases (per 100.000 inhabitants)	0,0068	0,0059	0,0058

Source: prepared by the authors based on data from SNIS, IBGE and DATASUS.

Note: NA = Not applicable.

Discussion

Here we have proposed eight complementary indicators to assist monitoring SDG 6: Drinking Water and Sanitation, of the 2030 Agenda, in particular target 6.1 - By 2030, reach the universal and equitable access to safe drinking water for all. We presented for illustration the situation of the city of Rio de Janeiro from 2016 to 2018, along with contextual information on the situation in Brazil using indicator 6.1.1 as presented by the United Nations.

The proportion of population using safe managed drinking water services in Brazil, as measured by indicator 6.1.1, was close to 100%. Nevertheless, the absolute number of people left out in 2018 was above 3.8 million. To achieve universal access, governments at all levels must join efforts and invest more in public programs and policies, especially by observing the new Legal Framework for Basic Sanitation, which involves the commitment to the SDG 6.

Due to the socioeconomic disparity between Brazilian regions, analyzing what happens at the most disaggregated geographic levels is essential to understand the social reality experienced by the population. Therefore, the use of complementary indicators that can be compared with those of international sustainable development agendas and that allow monitoring their goals at the local level, as discussed here, are of great importance.

The illustration of Rio de Janeiro using eight complementary indicators proposed allowed analyzing particularities involving access to safely managed drinking water not covered by indicator 6.1.1 presented by the UN. In addition to the quantity of water supplied, the way it is delivered to the population, compliance with the recommended consumption and quality parameters can be calculated at local level. Therefore, they help to improve discussion within the scope of incentive policies for the implementation and monitoring of the SDGs at the local level and can be of great relevance to assist managers and authorities, in addition to encouraging new strategies to contribute to local development.

Data and variables related to access to drinking water were not fully explored in the present work. Therefore, more complementary indicators can be created to help monitor target 6.1. The SNIS database provides more variables and data for other municipalities in Brazil, which can be used as inputs to investigate differences in access to water between municipalities in the same state or comparisons between municipalities in different regions of Brazil. Such data can also be used to calculate complementary indicators to monitor, at disaggregated geographic levels, other targets of other SDGs, which can provide new information and indicate policies to contribute to sustainable development at local level.

Finally, given the complexity involved in the discussion of universal and equitable access to safe drinking water for all, the main contribution of this study is the proposal of indicators that can be calculated at the local level and contribute to the 2030 Agenda and the SDGs. It is

also noteworthy that the need for data production and monitoring of water access in Brazil goes far beyond the 2030 Agenda, as it promotes social control and monitoring of public policies, enabling justified public pressure for more and better services.

Acknowledgements

The first author acknowledges the support of the Office to Coordinate Improvement of Higher Education Personnel (CAPES) for the research that led to this paper. She also thanks the contribution of professors Letícia de Carvalho Giannella from the National School of Statistical Sciences (ENCE/IBGE) and Ana Lúcia Nogueira de Paiva Britto from Federal University of Rio de Janeiro - UFRJ.

The authors report there are no competing interests to declare.

References

- ANA – Agência Nacional das Águas. *ODS 6 no Brasil: Visão da ANA sobre os indicadores*. Brasília, DF. <https://www.gov.br/ana/pt-br/aceso-a-informacao/institucional/publicacoes/ods6#:~:text=A%20fim%20de%20contribuir%20para%20o%20monitoramento%20e,Brasil%20-%20Vis%C3%A3o%20da%20ANA%20sobre%20os%20indicadores>. Access: 19/11/2024.
- BRASIL. *Lei nº 14.026, de 15 de julho de 2020. Atualiza o marco legal do saneamento básico*. Diário Oficial da República Federativa do Brasil. https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2020/lei/114026.htm Access: 19/11/2024.
- CRUZ, Ana Clara Cancelo da. 2021. *Acesso à água potável no Brasil: proposta e aplicabilidade de indicadores complementares para auxiliar o monitoramento da meta 6.1 da Agenda 2030*. Dissertação de mestrado. Escola Nacional de Ciências Estatísticas. Available at: <https://ence.ibge.gov.br/images/ence/doc/mestrado/dissertacoes/2021/AnaClaraCancelo.pdf> Access: 19/11/2024.
- DATASUS - Brazilian Integrated System of Health. Available at: <http://www2.datasus.gov.br/DATASUS/index.php?area=0203> Access: 19/11/2024.
- FUNASA. *Impactos na saúde e no Sistema Único de Saúde decorrentes de agravos relacionados a um saneamento ambiental inadequado*. Available at: https://bvsms.saude.gov.br/bvs/publicacoes/impactos_saude_SUS_decorrentes_agravos_saneamento_ambiental_relatorio_final.pdf Access: 19/11/2024.
- IBGE, Instituto Brasileiro de Geografia e Estatística. 2017. *Classificação e caracterização dos espaços rurais e urbanos do Brasil: uma primeira aproximação*. Rio de Janeiro, Brasil.
- IBGE, Instituto Brasileiro de Geografia e Estatística. Brasil. Available at: <https://cidades.ibge.gov.br/> Access: 19/11/2024.
- IPEA, Instituto de Pesquisa Econômica e Aplicada. 2018. *2030 Agenda – ODS - Metas Nacionais dos Objetivos de Desenvolvimento Sustentável*. Available: https://www.ipea.gov.br/portal/index.php?option=com_content&view=article&id=33895&Itemid=433 Access: 19/11/2024.
- IPEA, Instituto de Pesquisa Econômica e Aplicada. 2019. *Cadernos ODS: ODS 6*.
- ODS Brasil. *Objetivos de Desenvolvimento Sustentável: Objetivo 6 - Água Potável e Saneamento*. Available at: <https://odsbrasil.gov.br/objetivo/objetivo?n=6> Access: 19/11/2024.
- ONU Brasil - Organização das Nações Unidas Brasil. *2030 Agenda*. Available at: <https://nacoesunidas.org/pos2015/> Access: 19/11/2024.
- OPAS - Organização Pan-Americana da Saúde. 2020. *Recomendações importantes sobre água, saneamento e higiene Covid – 19 na comunidade*. Available at: <https://www.paho.org/pt/documents/key-recommendations-water-sanitation-and-hygiene-covid-19-community> Access: 19/11/2024.
- SNIS – Sistema Nacional de Informações sobre Saneamento. SNIS. Outubro, 2019. Available at: <http://www.snis.gov.br/institucional> Access: 19/11/2024.
- SNIS – National System of Information on Sanitation. Brazil. Available at: <http://app4.mdr.gov.br/serieHistorica/> Access: 19/11/2024.
- STS, Sala Técnica de Saneamento. 2020. *Recomendações para prevenção do contágio da Covid-19 (Novo Coronavírus – SARS-COV-2) pela água e por esgoto doméstico*. Available at: https://comiteat.sp.gov.br/wp-content/uploads/2020/04/Cartilha_Sala_de_Saneamento_COVID-19.pdf.pdf Access: 19/11/2024.

Alternativas com base em dados abertos para o cálculo do Indicador ODS 11.7.1: Proporção da área construída das cidades que é espaço público aberto para uso de todos

Isis Gonçalves Peixoto*

Resumo

A crescente demanda por dados e as limitações das pesquisas tradicionais, como aumento das taxas de não resposta e altos custos, tornam necessárias fontes alternativas de informação. Entre essas fontes, destacam-se dados abertos, incluindo imagens de satélite e *big data*, especialmente relevantes para a construção de indicadores sociais, econômicos e ambientais. Este estudo foca no indicador ODS 11.7.1, que mede a proporção de áreas urbanas destinadas a espaços públicos abertos para uso universal, conforme a Agenda 2030 da ONU. No Brasil, esse indicador ainda não é calculado devido à ausência de dados, mas é essencial para avaliar a qualidade de vida e o bem-estar populacional. O trabalho propõe alternativas para seu cálculo, utilizando dados abertos e técnicas já empregadas com sucesso em outros países, oferecendo uma contribuição significativa para o planejamento urbano inclusivo e sustentável.

Palavras-chave: Objetivos de Desenvolvimento Sustentável, ODS 11, Agenda 2030, Indicador ODS.

* Mestre pelo programa de pós-graduação da ENCE.

Introdução

O problema, motivo deste artigo, está relacionado à falta de dados disponíveis para calcular o indicador dos Objetivos de Desenvolvimento Sustentável denominado ODS 11.7.1 no Brasil, que mede a proporção de áreas urbanas dedicadas a espaços públicos abertos e acessíveis para todos, conforme critérios de sexo, idade e deficiência. Este indicador é essencial para a avaliação da qualidade de vida nas cidades, mas sua ausência impede o monitoramento adequado da meta 11.7 da Agenda 2030. O objetivo central deste estudo é propor alternativas metodológicas para calcular esse indicador com base em dados abertos, utilizando como estudo de caso o município do Rio de Janeiro. Para isso, os objetivos específicos incluem: contextualizar o indicador ODS 11.7.1, explorar abordagens internacionais para seu cálculo, aplicar duas metodologias no caso do Rio de Janeiro — uma baseada em mapas de uso do solo e dados do OpenStreetMap®, e outra usando imagens de satélite e técnicas de classificação supervisionada — e comparar os resultados obtidos por essas abordagens.

A demanda crescente por dados conjugada com a dificuldade em realizar pesquisas de forma tradicional, em parte devido ao aumento das taxas de não resposta e dos custos que essas pesquisas possuem, trouxe a necessidade de se buscar fontes alternativas de dados, como é o caso de registros administrativos e fontes de *big data* (McFeely, 2019), em especial aquelas apresentadas na forma de dados abertos. Censos e pesquisas por amostragem são usualmente as principais fontes de dados para a produção estatística, porém possuem custo elevado e demandam operações mais complexas (Silva, 2018).

Assim, dados abertos são blocos de construção do conhecimento aberto (OKFN, 2022), ou seja, podem ser acessados de forma gratuita. Suas principais características são a disponibilidade conveniente e modificável, poder ser reutilizados, redistribuídos e legíveis por máquina, dessa maneira, são dados multivalentes e universais.

No fim dos anos 2000, dados abertos eram somente uma ideia que surgiu de uma crítica entre os defensores de código aberto e da democratização do acesso a dados. Esta crítica era direcionada especialmente aos governos que acumulavam dados valiosos, pagos pelos contribuintes, que poderiam ser bem utilizados se fossem disponibilizados, trazendo benefícios econômicos e mudanças sociais (Davies *et al*, 2019).

Logo, dados abertos estão associados com a ideia de que todos os dados devem estar livres e disponíveis para uso de todos. Dentro desse conceito existem diversas fontes e categorias, essas fontes podem ser do governo, universidades ou de empresas privadas, e esses dados podem ser provenientes de pesquisas, sensores ou imagens de satélite, por exemplo.

Com isso, a oportunidade de adquirir imagens de satélite gratuitamente, em diversas resoluções, abriu as portas para a sociedade se aproximar dessa alta tecnologia. Ao mesmo tempo revelou que para analisar esses dados provenientes de imagens, é necessário o desenvolvimento de poderosos algoritmos, que possibilitem a extração das informações nelas contidas. Atualmente sua utilização é essencial no agronegócio, planejamento urbano e outras áreas de interesse público (Zanotta e Ferreira, 2019).

Além dos temas já citados, cresce a necessidade de construir indicadores sociais, econômicos, ambientais, dentre os quais destacam-se aqueles relacionados à Agenda 2030 da ONU, onde é possível elencar inúmeros desses indicadores que ainda não podem ser calculados por falta de dados. Por essa razão, explorar dados abertos como os encontrados na forma de *big data*, incluindo imagens de satélite, é tão relevante. Em especial, quando se trata da obtenção de informações para a construção e manutenção de políticas públicas frente à dificuldade de obtenção de dados em fontes tradicionais, como pesquisas amostrais e censos (SAVIO *et al*, 2020).

Espaço público aberto, segundo o *Guidebook for City Leaders* (UN-Habitat, 2020), são áreas urbanas acessíveis ao uso pela população, com ou sem edificações, que devem estar disponíveis gratuitamente para todos. Esses espaços incluem áreas de lazer, como parques e praças, bem como áreas de circulação, como ruas e calçadas.

No Brasil, o indicador ODS 11.7.1 mede a proporção da área construída nas cidades que é dedicada a esses espaços públicos abertos, considerando critérios de sexo, idade e pessoas com deficiência. A meta 11.7 visa, até 2030, garantir acesso universal a espaços públicos seguros,

inclusivos, acessíveis e verdes, com especial atenção a mulheres, crianças, idosos e pessoas com deficiência.

A importância desse indicador está no valor dos espaços públicos abertos para a qualidade de vida e bem-estar da população, que, infelizmente, é com frequência desprivilegiado quando da elaboração de políticas públicas e do planejamento urbano, seja por falta de recursos ou capacidade técnica de criar um espaço urbano multifuncional. Apesar disso, a Agenda 2030 se constituiu em um instrumento para que o aumento da oferta de tais espaços públicos possa ser monitorado em escala global. É como esse indicador mede o terreno alocado para espaços públicos abertos e a população que tem acesso a esses espaços por idade, gênero e deficiência, contribui para as dimensões sociais, de saúde e qualidade de vida da população, se mostrando um excelente indicador de prosperidade compartilhada (UN-Habitat, 2018).

Finalmente, em 2018, o Indicador 11.7.1 junto com outros 12 indicadores foram reclassificados de *Tier III* para *Tier II*, para indicadores conceitualmente claros, com método estabelecido, mas que não possuem produção de dados regular pelos países.

Adicionalmente, como o Brasil ainda não produz esse indicador por falta de dados (ODS-BRASIL, 2022b), a relevância desse trabalho se torna ampliada, pois provê alternativas para o seu cálculo com base no uso de dados abertos e com métodos já experimentados com sucesso em outros países. Por fim até 9 de agosto de 2022, o IBGE, produziu um gráfico com os resultados atualizados para o Brasil. Dos 254 indicadores atuais, 111 foram produzidos, 79 estão em análise ou construção, 54 não possuem dados e 10 não se aplicam ao país (ODSBRASIL, 2022a).

A relevância desse trabalho se dá porque, no Brasil, os dados necessários para o cálculo dos indicadores da meta 11.7 ainda não estão disponíveis (ODSBRASIL, 2022a). A ausência desses dados impede uma avaliação precisa do progresso em direção ao ODS 11, que busca cidades e comunidades mais inclusivas, seguras, resilientes e sustentáveis. Ao fornecer uma metodologia alternativa para o cálculo do Indicador ODS 11.7.1, este estudo contribui para a literatura ao preencher essa lacuna de informações e oferece uma base importante para decisões que promovam o desenvolvimento urbano sustentável.

Este artigo está estruturado em cinco seções principais, além da introdução e das considerações finais. A próxima seção aborda os fundamentos teóricos e conceituais do Indicador ODS 11.7.1, destacando a importância dos espaços públicos nas áreas urbanas. Em seguida, são apresentados os materiais e métodos, com ênfase nas fontes de dados e ferramentas empregadas para o cálculo do indicador. A quarta seção aplica a metodologia no município do Rio de Janeiro. Finalmente, a seção de conclusões resume os resultados obtidos e discute as implicações da pesquisa.

Materiais e métodos

Com base na experiência dos países pode-se concluir que fontes de imagens de satélite abertas como OpenStreetMap¹ e Google Earth² são essenciais para minimizar o tempo de processar essas imagens e úteis na identificação de espaços públicos. Essas imagens de satélite quando possível, devem ser do mesmo ano que os dados populacionais (UN, 2018). Já os registros administrativos, são bastante úteis para alguns países calcularem seus dados populacionais e a geocodificação desses registros pode ajudar no cálculo das proporções populacionais por área. Ao mesmo tempo, dados sobre cadastros de terra e classificação do uso da terra se mostraram necessários (UN, 2018).

Por fim, os metadados são essenciais para assegurar a adequação dos mesmos a um objetivo específico. É crucial ressaltar que metadados não envolvem recursos voltados para realizar cálculos, mas sim descrições que esclarecem a natureza e a procedência das informações. Para apoiar a gestão eficiente desses metadados, existem diversas ferramentas e softwares especializados. Além disso, é primordial que as versões e atualizações publicadas pela Divisão de Es-

¹ <https://www.openstreetmap.org/#map=18/-22.508410/-41.921430&layers=D>

² <https://www.google.com.br/earth/>

tatística das Nações Unidas (UNSD) e pelos institutos nacionais de estatística estejam sincronizadas. Essa harmonização garante que os usuários tenham acesso a alternativas consistentes e igualmente atualizadas, evitando discrepâncias na interpretação das informações (UN-Habitat, 2018).

Na subseção a seguir, será apresentada metodologia para o cálculo do Indicador ODS 11.7.1, que mensura a proporção de espaço público aberto nas cidades, somando a área de ruas, calçadas e outros espaços públicos acessíveis gratuitamente, em relação à área urbana total. A unidade de análise escolhida neste estudo é a cidade do Rio de Janeiro, cuja complexa estrutura administrativa e distribuição populacional permitem uma avaliação detalhada da disponibilidade de espaços públicos abertos. Neste trabalho não serão calculadas as áreas de serviço, apenas a proporção de espaço público aberto.

Indicador 11.7.1 – Módulo de Treinamento

O módulo de treinamento da UN-Habitat do Indicador SDG 11.7.1 (UN-Habitat, 2018), define alguns conceitos de modo a facilitar a construção desse indicador pelas cidades e países.

Esse indicador adota como conceito de Cidade a entidade operacional que incorpora áreas construídas e abertas, os espaços, que muitas vezes se estendem além dos limites administrativos formais. Essa definição, segue o conceito de “extensão urbana” desenvolvida pela Universidade de Nova Iorque, que foi elaborado com base em uma amostra global de 200 cidades (UN-Habitat, 2018).

Área Construída é a porção contínua de uma cidade ocupada por edifícios e outras superfícies impermeáveis, ou sua área funcional. Já o Espaço Público abrange locais de propriedade ou uso público, acessíveis a todos gratuitamente, e é classificado em quatro tipos: ruas, espaços abertos, equipamentos públicos e mercados. O cálculo do indicador envolve somar as áreas destinadas a ruas, calçadas e outros espaços abertos, como parques e praças, e calcular sua proporção em relação à área urbana total da cidade (UN-Habitat, 2018).

A identificação de espaços públicos abertos entre as cidades pode ser feita por meio de, entre outras formas, análise de imagens de satélite de alta a muito alta resolução, de mapas básicos fornecidos por diferentes organizações, ou por fontes de crowdsourcing, termo recente utilizado para denominar colaboração coletiva, como por exemplo OpenStreetMap®, que é desenvolvido por uma comunidade voluntária de mapeadores (OSM, 2022).

Embora essas fontes forneçam dados de linha de base importantes para o Indicador SDG 11.7.1, alguns dos espaços identificáveis podem não atender aos critérios de serem acessíveis ao público gratuitamente.

O termo espaço público aberto é, portanto, usado para se referir a espaços públicos abertos que são extraídos das fontes acima mencionadas com base em seu caráter espacial, mas que ainda não foram validados para confirmar se são acessíveis ao público gratuitamente.

Para obter a proporção de área da cidade alocada para ruas e calçadas - PEPARC, a área total de ruas e calçadas foi dividida pela área urbana total e multiplicada por 100, conforme (1).

$$PEPA_{RC} = \left[\frac{EPA_{RC}}{\text{Área urbana total}} \right] \times 100 \quad (1)$$

Onde:

$PEPA_{RC}$ = Proporção de área da cidade alocada para ruas e calçadas;

EPA_{RC} = Área da cidade alocada para ruas e calçadas.

Ao contrário do indicador de áreas alocadas para ruas e calçadas, que pode ser calculado exclusivamente com base em imagens de satélite e análise espacial, a coleta de dados sobre a quantidade de áreas urbanas alocada para espaços públicos abertos usa uma combinação de interpretação de imagem e validação por especialista (UN-Habitat, 2018).

Os Espaços Públicos Abertos (EPA), no contexto deste indicador, compreendem os terrenos não urbanizados, ou com o mínimo de edifícios, que são acessíveis ao público e que for-

necem áreas de lazer aos residentes e ajudam a valorizar a beleza e a qualidade ambiental dos bairros (UN-Habitat, 2018).

Para calcular a área de EPA, são somadas a área total de ruas e calçadas (EPARC) e a área de outros espaços públicos abertos (EPAO), divididos pela área urbana total, conforme (2).

$$EPA = \left[\frac{EPA_{RC} + EPA_O}{\text{Área urbana total}} \right] \quad (2)$$

Onde:

EPA = Espaço público aberto;

EPA_{RC} = Área da cidade alocada para ruas e calçadas;

EPA_O = Área da cidade alocada para outros espaços públicos abertos.

Para calcular o indicador principal, a Proporção de área construída de cidades que é o Espaço Público Aberto para uso de todos (PEPA), divide-se a EPA pela área urbana total e multiplica-se por 100, conforme (3).

$$PEPA = \left[\frac{EPA}{\text{Área urbana total}} \right] \times 100 \quad (3)$$

Onde:

PEPA = Proporção de espaço público aberto – Indicador ODS 11.7.1;

EPA = Espaço público aberto.

Este indicador ODS é a PEPA para regiões delimitadas com um raio de 400m uma vez que a identificação das populações com acesso aos espaços abertos públicos baseia-se no argumento genérico de que, pessoas que vivem dentro de uma distância pré-definida de caminhada para um determinado espaço e podem acessar facilmente e usá-lo sem quaisquer restrições. A distância considerada aceitável para ser percorrida a pé foi definida, pela ONU-Habitat, como sendo de 400m, ou 5min de caminhada. Com base nisso uma área de serviço baseada em rede viária é desenhada ao redor de cada espaço aberto público, usando a soleira de acesso de 400 metros (UN-Habitat, 2018).

Unidade de análise

Segundo os resultados preliminares do Censo Demográfico Brasileiro de 2022, a cidade do Rio de Janeiro abriga uma população de 6.625.849 habitantes, consolidando-se como a segunda maior cidade do país, atrás apenas de São Paulo, e com uma área total de 1.255,3 Km² (IBGE, 2022).

Em 1961, foi iniciada a divisão administrativa do município do Rio de Janeiro, por meio da criação de regiões que visavam coordenar as atividades dos estabelecimentos e serviços de natureza local ou distrital. Desde então, ocorreram diversas modificações na divisão administrativa da cidade, culminando na atual estrutura que conta com 5 áreas de Planejamento, 16 Regiões de Planejamento, 33 Regiões Administrativas e 163 Bairros (Rio de Janeiro, 2020). Até o momento, não foram registradas mais alterações nos limites administrativos da cidade do Rio de Janeiro.

Segundo o Plano Diretor de Desenvolvimento Urbano Sustentável do Município do Rio de Janeiro, Área de Planejamento (AP) é uma unidade territorial para efeitos de planejamento e desenvolvimento urbano. O município é dividido em 5 Áreas de Planejamento a partir de critérios de compartimentação ambiental, de características histórico-geográficas e de uso e ocupação do solo. Comparativamente, as 5 APs correspondem as seguintes Zonas: AP1 a Zona Central, AP2 engloba a Zona Sul e a Tijuca, AP3 a Zona Norte, AP4 compreende as regiões administrativas da Barra da Tijuca, Jacarepaguá e Cidade de Deus, e a AP5 abrange os demais bairros da Zona Oeste (Rio de Janeiro, 2008).

A Área de Planejamento 1 é constituída por 15 bairros, com cerca de 4,6% da população carioca e 2,8% do território. Corresponde ao centro histórico da cidade e é a região que mais sofreu com transformações urbanas na cidade (Rio de Janeiro, 2006).

A Área de Planejamento 2 é formada por 25 bairros, representa 17% dos habitantes e ocupa 8,2% do território. Corresponde à área de expansão da cidade que ocorreu decorrente do estabelecimento do sistema de bondes no século XIX. Localizada entre o mar e o Maciço da Tijuca, é a região mais conhecida como a imagem da Cidade do Rio de Janeiro (Rio de Janeiro, 2006).

A Área de Planejamento 3 possui 80 bairros e 16,6% do território municipal, além de 40,2% do total da população residente no Rio de Janeiro. Por um longo período, era um espaço com pouca identidade, comandado pela perspectiva de expansão territorial. A construção da Linha Amarela, em 1997, melhorou a conexão da região, porém a via expressa segmentou os bairros cortados por ela (Rio de Janeiro, 2006).

Já a Área de Planejamento 4 é constituída por 19 bairros e corresponde a 24% da área e 11,6% da população do município. Sua configuração geográfica é limitada pelos maciços da Tijuca e Pedra Branca, e pelo Oceano Atlântico. Durante muito tempo não estava inserida na malha urbana em função das características geográficas, garantindo um caráter mais rural com predominância de sítios e chácaras (Rio de Janeiro, 2006).

Por fim, a Área de Planejamento 5 é formada por 20 bairros, e possui 48,4% do território do município e abriga cerca de 26,6% da população carioca. A Zona Oeste, como é conhecida, foi tratada como última fronteira da urbanização do Rio de Janeiro. Nela, foram mantidos, durante muito tempo, os usos agrícolas e as extensas propriedades, que foram se extinguindo com a pressão da urbanização, a partir da década de 1960 (Rio de Janeiro, 2006).

Abordagem por meio de mapas

O Instituto Pereira Passos (IPP) se especializou em projetos de desenvolvimento econômico, coordenando diversos projetos urbanos como: o Rio Cidade, o Favela Bairro, o Rio Orla e o Porto Maravilha. É referência em fornecimento de dados e conhecimento sobre a cidade do Rio de Janeiro e largamente consultado para formulação de políticas públicas. Esses dados são abertos para uso público por meio do site Armazém de Dados³ criado em 2001 (Rio de Janeiro, 2022).

Neste trabalho foram consultados diversos mapas do Armazém de Dados. Os principais mapas que orientaram a produção do indicador são os mapas de edificações, lotes e quadras; o mapa de uso do solo; os mapas de limite municipal e limite dos bairros do município; e os mapas de cobertura vegetal e uso da terra. Além da alta qualidade dos mapas produzidos pelo IPP⁴, estes dados e informações estão em constante atualização, portanto refletindo a realidade presente.

Dados

Os dados de entrada foram os arquivos com os limites do município, imagens de alta resolução cobrindo a mesma área e dados do OpenStreetMap⁵, sobre o uso da terra e nomes de lugares. Foi realizada uma análise de caráter dos espaços como tamanho, forma e cobertura do solo através do mesmo banco de imagens. Os mapas provenientes do IPP são atualizados com frequência no mínimo anual, logo esses mapas refletem a dinâmica urbana da cidade, assim como as imagens de satélite. Se replicado em outras localidades deve-se levar em conta a velocidade do crescimento urbano e das transformações urbanísticas.

Os mapas utilizados para o cálculo do indicador extraídos do Data.Rio do IPP foram:

- Mapas de Edificações, lotes e quadras de 2013, atualizados em 2022;
- Mapa de Uso do Solo de 2019, atualizado em 2022;

³ <https://apps.data.rio/armazenzinho/>

⁴ <https://www.data.rio/>

⁵ <https://www.openstreetmap.org/#map=18/-22.50841/-41.92143&layers=D>

- Mapa de Limite Municipal e de Bairros do Município do Rio de Janeiro, atualizados em 2022;
- Mapa de Cobertura Vegetal e Uso da Terra de 2018, atualizado em 2022.

Complementarmente, foram utilizadas imagens de satélite do Google Earth® de 2021 e mapas do OpenStreetMap® de 2021 com as referências de nomes de lugares caracterizados como Espaços Públicos Abertos (EPA).

O recorte geográfico foi a área urbana do município do Rio de Janeiro, a qual inclui a totalidade do município, já que este é composto apenas por área urbana, segundo o Plano Diretor (Rio de Janeiro, 2011).

Os dados populacionais utilizados são do Censo 2010. Dessa fonte, foram utilizados: total da população do município, quantidade de mulheres na população, quantidade de crianças menores de 15 anos, quantidade de pessoas com idades entre 15 e 60 anos, quantidade de pessoas maiores de 60 anos e quantidade de pessoas com pelo menos uma deficiência (IBGE, 2010a).

Calculando o Indicador SDG 11.7.1 por meio de mapas

As fontes de dados de ruas e espaços abertos da cidade usadas neste trabalho, são talvez, as fontes mais precisas para o cálculo desse indicador. Foram usados mapas da malha urbana disponibilizados pelo Instituto Pereira Passos (IPP), que é o Instituto de Pesquisa do Governo da Cidade do Rio de Janeiro, responsável pela produção de informações sobre o município e por sua divulgação. O mapa da malha urbana da cidade do Rio de Janeiro foi adquirido no portal Armazém de Dados⁶ do Data.Rio⁷ (Armazém, 2021).

Também foram utilizados dados de imagens aéreas do Google Earth® e OpenStreetMap® para comparar a configuração da malha urbana proveniente dos mapas com as imagens de satélite atuais e efetuar correções, tais como a incorporação de algumas ruas e loteamentos recentes. Estes dados serão apresentados na próxima seção.

Na análise realizada, os cálculos de área foram corrigidos para considerar uma projeção cartográfica equivalente, garantindo maior precisão na medição das superfícies. Inicialmente, os dados obtidos estavam em coordenadas geográficas e na projeção UTM, respectivamente, ambas inadequadas para cálculos de área devido às distorções inerentes. Para assegurar resultados mais precisos, os dados foram projetados para o sistema de projeção policônica brasileira, amplamente utilizado para mapeamentos oficiais no Brasil, garantindo que as áreas calculadas reflitam corretamente as dimensões reais do território (Dutra *et al*, 2017).

Será considerado que, todas as pessoas que vivem nas áreas de serviço são, por sua vez, identificadas como tendo acesso conveniente aos espaços públicos abertos. Todos os edifícios dentro da área de serviço são habitáveis, e que a população está igualmente distribuída em todos os edifícios/áreas edificadas. A população não foi delimitada por áreas de serviço como sugere no Manual da UN-Habitat, mas sim dividida por Áreas de Planejamento (AP), pois elas refletem a dinâmica da cidade e as diferenças territoriais, e de modo a delimitar parcelas populacionais adequadas à dimensão do objeto de estudo.

O trabalho foi realizado em nove etapas, iniciando com a delimitação da área de estudo, busca por mapas de malha municipal em fontes de dados abertos, busca por nomes e referências de lugares, inspeção visual com imagens de satélite, elaboração do mapa de ruas e calçadas, cálculo da área de Espaços Públicos Abertos (EPA), elaboração do mapa com os EPA's, associação de dados populacionais para cada área de planejamento do município, e por fim, cálculo do Indicador SDG 11.7.1.

⁶ <https://apps.data.rio/armazenzinho/>

⁷ <https://www.data.rio/>

⁸ <https://www.openstreetmap.org/#map=18/-22.50841/-41.92143&layers=D>

Abordagem por meio de classificação de imagens de satélite e amostragem

Nesta seção são mostrados os resultados da aplicação da técnica baseada no uso de mapas para a obtenção do Indicador ODS 11.7.1 para o município do Rio de Janeiro. São apresentados os resultados obtidos utilizando mapas de uso do solo disponibilizados publicamente pela Prefeitura Municipal e mapas obtidos em fontes de dados abertos como o OpenStreetMap® e o Google Earth®.

Dados

Também foram utilizados dados do Google Earth® e do Google Maps® por meio do QGIS® versão 3.22.10, devido à sua alta resolução e às informações atribuídas a cada local, como nomes de praças, escolas, condomínios, para colaborar na distinção de áreas públicas das privadas. A frequência da atualização das imagens do Google Earth® é anual e as informações atribuídas ao Google Maps® são atualizadas em 24h após a aprovação dessas mudanças (Moore, 2021).

Cadastro de Área

Uma das principais fontes de erro não-amstral está no uso de cadastros que apresentam problemas de cobertura da população. Uma boa cobertura pode ser obtida por meio de cadastros construídos com base em um Censo populacional recente. Todavia, quando a população-alvo está relacionada ao uso do solo, o território pode ser usado como suporte para construir um cadastro de área com cobertura completa.

Neste trabalho o cadastro de área é construído pela autora, pela sobreposição de uma grade regular de segmentos quadrados no território de interesse, no caso, o município do Rio de Janeiro. Optou-se por usar segmentos quadrados de dimensão 250m por 250m, ou seja, quadrados de 62.500m² ou 0,625km² de área, que continham mais de 50% da sua área dentro do limite municipal. Ao todo foram necessários 19.244 quadrados para cobrir todo o território do município.

Cadastros de área de segmentos quadrados tem sido aplicado com eficiência em pesquisas agropecuárias por amostragem probabilística (Gallego, 1995; Torres, 2018). Essa abordagem é vantajosa já que cadastros de baixo custo e alta cobertura, provenientes de censos sejam agropecuários ou populacionais, têm suas informações atualizadas em intervalos demasiadamente longos. E como esse trabalho inclui elementos territoriais e populacionais os cadastros de área de segmentos quadrados são um valioso recurso utilizado nesta abordagem.

Amostragem

Para obter informações sobre a população, utilizamos uma amostra probabilística. Neste trabalho, a área-alvo é o município do Rio de Janeiro, devido à sua extensão, foi realizada uma classificação de imagens de satélite com uma amostra aleatória simples de 19.244 segmentos quadrados. Para estimar a área dos espaços públicos abertos (EPA), uma amostra aleatória desses segmentos é analisada, e a área correspondente às EPAs é identificada, gerando uma estimativa da proporção de espaços públicos abertos na cidade.

Para o cálculo da amostra em uma população finita com intervalo de confiança de 95%, foi utilizada a seguinte fórmula (4).

$$n = \frac{N}{1 + \frac{N \times d^2}{Z^2 \times p \times q}} \quad (4)$$

Onde:

n = número de elementos da amostra;

N = número de elementos da população;

Z = valor da abscissa da curva normal associada ao nível de confiança fixado;

d = erro tolerável da amostra;

p e q = proporção de se escolher um quadrado aleatoriamente.

Para a tomada de decisão, foi elaborado um gráfico com tamanhos da amostra segundo a margem de erro esperada, para uma população de 19.244 (N), proporção de 0,5, intervalo de confiança fixado em 95%. Foram calculados tamanhos de amostra para os erros 0,01; 0,05; 0,10; 0,15; 0,20 e 0,25.

Considerando os recursos disponíveis para essa pesquisa e também que em classificação de imagens de satélite o ideal seja que o tamanho do erro não ultrapasse 10%, foi escolhido trabalhar com esse valor (GONG, P. *et al*, 2019). Logo, considerando tal erro, o tamanho da amostra foi de 97 unidades elementares (n) em uma população de 19.244 (N).

Os 97 segmentos quadrados da imagem de satélite, foram extraídos diretamente no software QGIS® onde a classificação dessas imagens foi realizada de forma visual, delimitando as áreas de ruas e calçadas, áreas de outros espaços públicos abertos e áreas que não se enquadravam em nenhuma das duas alternativas.

Para estimar a área total de EPAs no município foi utilizado o estimador Horvitz-Thompson (HT), que presume que um universo ou uma população sejam formados por N elementos, e estima o total populacional a partir de uma variável de interesse (Y), para amostras com “n” unidades amostrais selecionadas. É um estimador não viesado do total populacional.

$$\widehat{Y}_{HT} = \sum_{i=1}^n \pi_i^{-1} Y_i \quad (5)$$

Onde:

\widehat{Y}_{HT} = Estimador para o total;

π_i = Probabilidade de inclusão na amostra do i-ésimo elemento;

Y_i = Valor da variável de interesse no i-ésimo segmento;

n = Número de elementos da amostra.

Consideramos que o valor Y_i representa o valor da variável de interesse (EPA, EPARC ou EPAO) no i-ésimo segmento quadrado, conforme (5). O peso amostral é o inverso da probabilidade de inclusão do segmento na amostra. Dessa forma o município é totalmente coberto por N segmentos e a probabilidade de seleção de qualquer elemento é n/N (Torres, 2018).

Calculando o Indicador SDG 11.7.1 por meio de classificação de imagens de satélite por amostragem

A elaboração de mapas de uso e cobertura do solo, por meio de imagens geoespaciais é realizada por classificação automática ou análise visual. Essa interpretação visual é baseada no conhecimento técnico do especialista para identificar padrões da imagem, são utilizadas informações como cor, forma e textura. Porém a classificação automática é mais usada em larga escala, já que a análise visual é mais cara e demorada (Almeida *et al*, 2014).

O mapeamento de cobertura terrestre e uso do solo pode ser por vetorização de forma visual, ou automatizado, por meio de algoritmos supervisionados ou não supervisionados. Na forma visual, se interpreta as informações e atribui às imagens de satélite, manualmente as classes de interesse de cada porção territorial. Essa técnica é eficiente, porém quando aplicada a grandes áreas demanda muito mais tempo que a automática (Bravo e Sluter, 2015).

Já a classificação automática pode ser supervisionada ou não supervisionada, e depende das informações disponíveis, pois exige informações preliminares da área estudada como por exemplo, os nomes e usos dos lugares. Essas informações agregam-se ao processo de classificação e fazem parte das regras desse sistema de informação geográfica (Almeida *et al*, 2014). Neste trabalho a classificação foi realizada de forma visual em uma amostra de segmentos da

área estudada, devido à dificuldade em se conseguir classificar automaticamente áreas que são públicas ou privadas, característica fundamental neste indicador (Figura 1).

Figura 1 - Amostra de imagem de satélite classificada (quadrado 11.653)



Fonte: Elaborado pela autora.

O fluxo deste trabalho se constitui de nove etapas, começando pela delimitação da área de estudo, seguida pela busca por imagens de satélite em bancos de dados abertos. Na sequência, as amostras foram selecionadas e foi feita a classificação das áreas identificadas como espaços públicos abertos (EPA), desagregados por ruas e calçadas (EPARC) e outros espaços públicos abertos (EPAO), e áreas Não EPAs, e a avaliação dos resultados. Foram estimados os resultados para a área total do município e por áreas de planejamento com suas respectivas proporções populacionais. Por fim foi estimado o indicador para a área total do município.

Resultados e discussão

Para o município do Rio de Janeiro, foram identificadas a área total do terreno da cidade, as partes alocadas para ruas e calçadas e as partes alocadas para Espaço Público Aberto (EPA). Em seguida foi calculado o Indicador ODS 11.7.1 - PEPA, a qual representa a Proporção da área que é Espaço Público Aberto para uso de todos, os resultados estão apresentados na Tabela 1.

Tabela 1: Áreas e proporções de áreas da cidade, total e por tipo de uso – Rio de Janeiro

Descrição	Abordagem Mapas	Abordagem Imagens Satélite	ONU-Habitat
Proporção da EPA – PEPA, Indicador ODS 11.7.1	50,25%	53,44%	45-50%
Proporção da EPARC – PEPARC	18,20%	8,35%	30-35%
Proporção da EPAO – PEPAO	32,06%	45,09%	15-20%

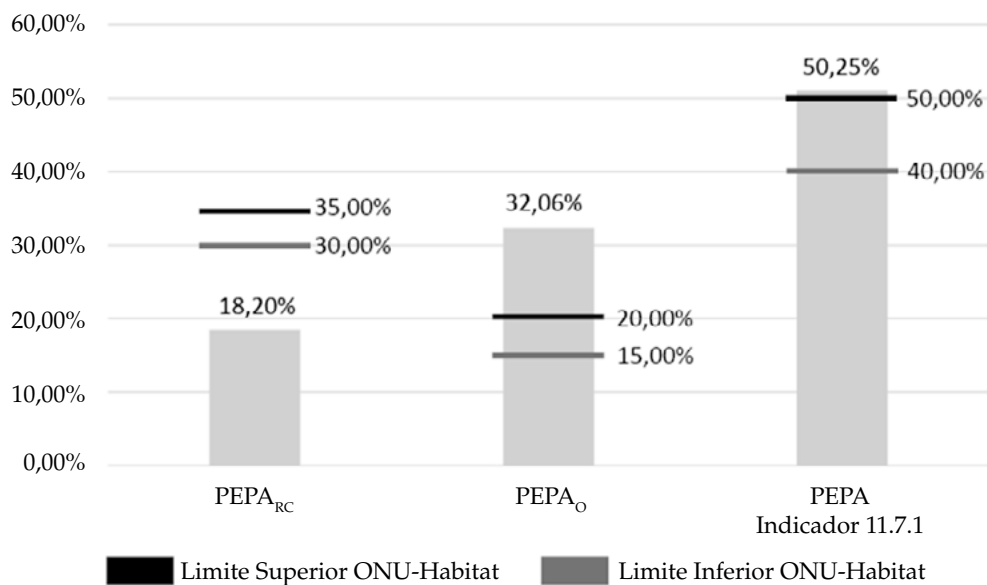
Fonte: Elaborada pela autora, utilizando dados e mapas do IPP e do OpenStreetMap® - 2021.

De acordo com os resultados apresentados na Tabela 1, para a abordagem com uso de mapas, observa-se que a proporção de terrenos alocados para ruas e calçadas (PEPARC) foi de 18,20%, o que representa uma porcentagem inferior à recomendação estabelecida pela ONU-Habitat, a qual sugere uma alocação de terrenos nessa categoria situada entre 30% e 35%. Em contrapartida, a proporção de terrenos destinados a outros espaços públicos abertos (PEPAO),

tais como praças, parques e áreas de lazer, foi de 32,06%, superando, assim, a recomendação mínima de 15% a 20% estabelecida pela mesma organização.

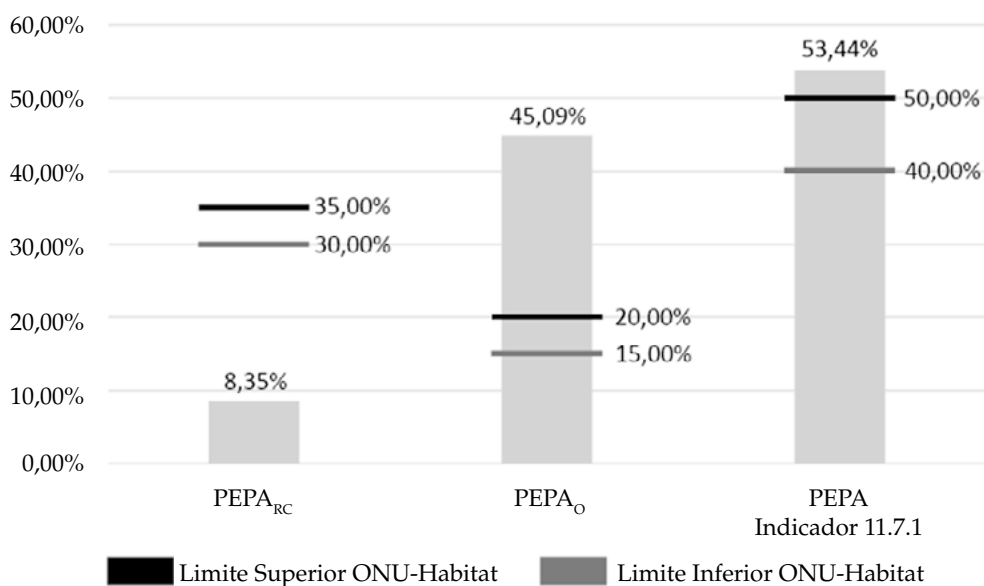
Já o Indicador ODS 11.7.1 - PEPA engloba tanto as áreas destinadas a ruas e calçadas quanto as demais áreas públicas abertas. No caso do município do Rio de Janeiro, o valor calculado para PEPA atingiu 50,25%, ultrapassando o limite mínimo recomendado de 45% a 50% estipulado pela ONU-Habitat conforme mostra o Gráfico 1.

Gráfico 1 - Proporções de espaços públicos abertos para o Município do Rio de Janeiro e Valores de Referência da ONU-Habitat – 2022 – Abordagem Mapas



Fonte: Elaborado pela autora, utilizando dados dos mapas do IPP, OpenStreetMap® e Metadados do Indicador, 2022.

Por outro lado, ao se considerar o Indicador ODS 11.7.1 - PEPA na segunda abordagem, por meio de imagens satélite, observa-se um resultado de 53,44% em relação à área total da cidade. Esse resultado assemelha-se ao obtido na primeira abordagem, apresentando valores ligeiramente acima do recomendado pela ONU-Habitat. No entanto, é importante destacar que as proporções de terrenos alocados para ruas e calçadas (PEPARC) e para outros espaços públicos abertos (PEPAO) (8,35% e 45,09%, respectivamente) diferem dos obtidos na primeira abordagem, bem como das recomendações da ONU-Habitat conforme indicado no Gráfico 2.

Gráfico 2 - Proporções de espaços públicos abertos para o Município do Rio de Janeiro e Valores de Referência da ONU-Habitat – 2022 – Abordagem Imagens de Satélite

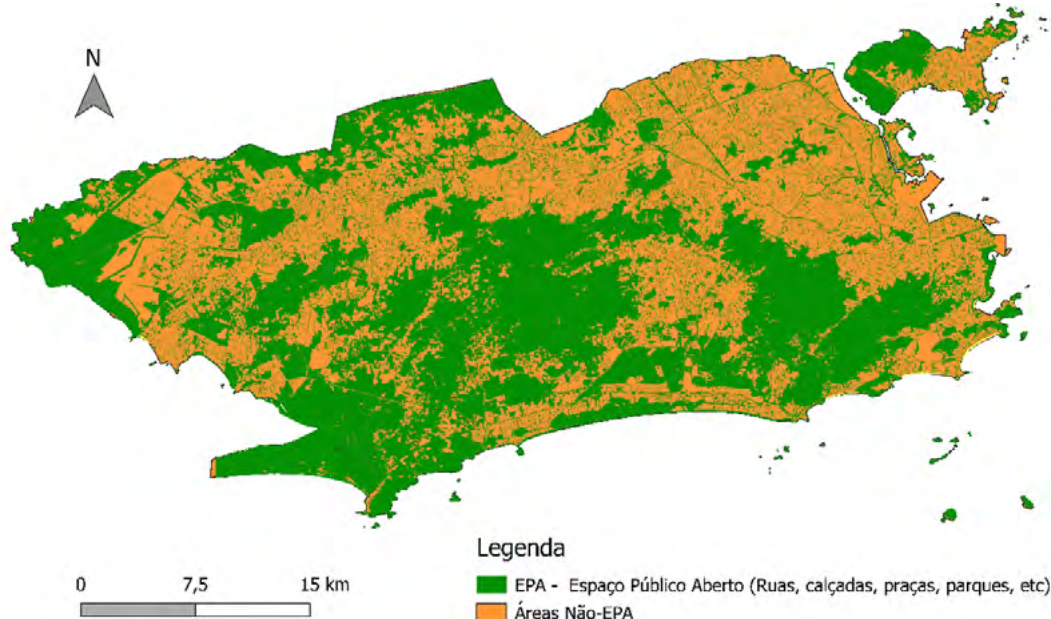
Fonte: Elaborado pela autora.

A discrepância identificada em relação à alocação de terrenos destinados a ruas e calçadas em detrimento das áreas públicas abertas, como parques e áreas verdes, pode ser vista como uma vantagem para os residentes da região. Este cenário proporciona benefícios consideráveis, incluindo um maior acesso a espaços públicos de lazer e recreação. Nesse contexto, é notável que a proporção de Espaços Públicos Abertos (PEPA) se apresenta de forma bastante satisfatória quando comparada às diretrizes estabelecidas pela ONU-Habitat.

Este sucesso na alocação de espaços públicos abertos é exemplificado pelo reconhecimento conferido ao Rio de Janeiro em 2021 pela Organização para Alimentação e Agricultura das Nações Unidas (FAO-ONU) e pela Arbor Day Foundation (ADF) como uma das "Cidades das Árvores" (Tree Cities of the World). Esse selo é concedido às cidades que demonstram comprometimento com a preservação e expansão da arborização urbana (ADF, 2022).

Adicionalmente ao indicador utilizado para avaliar o município, vale ressaltar a elaboração de um mapa abrangente que representa os espaços públicos abertos na cidade. Esse mapa foi desenvolvido com base nas informações cartográficas fornecidas pelo Instituto Pereira Passos (IPP), contribuindo significativamente para a compreensão e planejamento dos recursos de áreas públicas disponíveis na cidade, conforme a Figura 2.

Figura 2: Espaço Público Aberto (EPA), por Área de Planejamento (AP), Município do Rio de Janeiro - 2021



Fonte: Elaborada pela autora, utilizando dados dos mapas do IPP e do OpenStreetMap® - 2021.

Em resumo, a avaliação de diferentes abordagens para o cálculo do Indicador ODS 11.7.1 – PEPA, que mede a proporção de áreas públicas abertas em relação à área total, revelou resultados positivos e relevantes para o contexto do planejamento urbano. A utilização de mapas da malha municipal fornecidos pelo IPP, combinados com dados do OpenStreetMap®, bem como a abordagem com classificação de imagens de satélite, permitiu o cálculo desse indicador de forma eficaz, ágil e com custos reduzidos.

É importante salientar que, dado que o indicador em questão ainda não é calculado para o Brasil e suas subdivisões, não existe um parâmetro absoluto para verificar a exatidão dos valores obtidos por meio dessas abordagens. No entanto, a comparação entre as duas metodologias revelou uma concordância notável quanto ao indicador principal, demonstrando resultados bastante satisfatórios e uma aproximação detalhada.

No entanto, é relevante mencionar que as abordagens apresentaram divergências quanto aos indicadores desagregados. Nesse sentido, a escolha da abordagem a ser adotada deve ser orientada pelas necessidades específicas de cada estudo, considerando o equilíbrio entre precisão e custo. Em última análise, a aplicação dessas abordagens oferece uma base sólida para a avaliação e o planejamento de espaços públicos em áreas urbanas, contribuindo para um desenvolvimento mais sustentável e melhor qualidade de vida para os cidadãos.

Considerações finais

O presente trabalho teve como objetivo propor alternativas para calcular o Indicador ODS 11.7.1: proporção da área construída nas cidades que é espaço público aberto para uso de todos, com base em dados abertos. Para mostrar o desempenho das abordagens propostas, foi utilizado como ilustração o município do Rio de Janeiro. Foram utilizadas duas abordagens para o cálculo do indicador, a primeira baseada em mapas de uso aberto fornecidos pelo Instituto Pereira Passos (IPP) e a segunda utilizando imagens de satélite de uso gratuito.

Em ambas as metodologias empregadas, foi constatado que o valor do indicador 11.7.1 ultrapassou o limite superior recomendado pela ONU-Habitat. Essas proporções decorrentes de ambas as abordagens evidenciam de maneira notória a significativa presença de áreas verdes

e parques urbanos integrados à malha urbana da cidade. As duas abordagens, tanto a que se valeu de mapas quanto a que utilizou imagens de satélite, produziram resultados semelhantes para a medição global do indicador 11.7.1. No entanto, quando se procedeu ao cálculo das proporções desagregadas para ruas e calçadas ($PEPA_{RC}$) e outros espaços públicos abertos ($PEPA_O$), identificou-se uma discrepância significativa entre os valores obtidos, dependendo da abordagem adotada. A diferença observada entre as abordagens revelou-se mais acentuada no caso da $PEPA_{RC}$, totalizando pouco menos de 10 pontos percentuais, enquanto no caso da $PEPA_O$ a diferença alcançou ligeiramente mais de 13 pontos percentuais.

Para a abordagem com uso de mapas é necessário refletir sobre os propósitos para os quais estes foram elaborados, sobre as informações contidas e quais outras informações podem ser extraídas desses mapas. No caso deste trabalho os mapas não forneciam informações sobre ruas e calçadas, estas mesmas foram extraídas dos espaços vazios que pertenciam às quadras e lotes. Isso pode levar a erros de dimensionamento que devem ser mensurados e corrigidos por meio de outras técnicas, especialmente onde as calçadas não obedecem ao traçado oficial dos lotes. Tais procedimentos são recomendados em trabalhos futuros.

Uma das vantagens do uso de mapas no cálculo do indicador é possibilitar trabalhar com informações sobre a totalidade da área do município. Porém, este apresenta também limitações devidas ao fato de que os mapas não foram produzidos com o propósito de calcular o indicador, podendo haver dados e informações faltantes, que apesar de poderem ser obtidos por outras fontes ou por aproximação, pode causar inconsistências e imprecisão.

A abordagem com uso de imagem de satélite se baseou na classificação de uma amostra de subáreas, dada a limitação material para a classificação da totalidade da área do município, portanto está sujeita a erro amostral. Com a limitação da capacidade de trabalho, não foi possível fazer uma amostra grande o suficiente para manter o erro amostral abaixo de 10%. Por este motivo, o indicador desagregado pode ter retornado as proporções das ruas e calçadas e de outros espaços público mal dimensionadas.

Um diferencial do uso de imagem de satélite é permitir melhorar a qualidade da classificação utilizando amostras maiores e plano amostral mais elaborado que os adotados, como, por exemplo, estratificado. A estratificação é mais vantajosa quando os grupos analisados possuem grandes diferenças entre si e ainda permite reduzir a amostra dos estratos mais homogêneos e beneficiar os mais heterogêneos. Em contrapartida, esse método exige mais informações sobre a população estudada. Além disso, é possível utilizar classificador automático, o que possibilita classificar a totalidade da área ao invés de apenas uma amostra de subáreas, reservando a adoção de amostragem apenas na inspeção visual para estimar a qualidade da classificação. Tais melhorias são sugeridas em trabalhos futuros.

Neste trabalho foram utilizadas informações adicionais às imagens de satélite, como nome dos espaços e seus usos, nos mapas de uso aberto como OpenStreetMap® e Google Earth®. Essas alternativas permitiram mitigar os efeitos da sub-representação de algum estrato, permitindo diminuir, por exemplo, os efeitos da desigualdade da distribuição desses espaços na cidade. Grandes metrópoles e cidades com grande desigualdade social, possuem também desigualdade espacial, sendo assim regiões menos favorecidas economicamente, tendem a possuir uma

Vale ressaltar que indicadores são dinâmicos e replicam um contexto temporal. Os dados utilizados neste trabalho, provenientes de mapas, imagens de satélite e dados populacionais. A velocidade dessas transformações vai nos dizer qual a frequência e a necessidade de se atualizar e monitorar o indicador em questão. Logo, cada cidade deve considerar seu próprio contexto urbano e social, como cidades com urbanização recente ou acelerada, processos de desmatamento ou reflorestamento, criação de novos loteamentos urbanos, são vários fatores que devem ser considerados no monitoramento deste indicador.

Por fim, o indicador ODS 11.7.1 é capaz de revelar as desigualdades socioespaciais e se mostra importante para a crítica sobre a qualidade dos espaços públicos abertos, e outros es-

paços. Por isso é tão importante a investigação do acesso universal a esses espaços, não é por menos que o indicador avalia a proporção do espaço público aberto para o uso de todos, especialmente mulheres, crianças, idosos e pessoas com deficiência. A desagregação por características populacionais é indispensável para a elaboração de políticas públicas e planejamento urbano, pois seus resultados contribuem para o monitoramento nas dimensões sociais, de saúde e qualidade de vida.

Referências

ADF - Arbol Day Foundation – *Tree Cities of the World*. Disponível em: <https://treecitiesoftheworld.org/>. Acesso em: 17 de abril de 2022.

ALMEIDA, A. S.; WERNECKANA, G. L.; RESENDES, P. C. *Classificação orientada a objeto de imagens de sensoriamento remoto em estudos epidemiológicos sobre leishmaniose visceral em área urbana*. Caderno Saúde Pública 30 (8), Agosto 2014. Disponível em: <https://www.scielo.br/j/csp/a/6qkZdjzmkpsPQxdYZmrQHvk/?format=pdf&lang=pt>. Acesso em: 28 de abril de 2022.

ARMAZÉM – Armazém de Dados. *Estatísticas Municipais e Portal Geo*. Disponível em: <http://www.armazemdedados.rio.rj.gov.br>. Acesso em: 02 de dezembro de 2021.

BRAVO, J. V. M.; SLUTER, C. R. *O Problema da Qualidade de Dados Espaciais na Era das Informações Geográficas Voluntárias*. Bol. Ciênc. Geod. 21 (1), Mar 2015. Disponível em: <https://www.scielo.br/j/bcg/a/xkqHyRQxjvmsvkZjDTPFWBG/?lang=pt>. Acesso em: 21 de outubro de 2021.

DAVIES, T; WALKER, S; RUBINSTEIN, M; PERINI, F. *The State of Open Data: Histories and Horizons*. Cape Town and Ottawa: African Minds and International Development Research Centre. Publisher(s) African Minds, IDRC. e-ISBN 9781552506127, 2019. Disponível em: <https://www.idrc.ca/en/book/state-open-data-histories-and-horizons>. Acesso em: 20 de abril de 2022.

DUTRA, D. J.; TIMBÓ, M. A. E.; NERO, M. A.; RODRIGUES, E. L.; BORGES, A. F. *Sistema de coordenadas planas e projeção cartográfica para a representação da área territorial completa de Minas Gerais*. Geografias, Artigos Científicos. Belo Horizonte, v. 13, n. 2, jul./dez. 2017. Disponível em: <https://periodicos.ufmg.br/index.php/geografias/article/download/16060/30238/126157>. Acesso em 06 de fevereiro de 2025.

GALLEGO, F. J. P. *Sampling frames of square segments*. MARS Project, Institute for Remote Sensing Applications JRC, 21020 Ispra (Varese) Italy Report EUR 16317, 1995, Office for Publications of the E.C. Luxembourg. ISBN 92-827-5106-6. Disponível em: https://www.researchgate.net/publication/320766327_SAMPLING_FRAMES_OF_SQUARE_SEGMENTS. Acesso em: 28 de abril de 2022.

GIL, C. G. *Objetivos de Desarrollo Sostenible (ODS): una revisión crítica*. Papeles de relaciones ecosociales y cambio global N° 140 2017/18, pp. 107-118. Disponível em: https://www.cvongd.org/ficheros/documentos/ods_revisión_critica_carlos_gomez_gil.pdf. Acesso em: 15 de outubro de 2021.

GONG, P; LIU, H; ZHANG, M; LI, C; WANG, J; HUANG, H; CLINTON, N; JI, L; LI, W; BAI, Y; CHEN, B; XU, B; ZHU, Z; YUAN, C; PING H; SUEN, J; GUO, N; XU, W; LI, ZHAO, Y; YANG, J; YU, C; WANG, X; FU, H; YU, L; DRONOVA, I; HUI, F; CHENG, X; SHI, X; XIAO, F; LIU, Q; SONG, L. *Stable classification with limited sample: transferring a 30-m resolution sample set collected in 2015 to mapping 10-m resolution global land cover in 2017*. Science Bulletin (2019). Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S2095927319301380?via%3Dihub>. Acesso em 05 de junho de 2022.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Censo Demográfico 2010: Características Gerais da População, religião e pessoas com deficiência*. In: IBGE. Rio de Janeiro, 2010a. Disponível em: https://biblioteca.ibge.gov.br/visualizacao/periodicos/94/cd_2010_religiao_deficiencia.pdf. Acesso em: 02 de novembro de 2021.

IBGE - INSTITUTO BRASILEIRO GEOGRAFIA E ESTATÍSTICA. *Censo Demográfico 2010: Resultados Gerais da Amostra por áreas de ponderação*. 2010b. Disponível em: <https://censo2010.ibge.gov.br/apps/areaponderacao/index.html>. Acesso em: 21 de novembro de 2021.

IBGE - INSTITUTO BRASILEIRO GEOGRAFIA E ESTATÍSTICA. *Plano Geral de Informações Estatísticas e Geográficas*. Rio de Janeiro, 1992. Disponível em: <https://biblioteca.ibge.gov.br/biblioteca-catalogo.html?id=222898&view=detalhes>. Acesso em: 22 de novembro de 2021.

IBGE - INSTITUTO BRASILEIRO GEOGRAFIA E ESTATÍSTICA. *Censo demográfico 2022: Resultados preliminares*. Rio de Janeiro: IBGE, 2022. Disponível em: [<https://www.ibge.gov.br>] (<https://www.ibge.gov.br>). Acesso em: 30 set. 2024.

MCFEELY, S. *The Big (data) Bang: Opportunities and Challenges for Compiling SDG Indicators*. United Nations Conference on Trade and Development, Switzerland, and Centre for Policy Studies, University College Cork, Ireland, 2019. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1111/1758-5899.12595>. Acesso em: 29 de novembro de 2021.

MOORE, R. *O tempo voa na maior atualização do Google Earth em anos*. Director, Google Earth, Earth Engine & Outreach, GOOGLE EARTH, 2021 Disponível em: <https://blog.google/intl/pt-br/produtos/explore-e-encontre-respostas/timelapse-googleearth/#:~:text=O%20Google%20Earth%20ser%C3%A1%20atualizado,medida%20que%20recebermos%20mais%20imagens>. Acesso em: 16 de junho de 2022.

ODSBRASIL - Indicadores Brasileiros para os Objetivos de Desenvolvimento Sustentável. *Transformando Nosso Mundo - A Agenda 2030 para o Desenvolvimento Sustentável*. Disponível em: <https://odsbrasil.gov.br/home/agenda>. Acesso em: 25 de março de 2022a.

ODSBRASIL - Indicadores Brasileiros para os Objetivos de Desenvolvimento Sustentável. *Especialistas do IAEG revisam classificação de indicadores*. Disponível em: <https://odsbrasil.gov.br/Home/Noticia?id=45>. Acesso em: 25 de março de 2022b.

OKFN - Open Knowledge Foundation. *What is open?*. Disponível em: <https://okfn.org/pendata/>, Acesso em: 11 de abril de 2022.

OSM - OpenStreetMap. *Contribuidores do OpenStreetMap*. Disponível em: <https://www.openstreetmap.org/about>. Acesso em: 23 de março de 2022.

RIO DE JANEIRO – Prefeitura do Rio de Janeiro. *Anexo Técnico I: Informações Sobre Todas As Áreas De Planejamento*. Coordenação Operacional De Atendimento Em Emergências (Emergência Presente). 2006. Disponível em: <https://www.rio.rj.gov.br/dlstatic/10112/1529762/DLFE-220205.pdf/1.0>. Acesso em: 02 de novembro de 2021.

RIO DE JANEIRO - Prefeitura Municipal do Rio de Janeiro. *Instituto Municipal de Urbanismo Pereira Passos – IPP - Conheça o Instituto – História*. Disponível em: <http://www.rio.rj.gov.br/web/ipp/historia>. Acesso em: 4 de maio de 2022.

RIO DE JANEIRO - Prefeitura Municipal do Rio de Janeiro. *Plano Diretor do Rio de Janeiro. Lei de uso e ocupação do solo*. Rio de Janeiro: Prefeitura do Rio de Janeiro, 2011. Disponível em: <http://www2.rio.rj.gov.br/smu/buscafacil/Arquivos/PDF/LC111M.PDF>. Acesso em: 01 de março de 2022.

RIO DE JANEIRO – Prefeitura Municipal do Rio de Janeiro. *Relatório do Plano Diretor da Cidade do Rio de Janeiro*. Rio de Janeiro: 2008. Disponível em: <https://planodiretor-pcrj.hub.arcgis.com/pages/relatorios-antiores>. Acesso em: 12 de janeiro de 2022.

SAVIO, G; ANDREANO, M. S; BENEDETTI, R; PIERSIMONI, F. *Mapping Poverty of Latin American and Caribbean Countries from Heaven Through Night-Light Satellite Images*. Springer Nature, 2020. Disponível em: https://www.researchgate.net/publication/338658018_Mapping_Poverty_of_Latin_American_and_Caribbean_Countries_from_Heaven_Through_Night-Light_Satellite_Images. Acesso em: 12 de novembro de 2021.

SILVA, A. D. *Proposta de Método de Pareamento para Integrar Dados sobre a Agropecuária*. Tese (Doutorado em População, Território e Estatísticas Públicas) - Escola Nacional de Ciências

Estatísticas/IBGE, IBGE, 2018). Disponível em: <https://ence.ibge.gov.br/index.php/pos-grad-dout-teses/2018>. Acesso em: 21 de novembro de 2021.

TORRES, J. E. M. M. *Amostragem De Área E Aplicações Em Agropecuária*. Dissertação apresentada ao Programa De Pós-Graduação em Estatística do Centro de Ciências Exatas e da Natureza da Universidade Federal de Pernambuco – UFPE, 2018. Disponível em: <https://repositorio.ufpe.br/handle/123456789/32075>. Acesso em 29 de abril de 2022.

UN-Habitat. *City-Wide Public Space Strategies: a Guidebook for City Leaders*. Nairobi, 2020. Disponível em: https://unhabitat.org/sites/default/files/2020/03/cwpss_guidebook_20200116.pdf. Acesso em 15 de junho de 2022.

UN-Habitat. *Nova Agenda Urbana*. Conferência das Nações Unidas sobre Habitação e Desenvolvimento Urbano Sustentável. Quito, 2016. Versão em Português, 2019. Disponível em: <https://uploads.habitat3.org/hb3/NUA-Portuguese-Brazil.pdf>. Acesso em: 20 de fevereiro de 2022.

UN-Habitat. *SDG Indicator 11.7.1 Training Module: Public Space*. United Nations Human Settlement Programme (UN-Habitat), Nairobi, 2018. Disponível em: https://unhabitat.org/sites/default/files/2020/07/indicator_11.7.1_training_module_public_space.pdf. Acesso em: 01 de outubro de 2021.

UN-Habitat - *Earth Observations Toolkit for sustainable cities and human settlements*. Using Earth Observation Data to Calculate Open Spaces for Public Use in Colombia. Disponível em <https://eo-toolkit-guo-un-habitat.opendata.arcgis.com/pages/colombia-use-case-1171>. Acesso em: 20/04/2022.

UN – United Nations. *Kyoto Protocol to The United Nations Framework Convention on Climate Change*. 1998. Disponível em: <https://unfccc.int/resource/docs/convkp/kpeng.pdf>. Acesso em: 10 de março de 2022.

UN – United Nations. *Principles and Recommendations for Population and Housing Censuses, Revision 3*. Department of Economic and Social Affairs Statistics Division, New York. 2017. Disponível em: https://unstats.un.org/unsd/demographic-social/Standards-and-Methods/files/Principles_and_Recommendations/Population-and-Housing-Censuses/Series_M67rev3-E.pdf. Acesso em: 11 de março de 2022.

UN – United Nations. *Report of the United Nations Conference Environment and Development. Volume 1, Resolutions Adopted by the Conference*. Rio de Janeiro, 1992a. Disponível em: <https://documents-dds-ny.un.org/doc/UNDOC/GEN/N92/836/55/PDF/N9283655.pdf?OpenElement>. Acesso em: 15 de abril de 2022.

UN – United Nations. *Report of the World Summit on Sustainable Development, Johannesburg, South Africa, 26 August - 4 September 2002*. Johannesburg, 2002. Disponível em: <https://documents-dds-ny.un.org/doc/UNDOC/GEN/N02/636/93/PDF/N0263693.pdf?OpenElement>. Acesso em: 25 de janeiro de 2022.

UN – United Nations. *Resolution adopted by the General Assembly on 27 July 2012*. 2012. Disponível em: <https://documents-dds-ny.un.org/doc/UNDOC/GEN/N11/476/10/PDF/N1147610.pdf?OpenElement>. Acesso em: 15 de janeiro de 2022.

UN-Habitat. *SDG Indicator 11.7.1 Training Module: Public Space*. United Nations Human Settlement Programme (UN-Habitat), Nairobi, 2018. Disponível em: https://unhabitat.org/sites/default/files/2020/07/indicator_11.7.1_training_module_public_space.pdf. Acesso em: 28 de outubro de 2021.

UN-Habitat. *The New Urban Agenda*. United Nations Human Settlements Programme, 2020. Disponível em: <https://habitat3.org/the-new-urban-agenda/>. Acesso em: 15 de abril de 2022.

UN – United Nations. *United Nations Conference on Environment & Development Rio de Janeiro, Brazil, 3 to 14 June 1992 - AGENDA 21*. Rio de Janeiro, 1992b. Disponível em: <https://sustainabledevelopment.un.org/content/documents/Agenda21.pdf>. Acesso em: 28 de fevereiro de 2022.

ZANOTTA, D. C; FERREIRA, M. P. *Processamento de Imagens de Satélite*. Ed. Oficina de Textos, 2019. Disponível em: <https://pdfcoffee.com/processamento-de-imagens-de-satelite-2019-pdf-free.html>. Acesso em: 28 de outubro de 2021.

Expansão e interiorização da educação superior: Programas e políticas públicas de acesso da população brasileira

Lidiane Braga Sales de Souza*

Barbara Cobo Soares**

Denise Britz do Nascimento Silva***

Resumo

A educação superior no Brasil ganhou, nas duas últimas décadas, atenção especial no cenário político e econômico em prol da ampliação e diversificação do sistema educacional. O processo de expansão e interiorização da educação superior no país foi fundamentado em arcabouços legais, em especial na Constituição Federal de 1988, na Lei de Diretrizes e Bases da Educação de 1996 e nos Planos Nacionais da Educação. A ação se desenvolveu sob influência de mudanças estruturais, políticas e ideológicas, avanço das TICs e popularização da internet. Na primeira década do século XXI, a democratização do acesso à educação superior tornou-se uma realidade no país, com implementação de diversos programas e políticas públicas. Nesse contexto, este artigo visa mapear e analisar os programas e políticas públicas de expansão da educação superior com base em processos e contextos históricos, para compreender como as políticas foram formuladas e implementadas, utilizando uma abordagem qualitativa que examina documentos oficiais e literatura relevante, utilizando, para isso, os mapas de processos e resultados dos programas e políticas públicas implementadas e identificando suas sinergias e complementaridades. Os resultados sinalizam que os programas e políticas públicas de educação superior mostraram efeitos limitados para a população brasileira por não existir ainda um sistema educacional integrado no país..

* Mestre e doutoranda pelo programa de pós-graduação da ENCE.

** Professora do programa de pós-graduação da ENCE.

*** Professora aposentada do programa de pós-graduação da ENCE.

Introdução

A educação superior tornou-se tema muito recorrente em diversos trabalhos acadêmicos no país, em especial após ganhar atenção nas duas últimas décadas no cenário político e econômico a partir da implementação de diversas ações governamentais, como programas, políticas públicas e novas legislações, que visavam a ampliação e diversificação do sistema educacional (CAMARGO E ARAÚJO, 2018).

A importância desse nível educacional, considerado como um direito social fundamental deve-se, sobretudo, ao potencial que este possui de transformar a sociedade, atuando na redução das desigualdades sociais (OLIVEIRA, 2019). Tal vinculação deve-se à relação direta existente entre o nível educacional e o desenvolvimento econômico, e seus impactos sobre o aumento da renda individual, mobilidade socioeconômica, além da possibilidade de elevação do bem-estar, do consumo e do Produto Interno Bruto (PIB) (SILVA, MORAES, FISCHER e GRIBOSKI, 2021).

Essa associação é evidenciada empiricamente a partir do desenvolvimento de estudos, como os trabalhos de Schultz (1962), Lucas Jr. (1988) e Barro e Lee (2001) com base na teoria do capital humano, que revelam o ganho de benefícios privados com o aumento do investimento em educação, e, em sua extensão, na relação direta entre educação e o crescimento econômico. Entretanto, tal modelo apresenta falhas por se propor apenas a estudar uma decisão de investimento individual e não considerar outros impactos que o investimento em educação geraria, como o ganho social a partir da elevação da escolaridade de uma sociedade (VASCONCELLOS, 2012). As externalidades positivas geradas para sociedade vão além de ganhos individuais, com maiores efeitos no contexto social, o que reforça a importância da participação do Estado, seja na regulação, ou no fomento estatal à educação superior.

No Brasil, as ações governamentais de expansão e interiorização da educação superior nas primeiras décadas do século XXI resultaram em mudanças importantes na oferta desse nível educacional para a população brasileira, como o aumento expressivo de vagas, a oferta de cursos na modalidade a distância, flexibilização dos turnos, concurso unificado nacional para acesso, entre outras mudanças. Contudo, o panorama da educação superior internacional, em 2020, mostrou que a proporção da população adulta (25 a 65 anos) brasileira com educação superior (20%) ainda se encontra abaixo da média dos demais países (39%) da Organização para a Cooperação e Desenvolvimento Econômico (OCDE).

O desenvolvimento tardio do sistema educacional brasileiro, em relação aos demais países da OCDE, não foi o único fator que contribuiu para tal resultado. O processo de expansão e interiorização desenvolveu-se sob um contexto político econômico que influenciou as escolhas estratégicas, bem como a priorização dos programas e políticas públicas nesse sentido, configurando-se em um processo desafiador, complexo e indispensável devido à grande diversidade regional, a extensão continental do país, as limitações e fragilidades das políticas públicas existentes e nossas próprias condições históricas. O principal motor desse foi a rede privada de ensino, potencializada pela modalidade a distância.

Nesse contexto, o presente artigo apresenta uma revisão de literatura das políticas públicas e programas voltados à educação superior e implementados nas últimas décadas no País, com base em referências bibliográficas atuais que tratam tanto da conjuntura econômica e política na qual tais políticas e programas estavam inseridos, quanto de seus instrumentos legais e documentos oficiais. Com esse arcabouço, este artigo visa mapear e analisar os programas e

políticas públicas de expansão da educação superior, por meio da metodologia de Mapas de Processos e Resultados, com vistas a identificar suas sinergias e compreender como se deu o processo de expansão e interiorização da educação superior no Brasil, com destaque para os avanços recentes da oferta de vagas e para as limitações e os desafios ainda existentes para alcançar uma efetiva democratização do acesso à educação superior em todo o País.

Expansão e interiorização: elementos conceituais, contexto e instrumentos legais

Segundo Costa (2017), sob uma perspectiva de dimensão sociocultural e geográfica, a expansão e interiorização são fenômenos sociais que estão relacionados e visíveis em todos os âmbitos da sociedade. São ações equivalentes, contudo, seus termos não são sinônimos. A interiorização apresenta-se justaposta na dinâmica expansionista e ambas devem ser analisadas de forma conjunta. Historicamente, o termo interiorização surgiu na colonização do Brasil, com a ocupação inicial dos portugueses no litoral brasileiro, que concentrou “não só a economia, mas também o poder político e o progresso”, conforme Coelho (1999, p.82).

Embora não exista um consenso em relação à origem do fenômeno de interiorização na literatura, para Coelho (1999), a demanda surgiu e cresceu com a industrialização em distintas cidades do Brasil, gerando fluxos migratórios intensos, em especial da população do campo para as cidades, devido à falta de investimentos, políticas públicas, instrumentos legais e subsídios para o trabalho agrícola, característico do “interior”. Tal movimento produziu críticas sobre as condições sociais desses imigrantes, o que estimulou os investimentos estatais na interiorização de serviços sociais e econômicos.

Já no campo da educação, a interiorização se materializou com a expansão, de acordo com Costa (2017, p.255):

“...falar de interiorização da educação superior é também falar do seu processo de expansão. No caso em questão, representa o ato de levar e garantir a oferta e a manutenção de educação superior pública para o interior do país, seus estados e municípios, assegurando a oferta dessa educação como direito social a todos os brasileiros.”

Em consonância com Costa, a interiorização na área da educação tem-se configurado para Dourado (1998, p.539) como:

“...resultante da política de expansão/descentralização da Educação Brasileira, tendo se materializado no país de formas combinadas ou não e resultando na ampliação de vagas por meio da consolidação de instituições já instaladas, da criação e ou instalação de outras unidades ou pela criação de novas IES...”

A expansão do ensino superior é entendida neste estudo como o aumento das oportunidades educacionais, considerando para tal a ampliação do acesso, com o aumento de oferta de matrículas, a expansão de vagas e criação de novas instituições de educação superior (IES). A interiorização, por sua vez, será examinada como o produto das políticas de expansão adotadas, visando à descentralização da oferta de vagas, reestruturação de IES já instaladas com a implantação de polos avançados, criação de novas IES e instituição de polos regionais no interior dos estados e regiões interioranas.

Historicamente, a partir de meados do século XX, os países desenvolvidos passaram a experimentar uma nova forma de organização política, econômica e social do Estado, que passou a ter o papel fundamental de promover o bem-estar econômico e social da população e corrigir imperfeições de mercado (OLIVEIRA; POCHMANN, 2019). No Brasil, a promulgação da Constituição Federal em 1988 (CRF88), inspirada nos modelos de Estado de Bem-estar Social implementados nos países europeus (ESPING-ANDERSEN, 1991), estabeleceu direitos em prol

da inclusão social e redução das desigualdades para a população brasileira, representando a materialização da proteção social enquanto política pública, com a garantia de diversos direitos sociais, entre eles, assegurar o acesso à educação como direito de todos a ser provido pelo Estado e pela família (OLIVEIRA, 2019).

Segundo Zardo (2019), em relação à provisão do ensino, o “princípio de igualdade de condições para acesso e permanência” e a “gratuidade do ensino público em estabelecimentos oficiais” (Art. 206, incisos I e IV) e os dispositivos constitucionais (Art. 208), expressos na Carta Magna, visam assegurar para a população brasileira o acesso universal, gratuito ao ensino básico, assim como universalização paulatina do ensino médio, não incluindo a educação superior como um direito fundamental (DOTTA; BERTAZOLLI, 2020). Mesmo com a previsão na CRF88 da obrigatoriedade do Estado na garantia do acesso ao nível superior de ensino, tal acesso não está explicitado como propriamente público.

A cartilha neoliberal “Consenso de Washington”¹ trouxe ao Brasil, e a outros países em desenvolvimento da América Latina, mais do que diretrizes de cunho econômico. Acometeu mudanças estruturais, ideológicas e, no campo educacional, recomendações aos países em desenvolvimento para fomentar a oferta privada da educação superior. A Conferência Mundial sobre Educação para Todos, que ocorreu na Tailândia em 1990 com participação de diversos organismos multilaterais e 157 países, influenciou (e influencia) políticas e diretrizes para a educação no sentido de universalização do acesso escolar, garantia de repasse de recursos financeiros, descentralização da gestão, incentivo ao ensino à distância, sistemas nacionais de avaliação, estabelecimento de parâmetros curriculares nacionais, etc. (LIBÂNIO, 2012). Como resultado, no Brasil, a educação superior foi “privatizada indiretamente” (DOTTA; BERTAZOLLI, 2020), com crescimento e fortalecimento das instituições de educação superior privadas a partir de 2005, devido a isenções fiscais concedidas pelo Estado para o ingresso da população de baixa renda, que historicamente apresentou dificuldades de acesso às universidades públicas.

O processo de expansão da educação superior no Brasil encontra hoje, portanto, na rede privada o seu principal motor, com crescimento das matrículas no ano de 2019 em relação ao ano 2000 de 261% ante 134% na rede pública. Em 2019, os cursos com maior participação de novas vagas na rede privada eram relacionadas às áreas de Negócios, administração e direito (34,44%) e Educação (19,36%), em sua maioria essas áreas contemplam cursos de baixa duração e com menor custo de implementação (INEP, 2019).

A promulgação da Lei 9.394/1996 de Diretrizes e Bases da Educação (LDB) ocorreu em um momento oportuno. Considerando a progressiva introdução das novas tecnologias, a LDB representou um marco inovador para a educação superior, com reconhecimento legal dos cursos e programas de educação a distância (EaD) e por estabelecer as principais normas referentes à estrutura e à funcionalidade desse nível educacional no Brasil (GOMES, 2019). Foi regulamentada em 2005 pelo Decreto nº 5.622 (com alterações em 2007 pelo Decreto nº 6.303) atualizando, então, as diretrizes para educação superior a distância (ALVES, 2019). Concomitantemente, o Decreto nº 5.773/2006 entrou em vigor com o intuito de regular e estabelecer normas para a supervisão e avaliação das IES públicas e privadas, inclusive na modalidade a distância (BRASIL, 2006).

Por sua vez o Plano Nacional da Educação (PNE), instituído pela Lei nº 10.172 de 2001, caracterizou-se como uma política de Estado permanente. Inicialmente, realizou um diagnóstico da realidade do sistema educacional em todo o país, nos distintos níveis e modalidades de

¹ No artigo de Neto e Campos (2017, p.10990) é definido o termo “Consenso de Washington” que representa o nome do documento oficial resultado de uma reunião em Washington com “[...] um grupo de neoliberais representativos dos países centrais do capitalismo [...] com representantes do Banco Mundial, do FMI e do BID e com representantes da América Latina.” Com o intuito de “[...] eliminar toda a concorrência estrangeira da América Latina, a fim de manter a área como um mercado importante para a superprodução industrial norte-americana e seus investimentos privados, inclusive no setor educacional.”

ensino com metas a serem cumpridas em um período de dez anos (IFSP,2017). Em 2007, visando operacionalizar o PNE, o governo federal instituiu o Plano de Desenvolvimento da Educação (PDE), através de diversos instrumentos legais expressos nos Decretos nº 6.093, nº 6.094, nº 6.095 e nº 6.096, todos publicados no mesmo ano, além de três projetos de lei. O PDE teve como objetivo principal a educação básica, mas possuía algumas ações para educação profissional e ensino superior (IFSP,2017).

A Lei nº 13.005 de 2014 estabeleceu metas decenais associadas à expansão e interiorização do ensino superior no âmbito nacional no PNE de 2014-2024, com destaque para a meta 12, em especial a parte que se refere à elevação da taxa bruta de matrícula para 50% e da taxa líquida para 33% da população de 18 a 24 anos a ser atingida até 2024. Ambas as metas pavimentaram e favoreceram a formulação de políticas públicas educacionais voltadas à ampliação do acesso à educação superior pelo território brasileiro, com vistas à promoção da equidade de oportunidades desse nível educacional e à redução das desigualdades sociais e regionais (AZEVEDO, 2008).

Vale ressaltar, entretanto, que os processos de expansão e interiorização da educação superior no Brasil a partir de 2001, considerando a conjuntura no qual estavam inseridos, ocorreram voltados para o desenvolvimento econômico visando “formar mão de obra mais qualificada e barata e pela urgência da busca de mais matéria prima, para o trabalho no interior do país, seus estados e municípios”, podendo explicar os programas e políticas públicas escolhidos para os diferentes níveis educacionais com medidas determinantes, de mesma finalidade pelo Estado (COSTA, 2017).

Vale destacar que o processo de expansão e interiorização da educação superior no Brasil esteve ainda restringido após essas regulamentações em função do ainda baixo dinamismo do mercado de trabalho e da capacidade de gasto das famílias nos anos 90, assim como da própria capacidade fiscal do Estado Brasileiro que, após melhora na década de 2000, tornou a enfrentar reveses em função das crises de 2008 (impulsionado pela crise externa) e de 2015 (associada à crise política interna que resultou no impeachment da presidente Dilma Rousseff em 2016). O novo governo Temer foi marcado por uma agenda de ajuste fiscal, que contou com corte nos gastos públicos, em especial na educação, além de reformas nas áreas trabalhista e previdenciária. As medidas adotadas dificultaram o processo de inclusão educacional que estava em vigor, pois os gastos em educação deixaram de ter garantia de vinculação constitucional (OLIVEIRA,2019). Ademais, inviabilizou de imediato o cumprimento da meta 20 do PNE, que tem como previsão a ampliação do investimento público em educação pública, de forma a atingir até o 10º ano de vigência, o mínimo de 10% PIB do Brasil. Por sua vez, o cumprimento da meta 12 do PNE exige expansão de recursos das instituições federais no âmbito do ensino superior para viabilizar dobrar o número de matrículas (AMARAL, 2016).

O papel da modalidade a distância na expansão e interiorização da educação superior no Brasil

Antes de analisar o papel da educação à distância (EAD) na expansão e interiorização da educação superior, é essencial compreender a sua importância para um país como o Brasil, que geograficamente tem dimensões continentais e convive com muitas desigualdades sociais ao longo de todo o seu território.

A educação a distância é considerada por Alves (2018) como a modalidade mais democrática, por ser capaz de transpor barreiras com a utilização de tecnologias de comunicação e informação (TICs), em prol do conhecimento. O avanço das TICs nos países em desenvolvimento a partir da década de 1990 foi influenciado pelos termos “sociedade da informação” e “sociedade do conhecimento”, disseminados com frequência nas reuniões do grupo dos sete países economicamente mais desenvolvidos (G7) (SANTOS, 2018).

No que tange à nova configuração da sociedade a partir de então, Malanchen (2007) sinaliza que alguns organismos internacionais foram bem incisivos em prol da utilização das TICs, tal como a UNESCO que participou ativamente de vários projetos relacionados às “oportunidades criadas pelas TICs” em muitas áreas, em especial na educação com a modalidade a distância, enquanto o Banco Mundial exerceu a função de agente financiador desses projetos educacionais em vários países em desenvolvimento.

Com a popularização da internet, as TICs foram potencializadas em distintos campos e tornaram-se um marco na civilização mundial. Esse avanço propiciou formas inéditas de construção de conhecimentos e transmissão da informação, contribuindo para a criação de novos sistemas de comunicação. Particularmente na educação, as TICs viabilizaram a modalidade a distância, revolucionando o processo de ensino-aprendizagem de educar a população (SANTOS, 2018). A exigência por uma mudança de função e estrutura das instituições surgiu, de acordo com Preti (1996, p.18), com:

“... a crescente demanda por educação, devido não somente à expansão populacional como, sobretudo às lutas das classes trabalhadoras por acesso à educação, ao saber socialmente produzido, concomitantemente com a evolução dos conhecimentos científicos e tecnológicos”

O Brasil, para se adaptar ao mercado competitivo globalizado, sofreu interferências em especial na expansão da educação superior, que, segundo Santos (2011, p.1):

“... tem direcionado alguns programas de ampliação da rede pública, dentre eles o da intensificação de cursos à distância. Considerado como uma modalidade de ensino em crescente expansão, a EAD é vista como uma das possibilidades de promover o desenvolvimento da educação nacional. Com o advento da internet e dos avanços das Tecnologias de Informação e Comunicação o ensino a distância torna-se também hoje em dia um meio propício para a democratização do terceiro grau no Brasil.”

De acordo com Neves (2005) e Paulo e Costa (2017), o ensino na modalidade EAD adotou novos paradigmas e se manteve no cenário nacional mostrando que não era apenas um modismo, tornando-se parte de um processo amplo e contínuo, contribuindo na expansão da educação superior e na promoção da democratização do acesso à escolaridade superior e à atualização permanente. O processo de regulação da modalidade a distância na educação superior no Brasil é resultado não apenas de aplicação de instrumentos legais, mas também de orientações emanadas do poder público, da interação dos vários dispositivos reguladores, bem como da influência do mercado educacional (BARROSO, 2005).

A LDB também impulsionou a demanda por esse nível educacional com o reconhecimento legal dos cursos e programas de educação a distância (EaD). Com o intuito de regulamentar essa modalidade de ensino, foi implementado o Decreto nº 5.622/2005 que caracterizou o ensino na modalidade a distância, bem como apresentou de forma expressa os níveis educacionais que poderiam ofertar essa modalidade, como a duração dos programas de EAD, a exigência de momentos presenciais, além de validar a nível nacional os diplomas e certificados de cursos e programas a distância expedidos por instituições credenciadas e registrados na forma da lei.

Logo, o processo regulatório da modalidade a distância de ensino superior pavimentou e viabilizou o seu avanço no país, promovendo em especial a rede privada de ensino, que detém a maioria de participação do total de 2,4 milhões matrículas no Brasil. Por sua vez, a modalidade a distância dessa rede de ensino cresceu consideravelmente e, em 2019, superou a quantidade de alunos matriculados na modalidade presencial da rede pública (INEP, 2019).

Por outro lado, a rede pública de educação superior apresentava diversos problemas estruturais para expandir a oferta de cursos e vagas nessa modalidade, como falta de pessoal qualificado nas tecnologias de educação a distância e falta de financiamento para desenvolvimento tecnológico na implementação de políticas em EaD. Tais dificuldades foram enfrentadas com a criação de consórcios das universidades públicas, atendendo ao princípio de coordenação da

administração pública, assegurado pelo Decreto-lei 200 de 1967, em vigor. O principal potencial da EaD para os governos, segundo a leitura de Paulo e Costa (2017, p.5), consiste:

“... no aumento da capacidade dos sistemas de educação e treinamento, na possibilidade de alcançar grupos-alvo com acesso limitado à educação e ao treinamento convencional, na possibilidade de apoiar e melhorar a qualidade e relevância de estruturas educacionais existentes, na possibilidade de obter maior eficiência financeira na educação”

Contudo, a EaD também apresenta desvantagens. Em relação as que estão intrinsecamente relacionadas a seu modo de funcionamento, pode-se citar a baixa interatividade entre professores e alunos; a dependência total tecnológica; a dificuldade na retificação de erros nos materiais e avaliações; e docentes não inteiramente capacitados para exercer os novos papéis. Em relação às principais dificuldades que podem ser enfrentadas pelos alunos em EaD vale mencionar: a necessidade de disciplina e planejamento, pois a flexibilidade e autonomia dessa modalidade de ensino requerem gestão de tempo e rotinas para a obtenção de êxito; o acesso a aparelhos tecnológicos e à internet de boa qualidade, ou seja, sem interrupções constantes; e um local adequado para o acompanhamento das aulas.

Programas e políticas públicas para a expansão e interiorização da educação superior a partir do século XXI no Brasil

Antes de elencar as políticas públicas implementadas na primeira década de 2000 que contribuíram para a expansão e interiorização da educação superior no Brasil, cabe explicitar a definição instrumental de política pública apresentada por Jannuzzi (2021, p.24):

“Política Pública é o conjunto de medidas que resultam da interação de diferentes atores executivos eleitos, técnicos nos três níveis de governo, representantes parlamentares, movimento social, sindicatos, associações patronais ..., que, por meio de diferentes instrumentos –lei, incentivos, programas ... visam atender uma demanda coletiva, mitigar uma problemática social, econômica ou ambiental; ou ainda promover um valor público desejado.”

Em consonância com a definição supramencionada, serão apresentadas a seguir as políticas públicas implementadas na primeira década de 2000 que contribuíram para a expansão e interiorização da educação superior no âmbito nacional.

Programa Educacional de Financiamento Estudantil - FIES

O Programa Educacional de Financiamento Estudantil (FIES) foi criado pela Lei 10.260/2001, no governo de Fernando Henrique Cardoso, visando proporcionar aos alunos auxílio do Estado para o financiamento de cursos superiores não gratuitos. O Programa se propõe a financiar cursos prioritariamente de nível superior² a estudantes com renda per capita familiar de até 5 salários-mínimos, que não tenham nível superior completo e nem tenham sido beneficiados por essa modalidade de financiamento em outro momento. Vale ressaltar que o financiamento só é concedido a estudantes matriculados em cursos com avaliação positiva, ou seja, com nota maior ou igual a 3, no Sistema Nacional de Avaliação da Educação Superior (SINAES) (MEC³, 2021).

O processo de contratação do FIES pode ser solicitado a qualquer momento pois, desde 2010, opera com fluxo contínuo para os candidatos em cursos com avaliação positiva no SINAES e a inscrição ocorre no Sistema de Seleção do FIES (FIES Seleção). Os dados pessoais dos candidatos pré-selecionados informados no Sistema Informatizado do FIES (SisFIES) são validados

² Os cursos da educação profissional, técnica e tecnológica, e em programas de mestrado e doutorado também estão previstos nessa legislação, mas somente se houver disponibilidade orçamentária no Comitê Gestor do Fundo de Financiamento Estudantil (CG-Fies).

³ <http://portal.mec.gov.br/component/tags/tag/sisfies>.

pela Comissão Permanente de Supervisão e Acompanhamento (CPSA) da instituição de ensino escolhida pelo aluno. Por fim, a contratação do financiamento é feita através do Banco do Brasil ou a Caixa Econômica Federal, com o comparecimento dos interessados (estudante e fiador), para formalização do contrato do financiamento (MEC, 2021).

O FIES sofreu inúmeras alterações desde a sua criação em 2001, tais como o percentual do financiamento permitido, o tempo de carência, regras de amortização e renda per capita familiar para elegibilidade no programa. As regras em vigor indicam que os cursos de graduação poderão ser 100% financiados com taxa de juros mensal do financiamento definida pelo Conselho Monetário Nacional (CMN) e período de carência de 18 meses contados a partir do mês imediatamente seguinte ao da conclusão do curso.

A partir de 2018, com a justificativa de fortalecer o programa Novo FIES⁴, foram estabelecidas três modalidades de financiamento⁵. A primeira, que conta com recursos da União, não cobra juros dos estudantes, além de flexibilizar o pagamento de acordo com a capacidade de renda do estudante. Contudo, restringe a participação para estudantes com renda per capita mensal familiar de até três salários-mínimos. Na segunda modalidade, os recursos provêm de Fundos Constitucionais e de Desenvolvimento, contemplando estudantes com renda per capita mensal familiar de até cinco salários-mínimos, das regiões Centro-Oeste, Norte e Nordeste. A terceira modalidade, por sua vez, recebe recursos do Banco Nacional de Desenvolvimento Econômico e Social (BNDES) e engloba todas as regiões do país, contemplando estudantes com renda per capita mensal familiar de até cinco salários-mínimos.

Em relação aos resultados e contribuições do FIES (CARNIELLI, OLIVEIRA, 2010, p.5) afirmam que:

“... o FIES tem uma contribuição relevante na expansão das oportunidades de permanência e conclusão da educação superior. Segundo os alunos, sem o empréstimo concedido com os recursos do Fundo, não teriam concluído o curso superior. Nos depoimentos, foram também registradas críticas à disseminação das informações sobre as normas que regem a concessão e o pagamento do empréstimo e a exigência do fiador.”

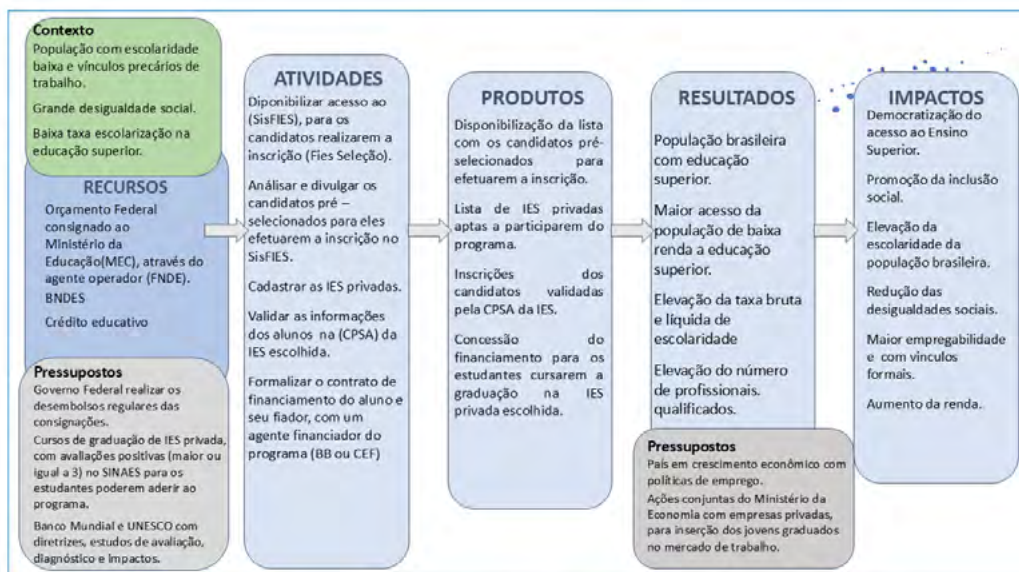
O FIES contribuiu para a ampliação do acesso à educação superior no Brasil concedendo um total de 2,7 milhões de financiamentos no período de 2009 a 2019 e atingiu a maior quantidade de concessão em 2014, com 732.673 novos financiamentos.

Com o intuito de mostrar a racionalidade do funcionamento do programa de forma sintética, foi elaborado o mapa de processos e resultados do FIES (Figura 1), de acordo com metodologia desenvolvida por Jannuzzi (2021). Nesta, são evidenciados, além dos processos envolvidos, o contexto de implementação do Programa e os pressupostos necessários para que sua implementação se dê conforme os objetivos e resultados explicitados no seu desenho institucional.

⁴ O Novo FIES foi instituído pela Lei nº 13.530/2017, para concessão de financiamento à estudantes de cursos superiores, não gratuitos, e com avaliação positiva nos processos conduzidos pelo Ministério da Educação, custeados pelo Fundo de Financiamento Estudantil abrangendo uma quantidade maior de estudantes que o modelo anterior.

⁵ <http://portal.mec.gov.br/component/tags/tag/44831-novo-fies>.

Figura 1 - Mapa de Processos e Resultados do FIES



Fonte: Elaborado a partir do Site MEC⁶ e Lei 10.260/2001.

Programa Universidade Para Todos - PROUNI

O Programa Universidade para Todos (PROUNI) foi criado em 2005 no governo de Luis Inácio Lula da Silva (Lei 11.096/2005) com o objetivo de promover o acesso ao ensino superior da população de baixa renda através do fornecimento de bolsas de estudos de 100% e 50%, liberadas de forma gradual para os alunos ingressos em Instituições de Ensino Superior (IES) da rede privada. Em contrapartida, foram concedidas isenções de impostos e contribuições para as IES privadas participantes do Programa.

Os candidatos ao programa não podem possuir diploma de curso superior e devem ter participado da última edição do Exame Nacional do Ensino Médio (ENEM) referente ao ano de concessão da bolsa, obtendo no mínimo na média das notas do exame de 450 e nota superior a zero na redação. Além disso, os solicitantes devem ter estudado todos os anos do ensino médio na rede pública, ou em escola da rede privada como bolsista integral da própria escola. Também podem concorrer a bolsas de estudo em cursos de licenciatura professores da rede pública de ensino, no efetivo exercício do magistério da educação básica e integrando o quadro de pessoal permanente da instituição, bem como pessoas com deficiência.

No ano de implementação do programa foram oferecidas 112.275 bolsas de estudos de 100% e 50%, sendo 2015 o ano com maior concessão (252.650 novas bolsas), conforme os dados do Ministério da Educação (MEC). De 2009 até 2019 o PROUNI concedeu cerca de 2 milhões de bolsas de estudo. Segundos dados do INEP, em 2019, o Brasil tinha 615.623 matrículas em cursos da educação superior na rede de privada com bolsa de 50% ou 100%.

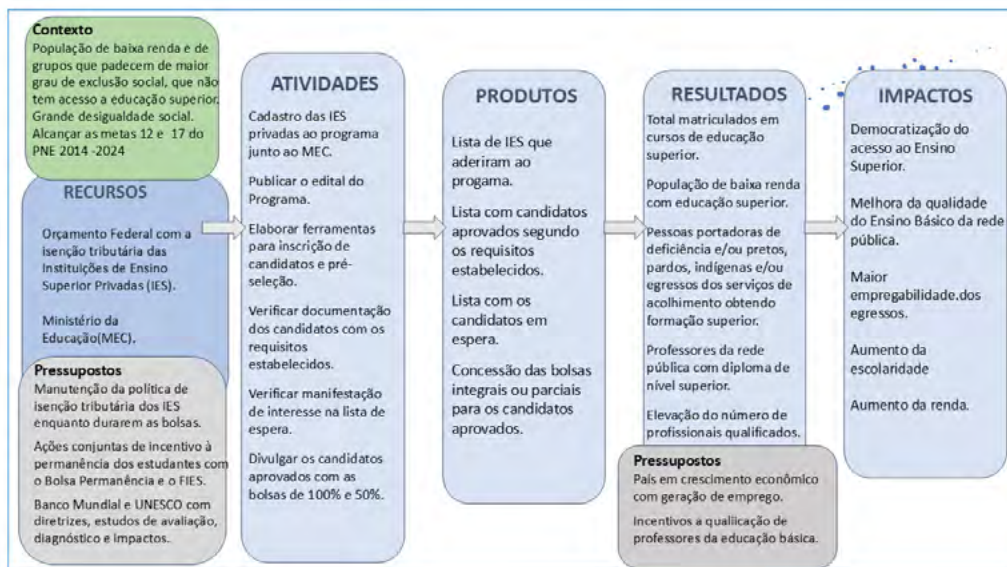
Ao permitir a inserção de um grupo populacional que por muitas gerações não teve acesso ao ensino superior e ao proporcionar a inclusão de diferentes realidades, diversidade étnica,

⁶ <https://acessounico.mec.gov.br/fies>

racial e cultural no ambiente acadêmico, o programa tem o potencial de contribuir de forma significativa na redução de desigualdades.

Assim como realizado para o FIES, a Figura 2 exibe o mapa de processos e resultados do PROUNI.

Figura 2 - Mapa de Processos e Resultados do PROUNI



Fonte: Elaborado a partir do Site MEC⁷, Lei 11.096/2005 e Jannuzzi (2021).

Universidade Aberta do Brasil - UAB

A Universidade Aberta do Brasil (UAB) foi implantada pelo Decreto nº5.800/2006 para ampliar e interiorizar a oferta de cursos e programas de educação superior pública, através da educação na modalidade a distância. A UAB faz parte de um sistema integrado constituído por parcerias entre consórcios públicos com entes governamentais nas esferas federal, estadual e municipal; universidades públicas e demais organizações interessadas, a fim de assegurar “o apoio acadêmico, tecnológico e administrativo às atividades de ensino-aprendizagem” (UAB,2023).

O público prioritário do programa são professores da educação básica que ainda não possuem educação superior e demais profissionais do magistério, conforme a Política Nacional de Formação de Professores prevista no Decreto 8.752 de 2016. Contudo, outros grupos também foram beneficiados com o programa, como os agentes públicos, conforme previsto na Política Nacional de Formação dos Agentes Públicos, e a população residente em regiões interioranas (MEC,2022). O funcionamento da estrutura operacional do programa é explicitado por (SANTOS, 2011, p.3):

“(...) funciona da seguinte maneira: nos polos de apoio presencial, local onde acontecem as atividades presenciais dos cursos, as instituições públicas ofertam os cursos. O MEC mantém, organiza e coordena as ações do Sistema UAB. Além desses elementos, existem os chamados atores: aluno, tutor presencial, tutor à distância, coordenador de polo, professor pesquisador, professor ou equipe conteudista e coordenador de curso.”

O programa seleciona as Instituições Públicas de Ensino Superior (IPES) por meio de editais semestrais. Essas instituições integram o sistema UAB com a oferta de cursos superiores na modalidade a distância, através de polos regionais, que desenvolvem atividades pedagógicas presenciais, com o apoio tutores e professores, além de terem acesso à biblioteca e laboratórios (MEC,2022).

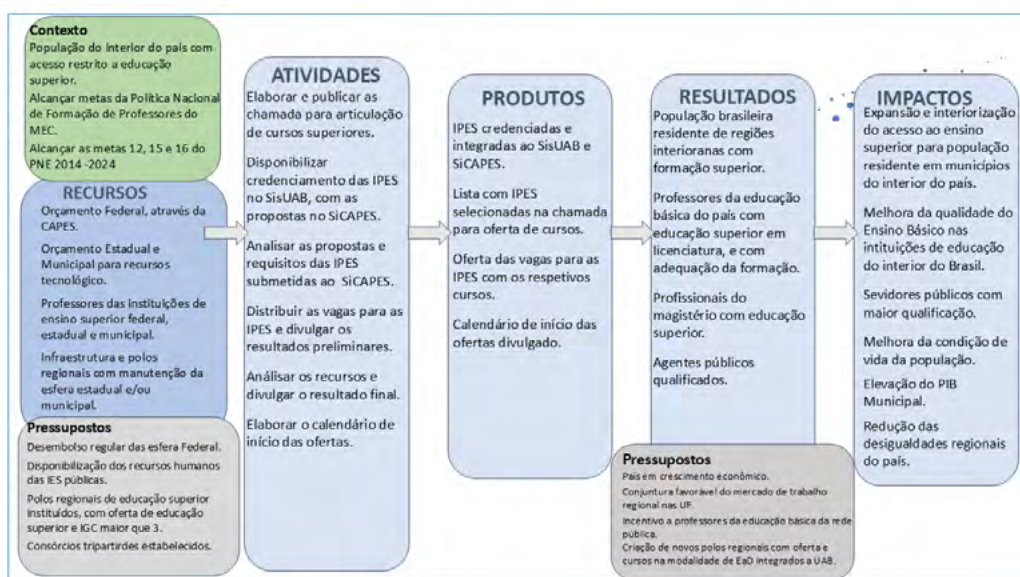
⁷ <https://acessounico.mec.gov.br/prouni>

As proponentes que podem submeter as propostas no Sistema Integrado da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (SiCAPES) são as instituições com credenciamento EAD no Sistema de Informações da UAB (SisUAB), a saber: instituições da rede pública de ensino superior (estadual, municipal e federal), as instituições integrantes do Sistema UAB e as instituições que possuem o Índice Geral de Cursos (IGC)⁸ maior ou igual a três (MEC,2022).

Segundo Santos (2011), a capacidade de promover a interiorização da educação superior, pela modalidade a distância, é uma das facetas mais importantes da UAB. Os dados do MEC indicam que o programa, em 2018, já possuía 777 polos no território nacional, com a oferta de 800 cursos. No ano de 2021 já possuía 121 alunos matriculados em 850 municípios brasileiros.

Assim como realizado para o FIES e PROUNI, a Figura 3 exhibe o mapa de processos e resultados da UAB.

Figura 3 - Mapa de Processos e Resultados da UAB



Fonte: Elaborado a partir do Site MEC⁹ e Decreto 8.752/2016.

Programa de Apoio a Planos de Reestruturação e Expansão das Universidades Federais - REUNI

A educação superior pública também foi beneficiada, em 2003, com o Programa de Apoio a Planos de Reestruturação e Expansão das Universidades Federais (REUNI) que, entretanto, só foi instituído pelo governo federal em 2007 por meio do Decreto 6.096. Tal programa fazia parte do Plano de Desenvolvimento da Educação (PDE) e visava “criar condições para a ampliação do acesso e permanência na educação superior da rede federal, com o “aproveitamento de estrutura física e recursos humanos existentes” (BRASIL, 2007).

Todas as universidades federais aderiram ao programa e seguiram as diretrizes e orientações estabelecidas pelos membros da REUNI. Foram elaborados pelas universidades o plano de reestruturação, com o “diagnóstico da situação antes da adesão ao programa, as metas a serem atingidas e quais as estratégias a serem utilizadas” (UFOP, 2018).

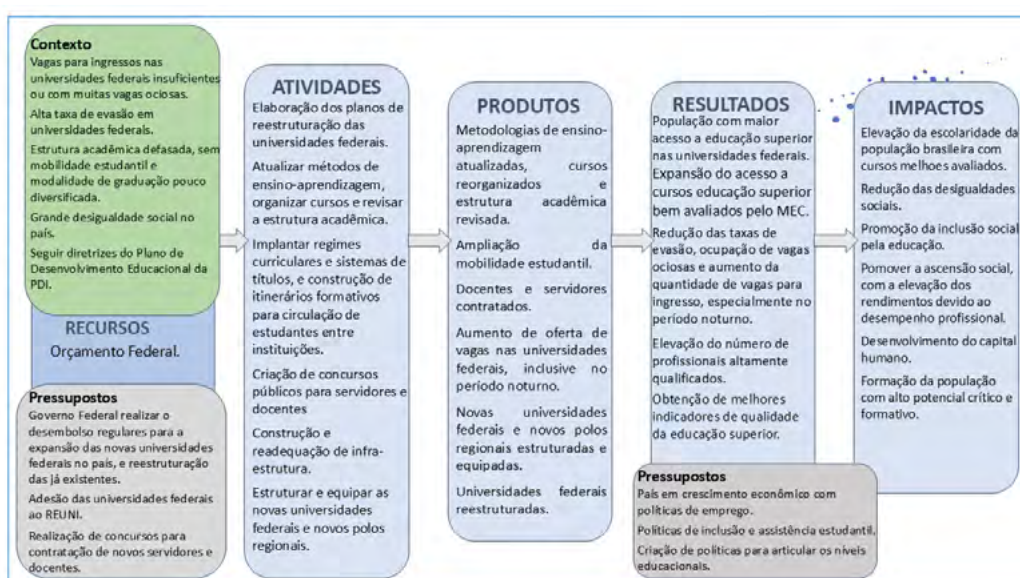
⁸ Segundo o site do MEC o Índice Geral de Cursos (IGC) mede o desempenho das instituições de ensino superior do país, o indicador “...é construído com base numa média ponderada das notas dos cursos de graduação e pós-graduação de cada instituição. Assim, sintetiza num único indicador a qualidade de todos os cursos de graduação, mestrado e doutorado da mesma instituição de ensino. O IGC é divulgado anualmente pelo Inep/MEC, imediatamente após a divulgação dos resultados do Enade.”

⁹ <http://portal.mec.gov.br/uab>

As ações do REUNI consideraram diversas dimensões com ações relacionadas ao aumento de vagas de ingresso (inclusive no período noturno); redução da evasão dos estudantes; diminuição de vagas ociosas; reestruturação acadêmico-curricular com a reorganização de cursos; atualização de metodologias, incluindo inovações pedagógicas no ensino aprendizagem; promoção da mobilidade estudantil interna e entre instituições de educação superior com o aproveitamento de créditos; comprometimento com as questões sociais, com adesão a políticas de inclusão social e assistência estudantil; entre outras ações (MEC,2022).

O programa contribuiu para a ampliação de grande parte das universidades federais. De acordo com dados do MEC, o número de municípios atendidos pela rede federal passou de 114 no ano de 2003, para 237 em 2011, por meio da criação de campi das universidades para interiorização e teve seu ápice orçamentário em 2013 (MEC, 2022). A Figura 4 apresenta o mapa de processos e resultados do REUNI.

Figura 4 - Mapa de Processos e Resultados do REUNI



Fonte: Elaborado a partir do Site MEC¹⁰ e Decreto 6.096/2007

Outros programas e políticas públicas

Para investigar a sinergia entre as políticas que propiciaram democratização e interiorização da educação superior para a população brasileira, é importante mencionar ainda outros programas, políticas públicas e ações governamentais que agiram conjuntamente com os já apresentados aqui.

A política de cotas, por exemplo, implantada pela Lei 12.711 em 2012, permeou todos os demais programas mencionados e foi importante na democratização do acesso ao nível superior e técnico para alunos oriundos de escolas públicas, de baixa renda, e de raça/cor preta, parda e indígena (JANNUZZI,2021). A legislação garantiu que 50% das vagas em universidades federais no país se destinassem estritamente a alunos que realizaram o ensino médio integralmente em escolas públicas. Parte das cotas também é destinada a alunos com renda familiar per capita de até 1,5 salário-mínimo, pretos, pardos, indígenas e pessoas com deficiência.

De acordo com o a pesquisa do Grupo de Trabalho de Políticas Etnorraciais da Defensoria Pública da União (DPU) e da ABPN (Associação Brasileira de Pesquisadores/as Negros/as), 36 das 59 universidades federais existentes até o ano de 2012 adotaram a Lei de Cotas em 2013, por

¹⁰ <https://reuni.mec.gov.br/o-que-e-o-reuni>

força da legislação. Pode-se concluir que dez anos após a sua implantação a Lei de Cotas transformou vidas:

“Pode-se dizer que no Brasil, como um todo, a Lei de Cotas impôs essa modalidade de ação afirmativa a 61% da rede de universidades federais. De tal maneira, a Lei de Cotas cumpriu o papel de universalizar as cotas raciais nas universidades federais”. (UFRGS,2022, p.5)¹¹.

Por sua vez, o Programa de Bolsa Permanência (PBP), criado através da Portaria nº 389/2013, se destina a conceder auxílio financeiro a estudantes de graduação de instituições de educação superior federais, com o intuito de “minimizar as desigualdades sociais, étnico-raciais e reduzir a evasão dos estudantes de graduação “em situação de vulnerabilidade socioeconômica”, em especial dos indígenas e quilombolas, contribuindo assim com a permanência e a diplomação dos beneficiados dessas instituições (MEC, 2018).

Segundo o site do MEC¹², em 2019, foram alocados para a execução do programa aproximadamente R\$ 139,5 milhões, atendendo 22.017 estudantes das seguintes categorias: 8.474 indígenas, 6.385 quilombolas e 7.158 em situação de vulnerabilidade socioeconômica, sendo que a oferta acumulada de bolsas de estudo concedidas atingiu 188.413 bolsas nesse mesmo ano.

Por fim, o Sistema de Seleção Unificada (SISU), normatizado pela Portaria nº 21/2012, é gerido pelo MEC visando a promoção da inclusão da população de todo o território brasileiro na participação do processo seletivo para ingresso em uma universidade pública. O SISU conta com um sistema eletrônico que possui informações relacionadas a oferta de vagas por cursos da educação superior de todas as universidades públicas do território nacional. Com as notas do ENEM, o sistema executa o ranqueamento dos estudantes por curso e modalidade, considerando a maior classificação das notas das edições do ENEM, até o limite da oferta de vagas. Em 2020, 87%¹³ das universidades federais existentes (MEC,2022) já participavam do programa.

Sinergia e complementaridades entre os principais programas e políticas públicas de expansão e interiorização da educação superior

Nesta seção, são analisadas a sinergia e complementaridades entre os diversos programas e políticas públicas implantadas a partir do século XXI no território nacional. Suas características, como o público atendido, finalidades e externalidades geradas são consideradas para compreensão de forma completa das singularidades e possíveis efeitos conjuntos produzidos.

Segundo Avila *et al.* (2013), ao abordar sinergias e complementaridades em políticas públicas, duas importantes questões devem ser consideradas: os atores envolvidos nos distintos momentos desde a concepção e formulação da política até o monitoramento e a sua efetiva implementação, além das externalidades geradas neste processo. As diversas externalidades geradas pelas políticas públicas e a interação de suas contribuições entre si geralmente ocorrem sob uma perspectiva da consequência. Contudo, a inter-relação de suas ações já deveria constar de uma estratégia no planejamento das políticas, ou ser decorrente do esforço conjunto e coordenado de várias instituições governamentais e da sociedade civil (NASCIMENTO, 2010).

Pereira e Teixeira (2013) defendem a tese de que a cooperação e parceria entre as instituições, inclusive as privadas, poderiam ser uma estratégia a contribuir com a sinergia e maior integração entre as políticas públicas, propiciando um melhor aproveitamento e enfoque das ações do Estado em busca prioritária do bem-estar social.

¹¹ <https://www.ufrgs.br/humanista/2022/09/22/apos-10-anos-transformando-vidas-lei-de-cotas-passa-por-revisao/>.

¹² http://portal.mec.gov.br/index.php?option=com_content&view=article&id=12498&Itemid=820

¹³ http://portal.mec.gov.br/index.php?option=com_content&view=article&id=12498&Itemid=820

De início, e por tudo que foi exposto até aqui, já é possível afirmar que o conjunto de ações dos programas e políticas de acesso à educação superior no Brasil que viabilizaram a expansão e interiorização desse nível educacional não ocorreram de forma integrada, mas sim sob uma perspectiva de consequência, com sinergia e complementaridades entre as suas ações, públicos-alvo, abrangências e finalidades, conforme suas principais características resumidas e relacionadas na Tabela 1.

Tabela 1 - Resumos das principais características dos programas e políticas públicas implementados no Brasil no período de 2001 até 2019

CARACTERÍSTICAS/ PROGRAMAS e POLÍTICAS PÚBLICAS	FIES	PROUNI	REUNI	UAB
PÚBLICO ALVO PRINCIPAL	Estudantes matriculados em cursos da educação superior que não o tenham concluído.	População de baixa renda sem educação superior.	Universidades federais no território nacional.	População brasileira com acesso restrito ao sistema público de ensino superior. Professores das redes de ensino da educação básica.
OBJETIVO PRINCIPAL	Conceder financiamento a estudantes de cursos superiores não gratuitos.	Conceder bolsa de educação.	Aumentar a rede de universidades federais no território nacional.	Ofertar vagas em cursos de licenciatura na rede pública em regiões interioranas do país.
FINALIDADE	Democratização do acesso a educação superior.	Democratização do acesso a educação superior.	Expansão e reestruturação das universidades federais.	Capacitar professores da educação básica com a interiorização cursos de licenciatura no território nacional.
PROCESSO	Expansão	Expansão	Expansão e Interiorização	Interiorização
ABRANGÊNCIA GEOGRÁFICA	Nacional	Nacional	Nacional	Nacional
NÍVEL EDUCACIONAL	Prioritário a Superior	Superior	Superior e Pós-graduação	Superior e Pós-graduação
MODALIDADE DE ENSINO	EaD ou Presencial	EaD ou Presencial	Presencial	EaD
REDE DE ENSINO	Privada	Privada	Pública federal	Pública
FINANCIAMENTO	Federal através de insenção tributária da IES privadas credenciadas	Federal	Federal	Federal

Fonte: Elaboração própria

O FIES e o PROUNI viabilizaram a expansão do acesso à educação superior no Brasil a partir da rede privada de ensino, e sinergicamente, desde a sua implantação, têm contribuído substancialmente para o alcance da meta 12 do PNE 2014 -2024, que trata da elevação das taxas bruta de matrícula e escolarização líquida na educação superior, em especial com a elevação da quantidade de matrículas. Os indicadores do Censo de Educação Superior do INEP refletem claramente a expansão desse nível educacional no país, entre 2000 e 2019, registrando um aumento de 219% no número de matrículas. Os programas podem ser utilizados de forma conjunta, em especial para alunos com bolsa de 50% no PROUNI. Tais programas se complementam ao atingirem públicos-alvo não idênticos, pois a concessão de bolsas integrais do PROUNI prioriza a população de baixa renda, impossibilitada de arcar com financiamentos, que muitas vezes não tem comprovação de renda familiar e/ou fiadores, conforme previsto no artigo 1º da Lei 11.096/2005.

A UAB e o REUNI apresentam sinergia e complementaridades, pois todas as políticas também contribuem para o alcance da meta 12 do PNE 2014 -2024 e asseguram a qualidade da oferta e expansão com novas matrículas no segmento público. Por outro lado, as suas distintas características as tornam complementares, por atingirem públicos-alvo distintos com a ampliação das opções de oferta a esse nível educacional, em diferentes modalidades de ensino e tipos de cursos.

O programa UAB distinguiu-se dos demais por interiorizar a oferta de cursos superiores na modalidade a distância do sistema público, nas redes federais de atendimento. Em especial, ao contribuir para o alcance da meta 15 do PNE 2014 -2024, que se refere a garantia da política nacional de formação dos profissionais da educação, em regime de colaboração entre a União,

os Estados, o Distrito Federal e os Municípios, no prazo de um ano de vigência do PNE¹⁴. Como promover a formação específica de nível superior em curso de licenciatura na área de conhecimento em que atuam os profissionais da educação e professores da educação básica, também cooperam com a meta 12. Ao assegurar a qualidade da oferta e expansão da educação de nível superior para, pelo menos, 40%¹⁵ das novas matrículas, no segmento público, esses programas ampliaram a oferta de vagas desse nível educacional, por meio da modalidade a distância na rede pública de ensino, viabilizados pelo regime de colaboração entre a União, os Estados, o Distrito Federal e os Municípios.

Considerações finais

Os processos de expansão e interiorização da educação superior no Brasil desenvolveram-se influenciados pela conjuntura internacional e nacional em que estavam inseridos profundamente conectados à política global neoliberal e às influências de organismos multilaterais como a OCDE, UNESCO e o Banco Mundial, assim como os avanços e potencialidades do uso das TICs na educação. As diretrizes econômicas resultantes desse processo levaram a mudanças estruturais e ideológicas significativas, em especial nas políticas educacionais no País, como o fortalecimento das instituições privadas e a intensificação de cursos à distância, entendidos como importantes ferramentas para democratizar o acesso ao ensino superior no Brasil respondendo às demandas de um mercado competitivo globalizado.

Todo o arcabouço legal implementado foi fundamental para a pavimentação desse processo, viabilizando a expansão e interiorização do nível superior. Destacam-se, nesse sentido, a CFB/88, que representou a possibilidade de acesso à população aos direitos sociais, entre os quais à educação superior, e a LDB/96 e o PNE, que permitiram que a expansão e a interiorização da educação superior no Brasil ganhassem força por meio da implementação dos programas e políticas públicas a partir no início do século XXI.

Os mapas de processos e resultados dos programas e políticas públicas de educação superior apresentados permitem um melhor entendimento do funcionamento de cada um desses programas, bem como a identificação das ações implementadas, a fim de avaliar a existência e identificar sinergias e complementariedades entre públicos-alvo, objetivos, abrangência e formas de financiamento que permitem, em seu conjunto, vislumbrar o real potencial e as principais restrições e limitações dos resultados alcançados. As principais políticas públicas e programas de educação superior implantados no âmbito federal, no período de 2001 a 2019, foram o FIES, PROUNI, REUNI, UAB, além da instituição do ENEM, e dos demais programas e políticas públicas transversais, como a Lei de cotas e o Programa de bolsa permanência.

Neste artigo, buscou-se mostrar que os programas e políticas públicas originaram-se em momentos distintos e se comunicam de forma restrita, com pouca gestão dessa interação, por não terem sido planejados e não possuírem coordenação com foco numa atuação integrada. Deve-se, entretanto, considerar que no período de análise o país sofreu diversos contingenciamentos e redução nos orçamentos educacionais, além de reviravoltas políticas, que impactaram os resultados esperados dos programas e políticas públicas voltados para educação superior.

As análises desenvolvidas mostraram que a modalidade de educação a distância ganhou força, em especial na rede privada de ensino, onde teve seu maior avanço, com maior quantidade de matrículas no período analisado, frente à rede pública que progrediu em menor escala, por diversas limitações. Logo, tal modalidade oportunizou o acesso da população residente em regiões interioranas ao ensino superior, de forma gratuita na rede pública e a um baixo custo na

¹⁴ <https://pne.mec.gov.br/18-planos-subnacionais-de-educacao/543-plano-nacional-de-educacao-lei-n-13-005-2014>.

¹⁵ <https://pne.mec.gov.br/18-planos-subnacionais-de-educacao/543-plano-nacional-de-educacao-lei-n-13-005-2014>.

rede privada em relação à modalidade presencial, com ganhos de resultados no curto prazo em termos de ingresso de alunos e oferta de vagas.

A evolução dos programas e políticas públicas indicam que os programas FIES e PROUNI apresentaram maior expansão em momentos próximos, respectivamente 2014 e 2015, enquanto a UAB é o único programa federal para cursos de nível superior da modalidade a distância, teve maior expansão mais recentemente, principalmente em função da difusão do uso das TICs. Embora apresentem algumas sinergias e complementaridades, as externalidades geradas pelos programas são limitadas, pois o acesso à educação superior ainda não é uma realidade para importante contingente da população brasileira espalhada no território. A ausência de planejamento e coordenação eficaz entre os programas e políticas públicas descritos deve-se a não implementação do Sistema Nacional de Educação (SNE), conforme previsão do art. 214 da CRF/88 e art. 13 do PNE, que estabelece a constituição de um sistema gestor articulador que promova a atuação dos entes federados (União, Estados, Municípios e o Distrito Federal) em regime de colaboração para efetivação das diretrizes, metas e estratégias do PNE, além de coordenar as diretrizes e ações para a educação no Brasil.

Referências

- ABRUCIO, F. L. .; FRANZESE, C. *Federalismo e Políticas Públicas: o Impacto das Relações Intergovernamentais no Brasil*. In: Tópicos de Economia Paulista para Gestores Públicos. FUNDAP. São Paulo: 2007. p. 13-31.
- ALONSO, K. M. *A EaD no Brasil: sobre (des)caminhos em sua instauração*. Educar em Revista, v. 37, n. 4, p. 37–52, 2014.
- ALVES, J. R. M. *A história da EaD no Brasil*. In: LITTO, F. M. L.; FORMIGA, M. M. M. (orgs.). *Educação a distância: o estado da arte*. São Paulo: Pearson Education do Brasil, 2009.
- AMARAL, N. C. PEC 241/55: a “morte” do PNE (2014-2024) e o poder de diminuição dos recursos educacionais. *Revista Brasileira de Políticas e Administração da Educação*, São Paulo, v. 32, n. 3, p. 653-673, set./dez. 2016.
- BARREYRO, G. B. e ROTHEN, J. C. *Avaliação e regulação da Educação Superior: normativas e órgãos reguladores nos 10 anos pós LDB*. *Revista Avaliação*, 9 Campinas: REDES, v. 12, p.133-147, mar. 2007. Disponível em <http://www.rothen.pro.br/artigos>. Acesso em 13 maio de 2012.
- BARROS, R. et al. *Confusões em torno da noção de público: o caso da educação superior (provida por quem, para quem?)*. *Sinais Sociais*, Rio de Janeiro, v. 2, n. 5, p. 76-99, dez. 2007.
- BERTOLIN, J.C.G., FIOREZE, C., & BARÃO, F.R. (2022). *Ensino Superior e Desigualdade Educativa no Brasil: Herança Elitista em um contexto de ampliação do acesso*. Em SciELO Preprints. <https://doi.org/10.1590/SciELOPreprints.3563>
- BRASIL. *Constituição Federal 1988*. Brasília, 1998. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao67.htm. Acesso em: 18 set. 2021.
- BRASIL. *Plano Diretor da Reforma do Aparelho do Estado*. [s.l: s.n.]. Acesso em: 20 out. 2021.
- _____. *Lei nº 9.394, de 20 de Dezembro de 1996*. Acesso em: 15 mar. 2021.
- _____. *Decreto no 2.494 de 10 de fevereiro de 1998 - Ministério da Educação, 1998*. Disponível em: <<http://portal.mec.gov.br/seed/arquivos/pdf/tvescola/leis/D2494.pdf>>. Acesso em: 15 abr. 2021.
- _____. *Decreto Federal nº 6017, de 17 de janeiro de 2007*.
- _____. *Lei N 10.172, de 9 De Janeiro de 2001. Plano Nacional de Educação*. Brasília -DF, 2001. Acesso em: 17 mar. 2021.
- _____. *Decreto nº. 42.765 de 22 de dezembro de 2010, 2010*. Acesso em: 15 abr. 2021.
- MEC, 2021. Disponível em: <http://portal.mec.gov.br/component/tags/tag/sisfies>. Acesso em: 7 dez. 2021
- MEC, 2022. Disponível em: <http://portal.mec.gov.br/politica-de-educacao-inclusiva?id=12265>. Acesso em: 25 fev. 2022.
- UAB, 2023. Disponível em: <https://uab.unifei.edu.br/sobre/>. Acesso em: 19 mar. 2023.
- Conceito.de, 2013 *Conceito de expansão*. Disponível em: <https://conceito.de/expansao>. Acesso em: 15 mar. 2022.
- BRITO, E. dos S. *Consórcios Intermunicipais: coordenação, colaboração e a nova governança pública*. In: CARNEIRO, J. M. B.; BRITO, E. S. (orgs.). *Consórcios Intermunicipais e Políticas Públicas Regionais*. 1ªed. São Paulo: Oficina Municipal, 2019.
- CAMPELLO, T.; FALCÃO, T.; COSTA, P. V. *O Brasil sem miséria*. Brasília. MDS, 2014.
- CASTRO, J.A. *Política Social no Brasil: marco conceitual e análise da ampliação do escopo, escala e gasto público*. *Revista Brasileira de Monitoramento e Avaliação*, Brasília, 1:66-95, 2011.

- CARVALHO, A. B. G. *O Papel Da Unirede na Construção das Políticas Públicas Para a EaD no Brasil*. Revista Brasileira de Informática na Educação, v. 22, n. 02, p. 59–75, 2014.
- CAMARGO, A. M. M., ARAÚJO, I.M. *Expansão e interiorização das universidades federais no período de 2003 a 2014: perspectivas governamentais em debate*. <http://periodicos.uem.br/ojs/acta>, ISSN on-line: 2178-5201 - Doi: 10.4025/actascieduc.v40i1.37659.
- COELHO, M. do S. da C. *Interiorização do Ensino Superior no Pará e o Banco Mundial*. In: Ver a educação, Belém, v.5, n. 1 e 2, p. 79-97, jan./dez., 1999.
- COSTA, M. S. *Contextualizando Expansão e Interiorização no Campo da Educação Brasileira*. RevistaExitus, Santarém/PA, Vol. 7, Nº 1, p. 250-276, Set/Dez 2017.
- COSTA, C. J. *Modelos de Educação Superior a distância e implementação da 110ª Universidade Aberta do Brasil*. Revista Brasileira de Informática na Educação, v. 15, p. 8, 2007.
- COSTA, C. J. ; DURAN, M. R. C. *A Política Nacional de Formação de Professores entre 2005 e 2010: a nova CAPES e o Sistema Universidade Aberta do Brasil*. Revista Brasileira de Pós-Graduação, v. 9, n. 16, p. 263–313, 2012.
- CUNHA, R. E. da. *Federalismo e relações intergovernamentais: os consórcios públicos como instrumento de cooperação federativa*. Revista do Serviço Público, v. 55, n. 3, p. 5–36, 2014.
- DOURADO, L. F. *Políticas e gestão da educação superior a distância: novos marcos regulatórios?* In: Educ. Soc., Campinas, vol. 29, n. 104 - Especial p. 891-917, out. 2008. Disponível em: www.cedes.unicamp.br
- DOURADO, L. F. *A interiorização do ensino superior e a privatização do público*. Goiânia: UFG, 2001.
- DOURADO, L. F. *Expansão e Interiorização da Universidade Federal de Goiás nos anos 80: a parceria com o poder público municipal*. In: IV Seminário Nacional de Estudos e Pesquisas História, Sociedade, Educação - Campinas SP. Anais do IV Seminário Nacional de Estudos e Pesquisas História, Sociedade, Educação, 1997. v. único. p.539-550.
- DOURADO, L. F., *O Ensino Superior em Goiás: regulamentação, políticas, perspectiva na reconstituição do campo universitário*. In: TOSCHI, M. S. (Org.). A LDB do Estado de Goiás Lei nº 26/1998: análises e perspectivas, Goiânia: Alternativa, 2001.
- DRAIBE, S. *Estado de Bem-Estar, Desenvolvimento Econômico e Cidadania: algumas lições da literatura contemporânea*. In: HOCHMAN, Gilberto; ARRETCHE, Marta; MARQUES, Eduardo (Org.) *Políticas Públicas no Brasil*, Rio de Janeiro, Fiocruz, 2007.
- FARIAS, A.J. *Difusão do neoliberalismo na América Latina*. Revista: Tensões Mundiais v. 4 n. 7 ano – 2008.
- FERNANDES, L. N., OLIVEIRA, C. A. H. S., *Educação Superior Espoliada e Seus Rebatimentos no Serviço Social*. Revista: Serviço Social & Realidade, Franca, v. 25, n. 1, 2016.
- IBARRA, David, *O neoliberalismo na América Latina*, Revista de Economia Política, vol. 31, nº 2 (122), pp. 238-248 abril-junho/2011.
- HARVEY, D. *O “Novo” Imperialismo: acumulação por espoliação*. Socialist Register, 2004, p. 95-126.
- INEP. *Relatório Financiamento da Educação no Brasil*. Brasília. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2018. Disponível em: <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/indicadores-educacionais/indicadores-financeiros-educacionais>. Acesso em: 7 Mar. 2023.
- INEP. *Censo da educação superior 2019: Microdados*. Brasília. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Acesso em: 5 Dez. 2021.
- INEP. *Censo da Educação superior 2019. Sinopse Estatística*. Brasília, 2020.

- JANNUZZI, P. M. *Monitoramento e Avaliação de Programas Sociais*. Campinas: Alínea, 2016.
- JANNUZZI, P. M. *Estado e Políticas Públicas I*. 2021. Apresentação em PDF. Disponível em: <<https://encecapacitacao.ibge.gov.br/ev/course/view.php?id=593>> Acesso em: 1 Abr. 2022.
- KERSTENETZKY, Célia Lessa. *O Estado do bem-estar social na idade da razão: a reinvenção do Estado social no mundo contemporâneo*. Rio de Janeiro: Elsevier, 2012. (capítulos 1, 2)
- LIBÂNEO, J. C. *O dualismo perverso da escola pública brasileira: escola do conhecimento para os ricos, escola do acolhimento social para os pobres*. Educação e Pesquisa, São Paulo, v. 38, n. 1, p. 13-28, 2012
- LOPES L.; VASCONCELLOS, M. (orgs.) *Manual de macroeconomia: nível básico e nível intermediário*. São Paulo: Atlas, 2008.
- OLIVEIRA, A.L.M. *Educação Superior brasileira no início do século XXI: inclusão interrompida?* Campinas, SP. 2019.
- MARKOSKI, A. S. *Obrigações de impacto social, um exercício sobre a utilização para a formação de professores*. R. BNDES, Rio de Janeiro, v. 26, n. 52, p. 109-152, dez. 2019.
- MARIANO, C. M. *Emenda constitucional 95/2016 e o teto dos gastos públicos: Brasil de volta ao estado de exceção econômico e ao capitalismo do desastre*. Rev. Investig. Const. v. 4, n.1, Curitiba, Jan./Apr., 2017.
- NETO, Filinto Jorge Eisenbach, CAMPOS, Gabriela Ribeiro, *O Impacto do Neoliberalismo na Educação Brasileira*, Educere – Congresso Nacional de Educação – 2017, Curitiba – PR.
- NEVES, C. M. C. *A educação a distância e a formação de professores*. In: Cap. 4. Rio de Janeiro: Editora PUC-Rio. 2005.
- OECD (2023), *Adult education level (indicator)*. doi: 10.1787/36bce3fe-en Acesso em: 13 fev. 2023.
- OECD (2023), *Public spending on education (indicator)*. doi: 10.1787/f99b45d0-en Acesso em: 13 fev. 2023.
- PINTO, I. F. A. *Educação e Desigualdade de Renda no Brasil*. FGV EPGE - Dissertações, Mestrado em Finanças e Economia Empresarial. 2017.
- PRETI, O. *Educação a Distância: uma prática educativa mediadora e mediatizada*. Cuiabá: NEAD/IE –UFMT. 1996.
- REVISTA HUMANISTA UFRGS, 2022. Disponível em: <https://www.ufrgs.br/humanista/2022/09/22/apos-10-anos-transformando-vidas-lei-de-cotas-passa-por-revisao/>. Acesso em: 25 mai. 2023.
- SANTOS, C. A. *Educação Superior a Distância no Brasil: democratização da oferta ou expansão do mercado*. RBPAAE - v. 34, n. 1, p. 167 - 188, jan./abr. 2018.
- SANTOS, F. C. . *UAB como Política Pública de democratização do ensino superior via EAD*. In: XXV Simposio Brasileiro, II Congresso Ibero-americano de política e administração da educação, 2011, São Paulo. Cadernos ANPAE. São Paulo: ANPAE, v.11, 2011.
- SANTOS, L. C., MENEGASSI, C. H. M. *A história e a expansão da educação a distância: um estudo de caso da UNICESUMAR* Revista Gestão Universitária na América Latina - GUAL, vol. 11, núm. 1, pp. 208-228, 2018
- SILVA, A. M., MORAES, M. C. B.; FISCHER, S. D., GRIBOSKI, C. M. *A importância da Educação Superior na redução das disparidades de renda em Santa Catarina*. Ensaio: aval. pol. públ. educ. [online]. vol.29, n.110, pp.263-287. Epub 13-Ene-2021. ISSN 1809-4465. 2021.
- SILVA, L. G. D. *Economia da Educação: o orçamento público com responsabilidade como instrumento de justiça social*. PUC-SP - Teses, São Paulo, 2021.

SILVA, M. A. P., COUTO, L. D. P., ELOI, M.O.S. *A regulação do ensino superior no Brasil: Uma breve reflexão*. Anais completos do VI Colóquio Internacional "Educação contemporaneidade" - Eixo 13 - Ensino superior no Brasil, 2012.

SILVEIRA, M. A. C. *Investimento em Capital Humano das Famílias Brasileiras: Evidência Microeconômica para o Ensino Superior* – 2020

VASCONCELLOS, Iraci Matos. *Educação superior no Brasil e desenvolvimento econômico: o peso do capital humano e do capital social*. 2012. 91 f. Dissertação (Mestrado em Economia Internacional; Políticas Públicas) - Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2012.

VONBUN, C.; MENDONÇA, J. L. O. *Educação Superior: uma Comparação Internacional e suas Lições para o Brasil*. Texto para discussão do IPEA, Rio de Janeiro, n. 1720, p.1-106, 2012.

ZARDO, L. A. R. *Demanda e oferta no ensino superior brasileiro: uma análise do (des)equilíbrio 1991 – 2015*. Tese de doutorado. 2019.

Indicadores de (des)equilíbrio da oferta e demanda no ensino superior brasileiro

Luiz André Ribeiro Zardo*

Pedro Luis do Nascimento Silva**

Joana Simões de Melo Costa***

Resumo

No Brasil, a partir de meados da década de 1990, ancorada pelo quadro legal fixado pela legislação e decretos governamentais associados, a educação superior experimentou um processo de expansão, apresentando crescimento expressivo de inscrições, vagas e ingressantes. No contexto de expansão da educação superior nas últimas décadas, esta pesquisa objetivou investigar a evolução e o equilíbrio da oferta e da demanda para ingresso nesse nível educacional. Para tanto, utilizou indicadores construídos a partir de estatísticas oficiais do Censo da Educação Superior. Os resultados apontam o desequilíbrio entre demanda e oferta, com diferenciais expressivos no padrão da trajetória dos indicadores associados às redes pública e privada.

Palavras-chave: Educação superior; Ingresso; Demanda e oferta.

* Doutor pelo programa de pós-graduação da ENCE.

** Professor aposentado do programa de pós-graduação da ENCE.

*** Instituto de Pesquisa Econômica Aplicada (IPEA), Brasil.

Introdução

A temática da educação manifesta-se relevante e de interesse em diversos segmentos da sociedade, em âmbito privado ou público. Envolve interesses de diferentes grupos ou classes sociais, tendo em conta que, em geral, seus processos pedagógicos e de formação profissional associam-se à capacidade organizacional, funcional, produtiva e de sobrevivência na e da vida coletiva em uma sociedade.

Para além da contribuição para a qualificação técnica e profissional de indivíduos – em que os processos educativos tenderiam a se adequar às necessidades e demandas dos setores produtivo e econômico, (Kuenzer, 2005; Pochmann, 2012), o sistema educacional é visto com o potencial de influenciar a transmissão intergeracional de capital humano e propiciar maior mobilidade socioeconômica. Assim, a melhoria de condições de vida como resultado da ascensão social/econômica e da inserção no mercado de trabalho de egressos do sistema educacional pode ser considerada como um indicativo de qualidade do processo educativo e formativo oferecido pelas instâncias educacionais formais (Dourado *et al.*, 2007).

Conjuntamente à consideração da relação entre educação e trabalho, cabe considerar o papel dos sistemas educacionais na perspectiva da formação de indivíduos de forma emancipatória, em processos educativos efetivos e formativos alicerçados em princípios fundamentais de cidadania (Gadotti, 2012; Pacheco, 2011).

Isso significa dispor de um sistema educacional que vá além da formação profissional ou de capacitação técnica para o exercício de alguma atividade laboral específica, dispondo de processos educacionais que englobem saberes necessários à formação de sujeitos autônomos, racionais, conhecedores de direitos e deveres e capazes de compreender a organização e a estrutura da sociedade na qual estão inseridos. Assim, se estabelece a educação como necessária para melhoria das condições de vida, da cultura e do desenvolvimento socioeconômico de uma sociedade.

No Brasil, a regulamentação do sistema educacional se apoia e opera com base na Constituição Federal de 1988 (CF-1988), na Lei de Diretrizes e Bases da Educação Nacional – Lei 9.394/96 (LDB-1996) e por decretos e resoluções do Conselho Nacional de Educação (CNE) (BRASIL, 2015).

A CF-1988 estabelece a educação como *direito de todos*, atribuindo como *competência* da União, Estados, Distrito Federal e Municípios proporcionar formas de *acesso* à cultura, à educação e à ciência (Art. 23). Embora a CF-1988, no que tange à provisão do ensino, faça referência a “princípio de igualdade de condições para acesso e permanência” e “gratuidade do ensino público em estabelecimentos oficiais” (Art. 206, incisos I e IV), parece haver distinção de abordagem nos níveis de ensino básico (infantil, fundamental e médio) e superior, quanto à *universalidade* e *obrigatoriedade*. É nítida a existência de dispositivos constitucionais (Art. 208) que buscam assegurar o acesso *universal, gratuito* (e *obrigatório*) ao ensino básico, e a *universalização* progressiva para o ensino médio *gratuito*. Já em relação à educação de nível superior, não dispõe de referência expressa para a sua oferta de forma *universal* e tampouco *obrigatória* (Cezne, 2006; Sguissardi, 2014).

O maior consenso quanto à universalização se dá com relação à educação básica. Relatório da Organização para a Cooperação e Desenvolvimento Econômico (OCDE) aponta que cerca de 15% dos indivíduos com idade entre 25 e 64 anos possuem ensino superior completo no Brasil, valor abaixo da média de 37% dos seus países membros. Ainda de acordo com a OCDE, a educação superior eleva a chance de emprego formal, pois indivíduos com ensino universitário apresentam taxa de desemprego 40% abaixo daqueles com ensino médio (OECD, 2017).

Pode-se, assim, considerar a educação de nível superior uma realização do ‘*acesso aos níveis mais elevados de ensino*’ e, além disso, como um percurso a ser cumprido a fim de se lograr maior ou melhor qualificação profissional, cultural, emancipação e ascensão social e econômica. Nessa linha, demanda por *acesso* à educação superior pode resultar em pressão social por *oferta* de vagas para esse nível educacional, sob a forma de pleito junto aos agentes governamentais para a implementação de políticas públicas destinadas à obtenção de equilíbrio na relação entre a *demanda* e a *oferta* de vagas.

No Brasil, a educação superior experimentou expressivo crescimento no número de instituições, cursos e vagas a partir da segunda metade década de 1990 (Ristoff, 2013; Sguissardi, 2014; Tachibana *et al.*, 2005). Na década seguinte, a educação superior foi acompanhada por formulação e implementação de diversas políticas e programas complementares à sua expansão, sob o argumento da necessidade de democratização do acesso (Brasil, 2007).

Cabe considerar que o termo acesso abrange aspectos associados ao *ingresso* no sistema educacional, à *permanência* e à *qualidade na formação* dos estudantes no percurso acadêmico, ver Silva e Veloso (2013). Tais aspectos associam-se à expansão da educação superior, segundo a evolução quantitativa do quadro de instituições, cursos, vagas etc., assim como as suas respectivas caracterizações quanto à forma de ingresso (processos seletivos), categoria administrativa (pública, privada), organização acadêmica (universidade, faculdade etc.), formas de financiamento, dentre outras (Silva e Veloso, 2013).

Adotou-se como foco do estudo o *ingresso*, ou seja, a etapa inicial de acesso à educação superior. O objetivo é investigar a evolução da *demand*a por e da *oferta* de vagas para ingresso na educação de nível superior no Brasil, e também se e como se articulam em equilíbrio (ou não) no sistema educacional ao longo do tempo.

O artigo é composto por 6 seções, incluindo esta introdução. A seção 2 trata do contexto institucional do ensino universitário no Brasil. A seção 3 apresenta as fontes e bases de dados utilizadas. A seção 4 descreve os indicadores educacionais considerados e a seção 5 apresenta e analisa a evolução de tais indicadores. Por fim, a seção 6 apresenta as conclusões.

Conceitos, marcos legais e políticas de expansão do acesso à educação superior no Brasil

Acesso e ingresso

Neste estudo, adotou-se a conceituação de acesso (à educação superior) abordada por Silva e Veloso (2013). No caso da educação superior, pode-se considerar o *ingresso* como a etapa da admissão de um indivíduo numa instituição formal, na qual o indivíduo estaria habilitado a frequentar/participar de algum curso, como o primeiro movimento com vistas a *acessar* esse nível educacional.

Silva e Veloso (2013) acrescentam que o *ingresso* em um curso pode não garantir efetivamente a sua continuidade ou conclusão por estar sujeito a atrasos, retenções, desistências ou evasões. Assim, há de se considerar a dimensão *permanência* a fim de assegurar sucesso na trajetória estudantil, que pode demandar *acesso* a, por exemplo, programas de apoio à permanência, tutoria, suporte, auxílio e/ou assistência estudantil, formulados e/ou implementados para essa finalidade. Simultaneamente, é importante considerar o acesso a processos formativos de qualidade, entendidos como *formação qualificada*, “[...] que favoreçam à apropriação do conhecimento crítico e à formação de sujeitos-protagonistas no processo educacional” (Silva e Veloso, 2013, p. 730).

Neste trabalho, em que se investigou a evolução da *demand*a e da *oferta* de vagas na educação superior no Brasil, considerou-se a dimensão *ingresso* como foco de estudo para, por meio de indicadores educacionais associados, tecer análises acerca da etapa inicial do acesso à educação de nível superior brasileira.

Políticas de expansão

Importante ter em mente que os processos de formulação de políticas educacionais, assim como os de implementação, continuidade etc., dependem das relações entre o Estado e a sociedade, dos arranjos institucionais, do contexto social, político, econômico, dos atores e agentes governamentais e sociais envolvidos e de seus interesses (Carvalho, 2011). Dessas relações podem decorrer ações que causem impactos na composição, manutenção ou modificação do quadro legal,

que incluem, por exemplo, reforma em dispositivos constitucionais, leis, diretrizes e regulamentações, na regulação do sistema educativo, assim como alterações nas formas de coordenação e controle estatal.

No Brasil, as políticas de reestruturação das universidades implementadas por meio da Reforma Universitária de 1968 com vistas à expansão universitária (Carvalho, 2011; Gonçalves & Vieira, 2015; Queiroz *et al.*, 2013; Sguissardi, 2014), resultaram em crescimento insuficiente. Foi a partir da década de 1990 que se pôde observar a formulação de diversas políticas, programas e ações no campo das políticas educacionais, apoiadas pela CF-1998 e pela LDB-1996, que proporcionaram crescimento mais expressivo no número de vagas para ingresso no ensino superior. Nesse período, o aumento do número de egressos do ensino médio e os desafios da educação no contexto da economia globalizada trouxeram a perspectiva de expansão do sistema de ensino superior.

O excedente de egressos do ensino médio, visto como "sintoma de desajuste do nosso sistema de ensino" para atender à necessidade de profissionais qualificados nas décadas de 1960 e 1970 (Franco, 1985, p. 16), fez com que a demanda pelo ensino superior se convertesse, na década de 1990, em componente do projeto educacional do governo, no qual se incluiu a reforma da educação superior.

Em 1995, por exemplo, cerca de 1,04 milhões de pessoas concluíram o ensino médio. Dessas, cerca de 743 mil o fizeram em instituições de ensino público (INEP, 1996). Entretanto, a política pública para a educação superior nos anos de 1990 expõe a opção pela expansão via crescimento privado da oferta de vagas para ensino superior. Dados da matrícula em instituições privadas do início dos anos 2000 mostram uma taxa de crescimento da ordem de 150,2% no período de 1994 a 2002, quase o triplo do crescimento das matrículas em instituições públicas, que foi de 52,3%. No período de 1994 a 2000 a proporção de matrículas nas instituições privadas subiu de 60% para 69,8% (Sguissardi, 2008b, p. 999).

Ristoff (2013, 2014) registra que as políticas de expansão da educação superior adotadas a partir do fim dos anos 1990 foram favoráveis à expansão e crescimento do setor privado, no qual "[...] o Estado brasileiro [assume], na segunda metade da década de 1990, um papel claramente incentivador desse crescimento" (Ristoff, 2013, p. 44), e cuja expansão no setor privado se dá com "[...] marcada aceleração de ritmo de 1999 a 2003" (Ristoff, 2014, p. 724).

Na segunda metade da década de 2000, verifica-se crescente investimento do governo federal na educação superior visando à expansão das universidades federais, como as políticas associadas aos programas Expansão e o REUNI (Amaral, 2008; Carvalho, 2015; Chaves & Amaral, 2014). No que tange ao ensino privado, intensificou-se o processo de mercantilização da educação, com as fusões de instituições, formando as empresas de ensino, que abriram o capital na bolsa de valores (Chaves, 2010).

Embora mudanças quanto à participação do público e do privado no ensino superior reflitam taxas de crescimento menores das instituições e das matrículas privadas ao final do período correspondente a 2003 e 2010, estudos como Carvalho (2006), Sguissardi (2008a), Martins (2009), e Chaves (2010), dentre outros, enfatizam a continuidade e forte participação do setor privado nesse nível de ensino.

Sguissardi (2014) aponta, com base em dados do Censo da Educação Superior¹, que no período de 1995 a 2002 não houve criação de nova universidade federal, enquanto o número de instituições privadas de educação superior apresentou um crescimento de cerca de 59% no período de 1999 a 2002. O autor destaca uma retração de 16% no número de instituições privadas categorizadas como *sem fins lucrativos* (confessionais e comunitárias) e crescimento de aproximadamente 113% das categorizadas como *com fins de lucro* (privadas particulares). Sguissardi (2014) aponta, ainda, a ocorrência da alteração de estatuto/regime jurídico em 58% das instituições privadas para categorização de instituições privadas com fins lucrativos.

No ano de 2003, foram desencadeados diversos programas relacionados à expansão das Universidades Federais, dentre eles o Programa Expansão (GENTIL; LACERDA, 2016) que teve

¹ O Censo da Educação Superior é realizado pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) e seus dados são disponibilizados em domínio público no sítio eletrônico do INEP.

como objetivo “expandir o sistema federal de ensino superior, com vistas a ampliar o acesso à universidade, promover a inclusão social e reduzir as desigualdades regionais”. No período de 2003 a 2006, o Programa Expansão promoveu a criação de nove universidades federais e resultou na consolidação ou instalação de 41 campus universitários, sobretudo no interior do país.

Em 2005, foi instituído o Programa Universidade para Todos (PROUNI), que concede bolsas de estudo integrais ou parciais a estudantes provenientes de escolas públicas, ou que foram bolsistas integrais em escolas privadas, para cursar a graduação e/ou cursos sequenciais em instituições privadas de ensino superior. Nesse cenário, as instituições privadas que aderem ao programa recebem isenção de tributos como forma de incentivo fiscal (Silva & Veloso, 2013).

Observa-se ainda, na década de 2000, a formulação de políticas educacionais e programas complementares à expansão da educação de nível superior sob o argumento da democratização do acesso ao ensino superior. Em 2007, foi elaborado o Plano de Desenvolvimento da Educação (PDE), que teve a educação superior como um de seus eixos norteadores, indicando para esse nível diretrizes como: expansão da oferta de vagas; garantia de qualidade; promoção de inclusão social pela educação; ordenação territorial; e desenvolvimento econômico e social.

No PDE (Brasil, 2007), alguns dos programas correlatos à expansão e democratização do acesso ao ensino superior incluíram: a) Reestruturação e expansão das universidades federais: Programa de Apoio a Planos de Reestruturação e Expansão das Universidades Federais (REUNI) e Programa Nacional de Assistência Estudantil (PNAES); b) Democratização do acesso: Programa Universidade para Todos (PROUNI) e Fundo de Financiamento ao Estudante do Ensino Superior (FIES); c) Avaliação como base da regulação: Sistema Nacional de Avaliação da Educação Superior (SINAES).

Em relação à democratização do acesso, o FIES (criado em 2001 e reformulado em 2010) destina-se à ampliação do acesso ao ensino superior privado por meio da concessão de financiamento a estudantes para cursarem em instituições privadas. Também em 2010 foi instituído o Sistema de Seleção Unificada (SiSU), por meio da Portaria Normativa n.º 2 do MEC, para selecionar candidatos a vagas em cursos das instituições públicas de educação superior participantes.

Em 2012 foi aprovada a Lei das Cotas (Lei n.º 12.711/2012), que dispõe que as instituições federais de educação superior reservem, em cada concurso seletivo para ingresso nos cursos de graduação, por curso e turno, no mínimo 50% (cinquenta por cento) de suas vagas para estudantes que tenham cursado integralmente o ensino médio em escolas públicas.

Em 2014 foi aprovado o Plano Nacional de Educação (Lei n.º 13.005/ 2014), estabelecendo para os dez anos subsequentes metas relacionadas à educação de nível superior, das quais se pode destacar: elevar a taxa bruta de matrícula na educação superior para 50% e a taxa líquida para 33% da população de 18 a 24 anos.

Pode-se notar que, sobretudo nas últimas duas décadas, ocorreu no Brasil a formulação e implementação de diversas ações no campo das políticas educacionais relacionadas à expansão do acesso à educação de nível superior, tais como as aqui apontadas.

Nas seções seguintes são apresentadas fontes e bases de dados, além de indicadores associados ao *acesso* (pelo *ingresso* à educação superior), obtidos a partir de estatísticas públicas, relacionados à *demand*a e à *oferta* de vagas na educação superior brasileira, a fim de desenvolver as análises de interesse.

Fontes e bases de dados

Neste trabalho foi utilizada uma das principais fontes de dados oficiais educacionais no Brasil com periodicidade anual: o Censo da Educação Superior (CES), realizado pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). Esta seção descreve os dados usados para obter indicadores da relação entre a *oferta* e a *demand*a por vagas para a educação superior no Brasil.

O CES tem como propósito elaborar as principais estatísticas da educação superior, levantando informações² referentes às Instituições de Educação Superior (IES), aos Cursos, Docentes e Alunos (INEP, 2015). A série histórica de dados do CES disponível para acesso e para baixar de forma *online* contém arquivos, dispostos por ano, referentes às *Sinopses Estatísticas* e aos *Microdados* para o período de 1995 a 2016.

As *Sinopses Estatísticas* do CES são formadas por tabelas-resumo com informações agregadas para diversos domínios de estudo, definidos, por exemplo, pela abrangência geográfica (Regiões, Unidades da Federação etc.), por esfera administrativa (pública, privada), por áreas de formação, dentre outras. Os *Microdados* se referem aos *dados brutos* obtidos na coleta das informações das IES.

Para examinar a evolução temporal dos indicadores relacionados à demanda e oferta de vagas para o *ingresso* no ensino superior, foram utilizadas as bases de dados do CES que pudessem fornecer, inicialmente, a maior série temporal possível e, posteriormente, o maior nível de desagregação para alguns domínios de interesse. Foram, assim, utilizadas as seguintes bases de dados como insumo para obtenção de estatísticas e construção de indicadores de interesse:

- Sinopse Estatística do Censo da Educação Superior, arquivo “MUNICIPIOS_ORG_DEP_1991_2010.xls”, constituindo série histórica harmonizada para o período de 1991 a 2010, e;
- Microdados do Censo da Educação Superior, arquivos “CURSO.TXT”, referentes aos anos 2011 e 2012, e arquivos “DM_CURSO.csv”, referentes aos anos 2013, 2014 e 2015.

Convém ressaltar que o período de referência (1991 a 2015) abordado neste estudo ficou limitado à disponibilidade de dados no sítio eletrônico do INEP, na ocasião da obtenção das bases de dados.

A partir desses arquivos foi construída uma base de dados que considerava o conjunto de variáveis comuns a ambas as bases (*Sinopse Estatística* e *Microdados*), gerando uma série histórica de 25 anos, abrangendo o período de 1991 a 2015. Essa base de dados, em nível agregado, é aqui denominada *SinopseCES*.

A escolha dessas bases de dados impõe uma limitação analítica que se refere à disponibilidade de informações referentes apenas aos cursos de graduação ofertados na modalidade presencial. Isso implica a impossibilidade de obtenção de resultados para alguns domínios de interesse para o período de 1991 a 2010, como, por exemplo, cursos em nível superior ofertados à distância ou sobre a oferta de cursos sequenciais, grau acadêmico (bacharelado, licenciatura, tecnológico), tipos de Áreas ou Cursos de formação, dentre outros.

Métodos

As próximas subseções apresentam indicadores associados à demanda e à oferta de vagas, que se relacionam à dimensão ingresso na educação superior no Brasil, utilizando estatísticas públicas oficiais provenientes do CES.

Estatísticas como insumos para obtenção de indicadores de oferta e demanda de vagas

De forma resumida, o Quadro 1 apresenta o nome e a descrição das variáveis da base de dados *SinopseCES* e utilizadas como suporte para a construção e o acompanhamento temporal de indicadores referentes à relação entre a demanda e a oferta de vagas em cursos de graduação presenciais no Brasil.

² Podem ser acessados, via sítio eletrônico do INEP diversos documentos e arquivos do CES (Resumos Técnicos, Sinopses Estatísticas, Microdados etc.).

Quadro 1. Nome e descrição das variáveis de interesse contidas na base de dados SinopseCES.

Nome da variável	Descrição da variável
ANO	Ano de Referência do CES
CO_REGIAO_CURSO	Código da Região
NO_REGIAO_CURSO	Nome da Região
CO_UF_CURSO	Código da Unidade da Federação
SGL_UF_CURSO	Sigla da Unidade da Federação
NO_MUNICIPIO_CURSO	Nome do Município
CO_LOCALIZACAO	Código da Localização
DS_LOCALIZACAO	Descrição da Localização
CO_ORG_ACAD_AGREG	Código de classificação de Organização Acadêmica
DS_ORG_ACAD_AGREG	Descrição da classificação quanto a Organização Acadêmica
CO_CATEG_ADM	Código de classificação de Categoria Administrativa
DS_CATEG_ADM	Descrição da classificação de Categoria Administrativa
Y_1	Nº de Cursos de Graduação Presenciais
Y_2	Nº de Vagas oferecidas em Cursos de Graduação Presenciais
Y_3	Nº de Inscritos por vestibular e por outros processos seletivos
Y_4	Nº de Ingressos por vestibular e por outros processos seletivos
Y_5	Nº de Ingressos por processo seletivo
Y_6	Nº de Ingressos por outras formas

Fonte: Base de dados elaborada mediante combinação da base de dados na Sinopse Estatística do Censo da Educação Superior (INEP), série histórica por município 1991-2010, arquivo "MUNICIPIOS_ORG_DEP_1991_2010.xls" e, Microdados do Censo da Educação Superior, arquivos "CURSO.TXT" 2011 e 2012, e "DM_CURSO.csv" 2013, 2014 e 2015. (INEP).

Considerando os registros provenientes da série histórica da *Sinopse Estatística* do CES, período de 1991 a 2010, as quantidades de interesse foram obtidas de forma direta, ou seja, assumindo os valores registrados em cada variável de estudo. Quando considerados os registros provenientes dos arquivos de *Microdados* do CES algumas quantidades de interesse foram derivadas da agregação de duas ou mais variáveis relacionadas, de acordo com a forma como estão organizadas e disponibilizadas as informações nos arquivos associados aos respectivos anos de referência (Anexo 1).

Indicadores associados à relação entre demanda e oferta de vagas

Considerando as informações da base de dados *SinopseCES* para um ano a de interesse, com $a \in \{1991, \dots, 2015\}$, foram utilizados os seguintes indicadores do tipo *razão*:

Razão de Demanda Efetiva (RDE)

$$RDE_a = \frac{Y_{3,a}}{Y_{2,a}}$$

A RDE visa avaliar em que medida a demanda objetiva, aqui entendida como a procura efetiva pelo ingresso em cursos presenciais de nível superior, por meio da inscrição em processos seletivos ou outros tipos de seleção, excede ou não a quantidade de vagas ofertadas.

Taxa de Aproveitamento de Vagas (TAV)

$$TAV_a = 100 \frac{Y_{4,a}}{Y_{2,a}}$$

A TAV mede a proporção de vagas ofertadas que foram de fato preenchidas. O seu complemento (isso é, $100 - TAV$) representa o percentual de vagas não-preenchidas.

Taxa de Sucesso de Ingresso (TSI)

$$TSI_a = 100 \frac{Y_{4,a}}{Y_{3,a}}$$

A TSI corresponde à proporção entre os inscritos daqueles que de fato acabam por ocupar uma vaga ofertada. Assim, essa medida indica a proporção da demanda atendida, e o seu complemento mensura a demanda reprimida.

Resultados

Demanda Efetiva

No que tange à relação entre a *demanda* e a *oferta* de vagas em cursos de nível superior para o período de 1991 a 2015 pode-se observar, por meio da Figura 1, que houve mais demandantes por vagas na modalidade presencial nas instituições de ensino superior do que vagas ofertadas (RDE acima do valor unitário).

No primeiro ano da série temporal, a RDE assumia o valor de 3,8 e, em 1995, chegou a 4,3. Tais valores indicam que a quantidade de inscritos alcançava aproximadamente quatro vezes o total de vagas, o que revela um descompasso entre demanda e oferta. Os 12 anos seguintes, entre 1995 e 2007, se caracterizaram por uma aproximação entre demanda e oferta. Ao final desse período, a RDE alcançou o valor de 1,8 inscritos por vagas. A Figura 1 também apresenta o padrão das curvas das quantidades de vagas e de inscritos. Esse período se caracterizou por uma expansão importante do número de vagas. Enquanto a taxa média anual de crescimento da quantidade de vagas foi de aproximadamente 13,6% ao ano, a de inscrições girou em torno de 5,7% ao ano.

Entre 2007 e 2013 houve aumento do RDE, que chegou ao nível de cerca de 3,5 inscritos por vagas, em 2013. O afastamento entre demanda e oferta deveu-se a uma desaceleração no ritmo de crescimento de vagas e ao aumento significativo do total de inscrições (Figura 1). Nesse período, a taxa de crescimento do número de vagas e de inscrições foi, em média, 3,3% e 14,9% ao ano, respectivamente.

Nos últimos dois anos da série, a curva do RDE exhibe padrão de decréscimo, apresentando, em 2015, aproximadamente 2,6 inscritos por vaga. O crescimento da oferta de vagas (em média 29,5% ao ano) foi mais acelerado do que o aumento da demanda (em média, 12,1% ao ano).

A Figura 2 exhibe as curvas da RDE segundo as categorias administrativas pública, Figura 2 (a), e privada, Figura 2 (b). Notam-se distintos padrões na evolução dos indicadores para as categorias pública e privada.

Quanto à categoria pública, verifica-se o padrão de crescimento na RDE no período 1991 a 2002, apresentando cerca de 6,2 inscritos por vagas, em 2001, e em torno de 8,9, em 2002, evidenciando crescimento mais expressivo da quantidade de inscritos (8,2% ao ano) diante de vagas ofertadas (5,1% ao ano). O período de 2003 a 2009 apresentou decréscimo da RDE, que caiu para 6,6 em 2009. Nesse período, o crescimento do número de inscritos se deu de forma mais moderada (1% ao ano), enquanto o número de vagas ofertadas manteve o padrão de crescimento

(cerca de 5% ao ano). No período subsequente, 2010 a 2015, a RDE inverte a trajetória, crescendo de cerca de 7,5 inscritos por vagas em 2010, para cerca de 12,2 em 2015. Nesse período, destaca-se expressivo aumento da quantidade de inscritos (17,2% ao ano), enquanto a taxa de crescimento da quantidade de vagas (8,1% ao ano) não acompanhou a de inscritos.

Houve crescimento de vagas em instituições públicas federais após a implementação dos programas Expansão e REUNI. O crescimento expressivo no número de inscritos na esfera pública na década de 2010 pode estar associado à mudança na forma de ingresso a partir da implementação do SiSU, em 2010, e, conseqüentemente, pela adesão crescente de instituições de ensino superior ao sistema, mesmo com algumas variações nos anos seguintes. Por exemplo, em 2010, 23 universidades públicas utilizaram o SiSU como forma de ingresso, já em 2015 foram 63 universidades públicas (Luz, 2017).

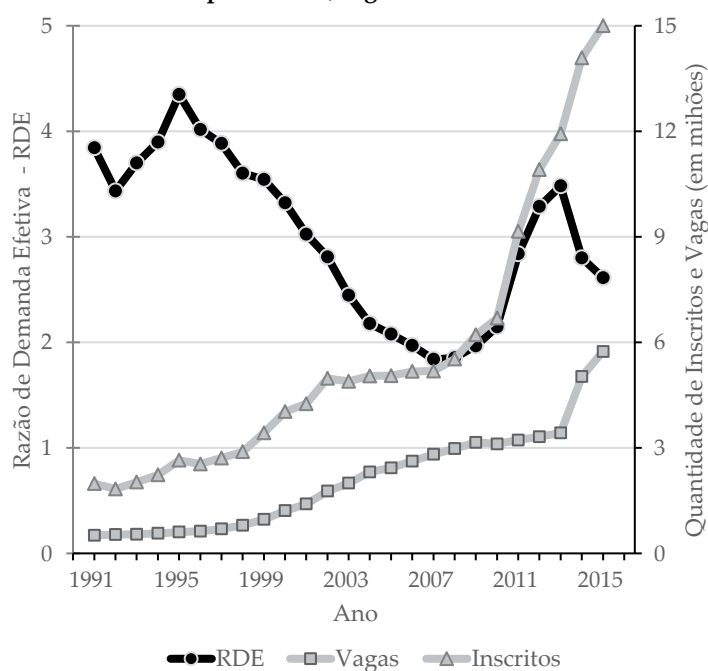
Para o ensino superior privado, o padrão da curva da RDE se mostra decrescente ao longo do período, partindo de cerca de 2,7 inscritos por vaga em 1991, para em torno de 1,2 em 2015. Nesse período, o crescimento da quantidade de vagas se mostrou superior ao de inscritos, registrando-se taxas de crescimento em torno de 11% e 8% ao ano, respectivamente.

Destaca-se o decréscimo mais expressivo na RDE privada no período de 1998 a 2002, durante o qual a quantidade de vagas ofertadas cresceu 20% e o número de inscritos apenas 12% ao ano. No período de 2003 a 2009, a RDE privada apresentou padrão de queda mais moderado.

No período de 2010 a 2013 observa-se alteração na trajetória da RDE privada, com crescimento de 1,25 inscritos por vagas, em 2010, para 1,62, em 2015. Nesse período, a quantidade de vagas diminuiu o ritmo de crescimento (2% ao ano em média) e a de inscritos teve o padrão de crescimento mais acentuado (9% ao ano). Esse fato pode associar-se ao aumento de vagas relacionado ao PROUNI, especialmente pela oferta de bolsas e pela reformulação do FIES. Tanto o PROUNI como o FIES tiveram crescimento na quantidade de beneficiados. Como aponta Chaves (Chaves, 2015), “com a injeção dos recursos do FIES, as IES privadas têm garantido o pagamento das mensalidades de uma grande parcela de estudantes, o que favorece a expansão desse setor”.

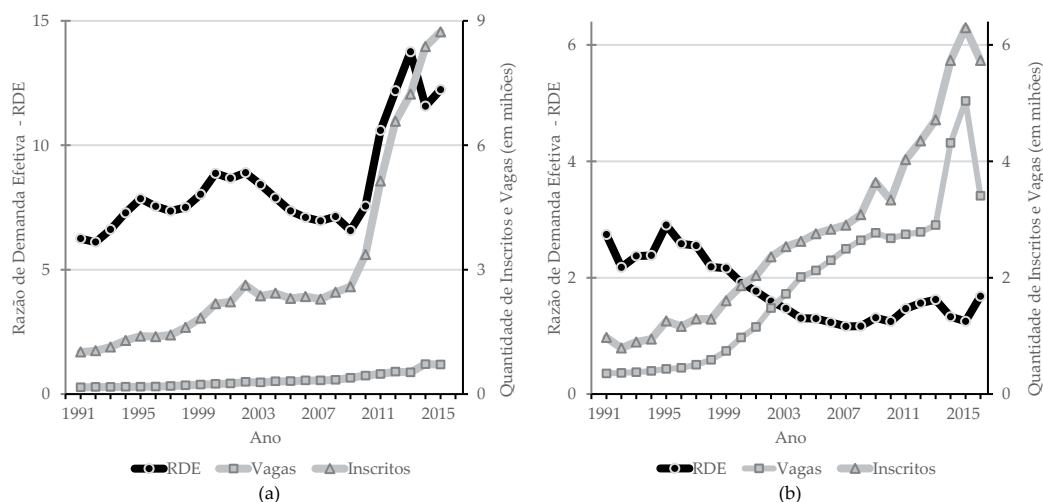
Pode-se destacar, ainda, a crescente e contínua oferta de vagas e de quantidade de inscritos, exibida de forma mais expressiva na esfera privada diante da pública. Por exemplo, enquanto a oferta de vagas referente ao setor privado registrou taxa média de crescimento em torno de 11% ao ano, para o setor público foi registrado crescimento de cerca de 6% ao ano, de 1991 a 2015.

Figura 1 - Razão de Demanda Efetiva – RDE para ingresso, quantidade de Inscritos e Vagas oferecidas em cursos presenciais, segundo ano de referência. Brasil. 1991-2015



Fonte: Com base na *SinopseCES* (vide seção 3.1)

Figura 2 - Razão de Demanda Efetiva – RDE para ingresso, quantidade de Inscritos e Vagas oferecidas em cursos presenciais, por categorias administrativas (a) pública e (b) privada, segundo ano de referência. Brasil. 1991-2015.



Fonte: Com base na *SinopseCES* (vide seção 3.1)

Aproveitamento de Vagas

A Figura 3 apresenta a evolução da TAV no período de 1991 a 2015. Em 1991, primeiro ano considerado, cerca de 82,6% das vagas foram preenchidas por ingressantes, indicando que menos de 20% das vagas ofertadas não foram ocupadas.

Pode-se observar o padrão de decréscimo da curva da TAV a partir de 1999 até 2009: variou de 80% em 1999 para 48% em 2009. A Figura 3 também apresenta a variação anual do total de vagas e de ingressantes. Nota-se que o aumento de vagas não preenchidas refletiu um crescimento do número de vagas superior ao do total de ingressantes. Nesse período, a quantidade de vagas cresceu cerca de 11% ao ano, enquanto a de ingressantes em torno de 6%.

Nos três anos seguintes, 2010-2012, verifica-se um padrão de crescimento da TAV, chegando a cerca de 60% em 2012. A redução da quantidade de vagas não preenchidas foi ocasionada por um período de baixo crescimento das vagas e elevado crescimento da quantidade de ingressantes.

Entre 2013 e 2015 há uma redução da TAV até atingir o nível de aproximadamente 40% de ingressantes em relação ao número de vagas ofertadas, o menor nível de toda a série histórica analisada. Nesse período, houve um aumento do número de vagas não acompanhado por um aumento equivalente no número de ingressantes.

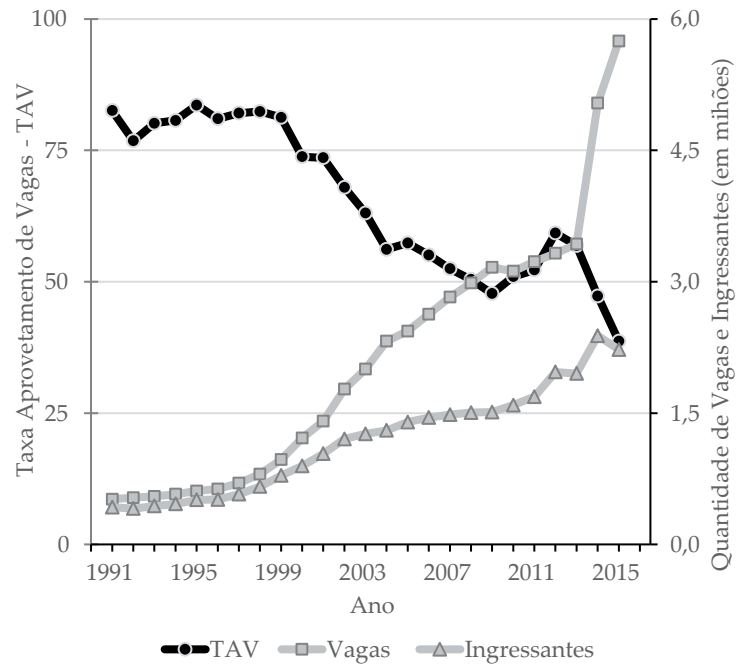
Destacam-se os distintos padrões das TAVs, quando examinadas segundo as categorias administrativas pública, Figura 4 (a), e privada, Figura 4 (b). É evidenciado que o padrão da curva das TAVs, em nível geral (público mais privado), foi expressivamente influenciado pelo padrão da curva referente à categoria administrativa privada. Isso se deu porque tanto a quantidade de vagas ofertadas quanto a de ingressantes pelo setor privado representaram, no período de 1991 a 2015, em cada ano, mais de 2/3 dessas quantidades, contabilizadas em nível geral (público mais privado).

Barros (2015) menciona que o baixo aproveitamento de vagas observado no setor privado pode relacionar-se à estratégia da realização de estoque (ou reserva) de vagas, em que instituições privadas buscam antecipar-se à demanda, esquivando-se de trâmites burocráticos para abertura de cursos e vagas junto ao Ministério da Educação (MEC).

No que tange ao ensino superior público, as TAVs apresentam trajetória estável, com valores entre 90% e 95%, até início da década de 2010. Entre 2011 e 2015 a TAV de instituições públicas caiu de 88% para 70%. Cabe registrar que a categoria administrativa pública associada à esfera Federal apresentou TAVs acima de 97% nas décadas de 2000 e 2010, enquanto que nas esferas Estadual e Municipal essas taxas giraram em torno de 95% e 68%.

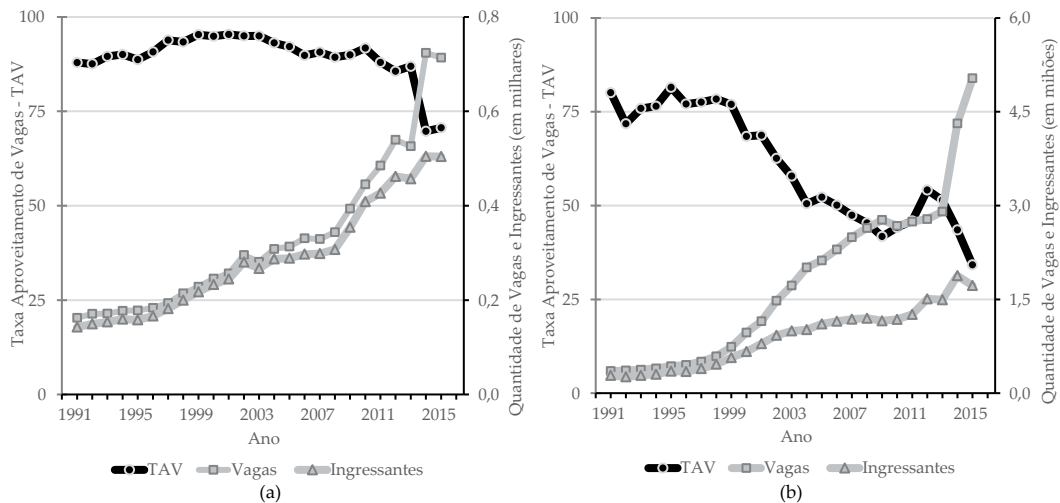
O decréscimo do aproveitamento de vagas em instituições públicas na década de 2010 pode ter relação, em alguma medida, com o favorecimento, por meio do SiSU, de poder concorrer a vagas em diferentes instituições pelo país, além de permitir, após o ingresso, a opção de mudar de curso ou universidade, aproveitando o resultado do processo seletivo, independentemente do local de residência do candidato/estudante.

Figura 3 - Taxa de Aproveitamento de Vagas – TAV, quantidade de Inscritos e Vagas oferecidas em cursos presenciais, segundo ano de referência. Brasil. 1991-2015



Fonte: Com base na *SinopseCES* (vide seção 3.1)

Figura 4 - Taxa de Aproveitamento de Vagas – TAV, quantidade de Inscritos e Vagas oferecidas em cursos presenciais, categorias administrativas (a) pública e (b) privada, segundo ano de referência. Brasil. 1991-2015



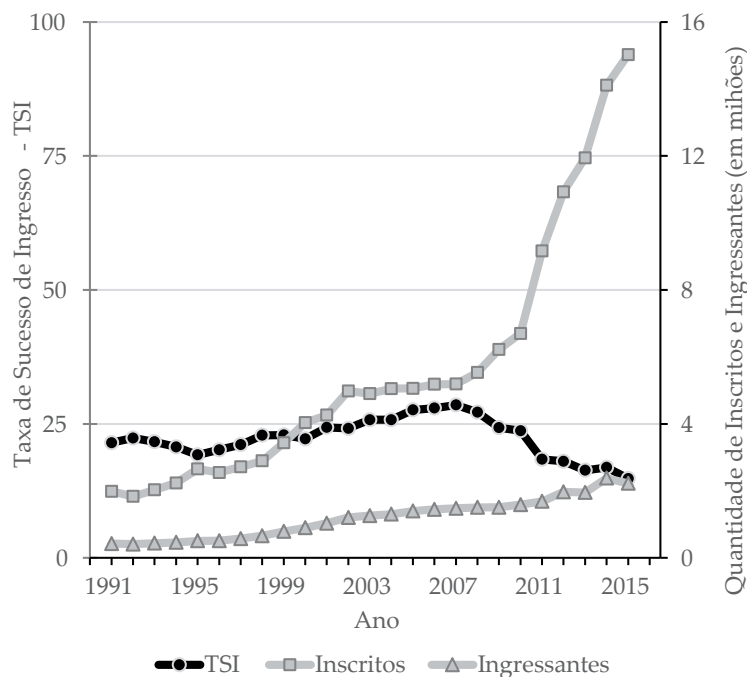
Fonte: Com base na *SinopseCES* (vide seção 3.1)

Ingressantes e inscritos

Por meio da Figura 5 observa-se a evolução da Taxa de Sucesso de Ingressos (TSI) no período de 1991 a 2015. Em 1991, a proporção da demanda por vagas que foi atendida correspondeu a 21%. Essa proporção oscilou ao redor de 20% até 1995. Entre 1996 e 2007, observa-se uma trajetória crescente: a TSI variou de cerca de 20%, em 1996, para cerca de 28% em 2007. A Figura 5 também apresenta a variação da quantidade de ingressantes e de inscritos. Assim, constata-se que esse maior acolhimento da demanda está relacionado com o fato de que nesse período houve uma estabilidade na quantidade de inscritos e um aumento no número de ingressantes.

Nos anos seguintes, observa-se um padrão de decrescimento na curva da TSI que, em 2015, registrou percentual de ingressantes em relação ao número de inscritos próximo de 15%. Nesses anos, o crescimento da quantidade de inscritos superou o do total de ingressantes, o que resultou nessa sensível elevação da demanda reprimida por vagas (complemento da TSI chegou a 85%).

Figura 5 - Taxa de Sucesso de Ingressantes – TSI, quantidade de Inscritos e Vagas oferecidas em cursos presenciais, segundo ano de referência. Brasil. 1991-2015

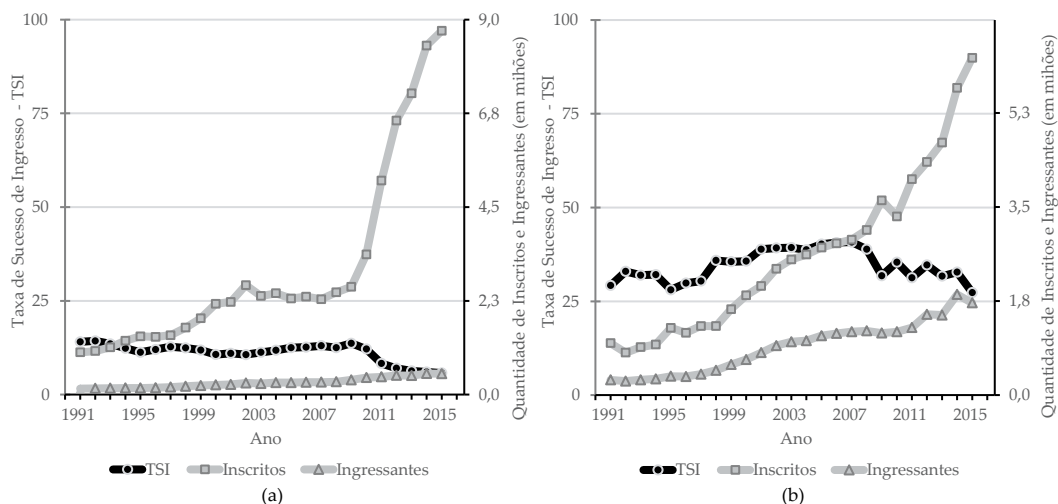


Fonte: Com base na *SimonseCES* (vide seção 3.1)

A Figura 6 apresenta a evolução da TSI desagregada segundo as categorias administrativas pública e privada mostrando maior volatilidade no comportamento da TSI privada. Além disso, a proporção da demanda atendida foi maior na rede privada que na pública, e há um aumento dessa diferença ao longo do período de estudo.

No período 1996-2007, caracterizado pelo aumento da proporção da demanda atendida em geral, enquanto a TSI pública oscila e termina se elevando de 12% para 13%, a TSI privada eleva-se consistentemente de 30% para 41%. De 2008 a 2015, a TSI se reduz tanto na rede pública quanto na privada, e chega a 6% e 27%, respectivamente.

Figura 6 - Taxa de Sucesso de Ingressantes – TSI, quantidade de Inscritos e Vagas oferecidas em cursos presenciais, categorias administrativas (a) pública e (b) privada, segundo ano de referência. Brasil. 1991-2015



Fonte: Com base na *SinopseCES* (vide seção 3.1)

Conclusão

Este trabalho visou investigar a evolução da *demanda* e da *oferta* de vagas para *ingresso* na educação superior no Brasil. O período analisado se caracterizou tanto pelo aumento da demanda (inscrições) quanto pelo aumento da oferta (vagas). Nos períodos 1996-2007 e 2013-2015, houve um crescimento maior da quantidade de vagas, reduzindo a relação inscritos por vagas e significando maior equilíbrio entre demanda e oferta. Já 2008-2013 corresponde a um movimento oposto. Comparando os pontos inicial e final do período, a quantidade de inscritos por vagas se reduziu de 3,8 para 2,6, o que significa que prevaleceu uma aproximação entre demanda e oferta. Ao fim do período, enquanto o total de inscrições por vaga praticamente dobrou na rede pública, na rede privada houve aparente equacionamento entre demanda e oferta.

Em relação ao aproveitamento das vagas oferecidas, destaca-se a ocorrência de redução da quantidade de ingressantes por vagas, notadamente nos períodos 1999-2009 e 2013-2015. Esse aumento de vagas não preenchidas ocorreu num cenário em que o seu crescimento superou o de ingressantes, e de forma mais expressiva na rede privada.

Já a demanda atendida (relação de ingressantes por inscritos) variou de cerca de 21%, em 1991, para 15%, em 2015. Chama a atenção a diferença entre as redes pública e privada que, em 2015, apresentavam TSI de 6% e 27%, respectivamente.

Em suma, o aumento do acesso como possibilidade de ingresso está caracterizado pela relevante ampliação da quantidade de vagas, de inscrições e de ingressantes do período. Esses eventos resultaram em uma aproximação da demanda (quantidade de inscrições) e da oferta (total de vagas). No entanto, alguns fatores chamam a atenção para um certo desajuste entre demanda e oferta. Em primeiro lugar, o aumento da oferta foi essencialmente conduzido pela rede privada. Nas instituições de ensino superior na rede pública, o crescimento de vagas não superou o de inscrições e, assim, houve uma elevação da demanda efetiva. Em segundo lugar, mesmo na rede privada, há um desequilíbrio entre oferta e demanda evidenciado pelo aumento de vagas não preenchidas.

Durante o período considerado neste estudo, aumentou a participação do setor privado na educação superior brasileira. Estudos no campo da educação superior com abordagem na expansão desse nível educacional e/ou na relação público-privado associada, como os encontrados, por exemplo, em Ristoff (Ristoff, 2013, 2014; RISTOFF, 2016), Sampaio (Sampaio,

2015) e Sguissardi (2014, 2015), expõem resultados semelhantes e indicam um processo de privatização/mercantilização no ensino superior no Brasil, favorecendo o expressivo aumento da participação do setor privado no provimento de vagas.

Importante ter em conta o papel da educação superior no desenvolvimento socioeconômico de uma sociedade, na relação educação, formação profissional e trabalho, e em proporcionar ascensão social, econômica e cultural de indivíduos. Conferências que tratam da educação superior em âmbito internacional, tais como a Conferência Regional de Educação Superior na América Latina e no Caribe e a Conferência Mundial Sobre Educação Superior (IESALC-UNESCO, 2008; Rosa, 2014; Unesco, 2009), além de reiterarem o papel transformador da realidade socioeconômica em uma sociedade pela educação, têm manifestado a educação superior como um bem público e de responsabilidade especial dos governos.

Nesse aspecto, cabe considerar que as ações em formulação e implementação de políticas e programas, regulação, bem como modelos de investimento e financiamento associados ao sistema educacional, estão sujeitas a disputas e arranjos no campo político e econômico (e do mercado) que impactam na forma de prover a educação superior.

Nesse ambiente, cabe refletir que as políticas educacionais implementadas durante um mandato eletivo podem e/ou necessitam dispor de meios que oportunizem de forma efetiva a equidade de *acesso* – em sua concepção mais ampla (ingresso, permanência e formação qualificada), de modo que desigualdades sociais, econômicas e culturais, as quais uma sociedade tenha sido ou esteja exposta e/ou experimentando, não sejam empecilhos para o *ingresso* ao ambiente acadêmico de indivíduos em situação socioeconômica menos favorecida, sobretudo, contemplando *acesso* a cursos e instituições de ensino superior de maior prestígio.

Referências

- AMARAL, N. C. (2008). *Financiamento da educação superior no Brasil: gastos com as IFEs – de Fernando Collor a Luiz Inácio Lula da Silva*. In M. Bittar, J. F. de; Oliveira, & M. Morosini (Eds.), *Educação Superior no Brasil: 10 anos pós-LDB*. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira.
- BARROS, A. da S. X. (2015). *Expansão da educação superior no Brasil: limites e possibilidades*. *Educação & Sociedade*, 36(131), 361–390. <https://doi.org/10.1590/ES0101-7330201596208>
- BOLFARINE, H., & BUSSAB, W. de O. (2005). *Elementos de amostragem*. Edgard Blucher.
- BRASIL. Congresso Nacional. Câmara dos Deputados. (2015). *Legislação brasileira sobre educação (3rd ed.)*. <http://www2.camara.leg.br/documentos-e-pesquisa/publicacoes/edicoes/paginas-individuais-dos-livros/legislacao-brasileira-sobre-educacao>
- BRASIL. (2007). *O plano de desenvolvimento da educação: razões, princípios e programas*.
- CARVALHO, C. H. A. de. (2006). *O ProUni no governo Lula e o jogo político em torno do acesso ao ensino superior*. *Educação & Sociedade*, 27, 979–1000. <https://doi.org/10.1590/S0101-73302006000300016>
- CARVALHO, C. H. A. de. (2011). *A política pública para a educação superior no Brasil (1995-2008): ruptura e/ou continuidade?* Universidade Estadual de Campinas.
- CARVALHO, C. H. A. de. (2015). *A política pública de expansão para a educação superior entre 1995 e 2010: uma abordagem neoinstitucionalista histórica*. *Revista Brasileira de Educação*, 20(60), 51–76.
- CEZNE, A. N. (2006). *O direito à educação superior na Constituição Federal de 1988 como direito fundamental* (Vol. 31, Issue 01, pp. 1–10). UFSM, Centro de Educação. <http://coralx.ufsm.br/revce/revce/2006/01/a8.htm>
- CHAVES, V. L. J. (2015). *Política de financiamento e a expansão da educação superior no Brasil: o público e o privado em questão*. *Etd Educacao Tematica Digital*, 17(2), 427–441.
- CHAVES, V. L. J. ., & AMARAL, N. C. (2014). *Política de financiamento da educação superior – análise dos Planos Nacionais de Educação pós-constituição/1988*. *Revista Eletrônica de Educação*, 8(1), 43–55.
- DOURADO, L., OLIVEIRA, J. F. De, & SANTOS, C. D. A. (2007). *A qualidade da educação: conceitos e definições*. In *Textos para Discussão N24 (INEP)* (pp. 1–68). Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira.
- FRANCO, M. A. C. (1985). *Acesso à universidade: uma questão política e um problema ideológico*. *Educação e Seleção*, 12, 5–8.
- GADOTTI, M. (2012). *Trabalho e educação numa perspectiva emancipatória*. II Fórum Mundial de Educação Profissional e Tecnológica, 1–7. http://2sitefmept.ifsc.edu.br/images/stories/pdf2012/apresentacao_debate4.pdf
- GONÇALVES, N. G., & VIEIRA, C. S. (2015). *Extensão Universitária no período da ditadura: concepções e relações com a Doutrina de Segurança Nacional e Desenvolvimento*. *Antíteses*, 8(15), 269. <https://doi.org/10.5433/1984-3356.2015v8n15p269>
- IESALC-UNESCO. (2008). *Declaração da Conferência Regional de Educação Superior na América Latina e no Caribe*. IESALC-UNESCO.
- INEP. (2015). *Censo da educação superior 2013: resumo técnico* (Vol. 1, p. 80). Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. <https://doi.org/10.1017/CBO9781107415324.004>

- KUENZER, A. Z. (2005). *Exclusão incluyente e inclusão excludente*. In J. C. Lombardi, D. Saviani, & J. L. Sanfelice (Eds.), *Capitalismo Trabalho e Educação* (3rd ed., pp. 77–95). Editora Autores Associados.
- LUZ, J. N. N. D. (2017). *Políticas de ingresso na educação superior pública no Brasil: contextos, concepções, movimentos e processos seletivos em perspectiva*. Universidade Federal de Goiás.
- MARTINS, C. B. (2009). *A reforma universitária de 1968 e a abertura para o ensino superior privado no Brasil*. *Educ. Soc. Campinas*, 30(106), 15–35.
- OECD. (2017). *Brasil*. In *Education at a Glance 2017: OECD Indicators*. OECD. <https://doi.org/10.1787/eag-2016-en>
- PACHECO, E. (Ed.). (2011). *Institutos Federais: uma revolução na educação profissional e tecnológica*. Ed. Moderna.
- POCHMANN, M. (2012). *Trabalho e formação*. *Educação & Realidade*, 37(2), 491–508. http://www.scielo.br/scielo.php?script=sci_arttext&pid=S2175-62362012000200009&lng=pt&nrm=iso
- QUEIROZ, F. C. B. P., QUEIROZ, J. V., VASCONCELOS, N. V. C. de, FURUKAVA, M., HÉKIS, H. R., & PEREIRA, F. A. B. (2013). *Transformações no ensino superior brasileiro: análise das Instituições Privadas de Ensino Superior no compasso com as políticas de Estado*. *Ensaio: Aval. Pol. Públ. Educ.*, 21(79), 349–370. <https://doi.org/10.1590/S0104-40362013000200009>
- RISTOFF, D. (2013). *Vinte e um anos de educação superior - expansão e democratização*. In *Cadernos do GEA* (Issue 3, p. 56). FLACSO, GEA; UERJ, LPP.
- RISTOFF, D. (2014). *O novo perfil do campus brasileiro: uma análise do perfil socioeconômico do estudante de graduação*. *Avaliação: Revista Da Avaliação Da Educação Superior (Campinas)*, 19(3), 723–747. <https://doi.org/10.1590/S1414-40772014000300010>
- RISTOFF, D. I. (2016). *Democratização do Campus: impactos dos programas de inclusão sobre o perfil da graduação*. *Gea*, 9.
- ROSA, C. de M. (2014). *Marcos legais e a educação superior no Século XXI*. *Revista Eletrônica de Educação*, 8(3), 236–250. <https://doi.org/http://dx.doi.org/10.14244/198271991029>
- SAMPAIO, H. (2015). *Novas dinâmicas do ensino superior no Brasil: o público e o privado*. In R. Heringer (Ed.), *Democratização da educação superior no Brasil: Novas dinâmicas dilemas e aprendizados* (7th ed., pp. 8–25). FLACSO, GEA, UERJ, LPP, 2012.
- SÄRNDAL, C.-E., SWENSSON, B., & WRETMAN, J. (1992). *Model Assisted Survey Sampling*. Springer.
- SGUISSARDI, V. (2008a). *Modelo de expansão da educação superior no Brasil: predomínio privado/mercantil e desafios para a regulação e a formação universitária*. *Educ. Soc. Campinas*, 29, 991–1022. <https://doi.org/10.1590/S0101-73302008000400004>
- SGUISSARDI, V. (2008b). *Modelo de expansão da educação superior no Brasil: predomínio privado/mercantil e desafios para a regulação e a formação universitária*. *Educ. Soc. Campinas*, 29(105), 991–1022. <https://doi.org/10.1590/S0101-73302008000400004>
- SGUISSARDI, V. (2014). *Estudo diagnóstico da política de expansão da (e acesso à) educação superior no Brasil: 2002-2012*.
- SGUISSARDI, V. (2015). *Educação superior no Brasil. democratização ou massificação mercantil?* *Educ. Soc.*, Campinas, 36(133), 867–889. <https://doi.org/10.1590/ES0101-73302015155688>
- SILVA, M. D. G. M. Da, & VELOSO, T. C. M. A. (2013). *Acesso nas políticas da educação superior: Dimensões e indicadores em questão*. *Avaliação*, 18(3), 727–747.

TACHIBANA, T. Y., FILHO, N. M., & KOMATSU, B. (2005). *O ensino superior no Brasil*. Insper, Centro de Políticas Públicas.

UNESCO. (2009). *Conferência Mundial sobre Ensino Superior 2009 As Novas Dinâmicas do Ensino Superior e Pesquisas para a Mudança e o Desenvolvimento Social* (pp. 1–9).

Anexos

Anexo A1: Variáveis derivadas como insumos para obtenção de indicadores

As variáveis derivadas $Y_{2,a,u}$ e $Y_{3,a,u}$, no ano de referência a , unidade elementar u , são obtidas a partir de variáveis originárias das bases de dados *Microdados* do CES (vide Quadro A1), e suas respectivas formas de obtenção foram dadas por:

- Para ano a , $a \in \{2011, 2012\}$:

$$Y_{2,a,u} = \sum_{k=2.1}^{2.4} Y_{k,a,u}; \quad Y_{3,a,u} = \sum_{k=3.1}^{3.4} Y_{k,a,u}$$

- Para ano a , $a \in \{2013\}$:

$$Y_{2,a,u} = \sum_{k=2.5}^{2.12} Y_{k,a,u}; \quad Y_{3,a,u} = \sum_{k=3.5}^{3.12} Y_{k,a,u}$$

- Para ano a , $a \in \{2014, 2015\}$:

$$Y_{2,a,u} = \sum_{k=2.13}^{2.24} Y_{k,a,u}; \quad Y_{3,a,u} = \sum_{k=3.13}^{3.24} Y_{k,a,u}$$

em que as descrições das variáveis Y_k de origem, com $k \in \{2.1, \dots, 2.24, 3.1, \dots, 3.24\}$, são apresentadas no Quadro A1.

Quadro A1. Ano de referência, nome e descrição das variáveis referentes às quantidades de vagas e de inscritos nas bases de dados Microdados do Censo da Educação Superior. 2001-2015

ANO	VARIAVEIS	DESCRICAÇÃO DAS VARIÁVEIS
2011 e 2012	Y _{2.1}	Nº de vagas oferecidas no turno integral
	Y _{2.2}	Nº de vagas oferecidas no turno matutino
	Y _{2.3}	Nº de vagas oferecidas no turno noturno
	Y _{2.4}	Nº de vagas oferecidas no turno vespertino
	Y _{3.1}	Nº de inscritos por processo seletivo no turno integral
	Y _{3.2}	Nº de inscritos por processo seletivo no turno matutino
	Y _{3.3}	Nº de inscritos por processo seletivo no turno noturno
	Y _{3.4}	Nº de inscritos por processo seletivo no turno vespertino
2013	Y _{2.5}	Nº de vagas oferecidas no turno integral para Outras Vagas
	Y _{2.6}	Nº de vagas oferecidas no turno matutino para Outras Vagas
	Y _{2.7}	Nº de vagas oferecidas no turno noturno para Outras Vagas
	Y _{2.8}	Nº de vagas oferecidas no turno vespertino para Outras Vagas
	Y _{2.9}	Nº de vagas no processo seletivo principal no turno integral
	Y _{2.10}	Nº de vagas no processo seletivo principal no turno matutino
	Y _{2.11}	Nº de vagas no processo seletivo principal no turno noturno
	Y _{2.12}	Nº de vagas no processo seletivo principal no turno vespertino
	Y _{3.5}	Nº de inscritos no turno integral para Outras Vagas
	Y _{3.6}	Nº de inscritos no turno matutino para Outras Vagas
	Y _{3.7}	Nº de inscritos no turno noturno para Outras Vagas
	Y _{3.8}	Nº de inscritos no turno vespertino para Outras Vagas
Y _{3.9}	Nº de inscritos no processo seletivo principal no turno integral	
Y _{3.10}	Nº de inscritos no processo seletivo principal no turno matutino	
Y _{3.11}	Nº de inscritos no processo seletivo principal no turno noturno	
Y _{3.12}	Nº de inscritos no processo seletivo principal no turno vespertino	

	Y _{2.13}	Nº de vagas oferecidas no turno integral para programas especiais
	Y _{2.14}	Nº de vagas oferecidas no turno matutino para programas especiais
	Y _{2.15}	Nº de vagas oferecidas no turno noturno para programas especiais
	Y _{2.16}	Nº de vagas oferecidas no turno vespertino para programas especiais
	Y _{2.17}	Nº de vagas novas oferecidas nos cursos de turno integral
	Y _{2.18}	Nº de vagas novas oferecidas nos cursos de turno matutino
	Y _{2.19}	Nº de vagas novas oferecidas nos cursos de turno noturno
	Y _{2.20}	Nº de vagas novas oferecidas nos cursos de turno vespertino
	Y _{2.21}	Nº de vagas remanescentes oferecidas nos cursos de turno integral
	Y _{2.22}	Nº de vagas remanescentes oferecidas nos cursos de turno matutino
	Y _{2.23}	Nº de vagas remanescentes oferecidas nos cursos de turno noturno
	Y _{2.24}	Nº de vagas remanescentes oferecidas nos cursos de turno vespertino
2014 e 2015	Y _{3.13}	Nº de inscritos nas vagas oferecidas para programas especiais no turno integral
	Y _{3.14}	Nº de inscritos nas vagas oferecidas para programas especiais no turno matutino
	Y _{3.15}	Nº de inscritos nas vagas oferecidas para programas especiais no turno noturno
	Y _{3.16}	Nº de inscritos nas vagas oferecidas para programas especiais no turno vespertino
	Y _{3.17}	Nº de inscritos no processo seletivo de vagas novas no turno integral
	Y _{3.18}	Nº de inscritos no processo seletivo de vagas novas no turno matutino
	Y _{3.19}	Nº de inscritos no processo seletivo de vagas novas no turno noturno
	Y _{3.20}	Nº de inscritos no processo seletivo de vagas novas no turno vespertino
	Y _{3.21}	Nº de inscritos no processo seletivo de vagas remanescentes no turno integral
	Y _{3.22}	Nº de inscritos no processo seletivo de vagas remanescentes no turno matutino
	Y _{3.23}	Nº de inscritos no processo seletivo de vagas remanescentes no turno noturno
	Y _{3.24}	Nº de inscritos no processo seletivo de vagas remanescentes no turno vespertino

Fonte: Com base nos arquivos de Microdados do Censo da Educação Superior (INEP), 2011, 2012, 2013, 2014, 2015.

Seção 4

Análises Setoriais e de Gênero sobre Mercado de Trabalho

Estimativas de desocupação trimestrais para estratos geográficos de Minas Gerais: estimação em pequenos domínios com séries temporais

Caio Gonçalves*

Luna Hidalgo**

Denise Britz do Nascimento Silva***

Jan van den Brakel****

Resumo

A produção de estatísticas cada vez mais granulares geograficamente, especialmente em anos intercensitários, é uma demanda recorrente para os órgãos produtores de estatística e se insere na agenda 2030 para acompanhamento dos Objetivos de Desenvolvimento Sustentável. O presente artigo apresenta modelos de séries temporais para produzir estimativas trimestrais de desocupados em dez regiões (estratos geográficos) do estado de Minas Gerais utilizando a Pesquisa Nacional por Amostra de Domicílios Contínua (PNADC). Para investigar se é possível produzir estimativas mais precisas do que aquelas baseadas no desenho amostral, adotou-se um modelo multivariado de séries temporais para pesquisas repetidas, que leva em consideração o erro amostral das estimativas, lida com a sazonalidade e estima as correlações entre os distúrbios das inclinações das séries como mecanismo de incorporar o efeito espacial. Os resultados indicam que é possível obter ganhos de precisão na ordem de 7,7 a 48,6%. Esses resultados evidenciam um potencial a ser explorado na PNADC, especialmente por ter completado dez anos em 2022, além de apontar a oportunidade de construção de um sistema de estatísticas baseado em modelos nos moldes adotados em programas internacionais de produção de indicadores de mercado de trabalho. pré-natal estiveram significativamente associadas à duração do aleitamento materno exclusivo. Conclui-se que é necessária a contínua intervenção do poder público no que diz respeito ao estímulo a tal prática.

Palavras-chave: Estimação em pequenas áreas; Modelos espaço de estados; Pesquisas amostrais repetidas; PNADC; Estatísticas regionais.

* Doutor pelo programa de pós-graduação da ENCE.

** Doutora pelo programa de pós-graduação da ENCE.

*** Professora aposentada do programa de pós-graduação da ENCE.

**** Maastricht University, Países Baixos.

Introdução

O Censo Demográfico, realizado pelo Instituto Brasileiro de Geografia e Estatística (IBGE), produz um conjunto amplo de informações em diversas temáticas, subsidiando políticas públicas e sendo base para pesquisas amostrais. Uma de suas principais vantagens é a produção de estatísticas em nível municipal. O total de pessoas desocupadas por municípios, por exemplo, só é conhecido nesse levantamento. No entanto, o IBGE segue, como outros países, um planejamento decenal do Censo Demográfico no país, conforme a lei nº 8.184, de 10 de maio de 1991 (Brasil, 1991).

A produção de estatísticas regionais é uma atividade que também está presente em alguns órgãos estaduais de estatística, com estudos e levantamentos específicos para seus estados. Porém, por questões orçamentárias, falta de recursos humanos, duplicidade das informações pesquisadas e aumento da carga de respondentes podem levar à não realização, como ocorreu em Minas Gerais. A Pesquisa de Emprego e Desemprego (PED), realizada pela Fundação João Pinheiro (FJP), órgão oficial das estatísticas do estado, foi descontinuada em maio de 2014 após quase 19 anos em atividade. A PED produzia informações mensais para o mercado de trabalho da região metropolitana de Belo Horizonte (RMBH) (Fundação João Pinheiro, 2014a). Além dessa pesquisa domiciliar, a Pesquisa por Amostra de Domicílios de Minas Gerais (PAD-MG) da FJP também foi interrompida na terceira edição realizada em 2013. Essa pesquisa era bianual e produzia indicadores de saúde, educação, trabalho e renda para a regionalização existente em Minas Gerais na época de sua execução (regiões de planejamento), abrangendo todo o território do estado (Fundação João Pinheiro, 2014b).

Adicionalmente, existe uma demanda para que governos locais e regionais se integrem efetivamente na agenda 2030, considerando sua realidade na definição das metas e indicadores. As iniciativas para acompanhamento dos 17 Objetivos de Desenvolvimento Sustentável (ODS) não se limitam ao nível nacional (United Nations Statistics Division, 2022), mas também alcançam o âmbito estadual e municipal no Brasil. Um exemplo é a rede do Observatório do Milênio de Belo Horizonte que acompanha os ODS para a capital do estado de Minas Gerais, elaborando relatórios e construindo banco de dados, entre outras atividades (Observatório do Milênio de Belo Horizonte, 2022). Alguns exemplos de indicadores monitorados incluem a taxa de desocupação total, a taxa de desocupação na faixa etária de 18 a 24 anos e a proporção de trabalhadores ocupados em atividades não agrícolas informais.

Diante da necessidade de produzir estatísticas municipais ou para agrupamentos de municípios, a capacidade de produzir informações a partir de pesquisas amostrais pode ser ampliada com o uso da abordagem baseada em modelos. A Pesquisa Nacional por Amostra de Domicílios Contínua (PNADC), a maior pesquisa domiciliar amostral em vigor do IBGE, possui estratos geográficos (grupos de municípios) para os quais podem ser investigadas possibilidades de produção de estimativas nas temáticas presentes na pesquisa. Isso não representa estatísticas ao nível municipal, mas sim em um nível intermediário entre Unidade da Federação (UF) e municípios, incluindo Regiões Metropolitanas (RMs), Regiões Integradas de Desenvolvimento (RIDEs)¹ e capitais brasileiras que já são contempladas nas divulgações da PNADC. Constitui um avanço na produção de informações e na compreensão das realidades locais para diversos estados, possibilitando uma recomposição parcial das pesquisas descontinuadas em Minas Gerais. O IBGE passou a divulgar, em 2022, indicadores da PNADC para 146 estratos geográficos como estatísticas experimentais². Cabe destacar, entretanto, que já era possível calcular indicadores por estratos geográficos desde o início da pesquisa, que foi em janeiro de 2012. No entanto, a pesquisa não foi desenhada para produzir resultados precisos nesse nível geográfico, o que acende um alerta para seu uso.

¹ Denomina-se Região Integrada de Desenvolvimento Econômico (RIDE) uma região metropolitana composta por municípios que pertencem a diferentes unidades da federação como a “Região Integrada de Desenvolvimento do Distrito Federal e entorno” que inclui o Distrito Federal e municípios de Goiás e Minas Gerais.

² Estatísticas experimentais são aquelas que ainda não atingiram um grau completo de maturidade em diferentes aspectos como harmonização, cobertura ou metodologia (IBGE, 2022a).

Conforme estudos do IBGE, embora tenham sido identificados apenas casos pontuais de estimativas com coeficiente de variação (CV) acima de 30% em trimestres específicos para três indicadores (taxa de desocupação, nível de ocupação e taxa de analfabetismo) (IBGE, 2022b), a precisão dos resultados é um fator que requer atenção, levando sempre em consideração o aproveitamento da amostra em cada trimestre. Além disso, a divulgação de indicadores por domínios de interesse como sexo, idade e atividades econômicas pode ser inviabilizada devido à perda de precisão, caracterizando um problema de estimação em pequenos domínios (Rao; Molina, 2015, Cap. 1). A abordagem baseada em modelos se aplica a esses casos devido à capacidade de reduzir o erro padrão das estimativas. Em especial, a PNADC, tendo completado dez anos de coleta de informações em 2022, permite o uso de métodos de séries temporais para pesquisas amostrais.

O Bureau of Labor Statistics enfrentou um problema semelhante para a produção de estatísticas mais desagregadas do mercado de trabalho (Local Area Unemployment Statistics - LAUS). Em 1989, publicou as primeiras estimativas baseadas em modelos de séries temporais para os estados americanos usando a Current Population Survey (CPS). A partir de 1996, deixou de publicar as estimativas diretas³ e passou a usar apenas as produzidas pelo modelo (U.S. Bureau of Labor Statistics, 2018). Outra experiência internacional é o caso do Statistics Netherlands que produz indicadores mensais por sexo e idade para total de ocupados, desocupados, força de trabalho e taxa de desocupação sob a abordagem baseada em modelo de séries temporais para pesquisas repetidas desde 2010.

Diante do cenário apresentado, tem-se o seguinte questionamento: é possível produzir estimativas mais precisas utilizando modelos de séries temporais levando em consideração a acumulação de informação não só no tempo, mas também no espaço?

A proposta deste estudo é desenvolver um modelo de séries temporais de estimação em pequenos domínios para produzir estimativas trimestrais do total de desocupados para os dez estratos geográficos do estado de Minas Gerais, preenchendo a lacuna de indicadores trimestrais regionalizados. Os resultados do modelo foram comparados com as estimativas diretas, bem como com os resultados de um modelo estrutural univariado, para verificação de ganhos ao usar uma estrutura multivariada. Gonçalves et al. (2022) encontrou resultados favoráveis em relação a modelos multivariados ao modelar em conjunto os estados da região Sudeste do Brasil.

A contribuição do presente artigo é a avaliação das possibilidades de produção estatística para grupos de municípios. Resultados positivos significam disponibilizar dados mais atualizados que as informações censitárias decenais e em maior frequência do que os dados anuais da PNADC para os estratos geográficos. Essas informações podem ser empregadas para o conhecimento das condições socioeconômicas e para a identificação de necessidades de emprego local que, por sua vez, pode subsidiar o planejamento e a execução de programas governamentais de emprego e renda.

Além desta introdução, o artigo está subdividido da seguinte forma: a seção 2 descreve os estratos geográficos em estudo e sua diversificação no território; a seção 3 introduz o modelo de séries temporais para pequenos domínios; a seção 4 apresenta os resultados encontrados e, por fim, a seção 5 encerra com as considerações finais.

Estratos geográficos em Minas Gerais

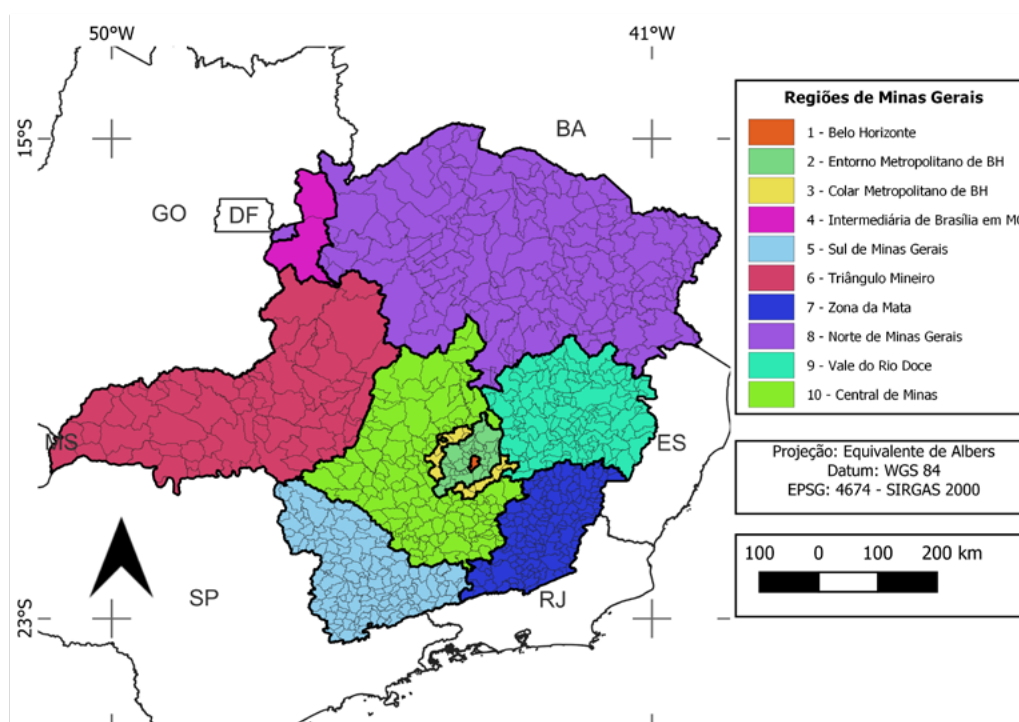
Minas Gerais é a terceira maior economia do país, representando 9% do Produto Interno Bruto (PIB) nacional, atrás apenas de São Paulo (31,2%) e do Rio de Janeiro (9,9%) em 2020 (IBGE, 2021). Com uma população de aproximadamente 21,5 milhões de pessoas em 2022 (IBGE, 2020a), Minas Gerais é o segundo estado mais populoso do Brasil, com o maior número de municípios (853) e ocupa cerca de 6,8% da área do território nacional. Em 2021, foi a oitava das 27 UFs no

³ Estimativas baseadas no desenho amostral.

ranking de menores taxas de desocupação, com 12,1%, enquanto a taxa nacional foi de 14% para o mesmo ano (IBGE, 2020b).

O território de Minas Gerais é dividido em dez regiões, seguindo o agrupamento de estratos geográficos da amostra mestra do Sistema Integrado de Pesquisas Domiciliares (SIPD) (Freitas; Antonaci, 2014; Albieri, Bianchini, 2015) do IBGE. Essa regionalização abrange a capital e nove grupos de municípios. A Figura 1 exibe as dez regiões na malha municipal de Minas Gerais agrupadas por estratos geográficos.

Figura 1 – Mapa das regiões construídas com os estratos geográficos - municípios de Minas Gerais



Fonte: IBGE (2010) e dados municipais do SIPD. Elaboração própria.

A RMBH é a junção da capital (Belo Horizonte - BH) com a região do Entorno Metropolitano. A combinação da RMBH com o Colar Metropolitano representa outra agregação geográfica referente à área metropolitana de Belo Horizonte. Apesar de não estar incluída nas desagregações divulgadas pelo IBGE, é possível obter estimativas para o restante do estado usando os microdados da PNADC. Isso abrange as demais regiões atualmente divulgadas no painel da PNADC com estratos geográficos (IBGE, 2022c), com exceção da parte de Minas Gerais pertencente à RIDE de Brasília e do Entorno Metropolitano de BH.

Em termos populacionais, as regiões têm uma distribuição quase uniforme, com percentuais variando de 10 a 13,7%, o que equivale a aproximadamente 2,6 milhões de pessoas em cada estrato geográfico. As exceções são as regiões Intermediária de Brasília e o Colar Metropolitano, que representam 0,5% e 2,8% da população estimada de Minas Gerais em 2019⁴. As razões para essa diferenciação estão relacionadas à construção dos estratos para compor a área metropolitana de Belo Horizonte e a RIDE de Brasília em Minas Gerais. Em termos econômicos, as três

⁴ As estimativas de população utilizadas na PNADC são baseadas nas Projeções Populacionais que são revisadas após a realização do Censo Demográfico e da Contagem da População prevista para acontecer no meio da década. Essa última operação estatística é relevante para captar alterações do total e da distribuição da população ao longo do território brasileiro e permite melhorar as estimativas populacionais intercensitárias.

regiões que mais contribuíram para a economia do estado em 2019, de acordo com o PIB, foram o Triângulo Mineiro, o Entorno Metropolitano de BH e a capital Belo Horizonte, seguidas pelas regiões Sul de Minas e Central de Minas. Essas cinco áreas concentraram 72,3% da produção em 2019.

Ao analisar o indicador PIB *per capita*, observa-se que o valor mais alto pertence ao Triângulo Mineiro (R\$ 38,7 mil). Belo Horizonte (BH), o Colar Metropolitano⁵ e o Entorno de BH ocupam a segunda, terceira e quarta posições, respectivamente, no ranking do PIB per capita em 2019. Em seguida, estão a região Intermediária de Brasília em Minas, composta por apenas dois municípios (Unaí e Buritis⁶), e o Sul de Minas. Abaixo da média de Minas Gerais (R\$ 30,8 mil) estão as regiões Vale do Rio Doce, Central de Minas, Zona da Mata e, por fim, a região com o menor valor do PIB per capita, Norte de Minas.

Observou-se recentemente uma maior desagregação das estatísticas publicadas a partir da PNADC para a área definida como Resto da UF. Essa área foi subdividida em sete regiões: Sul de Minas, Triângulo Mineiro, Norte de Minas, Zona da Mata, Vale do Rio Doce, Central de Minas e Colar Metropolitano. Esses sete estratos geográficos representam uma desagregação de áreas com economias distintas, abrangendo a região com o mais alto PIB *per capita* e a região com o menor valor. Isso ressalta a importância de publicar indicadores desagregados para essas novas áreas de interesse. A Tabela 1 ilustra as diferenças na taxa de desocupação em estratos geográficos do estado de Minas Gerais em relação às estimativas diretas da PNADC.

Tabela 1 – Estimativas trimestrais da taxa de desocupação baseadas no desenho amostral e respectivos coeficientes de variação - Minas Gerais e estratos geográficos - 4º trim. 2019 e 2º trim. 2020

Indicador e nível geográfico	4º trimestre de 2019		2º trimestre de 2020	
	Taxa (%)	CV (%)	Taxa (%)	CV (%)
Minas Gerais	9,5	3,2	12,9	3,1
1 - Belo Horizonte (BH)	10,2	6,8	14,2	6,7
2 - Entorno Metropolitano de BH	11,8	7,0	18,6	6,7
3 - Colar Metropolitano de BH	12,6	12,4	15,8	14,0
4 - Integrada de Brasília em MG	3,1	23,7	10,3	30,9
5 - Sul de Minas	7,5	9,5	12,2	8,8
6 - Triângulo Mineiro	7,2	9,6	9,5	11,2
7 - Zona da Mata	9,2	11,5	10,6	10,3
8 - Norte de Minas	13,0	8,9	14,5	8,8
9 - Vale do Rio Doce	10,0	12,2	13,2	12,2
10 - Central de Minas	6,3	10,9	9,2	9,7

Fonte: PNADC (IBGE, 2022d). Elaboração própria.

A PNADC foi criada para permitir o acompanhamento de estatísticas trimestrais e a evolução do mercado de trabalho no curto, médio e longo prazos, entretanto é multitemática e inclui suplementos com outras informações necessárias para retratar o país (IBGE, 2022e). Sua amostra é probabilística, com um plano amostral conglomerado de dois estágios de seleção. No primeiro estágio, selecionam-se as unidades primárias de amostragem (UPAs), que são os setores censitários, com probabilidade proporcional ao número de domicílios dentro de cada estrato definido. A definição dos estratos está presente no Sistema Integrado de Pesquisas Domiciliares (SIPD),

⁵ Em 2012, dois novos municípios foram acrescentados à região, Bom Jesus do Amparo e São Gonçalo do Rio Abaixo. Essa modificação não foi contemplada no desenho amostral das pesquisas domiciliares do IBGE dado que esse foi planejado anteriormente a modificação mencionada.

⁶ Os municípios Cabeceira Grande e Arinos passaram a fazer parte da RIDE do DF em 2018, porém o SIPD do IBGE produz informações para os dois municípios inicialmente considerados durante a construção da amostra mestra.

sendo o estrato geográfico um deles. No segundo estágio, são selecionadas 14 unidades domiciliares dentro de cada UPA da amostra, usando amostragem aleatória simples, com base na listagem do Cadastro Nacional de Endereços para Fins Estatísticos (CNEFE) atualizada (IBGE, 2022e).

Com exceção das regiões Integrada de Brasília em Minas Gerais e um caso no Colar Metropolitano de Belo Horizonte, as estimativas para todos os outros estratos geográficos apresentaram coeficientes de variação abaixo de 15% no 4º trimestre de 2019 e no 2º trimestre de 2020, que foi o primeiro trimestre após o início da pandemia de COVID-19 e teve uma redução na precisão das estimativas. De acordo com o IBGE (2014, 2021) e o estudo de Albieri (1999), estimativas com CV de até 15% são passíveis de publicação. No entanto, é sempre desejável melhorar a precisão das estatísticas publicadas. Por exemplo, informar que o total de desocupados no 4º trimestre de 2019 no Vale do Rio Doce tem seu valor no intervalo de 85,8 a 140,2 mil pessoas pode resultar em falta de confiabilidade para um gestor público durante a elaboração de programas de geração de emprego. O ideal é produzir estimativas com melhor nível de precisão possível, algo que foi inviabilizado nos trimestres após o início da pandemia, dada a redução do aproveitamento da amostra devido às dificuldades de coleta e à não resposta. Destacam-se os casos dos estratos Região Integrada de Brasília em Minas e Colar Metropolitano, com valores acima de 15%, e casos pontuais do Vale do Rio Doce, Central de Minas, Zona da Mata e Triângulo Mineiro que ultrapassaram essa marca.

As maiores taxas de desocupação estão presentes no Norte de Minas, Vale do Rio Doce e na área metropolitana de Belo Horizonte (Belo Horizonte, Entorno e Colar Metropolitano de BH). Em contrapartida, as áreas Central, Sul de Minas e Triângulo apresentaram as menores taxas de desocupação no 4º trimestre de 2019. Além disso, houve variações de curto prazo expressivas, conforme a Tabela 1, como no caso do Entorno Metropolitano de BH, que aparentemente sofreu uma rápida elevação da desocupação no período analisado. Portanto, a produção de estimativas que possam captar alterações significativas desses indicadores no curto prazo permitiria uma melhor análise da conjuntura local. Além disso, a produção de séries de tendência e o conhecimento do comportamento sazonal nas regiões, se existente, completariam um panorama valioso de evidências. A publicação dessas informações é possível com a utilização de modelos de séries temporais adequados para pesquisas amostrais repetidas na produção de estimativas em pequenos domínios.

É importante observar que o tamanho da amostra efetiva⁷ para os estratos geográficos de Minas Gerais variou de aproximadamente 4.000 a 5.000 pessoas investigadas e de 1.500 a 2.000 domicílios, com exceção das regiões Integrada de Brasília em Minas Gerais, que contaram com cerca de 850 pessoas e 300 domicílios, e o Colar Metropolitano de BH, que teve valores próximos a 1.900 pessoas investigadas e 600 domicílios entrevistados. Foi observado que, no segundo trimestre de 2022, o tamanho da amostra efetiva não retornou completamente ao patamar anterior à pandemia de COVID-19 em todos os estratos geográficos de Minas Gerais.

Modelos de séries temporais para estimação em pequenos domínios

O estudo de séries temporais, e de sua evolução, é de interesse para melhor compreensão do fenômeno medido. As séries temporais podem originar-se de fontes de dados como censos, pesquisas amostrais, registros administrativos e big data. Dependendo da fonte, é necessário considerar aspectos do processo gerador dos dados que podem influenciar na dependência temporal.

No caso de pesquisas amostrais repetidas com sobreposição de amostra, adiciona-se ao verdadeiro fenômeno um ruído intrínseco à formulação do plano amostral da pesquisa, conhecido como erro amostral. Portanto, o modelo de extração de sinal tem sido utilizado para

⁷ Entende-se por amostra efetiva aquela que efetivamente foi coletada, validada e que gerou os resultados finais publicados.

decompor uma série de estimativas diretas de uma pesquisa em dois processos: um associado à evolução do verdadeiro valor do indicador que representa o fenômeno de interesse e outro relacionado ao erro amostral, como sugerido nos estudos iniciais de Scott e Smith (1974) e Scott, Smith e Jones (1977).

O modelo estipulado para o verdadeiro valor do fenômeno também é escolhido com objetivos específicos na análise de séries temporais. Geralmente, séries com alta frequência, como mensais e trimestrais, são dessazonalizadas para facilitar a compreensão de seu comportamento a curto prazo. Portanto, modelos estruturais de séries temporais foram propostos na literatura, nos quais componentes mutuamente independentes são definidos para diferentes tipos de variações temporais, incluindo tendência-ciclo, movimento sazonal e variações residuais.

Grande parte dos sistemas de estatística produzem séries temporais referentes a domínios. Por exemplo, indicadores de desocupação são calculados por estado, por idade, por sexo, entre outros. A desagregação e a combinação desses fatores geram novos domínios para os quais o tamanho da amostra é reduzido, o que afeta a precisão de suas estimativas. Isso se configura em um típico problema de estimação em pequenos domínios ou, também chamado, estimação em pequenas áreas.

Há vasta literatura sobre estimação em pequenos domínios, sendo Rao e Molina (2015) uma das principais referências no assunto. De forma geral, os modelos de estimação em pequenos domínios são modelos multiníveis, que congregam informações amostrais dos vários domínios de pesquisa transversal (cross-section) em um mesmo instante no tempo (Brakel, 2019).

Entretanto, também é possível aproveitar as informações passadas nos casos de pesquisas amostrais repetidas, ou seja, o modelo ganha força tanto no espaço quanto no tempo. Uma forma de incorporar o poder da informação ao longo do tempo na estimação em pequenas áreas/domínios é utilizando modelos de séries temporais. Pfeffermann e Tiller (2006) desenvolveram um modelo estrutural de séries temporais para estimativas do total de ocupados e desocupados em domínios geográficos dos Estados Unidos. Foram elaborados modelos estruturais de séries temporais univariados para cada domínio, e modelos multivariados para todos os domínios, sendo esses estimados conjuntamente incorporando a correlação entre componentes não observáveis (Pfeffermann; Burck, 1990; Pfeffermann; Bleuer, 1993; Krieg; Brakel, 2012; Brakel; Krieg, 2015; Brakel; Krieg, 2016; Boonstra; Brakel, 2019).

Os modelos estruturais multivariados têm-se mostrado uma opção para a redução do erro padrão das estimativas baseadas em modelo comparativamente aquelas baseadas no desenho amostral. O estudo de Bollineni-Balabay, Brakel e Palm (2016) identificou que, para estimativas de transporte de carga rodoviário nacional por conta própria (medida em toneladas) registrados na autoridade holandesa responsável, a redução ocorreu na ordem de 40 a 70%. Nesse caso, foi utilizado um modelo multivariado estrutural de nove domínios (categorias de produtos), além de incluir aspectos específicos como o redesenho da Dutch Road Transportation Survey.

No caso do programa americano de produção de indicadores de mercado de trabalho em pequenas áreas (U.S. Bureau of Labor Statistics, 2018), modelos estruturais são ajustados para o total de ocupados e desocupados de forma independente para os 50 estados e o Distrito de Columbia, e então seus resultados são somados originando estatísticas da força de trabalho e da taxa de desocupação calculada como percentual de pessoas desocupadas na força de trabalho. Nos modelos são utilizadas as estimativas diretas da CPS e os registros administrativos de pagamentos aos trabalhadores e requerimentos de seguro-desemprego.

As estimativas baseadas em modelos podem se diferenciar bastante das estimativas diretas, especialmente nos casos de áreas com tamanho de amostra muito pequenas. Isso ocorre porque a pesquisa amostral originária não foi planejada para alcançar precisão aceitável no nível das pequenas áreas e sim, para domínios ou áreas mais agregados. Por mais que isso possa ocorrer, existe uma necessidade de que as estimativas em pequenas áreas possam ser usadas e, portanto, é necessário que a soma das estimativas das pequenas áreas (ou domínios) esteja compatível, ao menos próxima, da estimativa direta da área maior. Isso se justifica porque, em geral,

as estimativas diretas para áreas maiores são aproximadamente não viciadas e apresentam boa precisão.

Diante disso, a modelagem adotada neste artigo se configura como um modelo de extração de sinal, com um componente para representar o erro amostral, e considera um modelo estrutural que permite a estimação da sazonalidade. Possui um caráter multivariado, ao estimar em conjunto as componentes estruturais de todas as séries dos domínios estudados ganhando força das informações no espaço ao modelar a correlação entre as tendências das séries dos vários estratos geográficos.

Modelo de extração de sinal

Ao considerar um modelo estrutural para uma série temporal advinda de uma pesquisa amostral repetida, é necessário considerar a estrutura de correlação das observações imposta pelo desenho da pesquisa. O modelo de extração de sinal, conforme abordado inicialmente por Scott e Smith (1974) e Scott, Smith e Jones (1977), decompõe uma série de estimativas diretas $\{\widehat{y}_t\}$ em dois processos: o verdadeiro valor associado ao fenômeno $\{\theta_t\}$ e o distúrbio representado pelo erro amostral $\{e_t\}$ tal que

$$\widehat{y}_t = \theta_t + e_t \quad \forall t \quad (1)$$

O parâmetro populacional não observável θ_t pode ser representado pelos componentes estruturais clássicos de séries temporais: tendência (L_t), sazonalidade (S_t) e irregular (I_t).

$$\theta_t = L_t + S_t + I_t, \quad I_t \sim N(0, \sigma_I^2) \quad (2)$$

L_t representa um componente de tendência-ciclo, dado que se trata de um modelo para uma série temporal curta (42 observações trimestrais). Com intuito de apresentar esse componente com menores oscilações, optou-se por um modelo que apenas inclui um distúrbio na equação de inclinação ($\eta_{R,t}$). Esse modelo é referenciado na literatura como modelo de tendência suavizada (Durbin; Koopman, 2012). Assim, a tendência depende do seu comportamento no período anterior mais um termo que representa a inclinação (R_t) que é modelada por um passeio aleatório

$$L_t = L_{t-1} + R_{t-1} \quad (3)$$

$$R_{t-1} = R_{t-1} + \eta_{R,t}, \quad \eta_{R,t} \sim N(0, \sigma_R^2) \quad \text{cov}(\eta_{R,t}, \eta_{R,t'}) = 0, \quad \forall t \neq t' \quad (4)$$

A sazonalidade está expressa em um formato trigonométrico, sendo que o somatório de duas frequências ($S_{i,t}$) gera o componente sazonal ($S_{i,t}$) conforme as equações 5-8 (Durbin; Koopman, 2012).

$$S_t = \sum_{i=1}^{\frac{s}{2}=6} S_{i,t} \quad (5)$$

$$\begin{pmatrix} S_{i,t} \\ S_{i,t}^* \end{pmatrix} = \begin{bmatrix} \cos(\pi i/6) & \sin(\pi i/6) \\ -\sin(\pi i/6) & \cos(\pi i/6) \end{bmatrix} \begin{pmatrix} S_{i,t-1} \\ S_{i,t-1}^* \end{pmatrix} + \begin{pmatrix} \eta_{S,i,t} \\ \eta_{S,i,t}^* \end{pmatrix} \quad (6)$$

$$\begin{pmatrix} \eta_{S,i,t} \\ \eta_{S,i,t}^* \end{pmatrix} \sim N(0, \sigma_S^2 \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}), \quad \text{onde } i = 1, \dots, 6 \quad (7)$$

$$\text{cov}(\eta_{S,i,t}, \eta_{S,i,t'}) = 0, \quad \text{cov}(\eta_{S,i,t}^*, \eta_{S,i,t'}^*) = 0, \quad \forall t \neq t' \quad (8)$$

O erro amostral é expresso de forma proporcional ao erro padrão das estimativas diretas conforme proposta de Binder e Dick (1989), resultando em \tilde{e}_t como erro amostral escalonado

na equação 9. Esse mecanismo permite que efeitos sobre a pesquisa também sejam captados no modelo, por exemplo, oscilações no tamanho da amostra. Portanto, além da própria série de interesse baseada no desenho amostral, a série temporal do desvio padrão das estimativas também é incorporada na modelagem.

$$e_t = \widehat{c}_t \tilde{e}_t, \quad \widehat{c}_t = \sqrt{\text{var}(\widehat{y}_t)}, \quad (9)$$

Além disso, a PNADC possui um esquema de rotação da amostra da forma 1-2(5) no qual os domicílios são visitados em um mês, não ocorrem entrevistas pelos dois meses seguintes, e esse ciclo se repete cinco vezes. Em termos trimestrais, os domicílios são visitados 5 trimestres consecutivos, uma vez a cada trimestre, até o domicílio ser retirado definitivamente da amostra. A amostra é então composta por cinco grupos de rotação de forma que, a cada trimestre, um painel contendo 20% dos domicílios da amostra deixa de ser investigado e é substituído por um painel com novos domicílios para sua primeira entrevista (IBGE, 2022e). O padrão de rotação e sobreposição da amostra induz uma autocorrelação do erro amostral de acordo com um modelo AR(1), conforme 10. Isso é devido ao fato de que há 80% de sobreposição da amostra em trimestres consecutivos (lag 1), 60% em trimestres cujo lag é 2, 40% nos casos de lag 3 e 20% quando o intervalo entre os trimestres corresponde ao lag 4. Esse último, refere-se à autocorrelação gerada pela existência de domicílios em sua 5ª visita e que foram entrevistados pela primeira vez no mesmo período do ano anterior.

$$\tilde{e}_t = \widehat{\phi} \tilde{e}_{t-3} + \eta_{\tilde{e},t}, \quad \eta_{\tilde{e},t} \sim N(0, \sigma_{\tilde{e}}^2) \quad (10)$$

A identificação do modelo do erro amostral é possível a partir da estimação das funções de autocorrelação e autocorrelação parcial. Para isso, aplicou-se o método desenvolvido por Pfffermann, Feder e Signorelli (1998) que constrói séries de pseudo-erros a partir das estimativas calculadas por painéis que compõem as amostras de pesquisas repetidas. Após a estimação das autocorrelações dos erros amostrais, os parâmetros do modelo para representar o processo do erro amostral são obtidos pela resolução das equações de Yule-Walker. Para uma descrição completa do método, ver os estudos de Silva e Cruz (2002) e Rosseti e Silva (2017). No caso da PNADC, a amostra é composta de 5 painéis associados aos grupos de rotação (IBGE, 2022e) e o método descrito foi utilizado para obter uma estimativa para o parâmetro $\widehat{\phi}$.

Por fim, todos os distúrbios \mathbf{I}_t , $\eta_{R,t}$, $\eta_{S,t}$ e $\eta_{\tilde{e},t}$, são considerados mutuamente não correlacionados.

O modelo univariado pode ser generalizado para incluir várias séries temporais. No caso específico, há interesse em estimar conjuntamente as séries de desocupados (\widehat{Y}_t) para todas as $J=10$ regiões de Minas Gerais. Um modelo multivariado em que aparentemente as séries são correlacionadas pode ser usado nesse caso. Referenciado na literatura como Seemingly Unrelated Time Series Equations (SUTSE) (Harvey, 1989; Durbin; Koopman, 2012), está incluída nessa modelagem a correlação dos distúrbios dos componentes não observáveis sobre os quais é suposto existir uma relação de longo-prazo. O modelo pode ser expresso generalizando a Equação 11 da seguinte forma:

$$\begin{pmatrix} \widehat{y}_{1,t} \\ \vdots \\ \widehat{y}_{J,t} \end{pmatrix} = \begin{pmatrix} \theta_{1,t} \\ \vdots \\ \theta_{J,t} \end{pmatrix} + \begin{pmatrix} e_{1,t} \\ \vdots \\ e_{J,t} \end{pmatrix}, \quad j = 1, \dots, J \quad (11)$$

onde j representa o estrato geográfico, tal que $J=10$.

O modelo anterior pode ser reescrito no formato de espaço de estados de acordo com Harvey (1989) e Durbin e Koopman (2012). As covariâncias referentes aos termos de distúrbio das inclinações serão estimadas com a seguinte representação:

$$\text{COV}(\eta_{R,y_j,t}, \eta_{R,y_{j'},t}) = \rho_{y_j,y_{j'}}^R \times \sigma_{R,y_j,t} \times \sigma_{R,y_{j'},t}, \quad j \neq j' \quad (12)$$

onde $\rho_{y_j,y_{j'}}^R$ indica a correlação entre os distúrbios dos termos de erros da inclinação da região j e j' .

Após incorporar as estimativas obtidas para \widehat{Y}_t , \widehat{C}_t e os $\widehat{\phi}_t$ sob a abordagem baseada no desenho amostral, aplicou-se uma rotina de otimização da função de verossimilhança do modelo para estimar os hiperparâmetros. O filtro de Kalman é, então, empregado para obtenção das estimativas dos componentes não observáveis, de suas variâncias e das covariâncias entre os termos do vetor de estados que são necessárias para o cálculo da precisão das estimativas expressas pelos coeficientes de variação e intervalos de confiança. Foram produzidas estimativas filtradas para os componentes estruturais das séries (usando as equações de atualização do filtro de Kalman com base nas informações disponíveis até o tempo t). Para mais detalhes sobre o filtro de Kalman ver Harvey (1989).

Testes de diagnósticos foram utilizados para checar as hipóteses de normalidade dos resíduos (teste de Shapiro Wilk), a verificação de evidências de correlação serial (teste de Ljung-Box) e de heterocedasticidade dos resíduos (teste H). Para a comparação dos modelos e escolha da melhor especificação foram realizados testes da razão de verossimilhanças (Durbin; Koopman, 2012).

Dois principais resultados do processo de modelagem são as estimativas dos componentes não observáveis, a saber: as séries da tendência e do sinal. Destaca-se que, sob essa abordagem, o sinal é obtido a partir da soma das séries de tendência e sazonalidade, após extrair os dois componentes de erro: o componente irregular e o erro amostral. Adicionalmente, como na análise de conjuntura econômica a estimação da variação de indicadores (entre trimestres) é também relevante para interpretação dos resultados, decidiu-se calcular estimativas de diferença, especificamente da tendência para responder sobre sua significância estatística no curto prazo. A diferença da tendência em relação ao mês anterior é definida por $\Delta_t = L_t - L_{t-1} = R_{t-1}$ conforme equação 4. A construção dos intervalos de confiança para Δ_t permitem avaliar a significância estatística das variações trimestrais do total de desocupados das regiões estudadas.

Existem dois objetivos de avaliação que orientam a análise dos resultados: o primeiro é verificar se as estimativas baseadas em modelo são mais precisas que as estimativas diretas e o segundo é verificar se as estimativas baseadas em modelo estão próximas das estimativas diretas. O primeiro possui relação com a prática dos órgãos oficiais de estatísticas de publicação das estimativas e seus erros-padrões, sendo o coeficiente de variação uma medida com níveis recomendados para considerar uma estimativa publicável. Nesse caso, computou-se a diferença relativa média entre os erros-padrões sob as abordagens baseadas em modelo e no desenho amostral. Quanto mais alta e positiva essa diferença, maior a redução do erro padrão e maior o ganho de precisão registrado para as estimativas resultantes do modelo.

Para o segundo objetivo de avaliação, estabelece-se, conforme literatura de estimação em pequenos domínios (Tzavidis et al., 2018; Brown et al., 2001), as estimativas diretas como sendo aproximadamente não viciadas. Diante desse pressuposto, diagnósticos são realizados e no presente artigo computou-se o vício relativo calculado pela diferença das estimativas sob as abordagens baseadas em modelo e no desenho amostral. Quanto menor o vício relativo, mais próximas estão as estimativas baseadas em modelo das estimativas diretas.

Utilizou-se o software R (R Core Team, 2020) e funções disponíveis no pacote dlm (Petris, 2010) para estimação dos hiperparâmetros e para o filtro de Kalman.

Análise dos Resultados

Os resultados apresentados foram obtidos com os dados trimestrais do total de pessoas desocupadas com 14 anos ou mais de idade no estado de Minas Gerais e de seus dez estratos geográficos.

cos do primeiro trimestre de 2012 ao segundo trimestre de 2022, compondo as séries temporais de 42 observações. Os testes de diagnóstico com os resíduos padronizados revelaram que a normalidade, ausência de autocorrelação e erros homocedásticos foram comprovados ao nível de 1% de significância, concluindo pela adequação dos modelos univariados e do multivariado. A única exceção foi o modelo univariado para o Triângulo Mineiro, que apresentou evidências de erros autocorrelacionados.

Destaca-se também os valores da correlação entre os distúrbios dos termos de erros das inclinações da série de desocupados das dez regiões de Minas Gerais. Na maioria foram estimadas correlações altas, inclusive acima de 0,95. Esse conjunto de correlações foi avaliado com um teste de razão de verossimilhanças em que o modelo multivariado foi estimado considerando todas as correlações iguais a zero. Nesse caso, a hipótese nula de que todas as correlações são iguais a zero foi rejeitada, indicando que há evidências de correlações estatisticamente significativas e a superioridade do modelo multivariado.

O fato do modelo multivariado apresentar um conjunto de correlações significativas indicou que é possível aprimorar o processo de modelagem ao combinar informações das regiões em uma estimação conjunta. Para verificar se esse resultado implicou em ganhos de precisão, computou-se uma medida de diferença relativa média do erro padrão das estimativas baseadas em modelo com as estimativas diretas. A Tabela 2 apresenta esses resultados evidenciando que, para todas as regiões, ocorreu redução do erro padrão, porém em diferentes magnitudes.

Tabela 2 – Medidas relativas para comparação da precisão e vício entre as estimativas baseadas em modelo e as estimativas diretas por estrato geográfico

Nível geográfico	Diferença relativa média do erro padrão (%)		Vício relativo (%)	
	Modelo univariado	Modelo multivariado	Modelo univariado	Modelo multivariado
1 - Belo Horizonte (BH)	21,6	28,2	0,2	-0,2
2 - Entorno Metropolitano de BH	40,8	30,7	0,3	0,7
3 - Colar Metropolitano de BH	31,3	34,6	-0,1	0,3
4 - Integrada de Brasília em MG	22,7	25,5	-4,3	-3,4
5 - Sul de Minas Gerais	4,0	21,3	-1,1	-1,3
6 - Triângulo Mineiro	30,0	48,6	-1,7	-1,4
7 - Zona da Mata	12,3	21,9	-1,4	-2,4
8 - Norte de Minas Gerais	0,4	7,7	-0,5	-0,9
9 - Vale do Rio Doce	18,6	22,7	-0,7	-1,8
10 - Central de Minas	15,5	26,2	-1,0	-2,4

Fonte: PNADC (IBGE, 2022d). Elaboração própria.

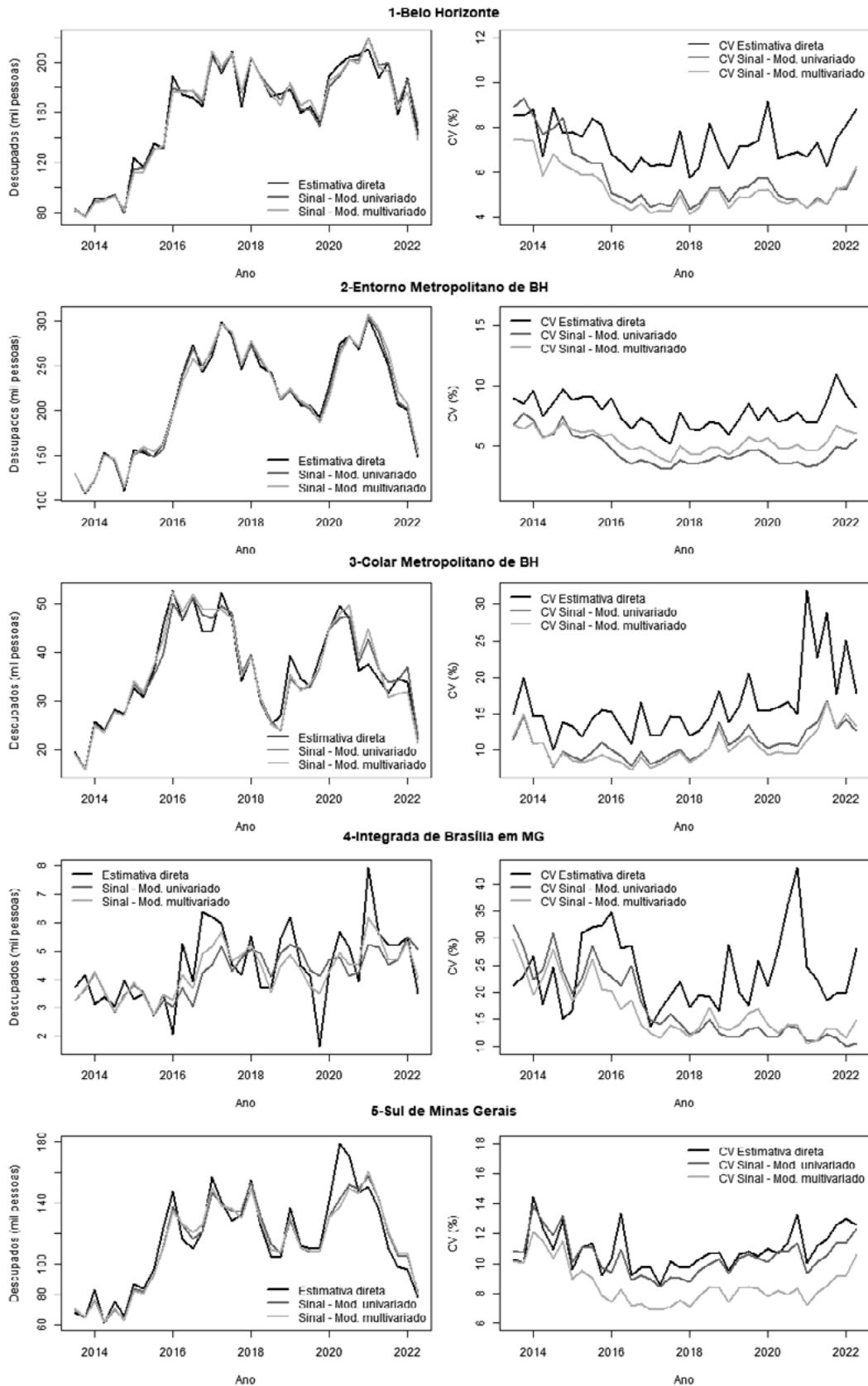
Além disso, não foi unânime a vantagem do modelo multivariado, pois na região do Entorno Metropolitano de BH, o modelo univariado teve maiores ganhos de precisão, representando a segunda maior redução observada, 40,8%. No entanto, a redução dentro do modelo multivariado para o Entorno Metropolitano de BH manteve-se em um patamar alto de 30,7%. De forma geral, ao usar o modelo multivariado, os ganhos de precisão para as estimativas baseadas em modelo, quando comparadas às estimativas diretas, variaram de 7,7 a 48,6%.

A Tabela 2 também apresentou o vício relativo, indicando o quanto a estimativa do componente de sinal dos desocupados se distancia da respectiva estimativa direta. Ambos os modelos apresentaram resultados próximos das estimativas diretas. O modelo univariado mostrou-se mais próximo para o Entorno metropolitano de BH, Colar Metropolitano, Sul de Minas, Mata de Minas, Vale do Rio Doce e Central de Minas, enquanto o modelo multivariado para a Integrada de Brasília em Minas Gerais e o Triângulo Mineiro. Por fim, os modelos univariados e

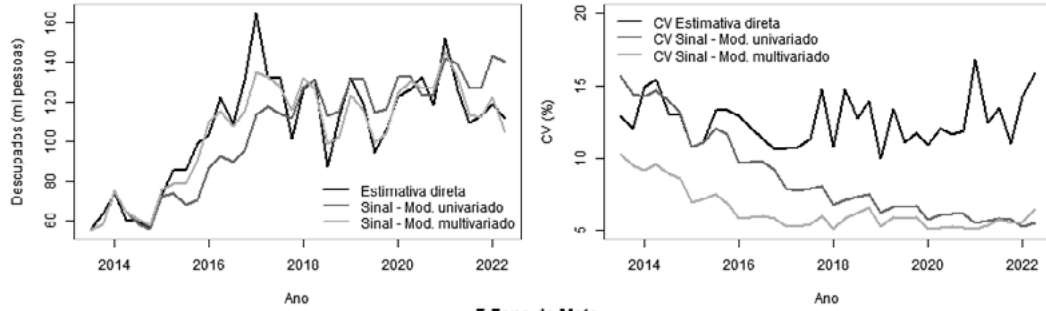
multivariado tiveram a mesma distância na capital, porém em direções opostas. Realizando essa mesma comparação com a estimativa direta divulgada para o estado de Minas Gerais e a soma dos sinais estimados, o vício relativo foi de -0,57% para os modelos univariados e -0,94% para o modelo multivariado, sendo baixos e próximos. Conforme Bell, Datta e Ghosh (2013), diferenças substanciais entre o total das estimativas baseadas em modelo para os domínios e o valor da estimativa direta da área maior indica falha no modelo sugerido. Esse problema não foi observado nos modelos estimados.

Pode-se ainda avaliar a evolução temporal para cada um dos estratos geográficos. A Figura 2 apresenta os gráficos das estimativas diretas de desocupados e do sinal estimado por modelos univariados e multivariado. De forma geral, as estimativas dos sinais são próximas das estimativas diretas. Existem exceções, como o caso da região Integrada de Brasília em Minas Gerais e para o Triângulo Mineiro e região Central o caso do modelo univariado. Em termos da precisão, diferentes situações ocorreram como a superioridade da abordagem baseada em modelos em relação às estimativas diretas em todo o período ou em grande parte do tempo como no Entorno Metropolitano de BH, no Colar Metropolitano e no Vale do Rio Doce. Houve ainda casos em que, no início da série, a abordagem baseada no desenho amostral apresentou menores CVs ou valores similares aos das estimativas baseadas em modelos: Belo Horizonte, Triângulo Mineiro, Central e Integrada de Brasília. Por fim, casos em que os valores do CV das estimativas diretas foram próximos dos encontrados para pelo menos um dos modelos estimados: Sul de Minas, Mata de Minas e Norte de Minas Gerais.

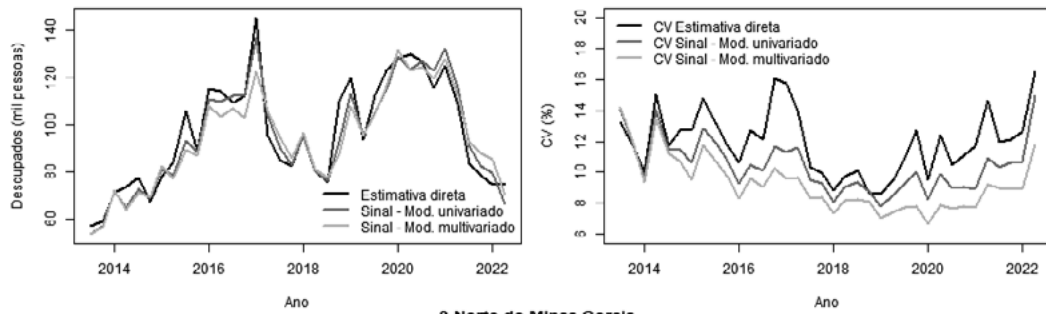
Figura 2 – Estimativas diretas e baseadas em modelo (sinal) do total de desocupados e respectivos coeficientes de variação - 3º trim. 2013 a 2º trim. 2022



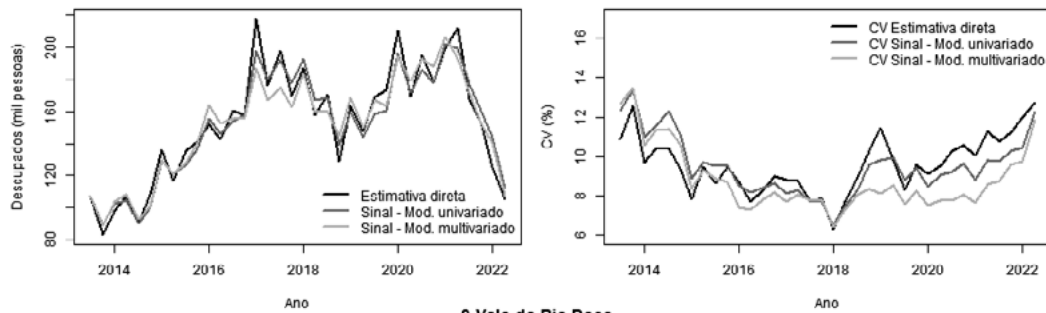
6-Triângulo Mineiro



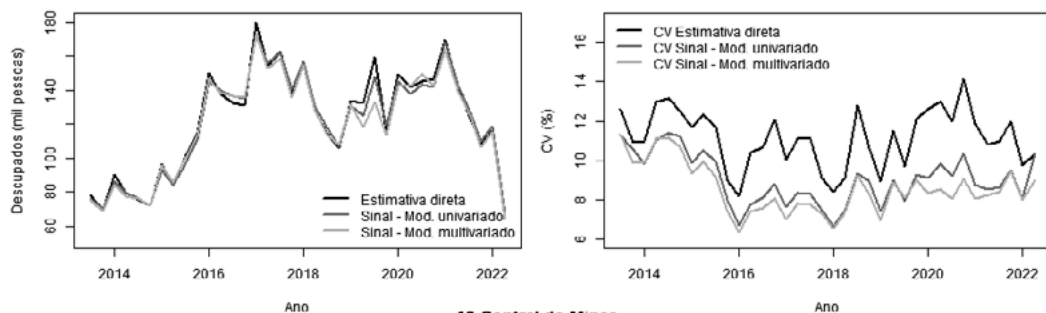
7-Zona da Mata



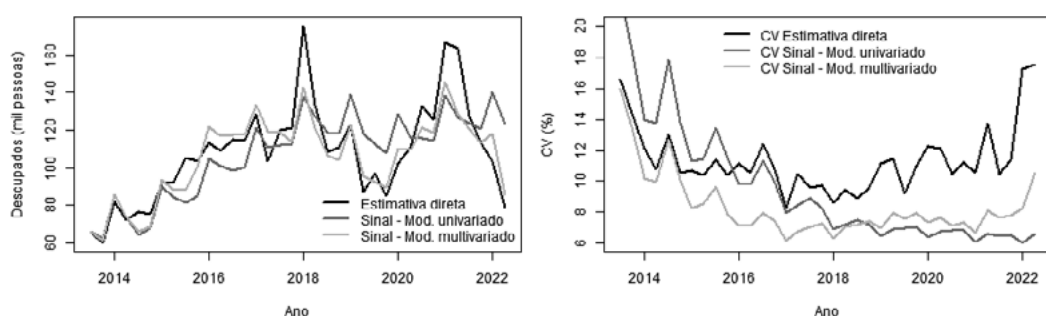
8-Norte de Minas Gerais



9-Vale do Rio Doce



10-Central de Minas



Em termos da precisão estatística das estimativas, existem vantagens na produção das estimativas baseadas em modelo comparativamente às estimativas diretas. Apesar de não aparecerem no início da série para alguns casos, as vantagens cresceram com o tempo.

Ressalta-se que o uso de modelos estruturais permitiu a produção da série de tendência com menor ruído dada a remoção da sazonalidade e dos resíduos (componentes irregular e de erro amostral). Ao analisar o comportamento ao longo do tempo da diferença da série de tendência, e seu respectivo intervalo de confiança de 95%, pode-se indicar em quais trimestres a variação foi estatisticamente diferente de zero. Observou-se a similaridade da dinâmica entre as regiões selecionadas com algumas diferenças nos períodos da crise econômica e política de 2015-2017, e a partir de 2020 com a pandemia de COVID-19 quando há evidências de diferenças significativas trimestre-a-trimestre. Destaca-se o período mais recente da série de reduções significativas no total de desocupados para todos os casos apresentados, indicando um cenário de recuperação econômica. Esse resultado ilustra a pertinência da produção de estatísticas trimestrais para estratos geográficos.

Por fim, a Tabela 3 ilustra a diferenciação das estimativas baseadas no desenho amostral e em modelo quanto ao valor estimado, o coeficiente de variação e os intervalos de confiança de 95% para o 2º trimestre de 2022. A redução do erro padrão, e consequentemente do coeficiente de variação mencionado anteriormente, se traduz em intervalos de confiança menores que, para a comunicação e o uso da estatística, se refletem em maior confiabilidade.

Tabela 3 – Estimativas do total de desocupados, intervalos de confiança a 95% (mil pessoas) e coeficientes de variação (%) - estratos geográficos de Minas Gerais - 2º trim. 2022

Nível geográfico	\hat{Y}_t	$CV(\hat{Y}_t)$	IC 95%	$\hat{\theta}_t$	$CV(\hat{\theta}_t)$	IC 95%
1 - Belo Horizonte (BH)	142,6	8,8	[118,0 ; 167,3]	138,1	6,3	[121,1 ; 155,1]
2 - Entorno Metropolitano de BH	147,6	8,2	[123,9 ; 171,4]	151,9	6,0	[134,0 ; 169,8]
3 - Colar Metropolitano de BH	21,7	17,7	[14,1 ; 29,2]	21,3	13,2	[15,8 ; 26,9]
4 - Integrada de Brasília em MG	3,5	28,2	[1,6 ; 5,5]	4,0	14,9	[2,9 ; 5,2]
5 - Sul de Minas Gerais	77,8	12,6	[58,6 ; 97,0]	79,9	10,6	[63,3 ; 96,5]
6 - Triângulo Mineiro	111,9	15,8	[77,1 ; 146,6]	104,6	6,5	[91,3 ; 117,9]
7 - Zona da Mata	75,2	16,5	[50,9 ; 99,5]	70,4	11,8	[54,0 ; 86,7]
8 - Norte de Minas Gerais	105,2	12,7	[79,0 ; 131,4]	107,0	11,8	[82,2 ; 131,9]
9 - Vale do Rio Doce	63,5	10,3	[50,6 ; 76,3]	63,8	9,0	[52,6 ; 75,1]
10 - Central de Minas	79,0	17,6	[51,8 ; 106,2]	85,5	10,6	[67,8 ; 103,2]

Fonte: PNADC (IBGE, 2022d). Elaboração própria.

Considerações Finais

O presente artigo responde ao questionamento de que é possível beneficiar-se do uso de modelo multivariado de séries temporais para produzir estimativas mais precisas, aqui exemplificadas pelos estratos geográficos de Minas Gerais no período estudado.

Ao comparar o modelo multivariado, que incorpora e estima a correlação entre os distúrbios das inclinações das séries dos estratos, com um modelo sem esses hiperparâmetros, foram encontradas evidências da significância estatística das correlações. Esse resultado a favor da estrutura multivariada também resultou em redução dos coeficientes de variação na ordem de 7,7 a 48,6%, a depender da região. Com exceção de uma única região (Entorno Metropolitano de BH), todas as demais apresentaram menores CVs no modelo multivariado. Ressalta-se ainda que as estimativas baseadas em modelos têm valores próximos aos das estimativas diretas, indicando que não há falhas na especificação dos modelos.

Avaliando a evolução temporal da redução do CV, observou-se que para determinadas regiões os ganhos de precisão acontecem desde o início da série, algumas apresentaram níveis de precisão próximos das estimativas diretas durante todo o tempo, enquanto em outras a precisão melhorou com o passar do tempo. Ressalta-se que os modelos utilizaram 42 observações e essa modelagem se beneficia do tamanho da série temporal disponível, ilustrando o uso promissor dessa proposta na próxima década.

O presente trabalho pode ser considerado o início de uma série de estudos que permitam a divulgação de dados importantes para o país sem condicionar ao aumento ou mudança do plano amostral da pesquisa em vigor. Deve ser continuado com outros procedimentos que podem aprimorar não só a precisão das estimativas, mas também lidar com a questão de reconciliação de séries temporais, o procedimento de benchmarking (Pfeffermann; Tiller, 2006; Bell; Datta; Ghosh, 2013; Bollineni-Balabay; Brakel; Palm, 2016; Boonstra; Brakel, 2022) ou da utilização de modelos que incorporam termos de tendências comuns (Brakel; Krieg, 2016). Modelos multiníveis que incorporam o efeito do tempo (Rao; Yu, 1994), também podem ser testados, bem como o uso de estatística bayesiana (Boonstra; Brakel, 2022).

Também é possível estender a modelagem para outros indicadores dos quais se deriva a taxa de desocupação (U.S. Bureau of Labor Statistics, 2018), ou ainda, desagregá-la para outros domínios de interesse, sexo e idade, tal como realizado pelo Statistics Netherlands. E ainda, pode-se calcular a taxa de desocupação para as pessoas de 18 a 24 anos e da proporção de trabalhadores ocupados em atividades não agrícolas informais, como feito pelo Observatório do Milênio de Belo Horizonte, para além da capital. Por fim, com os resultados do Censo Demográfico de 2022 (IBGE, 2024) será possível comparar os indicadores produzidos com base nos estratos da PNADC com aqueles obtidos pelo Censo, sendo um marco relevante para avaliação da modelagem.

Com isso, destaca-se que o artigo revelou como é possível utilizar a modelagem de séries temporais para obter estimativas no nível de estratos geográficos, para as quais encontrou evidências de aumento da precisão. Estudos futuros são necessários para a construção sistemática de indicadores com essa modelagem, porém, por meio do presente artigo, demonstrou-se sua viabilidade e potencial de produção de estatísticas oficiais relevantes mais granulares

Referências

ALBIERI, S. *Apresentação da precisão de estimativas nas tabelas de pesquisas por amostragem do IBGE*. Rio de Janeiro, jul. 1999.

ALBIERI, S.; BIANCHINI, Z. M. *Principais Aspectos de Amostragem das Pesquisas Domiciliares do IBGE - Revisão 2015*. Texto para discussão, n. 15, IBGE, Rio de Janeiro, 2015. Disponível em: <https://biblioteca.ibge.gov.br/index.php/bibliotecacatalogo?view=detalhes&id=294403>. Acesso em 01 jul. 2024.

BELL, W. R.; DATTA, G. S.; GHOSH, M. *Benchmarking small area estimators*. Biometrika, Oxford University Press, v. 100, n. 1, p. 189–202, 2013.

BINDER, R.; DICK, J. *Modelling and estimation for repeated surveys*. Survey Methodology, v. 15, p. 29–45, 1989.

BOLLINENI-BALABAY, O.; BRAKEL, J. van den; PALM, F. *Multivariate state space approach to variance reduction in series with level and variance breaks due to survey redesigns*. Journal of the Royal Statistical Society Series A (General), Blackwell Publishing, v. 179, n. 2, p. 377–402, 2016.

BOONSTRA, H. J.; BRAKEL, J. van den. *Estimation of level and change for unemployment using structural time series models*. Survey Methodology, v. 45, p. 395–425, 2019.

BOONSTRA, H. J.; BRAKEL, J. van den. *Multilevel time series models for small area estimation at different frequencies and domain levels*. Annals of Applied Statistics, v. 16, n. 4, p. 2314–2338, 2022.

BRAKEL, J. van den. *New data sources and inference methods for official statistics*. 2019. (Discussion Paper).

BRAKEL, J. van den; BUELENS, B.; BOONSTRA, H. J. *Small area estimation to quantify discontinuities in repeated sample surveys*. Journal of the Royal Statistical Society Series A (General), Blackwell Publishing, v. 179, n. 1, p. 229–250, 2016.

BRAKEL, J. van den; KRIEG, S. *Dealing with small sample sizes, rotation group bias and discontinuities in a rotating panel design*. Survey Methodology, v. 41, p. 267–296, 2015.

BRAKEL, J. van den; KRIEG, S. *Small area estimation with state space common factor models for rotating panels*. Journal of the Royal Statistical Society Series A (General), Blackwell Publishing, v. 179, n. 3, p. 763–791, 2016.

BRASIL. *Lei nº 8.184, de 10 de maio de 1991*. 1991.

BROWN, G.; CHAMBERS, R.; HEADY, P.; HEASMAN, D. *Evaluation of small area estimation methods and application to unemployment estimates from the UK LFS*. In: Proceedings of Statistics Canada Symposium 2001. Achieving Data Quality in a Statistical Agency: A Methodological Perspective. Statistics Canada, 2001. (Statistics Canada International Symposium Series: Proceedings), p. 1–10.

DURBIN, J.; KOOPMAN, S. J. *Time Series Analysis by State Space Methods*. 2. ed. Oxford: Oxford University Press, 2012.

FREITAS, M. P. S.; ANTONACI, G. A. *Sistema Integrado de Pesquisas Domiciliares Amostra Mestra 2010 e Amostra da PNAD Contínua*. Rio de Janeiro, 2014. (Textos para Discussão, 50).

FUNDAÇÃO JOÃO PINHEIRO. *Pesquisa de Emprego e Desemprego (PED)*. 2014a. Fundação João Pinheiro.

FUNDAÇÃO JOÃO PINHEIRO. *Pesquisa por Amostra de Domicílios (PAD- MG)*. 2014b. Fundação João Pinheiro.

GONÇALVES, C.; HIDALGO, L.; SILVA, D. B. N.; BRAKEL, J. van den. *Single-month unemployment rate estimates for the Brazilian labour force survey using state-space models*. Journal of the Royal Statistical Society Series A (General), Wiley-Blackwell, v. 185, n. 4, p. 1707–1732, 2022.

HARVEY, A. C. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press, 1989.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Malha Municipal*. 2010.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Conjunto mínimo de indicadores padrão de qualidade a ser aplicado no MERCOSUL*. Rio de Janeiro, 2014. (Textos para Discussão, 52).

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Estimativas de população enviadas ao TCU*. 2020a. Instituto Brasileiro de Geografia e Estatística.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Síntese de indicadores sociais: uma análise das condições de vida da população brasileira*. 2020b. Instituto Brasileiro de Geografia e Estatística.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Sistema de contas regionais: Brasil 2020*. 2021. Instituto Brasileiro de Geografia e Estatística.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Guia para Divulgação de Erros Amostrais nas Pesquisas por Amostragem Probabilística Realizadas pelo IBGE*. 2021. Instituto Brasileiro de Geografia e Estatística. Disponível em: <<https://biblioteca.ibge.gov.br/visualizacao/livros/liv101898.pdf>>. Acesso em: 09 jul. 2024.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Estatísticas Experimentais*. 2022a.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Pesquisa Nacional por Amostra de Domicílios Contínua – Nota técnica 01/2022: Estimativas para estratos de Municípios na PNAD Contínua*. Rio de Janeiro, 2022b.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Painel PNAD Contínua*. 2022c.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *PNAD Contínua - Pesquisa Nacional por Amostra de Domicílios Contínua*. 2022d.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Pesquisa Nacional por Amostra de Domicílios Contínua – Notas técnicas – Versão 1.12*. Rio de Janeiro, 2022e.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Censo Demográfico*. Rio de Janeiro, 2024. Disponível em: <<https://www.ibge.gov.br/estatisticas/sociais/populacao/22827-censo-demografico-2022.html>>. Acesso em: 17 jun. 2024.

KRIEG, S.; BRAKEL, J. van den. *Estimation of the monthly unemployment rate for six domains through structural time series modelling with cointegrated trends*. Computational Statistics & Data Analysis, Elsevier Science, v. 56, n. 10, 2012.

OBSERVATÓRIO DO MILÊNIO DE BELO HORIZONTE. *Relatório de acompanhamento dos Objetivos de Desenvolvimento Sustentável: Belo Horizonte 2022*. Belo Horizonte, 2022.

PETRIS, G. *An R package for dynamic linear models*. Journal of Statistical Software, v. 36, n. 12, p. 1–16, 2010.

PFEFFERMANN, D.; BLEUER, S. R. *Robust joint modelling of labour force series of small areas*. Survey Methodology, v. 19, p. 149–164, 1993.

PFEFFERMANN, D.; BURCK, L. *Robust small area estimation combining time series and cross-sectional data*. Survey Methodology, v. 16, p. 217–237, 1990.

PFEFFERMANN, D.; FEDER, M.; SIGNORELLI, D. *Estimation of autocorrelations of survey errors with application to trend estimation in small areas*. Journal of Business & Economic Statistics, v. 16, n. 3, p. 339–348, 1998.

PFEFFERMANN, D.; TILLER, R. *Small area estimation with state space models subject to benchmark constraints*. Journal of the American Statistical Association, v. 101, p. 1387–1397, 2006.

R CORE TEAM. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2020.

RAO, J.; MOLINA, I. *Small Area Estimation*. 2. ed. Wiley, 2015. (Wiley Series in Survey Methodology).

RAO, J. N. K.; YU, M. *Small-area estimation by combining time-series and cross-sectional data*. The Canadian Journal of Statistics, v. 22, p. 511–528, 1994.

ROSSETI, E. S.; SILVA, D. B. N. *Modelos para séries temporais para pesquisas amostrais repetidas*. In: ALBIERI, S.; DIAS, A. J. R. (Ed.). Rio de Janeiro: IBGE, 2017, (40 anos da unidade de métodos estatísticos do IBGE: alguns passos. Memória institucional, v. 22). p. 171–189.

SCOTT, A. J.; SMITH, T. M. F. *Analysis of repeated surveys using time series methods*. Journal of the American Statistical Association, v. 69, p. 674–678, 1974.

SCOTT, A. J.; SMITH, T. M. F.; JONES, R. G. *The application of time series methods to the analysis of repeated surveys*. International Statistical Review, v. 45, p. 13–28, 1977.

SILVA, D. B. N.; CRUZ, M. M. *Séries temporais de pesquisas amostrais periódicas*. São Paulo: Associação Brasileira de Estatística, 2002.

TZAVIDIS, N.; ZHANG, L.; LUNA, A.; SCHMID, T.; ROJAS-PERILLA, N. *From start to finish: a framework for the production of small area official statistics*. Journal of the Royal Statistical Society Series A (General), Blackwell Publishing, v. 181, p. 927–979, 2018.

UNITED NATIONS STATISTICS DIVISION. *Countries: SDG platform*. 2022.

U.S. BUREAU OF LABOR STATISTICS. *Handbook of methods: Local area unemployment statistics*. 2018.

Uma perspectiva interseccional dos marcadores sociais da diferença de gênero e raça no mercado de trabalho brasileiro (2014-2018)

Rodrigo Santos Martins*
César Augusto Marques da Silva**
Paula Alves de Almeida***

Resumo

No Brasil, desigualdades de raça/cor e sexo/gênero possuem efeitos claros no acesso e permanência no mercado de trabalho. Como marcadores sociais da diferença, tais características geram desigualdades nas diversas formas de inserção no mercado laboral, considerando as chances de desocupação, desalento e subocupação. Além disso, há um efeito da interseccionalidade desses marcadores, ou seja, sobreposição entre os efeitos da discriminação sobre determinados subgrupos de grupos já marginalizados, com aprofundamento das desigualdades crescente quanto maior o número de marcadores da diferença. Na presente pesquisa foram analisados os dados da Pesquisa Nacional por Amostra de Domicílios Contínua (PNAD Contínua), entre 2014 e 2018, período de crise no mercado de trabalho brasileiro. Os resultados confirmam que existe um efeito interativo entre raça/cor e sexo/gênero, indicando que indivíduos com esses dois marcadores sociais tiveram um impacto maior do que a soma dos impactos negativos de cada marcador, quando analisados isoladamente.

Palavras-chave: Interseccionalidade, Gênero, Raça/Cor, Mercado de Trabalho.

* Mestre e doutorando pelo programa de pós-graduação da ENCE.

** Professor do programa de pós-graduação da ENCE.

*** Doutora pelo programa de pós-graduação da ENCE, pós-doutoranda.

Introdução

A interseccionalidade é tema de pesquisa relativamente recente, mas, como fenômeno social, remonta a múltiplas dinâmicas, que vão além do mercado de trabalho, incluindo a sociabilidade em uma perspectiva histórica. A personagem “Tituba”, a bruxa negra de Salém narrada no romance de Maryse Condé¹, é um retrato do fenômeno. Tituba, negra, praticante da religiosidade de seus ancestrais africanos, imigrante forçada e escravizada no norte dos Estados Unidos do século XVII, relata e expõe constantemente como a vida e a reprodução social como um todo é diferente para homens e mulheres, para a população preta e branca, para os locais e os estrangeiros (CONDÉ, 2020). Passados mais de 300 anos e em outros contextos, cor, raça, sexo/gênero, migração, religião e outros marcadores sociais continuam importando.

Nesse caso, a dinâmica do mercado de trabalho em geral, e especificamente a brasileira, mostra recorrentemente como homens e mulheres e brancos e negros possuem diferentes chances e trajetórias ocupacionais. Seria plausível, portanto, considerar que as intensas mudanças no mercado de trabalho do Brasil a partir do final de 2014, concomitante ao início da crise econômica, com taxas crescentes de demissões e aumento do desemprego generalizado (IPEA, 2019), tivessem efeitos diferenciados entre cor/raça e sexo/gênero.

É nesse contexto que o artigo analisa os efeitos dos dois fundamentais marcadores sociais da diferença no mercado de trabalho – cor/raça e sexo/gênero, tendo como base uma perspectiva interseccional e o período de referência 2014 a 2018.

Quanto às desigualdades de sexo/gênero, no Brasil, temos no mercado de trabalho tanto conquistas das mulheres e mudanças em suas posições sociais, quanto permanências e reproduções das desigualdades entre os sexos (ARAÚJO E GUEDES, 2010). Diversas ocupações precarizadas são preenchidas preponderantemente por mulheres, com participação em nichos com pouco reconhecimento social, baixas remunerações, baixa seguridade social e maior informalidade (BRUSCHINI E PUPPIN, 2004). Já o pertencimento racial é uma das questões centrais no contexto da estruturação e reprodução de desigualdades sociais no Brasil. A desigualdade racial brasileira, comumente destilada em conjunto com outras formas mais sutis de discriminação, é obstáculo claro ao progresso e evolução social da população negra (HENRIQUES, 2001). Especificamente no mercado de trabalho, a cor/raça tem influência nas inserções ocupacionais de tais grupos, e conseqüentemente, sobre seu rendimento.

Nesse contexto de reprodução de desigualdades no mercado de trabalho, nos guiamos pela lente da interseccionalidade. Adicionalmente, consideramos os marcadores sociais da diferença que se configuram como um ramo de estudo e uma construção teórico-metodológica que visa elucidar mecanismos que buscam naturalizar situações de profunda desigualdade socioeconômica, como o sexo/gênero, a cor/raça, a classe social, o porte de deficiência, o status migratório, dentre outros.

Todas essas diferenças são, na realidade, constructos sociais influenciados pelas relações de poder, mas que com aparente naturalização, ofuscam as perversidades dos mecanismos de perpetuação de desigualdades. Assim, uma das principais motivações do uso desse arcabouço teórico é analisar como tais marcadores estão interligados e são relacionais (HIRANO *et al.*, 2019; SCHWARCZ *et al.*, 2019).

O artigo analisa os dados da PNADC no tocante ao mercado de trabalho a partir da interseccionalidade, analisando a interação entre os marcadores sociais e seus potenciais efeitos deletérios sobre as condições dos indivíduos, buscando evidenciar as relações discriminatórias em termos de raça e gênero. Além disso, busca-se analisar detidamente o efeito da interseccionalidade desses marcadores, ou seja, como os efeitos da discriminação agem sobre determinados subgrupos de grupos já marginalizados.

A hipótese é que as populações com marcadores específicos da diferença, no caso sexo/gênero e cor/raça, são os que possuem maior probabilidade de pertencerem ao grupo das pessoas desocupadas, desalentadas ou subocupadas. Adicionalmente, a partir das discussões bi-

¹ Tituba foi uma mulher negra escravizada em Salém, ao final do século XVII. A narrativa de Condé, contudo, vai além e recria elementos não historicizados, embora baseados na perspectiva histórica.

bliográficas, supomos que dentro dos próprios grupos que sofrem discriminação há aprofundamento das desigualdades quanto maior o número de marcadores da diferença. Considera-se que existe um efeito interativo entre esses, ou seja, uma pessoa com dois marcadores sofreria um impacto maior do que a soma dos impactos negativos de cada marcador, quando analisados isoladamente.

Para tal aplica-se um modelo de regressão logística para a análise das associações entre os marcadores da diferença e seu impacto sobre a probabilidade de o indivíduo pertencer à determinada condição na ocupação: estar desocupado, desalentado ou subocupado.

Além dessa introdução o artigo traz outras quatro seções. Na primeira, é abordada a discussão teórica sobre as relações do mercado de trabalho com as questões de gênero e étnico-raciais. A segunda descreve a base de dados e a metodologia aplicada. A terceira traz os resultados da análise descritiva e do modelo de regressão logística, que descreve e mensura o impacto dos marcadores sociais de gênero e raça na probabilidade de um indivíduo pertencer à determinada condição na ocupação. Na última seção, estão as conclusões e apontamentos finais, ressaltando as potencialidades da análise.

Mercado de trabalho, raça e gênero no Brasil contemporâneo

Mercado de trabalho e economia

Se na literatura liberal o mercado de trabalho é usualmente tratado como um mercado em que as forças de oferta e de demanda de mão de obra seriam equilibradas (STADUTO *et al.*, 2006), pouco uso teria essa perspectiva para análises da dinâmica da distribuição salarial (SILVA, 1980) ou das ocupações entre diferentes grupos sociais. Nesse sentido, uma corrente que privilegia as características individuais que são apresentadas coletivamente por grupos sociais ganha espaço, mostrando que as discriminações raciais e de gênero, por exemplo, ajudam a explicar tais diferenças salariais (MOORE, 1983; ASHENFELTER e HECKMAN, 1974; FREEMAN, 1976).

Especificamente no Brasil temos um mercado de trabalho segmentado e excludente, influenciado pela concentração de renda, desigualdades regionais e preservação da propriedade latifundiária, com persistência da informalidade e da subutilização da mão de obra (ARANDIA, 1991).

Considerando o panorama recente da economia nacional, destaca-se que a partir da década de 1990, com a introdução de novas tecnologias, abertura econômica e estabilização monetária, ocorreram profundas alterações no mercado de trabalho, com realocação setorial do emprego, diminuição da matriz empregatícia industrial e aumento do peso do setor de serviços. Uma das consequências imediatas de tal movimento foi o aumento da precarização e da informalização nas relações de trabalho (CUNHA *et al.*, 2012), em uma década com estagnação da atividade econômica mais severa que a década anterior. A diferença fundamental entre os períodos é que a década de 1980 conviveu de maneira permanente com o problema da hiperinflação, enquanto na década seguinte a economia teve estabilização monetária após 1994 (PINHEIRO *et al.*, 1999). Assim, entre 1990 e 1992 o país passou por forte recessão, com redução da atividade econômica e desemprego, e a partir de 1993 houve reversão desse processo, com crescimento econômico até 1997, quando foi interrompido pela crise asiática (NERI *et al.*, 2000). Novamente, os reflexos mais evidentes no mercado laboral foram a redução do emprego industrial, o aumento da informalidade, o aumento do número de empregos nos setores de comércio e serviços e aumento do desemprego (NERI *et al.*, 2000).

Já nos anos 2000 houve reversão desse quadro, com saldo positivo na balança comercial ainda em 2002, aperfeiçoamento tributário e tratamento fiscal diferenciado aos microempreendedores (REMY *et al.*, 2011). A tendência foi intensificada durante os governos de Lula (2003-2010) e Dilma (2011-2014), com reestruturação das políticas de emprego, ampliação de programas so-

ciais e fomento do crescimento econômico. Entre 2003 e 2014 houve expansão de 68% dos postos de trabalho formais (MATTEI e HEINEN, 2019).

Contudo, a partir de 2015 houve uma forte retração das atividades econômicas, que surtiu efeitos diretos sobre o mercado de trabalho brasileiro, desfazendo a conjuntura favorável que predominou no período anterior. A combinação de taxas de juros elevadas com redução expressiva da taxa de investimento promoveu um forte processo recessivo, com taxas negativas do PIB em 2015 e 2016 (MATTEI e HEINEN, 2019).

A crise resultou de uma série de fatores que teria reduzido a produtividade da economia e o PIB potencial, gerando choques de demanda que afetaram o desenvolvimento econômico com o esgotamento do modelo econômico no final de 2014, a crise da dívida pública doméstica e a política monetária contracionista para controle inflacionário implementada em 2015 (BARBOSA FILHO, 2017).

O mercado de trabalho nacional vivenciou lenta recuperação dos indicadores econômicos a partir do final de 2017. Esse movimento de lenta recuperação foi interrompido pelos efeitos da COVID-19 (BARBOSA *et al*, 2020).

A partir de 2019, a Síntese de Indicadores Sociais do IBGE continua a destacar a desigualdade no mercado de trabalho brasileiro com base em sexo/gênero, cor e raça. As mulheres, especialmente as negras e pardas, continuam a enfrentar as piores condições de emprego, incluindo maior informalidade e subutilização da força de trabalho. A taxa composta de subutilização, que considera desemprego, subocupação e desalento, é consistentemente mais alta para mulheres e pessoas pretas ou pardas. Em 2022, essa taxa sofreu uma expressiva redução, mas ainda permanece mais elevada nesses grupos em comparação com homens brancos (IBGE, 2023).

Além disso, o estudo mostra que a recuperação econômica pós-pandemia não tem sido equitativa, com os grupos marginalizados, como mulheres negras e pardas, continuando a enfrentar dificuldades no acesso a empregos formais e bem remunerados. A renda média de pretos e pardos, assim como das mulheres em geral, ainda é significativamente inferior à de homens brancos (IBGE, 2023).

Cabe ressaltar que no período de 2014 a 2018, o Brasil enfrentou não apenas os efeitos das crises econômicas globais, mas também uma crise política interna de grandes proporções. O impeachment da presidente Dilma Rousseff em 2016, seguido por uma série de ajustes fiscais e reformas trabalhistas em 2017, teve impactos profundos sobre o mercado de trabalho. A aprovação da reforma trabalhista, em particular, flexibilizou as relações de trabalho (KREIN, 2019) e aprofundou desigualdades já existentes.

Além disso, o ambiente político polarizado que contextualizou as eleições de 2018 gerou um clima de instabilidade no país e na economia. Os efeitos dessas turbulências políticas se refletiram diretamente na estrutura do mercado de trabalho, com uma crescente desigualdade de acesso a oportunidades de emprego formal. Dessa forma, a crise política e as mudanças estruturais no mercado laboral se somaram aos desafios trazidos pelas crises econômicas globais, criando um contexto adverso (KREIN, 2019).

Para o IBGE (2016), a população de 14 anos ou mais de idade, que é a população em idade de trabalhar, está subdividida entre as pessoas que constituem a força de trabalho e as que estão fora da força de trabalho.

Os que estão na força de trabalho são as pessoas ocupadas e as pessoas desocupadas. As pessoas ocupadas são aquelas que possuem trabalho, remunerado ou não, na semana de referência da pesquisa. As pessoas ocupadas podem ser não subocupadas por insuficiência de horas trabalhadas ou subocupadas por insuficiência de horas (ou seja, não trabalham 40 horas semanais e gostariam de trabalhar por um período maior).

As pessoas desocupadas são as que, com 14 anos ou mais, estavam sem trabalho na semana de referência e tomaram alguma providência para conseguir um emprego em um período de 30 dias de referência.

Dentre os que estão fora da força de trabalho temos os que estão na força de trabalho potencial e os que estão fora da força de trabalho potencial.

A força de trabalho potencial é formada por pessoas que não estavam nem ocupadas tampouco desocupadas na semana de referência, mas que naquele momento possuíam um po-

tencial para compor a força de trabalho. Esse grupo é formado pelas pessoas que procuraram trabalho, mas não tinham disponibilidade para trabalhar no momento e pelas pessoas que não procuraram emprego, mas que gostariam de trabalhar e estavam disponíveis para tal.

Dentro desse último grupo (pessoas que não procuraram emprego, mas que gostariam de trabalhar e estavam disponíveis para isso) estão os desalentados, que são os que gostariam de trabalhar mas não procuraram emprego por considerar que eram muito novos, muito idosos, inexperientes ou acreditavam que não encontrariam emprego (IBGE, 2016).

Raça/cor e mercado de trabalho

Em termos gerais, no Brasil coexistem três correntes de pensamento no debate sobre as relações raciais. A primeira corrente afirma que há pouca ou irrelevante discriminação racial e grande fluidez entre as raças. A segunda denota a discriminação racial como presente e generalizada, mas apenas transitória, sendo uma simples etapa no desenvolvimento das relações raciais pós-abolição da escravatura. Por fim, há a corrente que defende que a discriminação é persistente e tem cunho estrutural (TELLES, 2012). De modo geral, temos assim uma oposição entre a primeira e as duas últimas correntes, com um grupo argumentando sobre a democracia racial, encampada por Gilberto Freyre na década de 1930 e, por outros que se posicionam ao lado de Florestan Fernandes que, na década de 1950, concluiu que a suposta democracia racial era, na verdade, um mito (TELLES, 2012).

Na primeira perspectiva, Freyre e outros autores, especificamente Harris (1952) e Wagley (1952), concluíram que no Brasil a discriminação estava baseada em classes socioeconômicas e não em grupos raciais, muito embora admitissem que preconceitos e estereótipos influenciavam a vida social cotidiana (HARRIS, 1952; WAGLEY, 1952).

A pesquisa de Florestan Fernandes, na década de 1950, foi fundamental para contrapor a ideia de democracia racial. Suas conclusões indicam que o racismo era generalizado na sociedade brasileira, relacionado à escravidão e a seus efeitos sociais e psicológicos, mas que, sendo a discriminação racial incompatível com a dinâmica capitalista, tenderia a desaparecer (TELLES, 2012).

No âmbito do mercado de trabalho, tanto a origem social como a discriminação racial dificultam as trajetórias dos negros, com persistências de diferenciais educacionais e de rendimentos (OSÓRIO, 2009). De um ponto de vista mais amplo, há confirmação da existência de desvantagens acumulativas, com as quais os negros passam por desvantagens por toda a vida laboral, desde a inserção escolar, nas profissões a que são apresentados, ao círculo de influência da família (que pode indicar uma possível área de formação), e nas negociações de aumentos salariais ou de financiamentos para empreendedores (ANDREWS, 1992; DA SILVA e LIMA, 1992; LEITE, 2004; PORCARO, 1988; SANTOS, 2005; SOARES, 2000).

De acordo com Telles (2012), a categorização racial não apresenta uniformidade e pode se alterar de acordo com o contexto, indicando que as definições de quem é considerado preto, pardo ou branco no Brasil são relativas e flutuam entre as diversas culturas.

Vale ressaltar que nas pesquisas domiciliares do IBGE, independentemente da palavra utilizada – raça, cor ou uma terceira, o termo é empregado na tentativa de entender a diversidade cultural e física, a ancestralidade (OSÓRIO, 2009):

E é nesse sentido que se entende raça aqui: uma categoria social usada para designar grandes grupos humanos que compartilham uma marca visível, corporal ou não, de sua progênie comum, cuja história é uma mistura de mito e realidade. Grupos aos quais se pode ou não atribuir determinadas características que podem ou não ser usadas para estabelecer hierarquias (OSÓRIO, 2009, p.51).

É nessa acepção que utilizaremos cor e raça nesse trabalho, como delimitante de grupos que compartilham marcas visíveis ou invisíveis de sua origem comum, aos quais se podem atribuir características específicas que podem ou não ser utilizadas para hierarquização entre os diferentes grupos.

O procedimento de categorização por cor e raça utilizado nesse artigo seguiu o método amplamente reconhecido na literatura, que agrupa pretos e pardos, devido à semelhança em

seus indicadores sociais (PAIXÃO, 2008a; 2008b) e por considerar que esses dois grupos sofrem discriminações semelhantes (embora não exatamente iguais) que limitam e impactam suas vidas, seja no mercado de trabalho ou no acesso a oportunidades, na vulnerabilidade social ou até que os colocam como potenciais vítimas de violências sociais, institucionais ou do Estado. Neste trabalho, apresentaremos os dados de pretos e pardos tanto separadamente quanto de forma agregada na categoria negros. De forma semelhante, também foram agregados os dados de amarelos, indígenas e as respostas em branco, sendo que esses não foram considerados nas conclusões deste trabalho porque representavam uma parcela pequena na amostra da PNADC e os resultados para esses grupos não eram estatisticamente significativos.

Gênero e mercado de trabalho

O Brasil passou, principalmente a partir dos anos 1950, por transformações culturais e sociais que impactaram suas relações de gênero com queda na taxa de fecundidade, redução do tamanho familiar, aumento da expectativa de vida, com valores superiores para as mulheres, aumento expressivo nos anos de estudo médios das mulheres e alterações nos padrões culturais sobre os papéis definidos a homens e mulheres. Em conjunto, tais fatores ajudam a compreender a maior inserção das mulheres no mercado de trabalho (BRUSCHINI E PUPPIN, 2004).

Contudo, ainda restam desafios importantes a uma participação social plena e mais justa das mulheres, com persistências das desigualdades no mercado de trabalho, baixa representatividade política e violência contra a mulher (ALVES, 2011).

Retrospectivamente, entre 1950 e 2010 houve sensível redução do hiato de gênero: a PEA masculina cresceu 3,6 vezes, enquanto a PEA feminina cresceu 16 vezes (ALVES, 2016). Assim, por um lado, a participação feminina no mercado de trabalho aumentou no período recente, mesmo naqueles de crise econômica (MOURÃO, 2006), mas por outro, as barreiras para a igualdade de gênero ainda são intensas, sendo expressas na concentração de oportunidades de trabalho em setores de atividades que sofrem com a informalidade, precariedade, descontinuidades de tempo, menor proteção social, menor chance de ascensão funcional e atividades associadas a salários menores (ALVES, 2011).

Em que se pese tais avanços, as relações familiares ainda são determinantes para as diferenças de gênero no acesso ao mercado de trabalho. Às mulheres ainda são atribuídos os afazeres domésticos, cuidados de pessoas e a responsabilidade de dividir o tempo entre o trabalho e a família, sendo penalizadas por terem filhos e cônjuges, enquanto homens seriam premiados por tal. Seria a revolução de gênero incompleta ou estagnada, com a saída das mulheres para o mercado de trabalho, mas com menor desenvolvimento de arranjos mais igualitários entre os sexos, com homens ainda sub-representados nas atividades domésticas e de cuidados (WAJNMAN, 2015).

A discussão expressa por Wajnman (2015) demonstra como o mercado de trabalho reproduz as iniquidades de gênero. Pelos dados avaliados pela pesquisadora se depreendeu que características pessoais como ter cônjuge, ter filho em idade escolar ou ter filho em idade pré-escolar afetam as razões de chance de participação do mercado de trabalho de forma muito heterogênea entre homens e mulheres. Enquanto a razão de chance de homens estarem no mercado de trabalho aumenta quando o mesmo apresenta uma das três características citadas, a probabilidade de uma mulher participar do mercado de trabalho estando ocupada diminui sensivelmente. Além disso, as três características citadas ainda aumentam a probabilidade da mulher de ocupar trabalhos precários, com jornada parcial ou em trabalhos autônomos. A mesma situação não ocorre de maneira significativa entre os homens (WAJNMAN, 2015).

Assim, mulheres são sobrecarregadas com os afazeres domésticos e cuidados de pessoas, o que interfere na sua inserção no mercado de trabalho remunerado. Essas dedicam até o triplo de horas semanais aos afazeres domésticos em relação aos homens, considerando todos os grupos de escolaridade, grupos etários, condição de ocupação (desocupados e ocupados), e em todos os grupos de renda. Além disso, a presença de cônjuge acrescenta cerca de 4 horas a mais de trabalho doméstico para as esposas (OLIVEIRA e MARCONDES, 2016).

Há ainda o fato de que, comumente, as mulheres que deixam de fazer trabalhos domésticos ou reduzem o tempo dedicado a tais atividades o fazem graças a sua substituição por outras mulheres, ou seja, por trabalhadoras domésticas, e não porque os homens passaram a fazer mais tais atividades (JESUS, 2018; OLIVEIRA e MARCONDES, 2016).

De toda forma, a inserção feminina no mercado de trabalho ocorre em dois polos distintos: de um lado mulheres mais escolarizadas se inserem no mercado de trabalho e conseguem repassar as atividades domésticas para outras mulheres (geralmente com baixa escolaridade) que vão arcar com o ônus dessas atividades mal remuneradas e não reconhecidas socialmente. Sendo assim há um aumento de mulheres escolarizadas em cargos de chefia e gerência ao mesmo tempo em que aumenta a concentração de mulheres pouco escolarizadas em ocupações mal remuneradas e pouco protegidas socialmente, como empregadas domésticas e diaristas (CASTRO e STAMM, 2017; POCHMAN, 2010).

Em termos mais específicos, as taxas de atividade feminina no Brasil aumentaram entre 2002 e 2009, foram constantes nos anos seguintes e caíram após 2014, voltando a subir entre 2015 e 2019. Na verdade, o hiato de gênero só não se encontra em uma situação pior na segunda década do século XXI por causa da queda na taxa de participação masculina durante o período.

Já a taxa de desocupação das mulheres entre 2002 e 2014 apresentou queda, embora tenha se mantido constantemente superior à de homens. A partir de 2015 os efeitos da crise econômica elevaram o nível do desemprego, tanto entre mulheres como entre os homens até 2017. Há uma pequena queda em 2018 e 2019, seguida de alta em 2020 e 2021, quando volta a cair tanto para as mulheres como para os homens.

A interseccionalidade no mercado de trabalho

O conceito de interseccionalidade é usado a partir da década de 1970, especificamente com os questionamentos feitos pelo movimento feminista negro sobre a centralidade da experiência da mulher branca dentro das reivindicações de toda a categoria “mulher”. Mas foi a partir do trabalho de Crenshaw (1989, 2012) que se passou a uma formulação mais precisa do que seria a interseccionalidade, envolvendo a interdependência das relações de poder de raça, sexo e classe (HIRATA, 2014), ou seja, na conjugação entre vetores de discriminação que atuam conjuntamente para criar desigualdades, em encontros e sobreposições de discriminações que geram situações de desempoderamento (COSTA, 2013).

Para exemplificar o funcionamento do conceito, Crenshaw recorre a uma analogia com avenidas e cruzamentos de avenidas, onde cada eixo de discriminação é representado por uma avenida diferente, e as avenidas eventualmente se cruzam. Sendo assim, em alguns casos os indivíduos se encontram exatamente nos cruzamentos desses vários eixos e são atingidos conjuntamente por tais fluxos, quando os efeitos são intensificados. Uma mulher, negra e homossexual, por exemplo, seria como um pedestre posicionado no cruzamento de três avenidas de discriminação, sendo atingida simultaneamente pelas três discriminações, de sexo, cor/raça e orientação sexual (CRENSHAW, 2002; COSTA 2013).

Nesse sentido, o enfoque interseccional não apenas reconhece a multiplicidade dos sistemas de opressão, mas também considera que há uma interação na produção e reprodução das desigualdades sociais (BILGE, 2009). Diversos autores o têm adotado para realizar análises sobre o mercado de trabalho, apresentando resultados variados, porém esperados. Em termos gerais, a desigualdade de renda aumenta quando cor/raça e sexo/gênero são avaliados em conjunto (CASTRO, 2017; MAIA *et al*, 2018; FAUSTINO *et al*, 2017), em especial quando se avalia a segregação ocupacional vivenciada pelas mulheres negras (COSTA, 2013; HIRATA, 2014; MACEDO 2015).

Metodologia e base de dados

O trabalho apresenta estatísticas descritivas para desocupação, subocupação e desalento para todo o país entre o período 2014 a 2018, quando houve reversão das tendências do mercado de trabalho brasileiro, com aumento das demissões e do desemprego. Também realizou-se modelagem estatística a fim de identificar e mensurar as possíveis associações entre a presença dos marcadores sociais da diferença (raça/cor e sexo) com a condição dos indivíduos no mercado de trabalho durante o período analisado, permitindo quantificar o efeito das variáveis explicativas (cor/raça e sexo) sobre a respectiva variável resposta (desocupação, subocupação e desalento).

A base de dados utilizada no trabalho é a PNADC – Pesquisa Nacional por Amostra de Domicílios Contínua, pesquisa domiciliar amostral sob responsabilidade do Instituto Brasileiro de Geografia e Estatística – IBGE. Implantada em todo o território nacional em janeiro de 2012, a pesquisa foi planejada para apresentar resultados para o Brasil, Grandes Regiões, Estados Federativos e Regiões Metropolitanas. Dentre as principais aplicações dos resultados da pesquisa está o acompanhamento trimestral do mercado de trabalho e outras informações socioeconômicas anuais, vitais para o planejamento estatal (saneamento, escolaridade, condições habitacionais, entre outras). Sua população alvo compreende todas as pessoas que moram em domicílios particulares permanentes dentro da área geográfica coberta.

O desenho amostral da PNADC é planejado de tal maneira que haja rotação dos domicílios selecionados pela amostra, mantendo uma parcela da amostra sobreposta entre dois períodos de divulgação diferentes. Como a divulgação é trimestral, o critério adotado é o 1-2(5): o domicílio é entrevistado em um mês, é retirado da amostra por dois meses e retorna no mês subsequente (no próximo trimestre). O procedimento é repetido cinco vezes seguidas, quando o domicílio é retirado da amostra selecionada (TEIXEIRA JUNIOR *et al*, 2019)². Como o processo de amostragem na PNADC envolve estratificação, conglomeração, probabilidades desiguais e ajustes de não resposta, o ajuste do modelo considerou o desenho amostral e as características do plano amostral já que, por padrão, os principais programas de análise estatística consideram que as amostras são aleatórias simples com reposição. A não consideração dessas particularidades pode levar a resultados enviesados, gerando estimativas viciadas para a variância dos parâmetros (PESSOA e SILVA, 1998). Nesse caso foi usado o Complex Samples, pacote *addon* do SPSS para lidar com amostras complexas e considerar os respectivos pesos.

Os resultados são apresentados para as categorias pretos e pardos separadamente, assim como de forma agregada na categoria negros. Este procedimento de agregação é consolidado na literatura, dada a proximidade nos seus indicadores sociais e a importância da separação entre brancos e não brancos (PAIXÃO, 2008a; 2008b; TELLES, 2012), conforme já explicado anteriormente.

Foram ajustados modelos de regressão logística e, além das variáveis explicativas de cor/raça e sexo, também foram adicionadas variáveis como a região geográfica do indivíduo e seu grau de instrução. As variáveis adicionais foram introduzidas para isolar os efeitos advindos dos marcadores sociais dos efeitos produzidos por esses dois outros aspectos, também importantes em relação às chances de um indivíduo estar desocupado, subocupado ou em situação de desalento.

A não inclusão da variável idade, como variável de controle no modelo de regressão se justifica pela identificação de multicolinearidade entre idade e outras variáveis explicativas, como o grau de instrução, e os efeitos dos próprios marcadores da diferença (cor/raça e sexo). A idade, por estar fortemente correlacionada com essas variáveis, pode gerar distorções nos coeficientes, comprometendo a significância estatística e aumentando a variância das estimativas, o que resultaria em um modelo menos robusto e de interpretação confusa.

Na construção do modelo, a escolha das variáveis depende de critérios qualitativos e quantitativos. Os critérios qualitativos dizem respeito à decisão de quais fatores serão conside-

² Os conceitos e métodos adotados pela PNADC estão disponíveis em: <https://www.ibge.gov.br/estatisticas/sociais/educacao/17270-pnad-continua.html?edicao=18386&t=conceitos-e-metodos>

rados importantes para explicar o fenômeno estudado e os critérios quantitativos são dados pela comprovação empírica da significância das variáveis no escopo do modelo. Quando o fenômeno em estudo é complexo e influenciado por múltiplos fatores, cabe ao pesquisador a decisão da melhor combinação entres os critérios quantitativos e qualitativos (ALVES, 2011).

A regressão logística é um método de modelagem estatística bastante utilizado em decorrência, principalmente, da facilidade de interpretação dos parâmetros de um modelo logístico e também pela possibilidade de interpretação direta dos coeficientes como medidas de associação (PAULA, 2010). O método de regressão logística é usado nos casos em que a variável resposta é categórica e binária. A variável é binária quando admite duas opções de resposta: sucesso ou fracasso de ocorrência do evento (sim ou não). A soma da probabilidade de ocorrência de sucesso e de fracasso é sempre igual a 1, sendo que a probabilidade de ocorrência do evento varia entre 0 e 1.

O modelo de regressão logística permite a realização de análises referentes não apenas às probabilidades estimadas, mas também das vantagens (ou chances) estimadas de ocorrência de sucesso do evento de interesse, comparando-se a probabilidade de sucesso com a de fracasso.

A vantagem em favor da ocorrência do sucesso de um evento nada mais é do que a razão entre a probabilidade de ocorrência do sucesso e de não ocorrência. O modelo irá relacionar o logaritmo neperiano da vantagem em favor da ocorrência, por exemplo, de o indivíduo estar desalentado, com fatores associados àquela condição.

Sendo assim, se as chances para o acontecimento do evento forem maiores do que 1, a probabilidade de ocorrência é maior do que a de não ocorrência. De maneira análoga, se a vantagem for menor do que 1, então a probabilidade de ocorrência é menor do que a de não ocorrência. Presumivelmente, se a vantagem for igual a 1 então a probabilidade de ocorrência e de não ocorrência são iguais.

Ressalta-se a adequação do modelo logístico quando a variável resposta tem características nominais. Sendo assim o modelo tentará responder o quanto cada marcador social e a interação entre eles (interseccionalidade) influencia na variável resposta. A variável resposta do modelo será dada pelas chances de um indivíduo pertencer a um determinado agrupamento de indivíduos. Os agrupamentos serão: indivíduos ocupados ou desocupados, em desalento ou não desalentados e indivíduos em subocupação ou não subocupados.

Assim, o modelo de regressão logística será do tipo:

$$P(S) = \text{Probabilidade de Sucesso} = p_i$$

$$P(F) = \text{Probabilidade de Fracasso} = 1 - p_i$$

Segundo Paula (2010) no modelo logístico, no qual p_i é a probabilidade de "sucesso", o valor de uma variável explicativa qualquer (X) é definido por:

$$\ln\left(\frac{p_i}{1 - p_i}\right) = \alpha + \beta X$$

Em que α e β são parâmetros desconhecidos:

$$\text{vantagem} = \left(\frac{p_i}{1 - p_i}\right) = e^{\alpha + \beta X} = \exp(\alpha + \beta X)$$

O modelo irá relacionar o log da vantagem em favor da ocorrência do evento com os fatores associados, da seguinte maneira:

$$\text{Ln}\left(\frac{p_i}{1 - p_i}\right) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki}$$

Onde p_i será a probabilidade de o indivíduo estar desalentado/subocupado ou desocupado. X_{ki} são as variáveis de interesse (sexo, raça/cor, região geográfica, grau de instrução, etc.). Já $[\beta_0, \beta_1, \beta_2, \dots, \beta_k]$ é o vetor de parâmetros desconhecidos (DOBSON, 2002; ALVES, 2011).

Por fim, a vantagem de um indivíduo pertencer à determinada condição na ocupação é dada por:

$$\text{vantagem} = \left(\frac{p_i}{1 - p_i} \right) = \exp \left(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} \right)$$

A razão entre a vantagem de, por exemplo, um indivíduo negro estar desalentado em relação a um indivíduo branco, é chamada de razão de chance. Essa será a medida principal aqui avaliada. A razão de chance nos dará o fator pelo qual a chance de um subgrupo é multiplicada quando comparada com o grupo de controle. Se a razão de chance de um negro estar desalentado for de 1,5, considerando os brancos como o grupo de referência, então esse resultado nos indica que os negros têm 1,5 chances de se encontrar em desalento em relação aos brancos, ou seja, a chance é 50% maior.

$$\text{Razão de chance} = \left(\frac{\text{vantagem do grupo de estudo}}{\text{vantagem do grupo de referência}} \right) = \frac{e^{\alpha + \beta X(x=1)}}{e^{\alpha + \beta X(x=0)}} = e^{\beta} = \exp(\beta)$$

A avaliação da qualidade do modelo foi feita pela estatística Wald, foi testado se ao menos uma variável que foi considerada explicativa foi significativa dentro do modelo de regressão. A hipótese nula do teste é que todos os parâmetros β_{ki} são iguais a zero.

$$H_0: \beta_0 = \beta_1 = \beta_2 = \dots = \beta_{ki} = 0$$

$$H_1: \text{pelo menos um } \beta_{ki} \neq 0$$

Onde $k = 0, 1, 2, 3, \dots$, k são as variáveis explicativas do modelo e da equação $i=1,2$.

Resultados

As condições de ocupação foram significativamente heterogêneas segundo as informações de raça/cor. Na tabela 1, verifica-se que em toda a série desde 2014 é clara a desvantagem dos negros em relação aos brancos. Brancos apresentam condições melhores em relação aos negros, e, entre os negros, os pretos se encontram em situação um pouco pior do que os pardos. Entre os sexos o padrão também segue desigualdades. Homens consistentemente apresentam desocupação menor do que as mulheres.

Em conjunto os dados por raça/cor e gênero indicam o efeito deletério da interseccionalidade, que atua diretamente sobre os grupos que acumulam mais de um marcador social da diferença. Em 2018 a desocupação entre os homens brancos foi de 8,3% e a de mulheres negras foi de 16,3%. Para mulheres pretas a desocupação atingiu 17,5%.

Tabela 1. Taxa de desocupação por cor/raça e sexo – Brasil, 2014-2018.

		2014(%)	2015(%)	2016(%)	2017(%)	2018(%)
Branco	Homens	4,9	6,4	8,4	8,7	8,3
	Mulheres	6,6	8,3	10,4	11,6	10,9
	Total	5,6	7,2	9,3	10,0	9,5
Negro	Homens	6,4	8,4	11,7	12,7	12,4
	Mulheres	10,4	12,6	15,7	17,3	16,3
	Total	8,1	10,2	13,4	14,6	14,1
Pardo	Homens	6,4	8,4	11,6	12,5	12,5
	Mulheres	10,2	12,6	15,6	17,1	16,0
	Total	8,0	10,1	13,3	14,5	14,0
Preto	Homens	6,8	8,6	12,4	13,2	12,1
	Mulheres	11,5	12,5	15,9	18,2	17,5
	Total	8,8	10,2	13,9	15,4	14,5
Total	Homens	5,7	7,5	10,2	10,9	10,6
	Mulheres	8,5	10,5	13,1	14,6	13,8
	Total	6,9	8,8	11,5	12,5	12,0

Fonte: IBGE – PNAD Contínua, 2014-2018.

Observa-se na tabela 2, que em 2018 a taxa de subocupados foi maior entre os negros (9,2%) do que entre os brancos (5,4%), e para as mulheres (9,3%) em relação aos homens (6,0%). Em uma escala crescente de subocupação: homens brancos apresentaram os melhores níveis, com 4,2% de subocupados, mulheres brancas apresentaram 6,9%, homens pretos 7,6% e mulheres pretas com 12,4%, índice três vezes maior do que dos homens brancos.

Tabela 2 – Taxa de subocupação por cor/raça e sexo – Brasil, 2015-2018³.

		2015(%)	2016(%)	2017(%)	2018(%)
Branco	Homens	0,6	3,1	4,0	4,2
	Mulheres	1,0	4,9	6,6	6,9
	Total	0,8	3,9	5,2	5,4
Negro	Homens	1,2	5,7	7,2	7,5
	Mulheres	2,0	8,7	10,6	11,5
	Total	1,5	6,9	8,6	9,2
Pardo	Homens	1,3	5,9	7,2	7,5
	Mulheres	2,0	8,6	10,4	11,3
	Total	1,5	7,0	8,6	9,1
Preto	Homens	1,1	4,8	7,4	7,6
	Mulheres	2,2	9,0	11,3	12,4
	Total	1,5	6,6	9,1	9,6
Total	Homens	0,9	4,5	5,8	6,0
	Mulheres	1,5	6,8	8,7	9,3
	Total	1,2	5,5	7,0	7,5

Fonte: IBGE – PNAD Contínua, 2015-2018.

³ Vale ressaltar que em 2015 a pergunta sobre subocupação sofreu alteração na PNADC, motivo pelo qual não são apresentados os dados de 2014 para esse índice.

Os resultados para desalento são apresentados e é caracterizado na tabela 3. Em 2018, entre a população branca, havia 2,7% de desalentados, situação melhor do que para a população negra, com 5,6% de desalentados. Já as mulheres apresentaram taxas de 4,3% e os homens de 3,6%. Considerando cor/raça e sexo simultaneamente, as menores taxas foram as de homens brancos (2,3%), seguidos pelas mulheres brancas (3,1%), homens negros (4,5%) e mulheres negras (6,8%).

Tabela 3 – Taxa de desalento por cor/raça e sexo – Brasil, 2014-2018.

		2014(%)	2015(%)	2016(%)	2017(%)	2018(%)
Branco	Homens	0,7	0,8	1,5	1,8	2,3
	Mulheres	1,1	1,6	2,4	2,9	3,1
	Total	0,9	1,2	1,9	2,3	2,7
Negro	Homens	1,6	1,9	3,2	3,9	4,5
	Mulheres	2,9	3,7	5,5	6,1	6,8
	Total	2,2	2,7	4,2	4,9	5,6
Pardo	Homens	1,7	2,0	3,3	4,1	4,7
	Mulheres	3,1	3,8	5,8	6,4	7,0
	Total	2,3	2,8	4,4	5,1	5,7
Preto	Homens	1,1	1,4	2,5	3,1	3,8
	Mulheres	1,8	3,1	4,1	4,6	6,2
	Total	1,4	2,1	3,2	3,8	4,9
Total	Homens	1,2	1,4	2,4	3,0	3,6
	Mulheres	2,0	2,7	4,0	4,6	5,2
	Total	1,6	2,0	3,2	3,7	4,3

Fonte: IBGE – PNAD Contínua, 2014-2018.

A aferição dos efeitos da interseccionalidade no presente artigo foi feita a partir de 3 modelos de regressão logística, sendo um modelo para cada aspecto ou taxa/medida do mercado de trabalho: desocupação, desalento e subocupação. Em todos os modelos a categoria de referência foi homem branco, residente na Região Sul do país e Com Superior completo. Além disso, são utilizadas duas variáveis de interação: a interação entre ser mulher e pertencer ao grupo preto e a interação entre ser mulher e parda. Assim pudemos evidenciar se existem efeitos cumulativos de ser portador de mais de um marcador da diferença simultaneamente.

Os grupos da variável cor/raça identificados como amarelos, indígenas e não identificados não tiveram seus resultados analisados separadamente e foram agregados na categoria Outros. Apesar de formarem subgrupos importantes estatística e socialmente, esses grupos demandam a análise de questões específicas e não debatidas no âmbito do artigo.

Além de Cor/raça (Pretos, Pardos, Outros ou Brancos) e Sexo (Homem ou Mulher), as variáveis de grau de instrução (Sem instrução e menos de 1 ano de estudo, Fundamental incompleto, Fundamental completo, Médio incompleto, Médio completo, Superior incompleto, Superior completo) e região geográfica (Norte, Nordeste, Sul, Sudeste e Centro Oeste) foram utilizadas como variáveis de controle do modelo, para isolar seus efeitos sobre os resultados. Também foram incluídas as Interações (Mulher-Preta, Mulher-Parda, Mulher-Outras, Mulher-Branca, Homem-Preto, Homem-Pardo, Homem-Outros, Homem-Branco).

É importante destacar que o período de 2014 a 2018 foi utilizado exclusivamente para a análise descritiva. A escolha específica do ano de 2018 para a aplicação da regressão logística foi motivada pelo fato de ser o último dado disponível no momento da realização do estudo, além

de simplificar a modelagem estatística. Essa decisão buscou garantir a consistência e precisão dos resultados, facilitando a interpretação dos achados dentro do escopo temporal mais recente.

Desocupação

Os resultados do modelo logístico evidenciaram o efeito deletério tanto dos marcadores sociais da diferença como das interações entre esses mesmos marcadores para a desocupação.

A população do sexo feminino apresentou a chance de pertencer ao grupo de pessoas desocupadas multiplicada por 1,44 quando comparada aos homens, ou seja, a chance das mulheres estarem desocupadas foi 44% maior do que a dos homens.

A população preta teve razão de chances de 1,30, enquanto a parda de 1,35, quando os dois grupos são comparados com os brancos. Portanto a chance de os pretos e pardos estarem desocupados é de respectivamente 30% e 35% superior quando comparados com os brancos.

Tabela 4 – Coeficientes do modelo logístico de desocupação, subocupação e desalento – Brasil, 2018.

	DESOCUPAÇÃO			SUBOCUPAÇÃO			DESALENTO		
	B	Teste de hipótese	Exp(B)	B	Teste de hipótese	Exp(B)	B	Teste de hipótese	Exp(B)
		Sig.			Sig.			Sig.	
(Interceptação)	-3,48	0,000	0,031	-3,861	0,000	0,021	-6,16	0,000	0,002
Mulher	0,37	0,000	1,44	0,615	0,000	1,849	0,55	0,000	1,74
Homem	,000 ^a		1,00	,000 ^a		1,000	,000 ^a		1,00
Pretos	0,26	0,000	1,30	0,314	0,000	1,368	-0,08	0,280	0,92
Pardos	0,30	0,000	1,35	0,264	0,000	1,302	0,05	0,292	1,05
Outros e ND	0,18	0,234	1,19	0,265	0,133	1,303	-0,12	0,589	0,89
Branco	,000 ^a		1,00	,000 ^a		1,000	,000 ^a		1,00
Norte	0,36	0,000	1,43	0,555	0,000	1,742	1,32	0,000	3,75
Nordeste	0,56	0,000	1,75	1,000	0,000	2,719	1,93	0,000	6,91
Sudeste	0,48	0,000	1,62	0,236	0,000	1,266	0,44	0,000	1,55
Centro-Oeste	0,09	0,050	1,09	-0,058	0,371	0,944	0,34	0,000	1,41
Sul	,000 ^a		1,00	,000 ^a		1,000	,000 ^a		1,00
Sem instrução e menos de 1 ano de estudo	0,37	0,000	1,45	0,795	0,000	2,214	2,31	0,000	10,05
Fundamental incompleto ou equivalente	0,69	0,000	2,00	0,779	0,000	2,180	2,18	0,000	8,82
Fundamental completo ou equivalente	0,89	0,000	2,44	0,623	0,000	1,864	2,02	0,000	7,57
Médio incompleto ou equivalente	1,36	0,000	3,90	0,742	0,000	2,100	2,09	0,000	8,10
Médio completo ou equivalente	0,85	0,000	2,35	0,228	0,000	1,256	1,25	0,000	3,48
Superior incompleto ou equivalente	0,88	0,000	2,41	0,332	0,000	1,394	0,67	0,000	1,95
Superior completo	,000 ^a		1,00	,000 ^a		1,000	,000 ^a		1,00
Mulher*Preto	0,10	0,125	1,11	0,058	0,498	1,060	0,18	0,054	1,20
Mulher*Pardo	-0,04	0,330	0,96	-0,034	0,540	0,967	0,09	0,115	1,10
Mulher*Outros	0,11	0,591	1,11	0,123	0,569	1,131	0,22	0,468	1,24
Mulher*Branco	,000 ^a		1,00	,000 ^a		1,000	,000 ^a		1,00
Homem*Preto	,000 ^a		1,00	,000 ^a		1,000	,000 ^a		1,00
Homem*Pardo	,000 ^a		1,00	,000 ^a		1,000	,000 ^a		1,00
Homem*Outros	,000 ^a		1,00	,000 ^a		1,000	,000 ^a		1,00
Homem*Branco	,000 ^a		1,00	,000 ^a		1,000	,000 ^a		1,00

Subpopulação: Ano de referência = 2018

Variáveis Dependentes: DESOCUPAÇÃO; SUBOCUPAÇÃO e DESALENTO (categoria de referência = ,00) Modelo: (Interceptação), SEXO, COR, REGIAO, INSTRUCAO, SEXO * COR

^a Definido para zero porque este parâmetro é redundante.

Fonte: IBGE – PNAD Contínua, 2018.

As mulheres pretas apresentaram uma razão de chance de 2,08, ou seja, a mulher preta possuiu 108% a mais de chances de estar desocupada em relação ao homem branco, considerando os controles de influências advindas do nível de instrução e da região de residência. De maneira similar, a mulher parda apresenta razão de chances de 1,88, ou seja 88% de chance maior de estar desocupada do que os homens brancos.

Realizamos também a análise das mulheres pretas e das mulheres pardas tanto em relação às mulheres brancas como também em relação aos homens pretos ou pardos. Em relação às mulheres brancas, as mulheres pretas apresentaram razão de chances de 1,44, que é a exponencial da soma do beta de ser negro com o beta da interação entre ser mulher e ser negra, ou seja, a mulher preta possuiu 44% a mais de chance de estar desocupada que a mulher branca. De maneira similar, a mulher preta teve razão de chances de 1,60 em relação aos homens pretos, que é a exponencial da soma do beta de ser mulher com o beta da interação de mulher e negra, ou seja, 60% a mais de chance de estar desocupada.

A razão de chances das mulheres pardas quando comparada com as mulheres brancas foi de 1,30, ou seja, 30% a mais de chance de estar desocupada e, quando comparada aos homens pardos, de 1,39, ou seja, 39% a mais de chance de estar desocupada do que o homem pardo.

Subocupação

Os resultados do modelo para a subocupação são similares aos encontrados para a desocupação.

As mulheres apresentaram um aumento de 85% nas chances de se encontrar em situação de subocupação, quando comparada com os homens. Para pretos e pardos as razões de chance foram estimadas em 37% e 30%, respectivamente, em relação à população branca.

O efeito da interseccionalidade, observado para mulheres pretas, foi significativamente superior. As razões de chance desse grupo foram de 2,68, indicado que essas possuíam 168% mais chances de estarem subocupadas do que os homens brancos.

Desalento

Novamente, o grupo das mulheres apresentou razões de chance de pertencimento ao grupo superior a de homens, com 74% de chances a mais de estarem desalentadas. Para os pretos a situação foi inversa: esses tiveram uma chance 8% menor do que os brancos de estar em desalento. Os pardos apresentaram razão de chance de 5% a mais em relação aos brancos. Contudo, os resultados devem ser relativizados, uma vez que os intervalos de confiança do modelo não permitem afirmar que as razões são efetivamente diferentes entre os grupos.

Quando comparadas aos homens brancos, as mulheres pretas apresentaram razão de chances de 1,91 vezes, as mulheres pardas 2,01, e, as mulheres brancas, de 1,74, ou seja, esses grupos possuem mais 91%, 101% e 74% de chances de estarem desalentadas do que os homens brancos.

As mulheres pretas, quando comparadas com os homens pretos, possuíam razão de chance de 2,08, ou seja, 108% a mais de chance de estar desalentada. A mulher parda quando comparada com o homem pardo apresentou razão de chance de 1,91, ou seja, 91% de chances a mais de estar desalentada.

Novamente fica evidente o efeito deletério dos marcadores sociais da diferença. No caso específico do desalento o marcador de ser mulher tem muito mais influência do que o marcador de cor/raça. Nesse caso, o trabalho doméstico não remunerado e as responsabilidades de cuidado com os filhos e com a família pode incidir significativamente para o desalento feminino. Por outro lado, homens pretos e pardos não parecem ser atingidos pelo desalento de maneira substancialmente diferente dos homens brancos, possivelmente indicando que esses grupos estão dispostos a aceitar vagas de menor qualidade para evitar ficar fora do mercado de trabalho.

Os resultados obtidos nos modelos estatísticos utilizados foram, na maior parte dos casos, os esperados. A tabela 5 sintetiza os resultados.

Tabela 5 – Tabela resumo das razões de chances dos modelos logísticos de desocupação, desalento e subocupação – Brasil, 2018

Subgrupo	Desocupação	Desalento	Subocupação
Homem	1,00	1,00	1,00
Mulher	1,44	1,45	1,85
Branco	1,00	1,00	1,00
Pretos	1,30	-	1,37
Pardos	1,35	-	1,30
Homem branco	1,00	1,00	1,00
Homem preto	1,30	-	1,37
Homem pardo	1,35	-	1,30
Homem preto / Homem branco	1,30	-	1,37
Homem pardo / Homem branco	1,35	-	1,30
Homem preto / Homem pardo	0,96	-	1,05
Mulher branca	1,44	1,74	1,85
Mulher preta	2,08	1,91	2,68
Mulher parda	1,88	2,00	2,33
Mulher preta/ Mulher branca	1,44	1,10	1,45
Mulher parda/ Mulher branca	1,30	1,15	1,26
Mulher preta/ Mulher parda	1,11	0,95	1,15

Fonte: IBGE – PNAD Contínua, 2018.

A maior parte dos indicadores aponta para uma hierarquia com os homens brancos com os melhores resultados, seguidos pelos homens negros e mulheres brancas, seguidos por um distante quarto lugar das mulheres negras. Essas sempre apresentam os piores resultados, e a intensidade dos efeitos é dramática, chegando ao absurdo de terem próximo do triplo de chances de estarem em condições de fragilidade empregatícia em relação aos homens brancos. De todos os resultados, a condição da mulher negra é o mais claro caso do efeito do acúmulo de marcadores sociais da diferença, independentemente do nível educacional e região geográfica.

Considerações finais

O quadro desenhado no artigo indica as configurações pregressas do mercado de trabalho sob as perspectivas de gênero e raça/cor. Apesar de um período recente positivo para a redução das desigualdades como um todo - muito influenciado pela implementação de políticas públicas de distribuição de renda e de facilitação de acesso à educação -, e por consequência nas desigualdades de gênero e raça/cor, não vemos tendências à priori de que tais desigualdades tenham se arrefecido. Muito pelo contrário, o período econômico conturbado pós-2014, incluindo também a dinâmica econômica durante a pandemia de Covid-19, pode ter contribuído para o recrudescimento dessas desigualdades, com notória precarização das relações de trabalho e do peso preponderante da informalidade (BACEN, 2019).

Evidenciando os efeitos das relações discriminatórias no acesso ao mercado de trabalho, aqui nos concentramos nas questões de raça, gênero e nas suas interseccionalidades. Para medir esses efeitos utilizamos três indicadores principais: desocupação, subocupação e desalento.

A partir dos dados da PNADC observamos que a desocupação, a subocupação e o desalento foram maiores nas populações discriminadas durante o período estudado, ficando também claro que a sobreposição de marcadores sociais da diferença impactou no resultado do

subgrupo populacional. Em todos os indicadores estudados a mulher negra (pretas e pardas) teve os piores resultados.

Os modelos aplicados aos dados confirmaram que o acúmulo de marcadores sociais da diferença produziu desvantagens acumulativas, principalmente para a desocupação e subocupação. As mulheres pretas estiveram nas situações mais precárias, com aproximadamente o dobro de chance de estarem desocupadas, quase o triplo de estarem subocupadas e, por fim, quase o dobro de estarem desalentadas em relação aos homens brancos.

Adicionalmente, a situação da mulher preta é dramática também em relação às mulheres brancas e aos homens pretos. As mulheres pretas possuem 44% mais chances de estarem desocupadas e 45% a mais de chance de estarem subocupadas do que as mulheres brancas, e quando comparadas aos homens pretos, possuem 60% de chance a mais de estarem desocupadas e 96% a mais de chance de estarem em situação de subocupação. Por fim, possuem pouco mais do dobro (108%) de chances do que os homens pretos de estarem em condição de desalento. O mesmo ocorre em relação à mulher parda.

Os resultados demonstram que os efeitos das discriminações e os hiatos entre os diferentes grupos de raça e cor ainda são relevantes, seja pelo acúmulo de pequenas discriminações por toda a vida do indivíduo, seja no acesso à educação ou em outros fatores, que acaba por se evidenciar nos indicadores aqui observados.

Apenas para o desalento não foi possível detectar o efeito deletério entre os homens pretos e pardos quando comparados aos homens brancos. Embora o fenômeno possa ter múltiplas explicações, uma possibilidade seria de que esses grupos ocupam vagas mais precárias de trabalho dada a impossibilidade da sobrevivência em situações de desemprego, e, por consequência, não aparecem nas estatísticas sobre desalento.

De maneira geral a hipótese de estudo foi confirmada pelos resultados. As populações com marcadores sociais da diferença estudados, cor/raça e gênero, possuem de fato mais chances de estarem desocupadas, desalentadas e subocupadas. Ficou claro também o efeito das interseccionalidades: as mulheres negras que acumulam os dois marcadores da diferença apresentaram os piores resultados em todos os indicadores estudados, sem uma única exceção. Chama a atenção a significativa disparidade entre os indicadores das mulheres negras em relação a outros subgrupos. É comum observar que as mulheres negras enfrentam chances muito maiores de enfrentar situações de subemprego.

Os resultados encontrados corroboram as pesquisas da área, confirmando que raça e gênero são questões centrais na estruturação e reprodução das desigualdades sociais no Brasil. A contribuição deste artigo é apresentar resultados de mensuração da interseccionalidade, que não se apresenta apenas como a simples soma dos efeitos de raça e gênero, mas produz um acréscimo às discriminações já sofridas pelos indivíduos.

É importante reconhecer que os resultados obtidos neste estudo refletem um período de crise política e econômica no Brasil, o que pode ter influenciado os indicadores analisados. Diante disso, para pesquisas futuras, é fundamental explorar se os efeitos da interseccionalidade de gênero e raça sobre o mercado de trabalho permanecem consistentes em contextos mais estáveis ou se são intensificados em tempos de crise. Compreender essa dinâmica permitirá uma análise mais profunda e contextualizada das desigualdades presentes no mercado de trabalho brasileiro.

Por fim, cabe ressaltar que novas medidas de proteção social e de distribuição de renda serão necessárias para que o país chegue à equalização de oportunidades entre os diferentes grupos de raça/cor. Políticas sociais e/ou afirmativas focadas nas mulheres negras seriam muito eficazes, já que seriam focadas sobre o substrato populacional mais atingido com os efeitos cumulativos.

Referências

- ANDREWS, G. R. *Desigualdade racial no Brasil e nos Estados Unidos: uma comparação estatística*. Estudos Afro-Asiáticos, n.22, p.47-83. 1992.
- ARANDIA, A. K. *O mercado de trabalho frente à crise dos anos 80 e os planos de estabilização*. Indicadores Econômicos FEE, Porto Alegre, v. 18, n.4, p.148-164, 1991.
- ALVES, P. *O cinema brasileiro de 1961 a 2010 pela perspectiva de gênero*. Instituto Brasileiro de Geografia e Estatística. Dissertação de Mestrado. Rio de Janeiro. 2011.
- ALVES, J. E. D. *Crise no mercado de trabalho, bônus demográfico e desempoderamento feminino*. In: ITABORAI, N. R.; RICOLDI, A. M. (Org.). *Até onde caminhou a revolução de gênero no Brasil?* 1ª ed. Belo Horizonte: ABEP, 2016, v. 1, p. 21-44.
- ARAÚJO, C.; GUEDES, M. *Igualdade de oportunidade: a distância entre proposições e ações*. Brasil. Presidência da República. Secretaria Especial de Políticas para as Mulheres. Revista do Observatório Brasil da Igualdade de Gênero. 1ª impressão. Brasília: Secretaria Especial de Políticas para as Mulheres, 2010.
- ARROW, K. *The Theory of Discrimination*. In: *Discrimination in Labor Markets*, 1973, Princeton, NJ: Princeton University Press.
- ASHENFELTER, O.; HECKMAN, J. *Measuring the effect of an antidiscrimination program*. Center for economic analysis of human behavior and social institutions, National Bureau of Economic Research, 1974.
- BANCO CENTRAL DO BRASIL - BACEN, *Estudos especiais n 76/2019*.
- BARBOSA FILHO, F. H. *A crise econômica de 2014/2017*. Estudos Avançados, São Paulo, v. 31, n. 89, p. 51-60, 2017.
- BARBOSA, A.L.N.H.; COSTA, J. S.; HECKSHER, M. *Mercado de trabalho e pandemia da COVID-19: ampliação de desigualdades já existentes?*. Mercado de trabalho conjuntura e análise, v. 26, p. 1-9, 2020.
- BELTRÃO, K. I.; ALVES, J. E. D. *A reversão do hiato de gênero na educação brasileira no século XX*. Cadernos de Pesquisa (Fundação Carlos Chagas), v. 39, p. 125-156, 2009.
- BRUSCHINI, C.; PUPPIN, A. B. *Trabalho de mulheres executivas no Brasil no final do século XX*. Cadernos de Pesquisa, v.34, n.121, p.105-138, jan./abr. 2004.
- CASTRO, B.; STAMM, C. *Diferenças Salariais de Gênero e Raça no Mercado de Trabalho Brasileiro: uma Análise Estatística e Econométrica*, 2017.
- CRENSHAW, K. *Documento para o encontro de especialistas em aspectos da discriminação racial relativos ao gênero*. Revista Estudos Feminista, Florianópolis, v. 10, n. 1, p. 171-188, 2002.
- CONDÉ, M. *Eu, Tituba: Bruxa Negra de Salem*. Rosa Dos Tempos, 2020.
- COSTA, J. *Colonialidade e interseccionalidade: o trabalho doméstico no Brasil e seus desafios para o século XXI*. In: SILVA, Tatiana Dias; GOES, Fernanda Lira (Org.). *Igualdade racial no Brasil: reflexões no Ano Internacional dos Afrodescendentes*. Brasília: Ipea, 2013.
- DASILVA, D. F. e LIMA, M. *Raça, gênero e mercado de trabalho*. Estudos Afro-Asiáticos, v.23, p.97-111. 1992.
- FAUSTINO, I.; ARAÚJO, E.; MAIA, K. *Mercado de trabalho e discriminação: uma análise das diferenças salariais e discriminação por gênero no Brasil e macrorregiões (2004-2014)*. ANPEC, 2017.
- FREYRE, G. *Casa-Grande & Senzala*, 50a edição. Global Editora. 2005.

HASENBALG, C. 1979. *Discriminação e Desigualdades Raciais no Brasil*. Rio de Janeiro: Graal. (tradução de Racerelation in post-Abolition Brazil: the Smooth preservation of racial inequalities. Departamento de Sociologia, University of California, Berkeley, 1978.)

HARRIS, M. 1952. *Race Relation in Minas Velhas, a Community in the Mountain Region of Central Brazil*. In: Wagley, Charles (ed.) *Race and Class in Rural Brazil*: 47-81. Paris: Unesco.

HIRANO, L. F. K.; ACUÑA, M.; MACHADO B. F. (Org.). *Marcadores Sociais das Diferenças: fluxos trânsitos e intersecções*. Editora Imprensa Universitária, 2019.

HIRATA, H.. *Gênero, classe e raça. Interseccionalidade e consubstancialidade das relações sociais*. *Tempo Social*, vol. 26, n.1, São Paulo, p. 61-74, 2014.

KREIN, J.D.; COLOMBI, A.P.F.; *A reforma trabalhista em foco: desconstrução da proteção social em tempos de neoliberalismo autoritário*. *Educ. Soc.*, Campinas, v. 40, e0223441, 2019.

LIMA, M. *Trajetória educacional e realização socioeconômica das mulheres negras*. In: HASENBALG, C. A., SILVA, N. D. V. e LIMA, M. (Eds.). *Cor e estratificação social*. Rio de Janeiro: Contra Capa Livraria, 1999. p.148-158.

MACEDO, R. *Trabalho doméstico, consumo e interseccionalidade: possibilidades de agência na trajetória de uma (ex)empregada doméstica*, *Revista Mediações*, vol. 20, Londrina, pp.184-207, 2015.

MAIA, K.; CASTELAN, D.; BACCHI, M.; FONSECA, M. *Diferenças salariais e discriminação por gênero e cor no mercado de trabalho da região Sul do Brasil*. *Revista Estudo & Debate*, Lajeado, v. 25, n. 1, 2018.

MEDEIROS, C. *Raízes estruturais da crise financeira asiática e o enquadramento da Coreia*. *Economia e Sociedade*, Campinas, (11): 151-72, 1998.

MINCER, J. *The Distribution of Labor Incomes: A Survey With Special Reference to the Human Capital Approach*. *Journal of Economic Literature*, vol. 8, no. 1, 1970, pp. 1-26.

MOORE, R. *Employer Discrimination: Evidence From Self-Employed Workers*. *The Review of Economics and Statistics*, 65(3), 496-501. 1983

MOURÃO, T. F. *Mulheres no Topo de Carreira Flexibilidade e Persistência*, 2006.

NERI, M.; CAMARGO, J.; REIS, M.; CORTEZ, M. *Mercado de Trabalho nos Anos 90: Fatos estilizados e Interpretações*. Rio de Janeiro: IPEA, TD n. 743, 2000.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE. *Estatísticas de Gênero. Indicadores sociais das mulheres no Brasil*. Estudos e Pesquisas. Informação Demográfica e Socioeconômica n° 38, 2018

_____. *Pesquisa Nacional por Amostra de Domicílios Contínua. Medidas de Subutilização da Força de Trabalho no Brasil*. Divulgação Especial, 2016.

_____. *Pesquisa Nacional por Amostra de Domicílios Contínua. Outras formas de trabalho*. Divulgação Especial, 2018.

_____. *Pesquisa Nacional por Amostra de Domicílios Contínua. Notas técnicas Versão 1.6*. Divulgação Especial, 2019.

_____. *Pesquisa das Características Étnico-Raciais da População: um estudo das categorias declassificação de cor ou raça: 2008*, Rio de Janeiro, 2011

_____. *O que é desemprego*. Disponível em: <<https://www.ibge.gov.br/explica/desemprego.php>>. Acessado em 01 de ago. 2020.

_____. *Síntese de Indicadores Sociais uma Análise das Condições de Vida da População Brasileira: 2023*, Rio de Janeiro, 2023

- JESUS, J. C. *Trabalho doméstico não remunerado no Brasil: uma análise de produção, consumo e transferência*. Tese de doutorado em Demografia, Belo Horizonte, UFMG, 2018.
- OLIVEIRA, M. C. F. A.; MARCONDES, G. S. *Intersecções família e gênero: a questão das assincronias*. In: ITABORAÍ, N. *As transformações da vida familiar e na inserção social feminina: avanços, desigualdades e desafios*. São Paulo, SP: ABEP, 2016.
- OSORIO, G. *Mobilidade social sob a perspectiva da distribuição de renda*. Departamento de Sociologia, Universidade de Brasília, Brasília, 2003a.
- _____. *O sistema classificatório de cor ou raça do IBGE*. Brasília: IPEA, 2003b.
- _____. *A mobilidade social dos negros brasileiros*. Brasília: IPEA, 2004.
- _____. *A desigualdade racial de renda no Brasil: 1976-2006*, 2009.
- PAIXÃO, M. *O ABC das desigualdades raciais: um panorama do analfabetismo da população negra através de uma leitura dos indicadores do Censo 2000*. Teoria e Pesquisa, São Carlos 42(3): 245 – 264, jan./jul.2003.
- _____; CARVANO, L. M. *Reflexão e espanto: a questão racial nas políticas de assistência social da criança e do adolescente*. Caderno Abong: 75-91, 2005.
- _____; CARVANO, L. M. *Relatório nacional das desigualdades raciais; 2007 – 2008*, Outubro 2008a.
- _____. *A dialética do bom aluno. Relações raciais e o sistema educacional brasileiro*, 2008b.
- PAULA, G. A. *Modelos de regressão com apoio computacional*. Instituto de Matemática e Estatística, Universidade de São Paulo: dezembro, 2010.
- PESSOA, D.; SILVA, P. *Análise de Dados Amostrais Complexos*, 1998.
- PINHEIRO, A. C., GIAMBIAGI, F. e GOSTKORZEWICZ, J. *O Desempenho Macroeconômico do Brasil nos Anos 90*, 1999.
- PORCARO, R. M. *Desigualdade racial e segmentação no mercado de trabalho*. Estudos Afro-Asiáticos, v.15, p.171-207. 1988.
- REMY, M.; QUEIROZ, S. e FILHO, L. *Evolução recente do emprego formal no Brasil: 2000-2008*. 10. 57-78, 2011.
- SANTOS, J. A. F. *Estrutura de posições de classe no Brasil*. Belo Horizonte: UFMG, 2000.
- _____. *Efeitos de classe na desigualdade racial no Brasil*. Dados, v.48, n.1, p.21-65.2005.
- SILVA, N. *O preço da cor: diferenciais raciais na distribuição de renda no Brasil*, 1980.
- SILVA, N. *Segregação entre brancos e negros no mercado de trabalho brasileiro*. 2010. 145 f. Dissertação (Mestrado em Economia) – Instituto de Economia, Universidade Federal de Uberlândia, Uberlândia, 2010.
- SOARES, S. S. D. *Perfil da Discriminação no Mercado de Trabalho - Homens Negros, Mulheres Brancas e Mulheres Negras*. Brasília: IPEA, 2000.
- STADUTO, J.; MALDANER, I.; SOUZA, K.; GRECO S.; SHIKIDA, P. *Diferenças salariais por cor no mercado de trabalho da região Sul do Brasil*, 2006.
- TEIXEIRA, A.; ROSSETI, E.; ALVES, P.; SILVA, D. *Pesos longitudinais para a pesquisa nacional por amostra de domicílios contínua (PNAD Contínua)*. Mercado de trabalho (Rio de Janeiro. 1996), v. 25, p. 79-90, 2019.
- TELLES, E. *Race in another America: the significance of skin color in Brazil*. Princeton. Princeton University Press, 2004.
- _____. *O significado da raça na sociedade brasileira*. Tradução para português de Race in another America: the significance of skin color in Brazil. Princeton. Princeton University Press, 2012.

WAJNMAN, S. *“Quantidade” e “qualidade” da participação das mulheres na força de trabalho brasileira*, 2015.

WAGLEY, C. *Introduction to Race and Class in Rural Brazil*. Paris: UNESCO, 1952.

A importância das características do trabalho feminino para a fecundidade: reflexões para o Brasil em 2015

Angelita Alves de Carvalho*

Ana Elisa Mello Britto**

Resumo

A inserção feminina no mercado de trabalho tem sido expressiva no Brasil nas últimas décadas e, paralelamente, tem-se observado uma constante redução do número médio de filhos pelas mulheres, existindo, portanto, uma relação negativa entre estas duas variáveis em estudos de nível agregado. O objetivo deste artigo é, por meio de uma análise micro, analisar a relação entre os diferentes tipos de ocupação feminina e suas jornadas de trabalho com o número de filhos tidos. Foram utilizados os dados da PNAD de 2015, estatística descritiva e o modelo de Poisson para observar a relação entre ter filhos e as variáveis de mercado de trabalho. Foi possível constatar que ocupações de maior flexibilidade de horário e estabilidade influenciavam positivamente a fecundidade.

Palavras-chave: Tipo de ocupação. Número de filhos. Conciliação trabalho-família.

* Professora do programa de pós-graduação da ENCE.

** Mestre pelo programa de pós-graduação da ENCE.

Introdução

Assim como em vários países desenvolvidos as mulheres brasileiras estão, cada vez mais, se inserindo no mercado de trabalho. Segundo Itaboraí e Ricoldi (2016) a taxa de participação das mulheres no mercado de trabalho era de menos de 20% na década de 50, e passou para perto de 50% a partir de 2010. Ou seja, as mulheres tiveram um papel muito mais importante quando se trata do crescimento da População Economicamente Ativa (PEA)¹ quando comparadas aos homens até a década de 90. Pinheiro et al. (2016) observam que, em 1995, é possível perceber que a taxa de atividade para mulheres era de 54,3%, e em 2014 não chegou a 58%, indicando que as brasileiras atingiram, naquele momento, um “teto” de participação difícil de ser ultrapassado. De modo mais abrangente o estudo de Carvalho et al. (2023) revelou que entre 2003 e 2018 houve três movimentos diferentes de inserção feminina no mercado de trabalho. O primeiro, entre 2003 e 2011, de alto crescimento (combinado com recuperação da crise de 1999); o segundo, entre 2011 e 2014, de crescimento mais lento; e o terceiro, entre 2014 e 2018, marcado pela estagnação. Entre 2003 e 2018, a participação das mulheres no mercado de trabalho formal cresceu 4 pontos percentuais, saindo de 40% em 2003 e chegando a 44% em 2018. Dados recentes do IBGE (2024) mostraram que esse percentual passou para 53% em 2022, contudo ainda marcado por diferenciais regionais e de gênero.

Além disso, essa inserção feminina continua se dando de maneira desigual como destaque por Teixeira (2017), especialmente em alguns setores e ocupações, os quais, na maioria das vezes, possuem menores rendimentos, piores condições de trabalho e prestígio social. Além disso, há predominância feminina nas ocupações do tipo conta-própria, comércio de mercadorias, emprego doméstico, prestação de serviços e funcionalismo público. Esses resultados foram confirmados recentemente pelo IBGE (2024) o qual revelou que em 2022, as mulheres representavam cerca de 85,2% contra 60,4% dos homens no setor de serviço. Adicionalmente, considerando ambas as jornadas de trabalho, em ocupação no mercado de trabalho e no doméstico não remunerado, a carga horária total semanal de trabalho das mulheres foi de 54,4 horas e, para homens, 52,1 horas.

Muitas vezes, essa inserção diferenciada das mulheres no mercado de trabalho decorre de suas decisões reprodutivas², as quais também têm passado por mudanças. Segundo o IBGE o número médio de filhos por mulher tem reduzido constantemente nas últimas décadas, passando de 6,2 filhos por mulher em 1960 para 1,9 em 2010 com previsão de atingir em torno de 1,6 em 2024 (IBGE, 2018). Diversos fatores contribuíram para a redução da fecundidade, dentre eles se destacam questões institucionais (acesso à métodos contraceptivos), econômicos (urbanização, acesso à crédito e renda, etc.) e sociais (maior escolaridade feminina, mudanças nos valores sociais à respeito do papel e das possibilidades das mulheres na sociedade), e segundo alguns autores, a inserção feminina no mercado de trabalho (MARTINE, 1996; CAVENAGHI; BERQUÓ, 2014).

Por outro lado, segundo Bruschini (2007), foi exatamente essa queda da fecundidade o fator mais importante para a entrada das mulheres no mercado de trabalho. Para a autora, a mulher foi “liberada” para trabalhar a partir do momento em que começou a ter menos filhos para cuidar. Contudo, a autora afirma que a manutenção do modelo patriarcal, em que a mulher ainda é responsável pelas atividades domésticas e o homem é o provedor do lar, continua sendo um dos grandes responsáveis por condicionar a participação feminina no mercado de trabalho à outras variáveis além de sua qualificação. Outros fatores como os processos de urbanização e industrialização também se destacam como propulsores da inserção feminina nas atividades produtivas, os quais proporcionaram à elas contato com outras ideias e oportunidades, além da

¹ Segundo definição do IBGE a PEA é composta por pessoas de 10 a 65 anos classificadas como ocupadas ou desocupadas na semana de referência da pesquisa realizada.

² É importante mencionar que estas decisões se dão em contextos bastantes distintos dependendo das características das mulheres e que não são necessariamente realizadas de forma autônomas e conscientemente por elas. E ainda, não são, necessariamente, apenas decisões das mulheres, pois envolvem também os seus parceiros.

expansão da educação. Além disso, as lutas e reivindicações do movimento feminista também contribuíram para a redefinição dos papéis sócio-familiares, acesso e expansão aos direitos civis (escolaridade, voto, etc.) e redução da fecundidade (BRUSCHINI, 2007).

Cunha e Vasconcelos (2016) mostraram que as brasileiras mais participativas no mercado de trabalho são aquelas que moram sozinhas, indicando que a ausência de filhos, ou idosos, ou outras pessoas dependentes possibilita maior participação no mercado de trabalho. Nesse sentido, fica evidente que a condição da mulher da família, especialmente seu papel de cuidadora é decisivo para a sua inserção no mercado de trabalho. E somado, a isso, a falta, praticamente completa, de maior atenção do Estado com políticas voltadas para o compartilhamento dos cuidados, faz com que as mulheres brasileiras apresentem jornadas de trabalho doméstico não remunerado muito superiores aos dos homens, o que acaba por fortalecer a desigualdade entre os gêneros (SORJ et al., 2007; CAVENAGHI; BERQUÓ, 2014). E é justamente a permanência de um modelo que empurra as mulheres para duplas e triplas jornadas de trabalho, ou seja, em que existe uma grande dificuldade na conciliação entre a vida reprodutiva e profissional, que faz com que as taxas de fecundidade brasileiras continuem reduzindo.

Dias-Junior (2007) revelou que existe uma diferença de número de filhos quando se comparam diferentes tipos de ocupações no Brasil. Ocupações que exigem mais especialização, como as de nível superior, tiveram uma média de filhos inferior ao nível de reposição, ou seja, menos que 2,1 filhos. Já nos setores chamados de manuais, como de trabalho doméstico e rural, apresentaram uma média de filhos maior, acima do nível de reposição. Não foram encontradas diferenças grandes entre as mulheres ocupadas e não ocupadas, levando o autor a crer que as diferenças entre os tipos de emprego são muito mais impactantes do que estar desempregada ou não (DIAS-JÚNIOR, 2007).

Em nível agregado, os estudos mostram que existe uma relação negativa entre ter filhos e trabalhar entre as mulheres para a maior parte dos países. Essa tendência foi observada nos países desenvolvidos até os anos 90, a partir de quando houve a inversão desta relação, a qual passou a ser positiva, como, por exemplo, o caso dos países nórdicos e da OECD (AHN; MIRA, 2002; ENGELHARDT, PRSKAWETZ, 2002; DA ROCHA; FUSTER, 2006; BREHM E ENGELHARDT, 2015; OSHIO, 2019), especialmente devido à políticas de conciliação trabalho-família³ e redução da desigualdade de gênero (GORNICK; MEYERS, 2003; KÖGEL, 2004; MATYSIAK; VIGNOLLI, 2008). Brehm e Engelhardt (2015) mostraram que para os países da OECD quanto mais recente o conjunto de dados usado, maior a probabilidade da correlação de séries temporais ser positiva entre a fecundidade e a taxa de participação no mercado de trabalho, mesmo após o controle da heterogeneidade entre países específicos. E concluem que o emprego feminino mais elevado pode tornar os contextos socioinstitucionais mais favoráveis à gravidez, levando a uma associação positiva entre os dois indicadores.

Estudos para estes países mostram que as diferenças de flexibilidade e segurança profissional encontradas entre o setor público e privado se mostram determinantes para a fecundidade das mulheres no mercado de trabalho. Além disso, o tipo de ocupação tem sido uma variável importante para explicar a fecundidade, pois a decisão por um tipo específico de emprego pode existir por conta de preferências pessoais de um estilo de vida almejado, dado que uma mulher que deseja uma família grande, por exemplo, precisa ter um bom salário para poder sustentá-la. A mulher que tem um desejo elevado por filhos pode ser direcionada para carreiras que sejam melhores de se conciliar com a maternidade, mais presentes no setor público ou também no mercado informal de trabalho. Logo, certos tipos de emprego podem influenciar positivamente ou negativamente a taxa de fecundidade. No caso do emprego público poderia relacionar po-

³ Medidas para facilitar a conciliação entre a vida laboral e familiar são indispensáveis para o alcance dos avanços na promoção da igualdade de gênero e do trabalho decente. Um novo modelo de conciliação entre o trabalho e a família se faz necessário, o que pressupõe uma nova forma de entender as responsabilidades familiares como uma função que diz respeito à manutenção e reprodução da sociedade. Este novo modelo baseia-se na noção de corresponsabilidade social pelas atividades de cuidado, ou seja, numa redistribuição das responsabilidades familiares entre famílias, mercado e Estado; e também entre homens e mulheres (OIT, 2009).

sitivamente com o número de filhos, e trabalhos informais negativamente (MARTÍN-GARCIA, 2010; MARTIN-GARCIA; CASTRO-MARTIN, 2013).

No contexto europeu, empregos de jornadas reduzidas (conhecidos como "part-time" ou de meio expediente) podem ser mais favoráveis para conciliar trabalho e maternidade (Schmitt, 2012 apud Bellani e Esping-Andersen (2013), pois estes além do tempo de dedicação ser menor, possuem os mesmos direitos legais e de proteção ao trabalhador, o que acaba por facilitar a criação dos filhos e se torna uma alternativa entre ficar desempregado ou ter um trabalho de tempo integral (CAVALLI, 2012; RONSEN; SKREDE, 2010). Por outro lado, alguns estudos (Bonet et al. 2013; Modena et al, 2011; Adsera, 2006) têm mostrado que a insegurança no trabalho faz a fecundidade cair, pois contratos temporários não dão a segurança econômica necessária que é preciso quando se planeja começar uma família (BAIZAN, 2005).

Apesar das mudanças ocorridas no cenário do mercado de trabalho brasileiro, ainda persistem desigualdades de gênero tanto no âmbito familiar quanto profissional que podem interferir nas decisões reprodutivas femininas. Além disso, a baixa oferta no Brasil de empregos com jornadas flexíveis, meio expediente e ausência de políticas públicas de conciliação trabalho-família, tornam a conciliação entre a vida laboral e as escolhas reprodutivas ainda mais difíceis.

Nesse sentido, questiona-se: de que forma o tipo de ocupação, o setor do trabalho e a jornada de trabalho se relacionam com o número de filhos tidos pelas mulheres brasileiras? O objetivo geral deste trabalho é analisar a relação entre inserção feminina no mercado de trabalho e fecundidade, e como as diferentes formas de inserção da mulher no mercado de trabalho (setor, ocupação e horas trabalhadas) se correlacionam com o número de filhos nascidos vivos destas no Brasil.

Metodologia

Foram utilizados dados da Pesquisa Nacional por Amostra de Domicílio (PNAD) do ano de 2015, a última realizada no modelo anual e com microdados sobre fecundidade disponibilizado pelo IBGE até o momento. Apesar de antigos, esses são os únicos dados disponíveis para esse tipo de análise no Brasil. Acredita-se que, possivelmente, muitas mudanças podem ter ocorrido desde então, especialmente devido ao impacto da pandemia de COVID e no regime de trabalho pós-pandemia. Os dados do Censo de 2022 bem como da PNADC, caso disponibilizados, serão essenciais para se atualizar os estudos desse tema.

A PNAD é uma pesquisa realizada através de uma amostra probabilística de domicílios e é obtida por intermédio de três estágios, sendo a pesquisa caracteriza por um desenho amostral complexo. Este tipo de amostra utiliza métodos de estratificação, conglomeração, ajuste e seleção proporcional ao tamanho, com o objetivo de alcançar uma precisão maior nas estimativas de interesse. Assim, exige-se a incorporação de algumas técnicas específicas de estimação (SILVA; PESSOA; LILA, 2012). Tais técnicas consistem em utilizar softwares com rotinas para permitir o uso de pesos amostrais (ANDRADE, 2006).

Nesse sentido, foi feita a expansão da amostra da PNAD 2015, através do pacote *survey*, no software R. As variáveis necessárias para a incorporação do desenho amostral complexo da PNAD 2015 foram: V4609 (projeção da população), V4610 (inverso da fração), V4617 (STRAT - Identificação de estrato de município auto-representativo e não auto-representativo), V4618 (PSU - Unidade primária de amostragem) e V4729 (Peso atribuído a pessoa, que é o valor do peso básico multiplicado pelo fator de ajuste. Cada domicílio recebe um peso e às pessoas daquele domicílio é atribuído esse mesmo peso).

A variável resposta utilizada foi o número de filhos por mulher, pois se quer ver com o fato de estar ou não no mercado de trabalho e as condições deste trabalho se correlacionam com a quantidade de filhos que a mulher tem. A variável de interesse, parturição da mulher foi criada a partir da junção da pergunta sobre se tinha filhos nascidos vivos e quantos filhos tinha (no domicílio e fora). Foram excluídos casos incompatíveis com a idade da mulher.

Como a parturição envolve o conceito que captura o número médio de filhos das mulheres ao final de sua vida reprodutiva, optou-se por analisar somente mulheres com mais entre 40 a 49 anos, as quais apresentariam menores chances de não terem filhos e já teriam, na maioria das vezes, sua descendência completada e possibilitaria perceber os efeitos do mercado de trabalho sobre a vida reprodutiva das mulheres. Sabe-se que o Brasil está passando por um processo de adiamento da fecundidade como mostrado por alguns estudos como Rios-Neto *et al* (2018) e Lima *et al.* (2018). Fica claro nestes estudos que o adiamento impacta mais especificamente os grupos de 15 a 24 anos, que adiam, e os de 30 a 39 anos, que fazem a maior parte da recuperação. A fecundidade das mulheres de 40 a 49 anos ainda continua sendo muito pequena em relação ao total da fecundidade dos demais grupos. Se esse adiamento se intensificar, será preciso repensar esse recorte para novos estudos, dado que essas mulheres podem ainda não ter finalizado suas vidas reprodutivas. Nesse sentido, o total de mulheres no banco de dados foi de 23335 observações. Além disso, selecionou-se somente mulheres que viveram ou viviam em união, pois a decisão por filhos, na maioria das vezes, envolve uma relação conjugal seja ela formal ou informal, pois mais de 84% das mulheres que nunca viveram em união não tem filhos

Adicionalmente, é preciso lembrar dos problemas de causalidade quando se utilizam estudos transversais para a análise da relação entre mercado de trabalho e fecundidade ou vice e versa (Cunha e Vasconcelos, 2016). Nesse sentido os modelos irão refletir somente a relação existente entre as variáveis, não sendo possível a análise de impacto ou causalidade. Uma variável criada a fim de se ter algum incídio sobre possíveis efeitos das condições de trabalho sobre a fecundidade foi se teve filho no trabalho. Essa variável é uma tentativa de controlar o fato de que não se sabe a correta direção da relação, pois não houve o acompanhamento dos eventos ao longo da vida da mulher.

As variáveis explicativas foram divididas em dois blocos. O primeiro é composto pelas variáveis de interesse desse trabalho, qual seja, aquelas relativas ao trabalho da mulher, e foram: Trabalhou na semana de referência; Número de trabalhos; Setor do emprego (privado e público); Horas trabalhadas (até 39 horas; 40 e 44 horas; acima de 44 horas); Posição na ocupação (carteira assinada, funcionárias públicas e militares, trabalhadoras sem carteira, conta própria empregadoras e trabalhadoras na construção/produção para o próprio uso/consumo); faixa de rendimento domiciliar per capita (sem rendimento, até 1 salário, de 1 até 4 salários mínimos e maior que 5 salários mínimos). O segundo bloco foi composto por variáveis de controle, e que continham informação do contexto sociodemográfico da mulher: Região Geográfica; Situação censitária; Idade; Raça/cor (branca, negras e outras); Grupos de anos de estudo (não tinham instrução e até 3 anos de estudo, de 4 a 7 anos de estudo, de 8 a 10 anos, de 11 a 14 e 15 anos ou mais).

Em função do evento de interesse ser o número médio de filhos de uma mulher ao final de sua vida reprodutiva, ou seja, um evento de contagem, utilizou-se do modelo Poisson, onde segundo Dobson (2001).

$$Y_i \sim \text{Poisson}(\mu_i) \quad e \quad g(\mu_i) = \ln(\mu_i) = \ln(\eta_i) + \ln\theta_i = \ln(\eta_i) + X_i'\beta$$
$$\mu_i = \eta_i e^{X_i'\beta}$$

onde $\mu_i = \eta_i\theta_i$

onde η_i é a população exposta ao “risco” de ter filhos e θ_i é a probabilidade de se ter filhos.

Para se encontrar o melhor modelo para o ajuste dos dados realizou-se testes estatísticos, mas por ser uma pesquisa de amostra complexa, segundo Feital (2011), os estimadores obtidos via máxima verossimilhança (MV) podem não ser tão precisos. Com isso, foi necessário utilizar os estimadores obtidos através do método da máxima pseudo-verossimilhança (MPV). O método da MPV leva em consideração as características do plano amostral e, com isso, produz um estimador para a variância que é menos sensível às observações atípicas que são encontradas na amostra (FEITAL, 2011).

Foram estimados 5 modelos, com mudanças em relação a variáveis do mercado de trabalho. O primeiro, e mais geral, teve como variável explicativa o fato da mulher estar ou não trabalhando na semana de referência. Dado que a mulher trabalha, os demais modelos tiveram o intuito de analisar de que forma essa inserção se associava com o número de filhos. Assim, o segundo modelo incluiu horas trabalhadas e o setor que estava inserida a mulher. O terceiro modelo refere-se ao tipo de ocupação que a mulher estava inserida. O quarto e o quinto modelo se referem aos modelos 2 e 3 com a inserção da variável se teve filho no trabalho.

Resultados

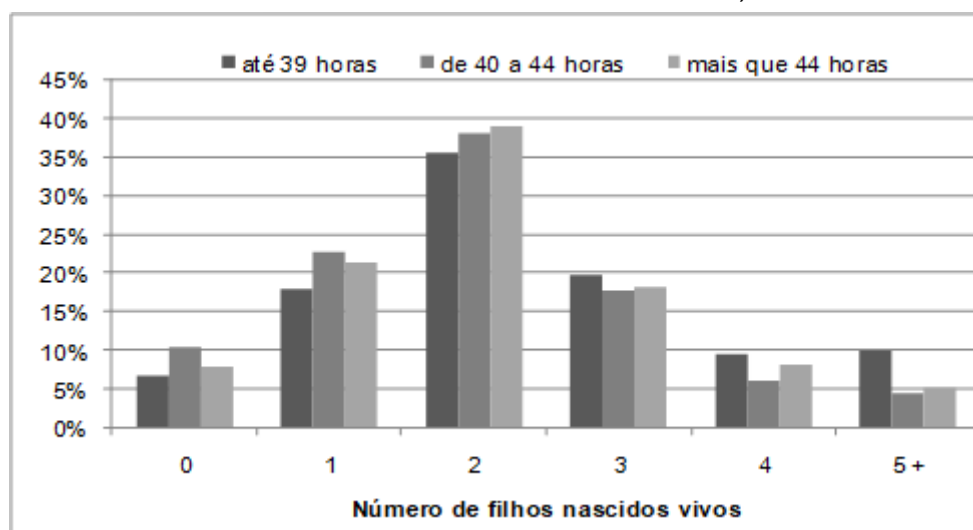
Essa seção está subdividida em duas outras subseções. A primeira traz uma análise descritiva univariada das principais variáveis de mercado de trabalho e sua relação com a parturição das mulheres. E a segunda apresenta os resultados dos 5 modelos estimados.

Análise descritiva

Ao analisar a primeira variável de interesse, que é o fato de estar ou não fora do mercado de trabalho obteve-se que para aquelas que não trabalharam na semana de referência o número médio de filhos ao final da vida reprodutiva é de 2,8 filhos em média, sendo 2,3 entre aquelas que trabalharam na semana de referência. Ou seja, parece que mulheres que trabalham apresentam menos filhos, contudo é preciso ter cuidado, pois esta média não está controlada por outras variáveis, as quais serão analisadas a seguir.

Pode-se verificar, no Gráfico 1, que há uma prevalência de mulheres, independentemente das horas trabalhadas, com dois filhos, sendo maior para as que trabalham mais de 44 horas. Interessante observar que no grupo daquelas que têm 3, 4 e 5 e mais filhos é maior o percentual de mulheres que trabalham menos de 40 horas semanais. Esses resultados podem indicar uma dificuldade dessas mulheres de conseguirem se inserir completamente no mercado de trabalho. Ao mesmo tempo, também pode ser um reflexo de que, por trabalharem menos horas, essas mulheres conseguiram ter uma quantidade de filhos maior.

Gráfico 1. Percentual de mulheres entre 40 a 49 unidas ou uma vez unidas segundo o número de filhos nascidos vivos e trabalho na semana de referência, Brasil 2015



Fonte: IBGE/PNAD (2015)

Dado que as mulheres estão inseridas no mercado de trabalho, sabe-se que o número de filhos pode variar dependendo do tipo de ocupação, setor e horas trabalhadas na semana, como mostrado para em estudos internacionais. Os dados para o Brasil (TAB.1) apresentaram poucas

diferenças, em que aquelas mulheres que estavam localizadas no setor privado apresentaram um percentual levemente maior entre aquelas que tinham um filho (24,7%). Já as mulheres do setor público apresentaram a mesma tendência para a categoria que tinham dois filhos (41,7%). Destaca-se o resultado da média de filhos superior para as mulheres que estão em outro setor, que mesmo não sendo o foco deste estudo, revela que profissões informais ou por conta própria, no caso do Brasil, acabam por atrair mulheres que possuem mais filhos, possivelmente por permitirem maior conciliação com os cuidados das crianças.

Tabela 1. Percentual de mulheres entre 40 a 49 unidas trabalharam na semana de referência segundo número de filhos nascidos vivos e setor do trabalho, Brasil, 2015

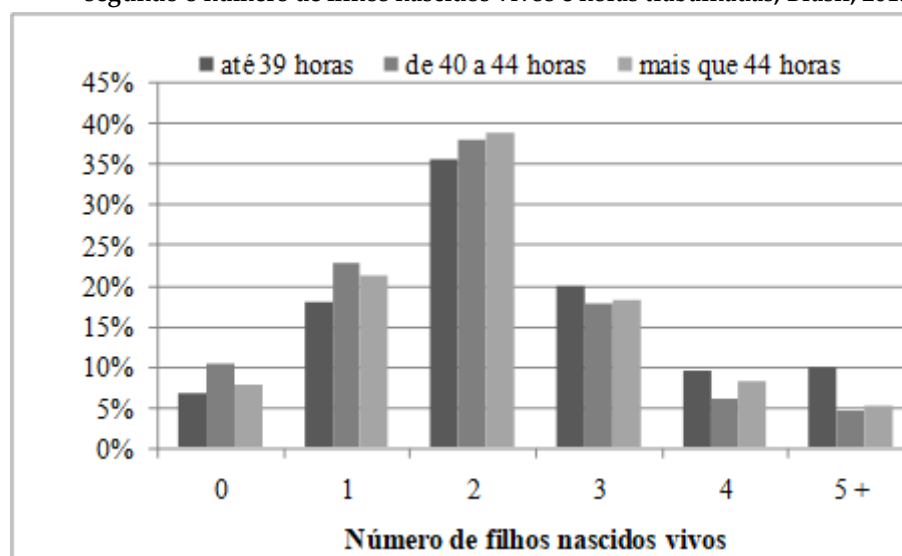
Setor do emprego	Número de filhos nascidos vivos						Total	Média de filhos
	0	1	2	3	4	5 e +		
Privado	10,0	24,7	37,5	16,4	6,9	4,5	100,0	2,0
Público	10,3	21,4	41,7	17,2	5,4	3,9	100,0	2,0
Outros*	6,5	16,7	33,9	21,3	10,1	11,5	100,0	2,6
Total	8,4	20,2	36,7	18,9	8,1	7,7	100,0	2,3

Fonte: IBGE/PNAD (2015)

Nota: * inclui trabalhadores domésticos, conta própria, empregadores e trabalhadores sem carteira assinada

Sabe-se que tempo disponível para o cuidado com os filhos é outra variável chave para as escolhas reprodutivas. Assim, torna-se importante conhecer o número médio de horas dedicadas ao trabalho remunerado exercido pela mulher. Nota-se que a média de filhos é mais elevada (2,6 filhos em média) para mulheres que trabalham menos de 40 horas semanais e é menor (2,1 filhos em média) para aquelas mulheres que possuem jornadas clássicas de trabalho semanal (40 a 43 horas semanais). Entre aquelas que trabalham mais de 44 horas semanais o número médio de filhos também foi ligeiramente maior (2,3 filhos em média). O Gráfico 2 mostra que, de forma geral, mulheres que trabalham menos de 40 horas semanais, ou seja jornadas menores de trabalho, possuem percentuais de 2 ou mais filhos maior, sendo que 40% delas possuem mais de 3 filhos. Enquanto que mulheres com jornadas de 40 a 44 horas, ou seja, tempo clássico de trabalho semanal, apresentam um elevado percentual de zero (10%) e um filho (22%).

Gráfico 2. Percentual de mulheres entre 40 a 49 unidas que trabalharam na semana de referência segundo o número de filhos nascidos vivos e horas trabalhadas, Brasil, 2015



Fonte: IBGE/PNAD (2015)

É importante dizer que não é muito claro se são mulheres que têm famílias maiores que apresentam menos disponibilidade de horas para trabalhar ou se procuram empregos de maiores jornadas para conseguirem sustentar suas famílias numerosas. O que parece ser fato é que trabalhos em jornadas clássicas não facilitam a conciliação com a maternidade, sejam não atraindo mulheres que possuem maior número de filhos, não contribuindo para que as mulheres trabalhadoras tenham filhos.

Outro fator importante é a posição da mulher na ocupação. Na Tabela 2 nota-se que a média de filhos das trabalhadoras com carteira assinada é uma das mais baixas de todas as ocupações. Possivelmente isto está associado ao fato de que se tratam de ocupações clássicas, com pouca flexibilidade e com jornada extensa de trabalho (40 a 44 horas semanais) o que dificulta a conciliação trabalho-família. Já as mulheres ocupadas produção/construção para o próprio consumo apresentam maior número médio de filhos (4,3 em média), seguidas daquelas que não são remuneradas e as que trabalham sem carteira assinada. Essas ocupações não seguem regulamentações trabalhistas, pois, muitas vezes são ocupações não remuneradas, e as mulheres ocupadas nestas atividades são, majoritariamente, mulheres rurais que trabalham de forma não remunerada sem muita distinção entre o que é o trabalho doméstico não remunerado e o que é o seu trabalho voltado para o mercado. E nesse caso, como os ambientes profissionais e familiares são quase inseparáveis e as linhas que os demarcam são quase inexistentes, ficando difícil dizer se realmente são trabalhos que podem ser mais flexíveis e de menores jornadas, o que possibilitaria maior equilíbrio entre vida laboral e familiar. Embora as empregadoras e funcionárias públicas sejam maioria quando se observava mulheres com 2 filhos, seu número médio de filhos está abaixo do nível de reposição. Quando avalia-se as empregadoras e funcionárias públicas, a maioria se concentra na categoria de 1 e dois filhos, especialmente esta última.

Tabela 2. Percentual de mulheres entre 40 a 49 unidas que trabalharam na semana de referência segundo número de filhos nascidos vivos e ocupação, Brasil, 2015

Posição na ocupação	Número de filhos nascidos vivos						Total	Média de filhos
	0	1	2	3	4	5 e +		
Empregada com carteira assinada	10,4	25,0	37,4	16,4	6,4	4,3	100,0	2,01
Militar	6,8	39,6	40,1	13,4			100,0	1,60
Funcionária pública	10,8	23,0	41,7	16,6	5,2	2,7	100,0	1,93
Empregada sem carteira assinada	7,0	16,8	38,9	18,7	9,4	9,2	100,0	2,44
Doméstica com carteira assinada	6,2	17,2	34,7	22,4	8,6	10,9	100,0	2,53
Doméstica sem carteira assinada	6,0	13,6	29,9	23,7	12,3	14,5	100,0	2,84
Conta própria	7,3	18,9	36,0	20,0	8,7	9,1	100,0	2,43
Empregadora	7,7	22,1	42,1	18,2	7,1	2,8	100,0	2,06
Trabalhadora próprio consumo	3,2	8,5	18,1	18,7	14,2	37,4	100,0	4,30
Não remunerado	4,6	12,2	36,1	22,4	11,9	12,9	100,0	2,85
Total	8,4	20,2	36,7	18,9	8,1	7,7	100,0	2,31

Fonte: IBGE/PNAD (2015)

Uma explicação para esses achados pode ser o fato da estabilidade, especialmente, no emprego público. Estas mulheres sabem que terão seus trabalhos quando voltarem de suas licenças-maternidade sem qualquer risco de os perderem (especialmente entre as estatutárias), permitindo melhor conciliação trabalho/família. É importante destacar ainda que as trabalhadoras

para o próprio uso/consumo possuem mais de 30% das mulheres concentrada na categoria de 5 filhos ou mais. Avaliando as mulheres que trabalham sem carteira assinada, mais de 18% tinha 3 filhos na data de referência da pesquisa. Esses resultados, podem indicar maior vulnerabilidade dessas mulheres, dado que não estariam com seus direitos trabalhistas garantidos.

Sabe-se que a renda também é um fator muito importante na decisão de ter ou não filhos, pois, mulheres que ganham mais terão mais dinheiro para investir e cuidar dos filhos. Por outro lado, essas mulheres costumam ser mais escolarizadas, as quais tendem a adiar a maternidade e apresentar mudanças em suas preferências, o que leva a uma redução do número de filhos tidos. Na Tabela 3 fica claro que as mulheres que têm um salário maior acabam apresentando menos filhos em média, pois 26% das mulheres entre 40 a 49 anos que tinham rendimento per capita domiciliar maior que 5 salários mínimos não tinha filhos, atingindo uma média de 1,3 filhos. Enquanto que aquelas com até um salário mínimo tinham em média 2,8 filhos e 47% tinham mais que três filhos.

Tabela 3. Percentual de mulheres entre 40 a 49 unidas que trabalharam na semana de referência segundo número de filhos nascidos vivos e rendimento domiciliar per capita (em salários mínimos), Brasil 2015

Faixa de rendimento	Número de filhos nascidos vivos						Total	Média de filhos
	0	1	2	3	4	5 ou +		
sem rendimento	12,7	16,6	35,4	21,5	12,7	1,2	100,0	2,15
até 1	5,0	14,1	33,1	22,4	12,4	12,9	100,0	2,78
de 1 a 4	10,3	24,6	40,9	16,8	5,3	2,1	100,0	1,90
5 ou+	25,8	28,9	34,6	8,1	2,0	0,6	100,0	1,34
Total	8,0	19,0	36,4	19,6	9,1	8,0	100,0	2,41

Fonte: IBGE/PNAD (2015)

Análise da modelagem

Após essa breve análise descritiva foram estimados os modelos. O modelo 1 busca explicar o impacto da variável “Trabalhou ou não na semana de referência” sobre o número de filhos. Nota-se, na Tabela 4, que essa variável não foi significativa, isso pode estar relacionado ao fato da base só conter mulheres de 40 a 49⁴, ou seja, essas mulheres já possivelmente tiveram todos os filhos que poderiam ter ao longo da vida, logo, estar trabalhando não faz diferença no número de filhos das mesmas. Quanto às variáveis de controle, observa-se (Tab. 4) que em todas as regiões, o número médio de filhos por mulher, é menor em relação à região Norte. Mulheres que viviam em área rural tinham um acréscimo de 14% no número esperado de filhos, em relação às que viviam em área urbana. A cada ano que a mulher envelhece ocorre um aumento de 1% no seu número médio de filhos. Tanto as mulheres negras quanto as de outras raças/cores, possuem número esperado de filhos maior do que em comparação com mulheres brancas. Analisando a faixa de rendimento declarada, nota-se que, as mulheres que recebiam até 1 salário mínimo tinham um número médio de filhos 17% maior do que as sem rendimento, já para as que recebiam mais de 5 salários esse número era 27% menor. Sobre a escolaridade, percebe-se que as mulheres de todas as categorias têm um número médio de filhos menor do que as que não tinham instrução ou até 3 anos de estudo. A importância da escolaridade tem sido apontada em diversos estudos como sendo uma a variável socioeconômica mais importante para explicar o número de filhos por mulher, visto que o aumento da educação feminina pode expandir o

⁴ Foram realizados os modelos utilizando mulheres de 20 a 49 anos e, neste caso, a variável de trabalho foi significativa, enfatizando a hipótese de que esta só não tem significância para mulheres mais velhas devido ao fato de que estas já tiveram todos os filhos que poderiam ter.

acesso à informação sobre os métodos contraceptivos, melhorar as oportunidades de emprego e “empoderar” as mulheres na defesa de suas escolhas sexuais e reprodutivas (POIATTI,2020). E segundo Duarte e Teixeira (2021) o impacto maior está ligado à conclusão do ensino médio, o qual é ciclo que mais contribui para a redução do número de filhos, seguida da finalização do ensino superior e ensino fundamental

No Modelo 2 (TAB.4), inseriu-se a variável de setor de trabalho, bem como as demais as variáveis referentes ao mercado de trabalho que foram escolhidas previamente. Nota-se que, as mulheres do setor privado possuem um número médio de filhos 5% menor do que as do setor público. Isso mostra que, o Brasil vem seguindo a tendência dos países europeus, onde, de acordo com Esping-Andersen e Bellani (2013); Martín-Garcia (2010); Castro-Martin e Martín-Garcia (2013), resultados similares foram encontrados. Quando se avalia o número de trabalhos (TAB.4) que a mulher tinha na data de referência da pesquisa, nota-se que as que possuem mais de 1 trabalho possuem um número médio de filhos maior em comparação com as que só trabalhavam em somente 1 emprego na semana de referência. Isso pode se explicar pelo fato de que, mulheres que possuem mais filhos costumam buscar por trabalhos informais, para dar conta das tarefas do lar e garantir seu sustento, pois essas possuem horas mais flexíveis, embora não tragam estabilidade. Por fim, quando se diz respeito às horas trabalhadas, somente uma categoria foi significativa. Logo, mulheres que trabalham de 40 a 43 horas, possuem um número médio de filhos 2% menor do que as que trabalham menos que isso.

Já o Modelo 3 (TAB.4) que é o modelo com todas as variáveis de trabalho, sem a variável de setor do emprego, e com a variável de tipo de ocupação. Nota-se que para a variável de ocupação, todas as categorias possuem um número de filhos maior do que a mulher de carteira assinada. Esses achados corroboram com resultados já vistos por Dias-Junior (2007) que, em sua tese, mostrou que o tipo de ocupação onde a mulher se encontra se torna até mais importante para o número de filhos, do que estar trabalhando ou não.

Para ter mais uma confirmação da hipótese que a estabilidade é importante para se tornar mãe, foi criada uma variável chamada “Teve filho no trabalho”, onde foram excluídas as mulheres sem filhos e, após isso, verificou-se se a idade do filho mais novo das que restaram era menor ou maior do que os anos no trabalho principal. Com essa variável, foram estimados novamente os modelos de setor e ocupação e, ambos os resultados (expostos através dos Modelos 4 e 5), indicaram que o número médio de filhos das mulheres que tinham tido o filho mais novo no trabalho era até 17% maior do que das que tinham tido o último filho fora do trabalho atual. Ou seja, a estabilidade do emprego atual pode ter levado essas mulheres a terem mais filhos do que teriam inicialmente.

Tabela 4. Resultados dos Modelos de regressão de Poisson para explicar o número de filhos nascidos vivos pelas mulheres de 40 a 49 anos unidas, Brasil 2015

Variáveis explicativas	Categorias	Modelos									
		Modelo 1		Modelo 2		Modelo 3		Modelo 4		Modelo 5	
		Sig.	exp ^β	Sig.	exp ^β	Sig.	exp ^β	Sig.	exp ^β	Sig.	exp ^β
Região Geográfica	Norte	0,000	0,835	0,000	0,827	0,000	0,830	0,000	0,836	0,000	0,836
	Nordeste	0,000	0,811	0,000	0,819	0,000	0,817	0,000	0,822	0,000	0,823
	Sul	0,000	0,835	0,000	0,827	0,000	0,828	0,000	0,828	0,000	0,829
	Centro-Oeste	0,000	0,861	0,000	0,878	0,000	0,875	0,000	0,874	0,000	0,877
Situação Censitária	Urbana	0,000	1,139	0,000	1,139	0,000	1,127	0,000	1,123	0,000	1,118
	Rural	0,000	1,010	0,000	1,010	0,000	1,010	0,000	1,016	0,000	1,016
Idade	continua	0,000	1,010	0,000	1,010	0,000	1,010	0,000	1,016	0,000	1,016
	Branca	0,000	1,083	0,000	1,073	0,000	1,069	0,000	1,078	0,000	1,080
	Negra	0,010	1,162	0,000	1,209	0,000	1,214	0,000	1,224	0,000	1,225
Raça/cor	Outras	0,000	1,083	0,000	1,073	0,000	1,069	0,000	1,078	0,000	1,080
	Sem	0,030	1,174	0,760	1,062	0,660	1,062	0,740	1,062	0,700	1,074
	Até 1 salário	0,340	0,932	0,340	0,827	0,410	0,827	0,480	0,878	0,520	0,889
Faixa de rendimento	De 1 a 4	0,000	0,733	0,020	0,638	0,030	0,638	0,028	0,803	0,260	0,811
	5 ou mais	0,000	0,887	0,000	0,887	0,000	0,887	0,000	0,872	0,000	0,874
	Sem instrução e 4 a 7	0,000	0,811	0,000	0,811	0,000	0,819	0,000	0,808	0,000	0,810
Grupos de anos de estudo	8 a 10	0,000	0,705	0,000	0,719	0,000	0,719	0,000	0,724	0,000	0,726
	11 a 14	0,000	0,691	0,000	0,684	0,000	0,684	0,000	0,696	0,000	0,701
	15 a ou mais	0,480	1,010	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Trabalhou na semana de referência	Sim	0,480	1,010	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	Não	0,060	1,051	0,040	1,051	0,234	1,029	0,160	1,034	0,210	1,150
Nº trabalhos na semana de referência	Um	0,060	1,051	0,040	1,051	0,234	1,029	0,160	1,034	0,210	1,150
	Dois	0,190	1,150	0,190	1,150	0,427	1,088	0,430	1,089	0,427	1,088
Três ou mais	Até 39	0,080	0,980	0,200	2,968	0,337	0,988	0,540	1,007	0,250	0,990
	De 40 a 43	0,250	0,990	0,320	0,990	0,334	0,988	0,410	1,010	0,320	0,990
Horas trabalhadas	Mais de 44	0,065	0,972	0,130	1,022	n/a	n/a	n/a	n/a	n/a	n/a
	Outros	0,065	0,972	0,130	1,022	n/a	n/a	n/a	n/a	n/a	n/a
Setor do trabalho principal	Público	0,150	1,030	0,000	1,062	0,010	1,041	0,000	1,083	0,680	0,994
	Privado	0,000	1,062	0,000	1,062	0,000	1,062	0,000	1,051	0,000	1,051
	Outros	0,010	1,041	0,010	1,041	0,010	1,041	0,110	1,020	0,010	1,041
	Outros	0,000	1,083	0,000	1,083	0,000	1,083	0,020	1,055	0,000	1,055
Tipos de ocupação	Carteira assinada	0,000	1,051	0,000	1,051	0,000	1,051	0,000	1,051	0,000	1,051
	Funcionária pública e militares	0,000	1,051	0,000	1,051	0,000	1,051	0,000	1,051	0,000	1,051
	Sem carteira assinada	0,000	1,051	0,000	1,051	0,000	1,051	0,000	1,051	0,000	1,051
	Conta própria e empregadoras	0,000	1,051	0,000	1,051	0,000	1,051	0,000	1,051	0,000	1,051
Trabalhadora para próprio	Trabalhadora para próprio	0,000	1,051	0,000	1,051	0,000	1,051	0,000	1,051	0,000	1,051
	Trabalhadora para próprio	0,000	1,051	0,000	1,051	0,000	1,051	0,000	1,051	0,000	1,051
Teve filho no trabalho	Não	0,000	1,171	0,000	1,171	0,000	1,171	0,000	1,171	0,000	1,171
	Sim	0,000	1,171	0,000	1,171	0,000	1,171	0,000	1,171	0,000	1,171

Fonte: IBGE/PNAD (2015)

Nota: Foi considerado o nível de significância de 10%

Considerações finais

Este trabalho buscou analisar como as diferentes formas de inserção da mulher no mercado de trabalho se correlacionam com o número de filhos das mesmas, no Brasil, em 2015. E conclui-se que, apesar da inserção feminina no mercado de trabalho não se mostrar significativa para sua parturição, as diferentes características do trabalho exercido pelas mulheres, especialmente, o setor, o tipo de ocupação e o número de horas trabalhadas possui forte associação com sua parturição final.

As características do trabalho das mulheres no mercado de trabalho se mostraram relevantes para o número de filhos tidos, visto que, quando foram avaliadas as diferenças entre os tipos de ocupação em que as trabalhadoras estavam inseridas, foi encontrado que todos os tipos de ocupação estudadas tinham um número médio de filhos maior do que das mulheres que trabalhavam com carteira assinada. Esse achado pode estar relacionado ao fato que empregos com carteira assinada não apresentam flexibilidade de horário e, ao mesmo, tempo, possuem uma jornada extensa de trabalho, para a maioria dos trabalhadores, de 44 horas semanais. Soma-se ainda o fato de que esse seria o setor que inclui as atividades mais tradicionais, com metas e pressões no ambiente de trabalho (como por exemplo comércio, empresas privadas, bancos, corretoras, etc.), que não podem dificultar a articulação entre a vida laboral e familiar, influenciando negativamente a implementação das preferências reprodutivas pelas mulheres.

Focando na relação entre empregos privados e públicos pode-se perceber os possíveis efeitos da estabilidade e do ambiente de trabalho no comportamento de fecundidade. Encontrou-se uma redução de 5% no número médio de filhos das funcionárias de empregos privados, quando comparado com as mulheres que trabalhavam no setor público. Esses achados corroboram estudos que mostram a importância da estabilidade financeira para a decisão de ter filhos, o que é característico do setor público.

A jornada de trabalho também foi um fator relacionado à parturição das mulheres, pois aquelas que trabalhavam entre 40 e 44 horas tinham um número médio de filhos menor do que as que trabalhavam até 39 horas. Jornadas com muitas horas extras (acima de 44 horas) também contribuíam para o aumento do número de filhos. No caso de trabalhos com menores jornadas é esperado que estas mulheres possam se sentir mais motivadas e seguras a ter filhos por dispor de mais tempo para se dedicarem a eles. Mas também podem estar sob o efeito do estudo ser transversal. Um caminho oposto poderia se dar pelo fato de que estas mulheres acabam optando por trabalhar menos para poderem articular o cuidado dessa quantidade maior de filhos com as atividades remuneradas. Em mesmo sentido, as mulheres que trabalhavam mais de 44 horas podem ser aquelas inseridas em ocupações como consumo próprio, sem carteira e não remunerada e que novamente podem estar nessa situação por justamente terem mais filhos e em condições de vulnerabilidade, especialmente no planejamento de sua fecundidade.

Desta forma, pode-se dizer que trabalhos públicos, que garantem estabilidade e que apresentem jornadas menores de trabalho agem positivamente sobre o número de filhos e, por outro lado, empregos de carteira assinada, com jornadas tradicionais de trabalho e do setor privado dificultam a realização da maternidade pelas mulheres. Adicionalmente, variáveis de contexto também se mostraram relevantes para determinar o número de filhos que a mulher, tais como escolaridade, renda, região geográfica e raça/cor.

Apesar de acreditar-se nos efeitos do trabalho da mulher sobre sua parturição, é importante destacar que existe o efeito da causalidade entre os fenômenos estudados, o qual não foi controlado neste trabalho. E assim não se tem como saber de fato, por exemplo, se é o emprego de carteira assinada e suas características que contribuem para que as mulheres tenham uma média de filhos melhor, ou se foi justamente por terem uma média de filhos mais elevada que levou as mulheres a se inserirem em ocupações que não eram de carteira assinada, a fim de tentarem conciliar sua vida reprodutiva já existente com sua vida laboral.

Adicionalmente, destaca a contribuição dos estudos que investiguem a relação entre mercado de trabalho e as preferências reprodutivas, uma vez que essas podem ajudar a prever o comportamento da fecundidade. Nesse nicho, cito a tese Leocádio, (2023) a qual concluiu que o maior o número de horas de trabalho remunerado ou quanto maior a parcela de horas de trabalho remunerado realizado pela mulher (entre o casal), maiores são as probabilidades de ter a intenção de ter o segundo filho. Ademais, estar ocupada (vs. não estar ocupada) aumenta esta mesma probabilidade. A autora afirma que seus resultados sugerem a existência de um efeito renda, de forma semelhante ao encontrado para os países desenvolvidos.

De qualquer maneira, estudos que mostrem estas relações são muito importantes, a fim de evidenciar os impactos das escolhas laborais sobre a vida reprodutiva, pois pode-se inferir

que a presença de um filho demanda tanto das mulheres – e apenas delas – para se readequarem à vida profissional e mesmo pessoal devido à inexistência de um entendimento de que o trabalho doméstico e de cuidados não remunerado é uma responsabilidade social, devendo ser compartilhado por famílias, Estado e mercado. A falta de envolvimento efetivo dos homens nestas atividades, bem como a baixíssima participação do Estado e do mercado neste campo faz com que reste às mulheres o peso quase que exclusivo pela reprodução social.

Referências

- ADSERA, A. *An economic analysis of the gap between desired and actual fertility: The case of Spain*. Review of Economics of the Household, v. 4, n. 1, p. 75-95, 2006.
- AHN, N. & MIRA, P., *A note on the changing relationship between fertility and female employment rates in developed countries*. Journal of Population Economics, n. 15, p 667-682, 2002.
- BAIZÁN, P. (2005). *The impact of labor market status on second and higher-order births. A comparative study of Denmark, Italy, Spain and United Kingdom*. DemoSoc (WorkingPaper 11: 1–28).
- BILAC, E. D. *Trabalho e Família*, Tempo Social, São Paulo, v.26, n.1, p. 129-145, 2014.
- BREHM, U; ENGELHARDT, E. *On the age-specific correlation between fertility and female employment: Heterogeneity over space and time in OECD countries*. Demographic Research, v. 32, p. 691-722, 2015.
- BRUSCHINI, M. C. *Trabalho e gênero no Brasil nos últimos dez anos*. Cad. Pesqui. São Paulo, v. 37, n. 132, 2007.
- CARVALHO, P. A. de; VIEGO, V. *Evolução do emprego feminino no mercado de trabalho brasileiro: uma análise shift-share entre 2003 e 2018*. Economia e Sociedade, v. 32, p. 207-224, 2023.
- CASTRO-MARTÍN, T.; MARTÍN-GARCÍA, T. (2013). *The fertility gap in Spain: Education, Employment, and Fertility*. In: ESPING-ANDERSEN, G. (Ed.) *The fertility gap in Europe: Singularities of the Spanish case* (Number 36 in Social studies collection, Chapter 2, p. 45-81). Barcelona: “la Caixa” Welfare Projects.
- CAVALLI, L. *Fertility intentions of employed mothers in Italy: does the choice of public versus private sector matter?* Working Paper Series, n. 27, Verona, 2012.
- CAVENAGHI, S.; BERQUO, E. *Perfil socioeconômico e demográfico da fecundidade no Brasil de 2000 a 2010*. In: CAVENAGHI, S.; CABELLA, W. (Org.). *Comportamiento reproductivo y fecundidade en América Latina: una agenda inconclusa*. Rio de Janeiro: ALAP, 2014, v. 1, p. 75-90.
- CUNHA, M. S.; VASCONCELOS, M. R. *Fecundidade e participação no mercado de trabalho brasileiro*, Nova Economia, v. 36, n.1, p. 179-206, Maringá, 2016.
- DA ROCHA, J.M.; FUSTER, L. *Why are Fertility Rates and Female Employment Ratios Positively Correlated across OECD Countries?* International Economic Review, v. 47, n.4, 1187-1222, 2006
- DEDECCA, C. S. *Tempo, Trabalho e Gênero*. In: COSTA, A. A.; OLIVEIRA, E. M. de.; BEZERRA, M. E.; SOARES, V. (Orgs.). *Reconfiguração das relações de gênero no trabalho*. São Paulo: CUT Brasil, 2004. Pg. 144p.
- DIAS- JÚNIOR, C. S., *Comportamento Reprodutivo: Uma análise a partir do grupo operacional das mulheres*. 2007. Tese (Doutorado em Demografia), UFMG/Cedeplar, Belo Horizonte, 2007.
- DOBSON, A. J. *An introduction to generalized linear models*. CRC press, 2001.
- DUARTE, H. F. F. L.; TEIXEIRA, E. C. *Efeito do nível de escolaridade sobre a fecundidade no Brasil*. Economia & Região, 9(1), 167-185, 2021
- ENGELHARDT, H., PRSKAWETZ, A., *On the Changing Correlation Between Fertility and Female Employment over Space and Time*. European Journal of Population, v. 20, p. 35-62, 2014.
- FEITAL, T. *Regressão Logística Aplicada a Dados Amostrais Complexos: Um estudo sobre a percepção dos jovens em relação a importância da educação*. Monografia apresentada a UFJF, 2011
- IBGE. *Estatísticas de Gênero Indicadores sociais das mulheres no Brasil*. Estudos e Pesquisas. Informação Demográfica e Socioeconômica, n.38, 2024. Disponível em: https://biblioteca.ibge.gov.br/visualizacao/livros/liv102066_informativo.pdf. Acesso em: 26/06/2024.

IBGE. *Projeções da população: Brasil e unidades da federação: revisão 2018*. Coordenação de População e Indicadores Sociais. – 2. ed. - Rio de Janeiro, 2018. Disponível em: <https://biblioteca.ibge.gov.br/visualizacao/livros/liv101597.pdf>. Acesso em: 26/06/2024.

ITABORAI, N. R.; RICOLDI, A. M. (Org.). *Até onde caminhou a revolução de gênero no Brasil?* Belo Horizonte: Abep, 2016.

KÖGEL, T. *Did the association between fertility and female employment within OECD countries really change its sign?* Journal of Population Economics, n. 17, p. 45-65, 2004.

LAVINAS, L. *Emprego feminino: o que há de novo e o que se repete*. Dados, , Rio de Janeiro, v.40, 1997.

LIMA, E. E. C., ZEMAN, K., SOBOTKA, T., NATHAN, M. AND CASTRO, R. *The Emergence of Bimodal Fertility Profiles in Latin America*. Population and Development Review, vol. 44, no. 4, 723–743, 2018 <https://doi.org/10.1111/padr.12157>

MARTÍN-GARCÍA, T., *The Impact of Occupational Sex-Composition on Women's Fertility In Spain*. European Societies, Madrid, v. 12, p. 113-133, 2010.

MARTINE, G. *Brazil's fertility decline, 1965-95: a fresh look at key factors*. Population and development review, p. 47-75, 1996.

MATYSIAK, A.; VIGNOLI D., *Fertility and Women's Employment: A Meta-analysis*. European Journal of Population, v. 24, issue 4, p 363-384, 2008.

MODENA, F. *et al.* (2011). *Economic In security and Fertility Planning - the Case of Italy*. Paper Prepared for the IARIW - OECD Conference on Economic Insecurity: Paris, 2011.

PAZELLO, E. T., *A Maternidade Afeta o Engajamento da Mulher no Mercado de Trabalho? Um Estudo Utilizando o Nascimento de Gêmeos como um Experimento Natural*. Estudos Econômicos, São Paulo, v. 36, n. 3, p. 507-538, 2006.

PINHEIRO, L. S.; LIMA JR., A. T.; FONTOURA, N. O.; SILVA, R. da. *Mulheres e trabalho: breve análise do período 2004-2014*. Brasília: Ipea, 2016. (Nota Técnica, n. 24).

POIATTI, N. D. *Escolaridade feminina e a sua importância na escolha voluntária da fecundidade*. Revista Ibero-Americana de Estudos em Educação, p. 1786-1798, 2020.

OIT. Organização Internacional do Trabalho Escritório. *Notas da OIT sobre trabalho e família, 2009*. Disponível em: https://www.ilo.org/wcmsp5/groups/public/---americas/---ro-lima/---ilo-brasilia/documents/publication/wcms_229649.pdf. Acesso em: 14 de julho de 2020.

OSHIO, T. *Is a positive association between female employment and fertility still spurious in developed countries?* Demographic Research, v. 41, p. 1277-1288, 2019.

RIOS-NETO, EDUARDO L. G., MIRANDA-RIBEIRO, A. AND MIRANDA-RIBEIRO, P. *Fertility Differentials by Education in Brazil: From the Conclusion of Fertility to the Onset of Postponement Transition*. Population and Development Review, vol. 44, no. 3, 489–517, 2018 <https://doi.org/10.1111/padr.12165>

RØNSEN, Marit; SKREDE, Kari. *Can public policies sustain fertility in the Nordic countries?: Lessons from the past and questions for the future*. Demographic research, Rostock , v. 22, p. 321-346, , 2010

SILVA, P. L. N; PESSOA, D. G. C.; LILA, M. F. *Análise estatística de dados da PNAD: incorporando a estrutura do plano amostral*. Ciência e Saúde Coletiva, v.7, p. 659-670, 2012.

SORJ, B., FONTES, A.; MACHADO, D. C. *Políticas e práticas de conciliação entre família e trabalho no Brasil*. Cadernos de Pesquisa, São Paulo, n. 37, set-dez, 2007.

TEIXEIRA, M. O. *Um olhar da economia feminista para as mulheres: os avanços e as permanências das mulheres no mundo do trabalho entre 2004 e 2013*. Tese (Doutorado)– Universidade Estadual de Campinas, Instituto de Economia, Campinas, 2017.

Características demográficas dos NEETs no Brasil relacionadas à (não) procura por ocupação

Marcos Filgueiras de Sousa*
Miguel Antonio Pinho Bruno**
Angelita Alves de Carvalho***

Resumo

A situação ocupacional da população jovem relacionada à educação e mercado de trabalho é um tema de crescente atenção na literatura. Os desdobramentos sociais, demográficos e econômicos são imensos. Aspectos como cidadania, oferta de mão de obra e igualdade de condições entre homens e mulheres aparecem com frequência quando é analisada a população jovem na condição NEET (*not in employment, education or training*). O presente estudo analisa a população entre 15 e 29 anos residente no Brasil que se encontra nessa condição a partir dos dados da PNAD Contínua e da PNAD, pesquisas domiciliares do Instituto Brasileiro de Geografia e Estatística. A PNAD foi encerrada, sendo os últimos dados disponíveis do ano de 2015. O uso desta pesquisa em paralelo com a PNAD Contínua visa comparar o resultado das respectivas pesquisas. Em relação à população NEET, esta pode ser subdividida naqueles que não estudam, não estão ocupados e, também: (1) não procuram ocupação; ou (2) procuram ocupação. A estratégia de análise consiste em utilizar uma regressão logística para compreender as características sociodemográficas e econômicas da população NEET relacionadas à não procura por ocupação. As conclusões quantitativas do estudo vão de encontro aos apontamentos da literatura de conciliação trabalho-família, uma vez que este grupo é formado majoritariamente por mulheres que não procuram ocupação devido aos “afazeres domésticos, cuidado do(s) filho(s) e/ou de outro(s) parente(s)”. O acesso a serviços estruturais como educação integral e creche então se fazem fundamentais para liberar tempo de cuidados dos filhos e facilitar o ingresso na força de trabalho. A precariedade ou ausência de tais serviços, assim, tende a penalizar proporcionalmente mais o ingresso no mercado de trabalho das mulheres, principalmente aquelas com as seguintes características sociodemográficas e econômicas apontadas na regressão logística: cônjuge, baixa escolaridade, com filhos, residentes em área rural e situadas em regiões de menor renda.

Palavras-chave: NEET. Nem-nem. Mercado de Trabalho. População jovem. Demografia da força de trabalho. Conciliação trabalho-família.

* Doutor pelo programa de pós-graduação da ENCE.

** Professor do programa de pós-graduação da ENCE.

*** Professora do programa de pós-graduação da ENCE.

Introdução

As motivações para compreensão do objeto de estudo NEET (*not in employment, education or training*) no Brasil são imensas. O conceito inclui quem estuda em tempo parcial ou integral, mas exclui atividades educacionais de curta duração e aquelas não ligadas ao sistema educacional formal. O recorte de ocupação é baseado nas orientações da OIT e da OECD e considera aqueles que têm trabalho remunerado por ao menos uma hora na semana de referência da pesquisa ou estavam temporariamente ausentes desse trabalho (OECD, 2016c). Uma das motivações para compreender em maiores detalhes esta população é a sua representatividade. Em 2016, o Brasil possuía 48,5 milhões¹ de jovens entre 15 e 29 anos, sendo 11,711 milhões NEETs, ou seja, 24,1% dos jovens da faixa etária citada. Destes, 7,252 milhões estavam na situação de NEETs inativos (não estudavam, não estavam ocupados e não procuravam ocupação) e 4,459 eram NEETs desocupados (não estudavam, não estavam ocupados, mas procuravam ocupação).

Outra motivação para maior compreensão do objeto de estudo é a diversidade de situações sociodemográficas e econômicas que esta situação pode comportar. A literatura abrange aspectos como cidadania, oferta da mão de obra e igualdade de condições entre homens e mulheres quando é estudada a população na condição NEET (*not in employment, education or training*). Vários trabalhos relacionados à situação da população jovem no que diz respeito à obtenção de qualificação e inserção no mercado de trabalho foram realizados sob o contexto pós-crise econômica de 2008 (EUROFOUND, 2012; KOVROVA; LYON, 2013; CARCILLO *et al.*, 2015; OECD, 2016c; SUSANLI, 2016). Segundo Eurofound (2012), seria preciso dedicar especial atenção ao futuro de 94 milhões de jovens europeus entre 15 e 29 anos que viverão em uma era globalizada com uma população mais envelhecida. Assim, Eurofound (2012) destaca a necessidade de analisar a situação em relação ao estudo e, também, ao fato de o jovem estar ou não na força de trabalho. Especificamente no caso brasileiro, a crise econômica e a dificuldade para obter ocupação se intensificou para toda a população e, em especial, para os jovens a partir de 2014 (ALVES, 2016).

Ademais, cabe destacar que a população NEET também possui grande parte na situação de inativos (não procuram ocupação). Um aspecto relacionado ao fenômeno NEET no Brasil que é sinalizado de forma recorrente na literatura é a forte composição de mulheres jovens com filhos que estão fora da força de trabalho (CAMARANO; KANSO, 2012; CABANAS *et al.*, 2014; COSTA; ULYSSEA, 2014). Este panorama relacionado à condição não economicamente ativa de muitas mulheres ao redor do mundo é fortemente associado aos cuidados dos filhos (CARCILLO *et al.* 2015). Assim, a limitação de possibilidades para conciliar profissão e atividades familiares torna-se, então, elemento fundamental para a contextualização do objeto de estudo NEET, especificamente no que diz respeito aos inativos sob a ótica do mercado de trabalho. Tal situação se insere no debate da OIT sobre Condições para o Trabalho Decente, especificamente no que diz respeito à conciliação trabalho/família (SORJ, 2004).

Uma vez apresentada essa segmentação entre NEETs desocupados e inativos, tal recorte também guiará a estratégia empírica utilizada neste estudo. Há especial interesse em entender o quanto que um modelo preditivo da situação NEET relacionada à não procura por ocupação dialoga com aspectos relacionados à conciliação trabalho/família. Visto isso, tal temática assume papel central na revisão de literatura do presente artigo.

Para contextualizar a política de conciliação trabalho/família, Sorj *et al.* (2007, p. 574) destacam que “o baixo desenvolvimento de serviços coletivos que permitem socializar os custos dos cuidados com a família penaliza a quantidade e qualidade da inserção feminina, sobretudo das mães, no mercado de trabalho”. As autoras citam a existência de três políticas mais comuns associadas à conciliação trabalho/família a partir dos apontamentos de Gornick e Meyers (2003):

- i) licenças do trabalho para cuidar dos filhos, sem perda do emprego e com a manutenção do salário ou de outros tipos de benefícios monetários equivalentes; ii) regulação do tempo do trabalho que permite aos pais reduzir ou realocar as horas de trabalho quando as neces-

¹ Dados da Pesquisa Nacional por Amostra de Domicílios Contínua (PNAD Contínua) do Instituto Brasileiro de Geografia e Estatística (IBGE).

sidades de cuidados com os filhos são mais prementes, sem custos econômicos e para o desenvolvimento da carreira profissional; iii) acesso a creches, pré-escolas e escolas em tempo integral como um arranjo alternativo para os cuidados dos filhos quando os pais estão no local de trabalho (SORJ *et al.*, 2007, p. 574).

Dessa forma, podem-se reunir três aspectos de bem-estar social ligados à conciliação trabalho/família: benefícios ligados à licença para cuidado dos filhos; flexibilidade do trabalho; e acesso aos serviços de educação e cuidado em tempo integral. Contudo, o acesso a benefícios como licença-maternidade no Brasil se depara com uma questão crucial: a formalização do emprego (GOLDANI, 2002; SORJ, 2004; SORJ *et al.*, 2007; BILAC, 2014). Sorj *et al.* (2007, p. 575) destacam o crescimento da informalidade no mercado de trabalho durante os anos 1990, o que resultou em um contingente maior de trabalhadores que não tiveram acesso aos direitos básicos da legislação, “notadamente, as cônjuges e chefes de famílias monoparentais que estão proporcionalmente mais representadas no segmento informal da economia.”

No entanto, o panorama na década de 2000 apresentou algumas mudanças relacionadas ao contexto da formalização. Bilac (2014) destaca que tal década, principalmente no período 2004-2009, foi caracterizada por uma conjuntura internacional favorável que teve impactos positivos no mercado de trabalho, sobretudo na geração e formalização dos empregos, assim como na melhoria de renda dos trabalhadores. Essa formalização, assim, favorece uma ampliação da cobertura de mulheres que conseguem exercer o direito à licença-maternidade.

Sobre os benefícios ligados à licença, Sorj *et al.* (2007) apontam que a legislação trabalhista brasileira focaliza essencialmente no momento inicial da procriação ao garantir a estabilidade para a gestante, a licença-maternidade e a creche nos primeiros meses de vida. Assim, tomando como perspectiva os apontamentos das políticas de conciliação trabalho/família de Gornick e Meyers (2003), é possível reparar que no Brasil garante-se a licença para cuidado dos filhos, no entanto, as limitações na difusão de jornadas flexíveis e no oferecimento de creches/pré-escolas/escolas em tempo integral tornam-se elementos de dificuldade para maior conciliação trabalho/família (SORJ *et al.*, 2007).

Em relação ao aumento da participação feminina no mercado de trabalho, Sorj *et al.* (2007) e Bilac (2014) destacam a ocorrência de um modelo de “delegação”², no qual parentes, vizinhos e empregada doméstica também teriam uma atuação junto aos cuidados dos filhos. Essa alternativa relacionada à “delegação” é uma forma de contornar a situação, tendo em vista as limitações das políticas públicas (conforme os apontamentos de Gornick e Meyers (2003)) ligadas à licença para cuidado dos filhos, da flexibilidade do expediente de trabalho e da cobertura de horário integral em creches, pré-escolas e escolas.

Bilac (2014) também aponta que a Europa e os Estados Unidos utilizaram as políticas públicas de forma que as mulheres pudessem contar, em certa medida, com um melhor ambiente para a conciliação trabalho/família, enquanto, no Brasil, houve busca por soluções alternativas, tal como a delegação dos cuidados, conforme anteriormente destacado. Ainda acerca da investigação do caso brasileiro, Simões *et al.* (2013) destacam que as políticas de incentivo à permanência na escola, em especial o Programa Bolsa Família, podem ter atuado no sentido de reduzir a taxa NEET de jovens inativos até 17 anos de idade. Ademais, Simões *et al.* (2013) apontam que a oferta de serviços de creche às famílias mais pobres poderia viabilizar condições para que as jovens mães ingressassem no mercado de trabalho e/ou retomassem os estudos em paralelo ao exercício da maternidade.

Por fim, um ponto de cautela em relação à delimitação do objeto de estudo é apresentado por Yates e Payne (2006). Os autores enfatizam que parte dos NEETs na situação fora da força de trabalho podem realizar uma opção consciente no sentido de ter filhos. Este ponto é especialmente importante no sentido de ressaltar que parte do evento NEET ligado à inatividade, sob a ótica do mercado de trabalho, pode comportar situações que não seriam exatamente eventos graves ou duradouros para serem alvo de políticas públicas específicas.

² Terminologia de Hirata e Kergoat (2007).

Uma vez visto os principais aspectos sociodemográficos e econômicos relacionados ao fenômeno, a abordagem do artigo prossegue com a análise empírica. Como próximos passos relacionados à investigação, são apresentadas as bases de dados e métodos de regressão logística utilizados, a análise descritiva dos dados e, posteriormente, a modelagem estatística. Por fim, são apresentadas as conclusões.

Bases de dados e métodos utilizados

Referente às bases de dados, é apresentada aqui uma caracterização destas a partir de suas semelhanças, diferenças e contexto de construção. Uma abordagem sobre as diferenças entre a PNAD Contínua (criada em 2012) e a PNAD (em vigor entre 1967 e 2015³) torna-se relevante pelo fato de que a primeira pesquisa citada foi criada para substituir a segunda. No período 2012-2015, as pesquisas estiveram paralelamente em operação e são disponibilizadas pelo IBGE, contudo a PNAD não se encontra disponível a partir de 2016⁴.

Parte do presente trabalho se propõe a realizar uma análise de regressão logística com dados de 2016 (3º trimestre) da PNAD Contínua. O especial interesse pelo 3º trimestre de 2016 é entender, em um momento de desocupação crescente no mercado de trabalho brasileiro, por que jovens com determinados perfis sociodemográficos apresentam maiores chances de (não) procurar ocupação. Também são analisadas regressões com dados de 2014 a partir da PNAD e da PNAD Contínua do 3º trimestre de 2014. Uma finalidade metodológica neste caso é entender semelhanças e diferenças nos resultados e observar como a PNAD Contínua se apresenta ao substituir a PNAD no sentido de gerar resultados similares.

Ademais, cabe enfatizar que a PNAD era realizada em setembro, ou seja, ocorria em um momento específico do 3º trimestre. Destacar esse ponto serve para sustentar a escolha da PNAD Contínua realizada no 3º trimestre na análise de regressão. Afinal, os resultados das regressões geradas pelos dados da PNAD Contínua e da PNAD são comparados ao longo deste artigo. A contribuição dessa análise, além de avançar no entendimento do fenômeno NEET, visa também a explorar diferenças nos sinais dos parâmetros estimados pela regressão via PNAD Contínua e PNAD para compreender se as principais conclusões se mantêm quando analisadas por meio dessas duas pesquisas distintas.

Diferenças na área de abrangência da pesquisa e na rotação da amostra são aspectos relevantes que podem gerar resultados variados para os mesmos indicadores calculados a partir da PNAD Contínua e da PNAD. As principais diferenças entre as duas pesquisas citadas são apresentadas em IBGE (2015a) e resumidas adiante:

³ Detalhamento sobre o histórico da PNAD em IBGE (2015b).

⁴ Essa substituição está inserida no contexto do Projeto de Reformulação das Pesquisas Domiciliares Amostrais do IBGE e do Sistema Integrado de Pesquisas Domiciliares (SIPD). Maiores detalhes estão presentes em IBGE (2015a).

Quadro 1: Comparações metodológicas entre PNAD Contínua e PNAD

Item	PNAD Contínua	PNAD
População investigada para o tema trabalho	Pessoas de 14 anos ou mais de idade	Pessoas de 10 anos ou mais de idade
Esquema de rotação dos domicílios	1 mês dentro – 2 meses fora (processo se repete 5 vezes)	Não possui
Abrangência de coleta da informação	Nacional (Amostra de 3500 municípios)	Nacional (Amostra de 1100 municípios)
Abrangência de divulgação dos resultados	Brasil, Grandes Regiões, UFs, RMs que incluem os municípios das capitais e municípios das capitais.	Brasil, Grandes Regiões, UFs e 9 RMs
Periodicidade	Trimestral, Anual e Mensal por meio de trimestres móveis, para um conjunto restrito de indicadores de trabalho para o Brasil.	Anual

Fonte: Adaptado de IBGE (2015a).

Os aspectos que resultaram na escolha da PNAD Contínua como principal base para a análise são: a quantidade de municípios existentes na amostra (mais do que o triplo de municípios da PNAD conforme apresentado no quadro 3.3); o refinamento na condição do domicílio (PNAD 2015 possui 8 categorias enquanto que a PNAD Contínua possui 20); e o fato de que, a partir de 2016, a PNAD Contínua torna-se a principal pesquisa amostral⁵ que permite o acompanhamento de curto e médio prazo de indicadores do mercado de trabalho brasileiro.

Referente ao método utilizado, o modelo de regressão logística é útil para estimar equações de regressão que forneçam o valor esperado de uma variável binária dependente Y com uma ou mais variáveis preditoras (HEERINGA *et al.*, 2010). A construção do modelo é realizada a partir do pacote *survey* presente no software R. Este pacote viabilizou a incorporação do peso e o plano amostral da pesquisa domiciliar utilizada. A base de dados não passou por edições e tratamentos adicionais, sendo trabalhada tal como disponibilizada pelo IBGE.

O modelo do presente trabalho permite analisar os NEETs no Brasil entre 15 e 29 anos de idade e atende ao objetivo de identificar as características sociodemográficas associadas à não procura por ocupação daqueles NEETs classificados como inativos (ou fora da força de trabalho).

A PNAD Contínua é uma pesquisa que fornece informações por indivíduo através dos microdados. Assim, pode-se prosseguir com um modelo que estime probabilidades preditas relacionadas à variável resposta binária Y associadas a uma distribuição Bernoulli. As probabilidades associadas aos valores 0 ou 1 do modelo binário possuem a seguinte interpretação:

- $P(Y_j = 1) = \pi_j$ é a probabilidade de ocorrência do evento para o j -ésimo indivíduo, onde o evento é ser NEET desocupado (procura ocupação);
- $P(Y_j = 0) = 1 - \pi_j$ é a probabilidade de não ocorrência do evento para o j -ésimo indivíduo, ou seja, o evento é ser NEET fora da força de trabalho (não está procurando ocupação).

O modelo citado apresenta os seguintes componentes:

$$\text{logit}(\pi_j) = \ln\left(\frac{\pi_j}{1 - \pi_j}\right) = \eta_j = \mathbf{x}_j^T \boldsymbol{\beta} \quad j = 1, \dots, n;$$

η_j é o preditor linear;

\mathbf{x}_j^T é o vetor de variáveis explicativas;

⁵ Detalhamentos adicionais em relação ao Plano Amostral da PNAD Contínua e diferenças em relação à PNAD são apresentados em IBGE (2015a, p.5).

β é o vetor de parâmetros.

A construção do modelo de regressão logística com a incorporação do peso e do plano amostral, segundo Heeringa *et al.* (2010), é composta por 4 etapas: i) Especificação do modelo; ii) Estimativa dos parâmetros do modelo e desvios padrão; iii) Avaliação do modelo ajustado; e iv) Interpretação e inferência.

Referente à parte 1 (processo de especificação), Heeringa *et al.* (2010) aconselham que os analistas sigam as seguintes recomendações de Hosmer e Lemeshow (2000) para selecionar as variáveis do modelo:

- (1.1) Realização de análises bivariadas iniciais de relação da variável resposta Y com as candidatas a variáveis de preditores individuais;
- (1.2) Seleção de preditores que têm uma associação bivariada com a variável resposta y com significância $p < 0,25$ como candidatos para efeitos principais em um modelo de regressão logística multivariada;
- (1.3) Análise da contribuição de cada preditor para o modelo multivariado utilizando o Teste de Wald;
- (1.4) Verificação da hipótese de linearidade para preditores contínuos;
- (1.5) Avaliação da existência de interações cientificamente justificadas entre preditores.

Em relação à parte 2, na qual ocorre a estimação dos parâmetros e desvios-padrão do modelo, Heeringa *et al.* (2010) apontam que o método da máxima verossimilhança não pode ser utilizado no caso de amostras complexas. A justificativa é que as probabilidades de seleção (e resposta) para as observações $j=1, \dots, n$ da amostra não são iguais.

Para superar essa limitação, o modelo construído neste artigo faz uso do Método da Pseudo Máxima Verossimilhança proposto por Binder (1983). Essa abordagem para estimação de parâmetros é combinada com um estimador linearizado da matriz de variância-covariância para as estimativas de parâmetros, levando em consideração características de amostras complexas (HEERINGA *et al.*, 2010). Contudo, a partir dessa abordagem, Pessoa e Nascimento Silva (1998, p. 83) destacam que os “*métodos usuais de diagnóstico de ajuste de modelos (tais como gráficos de resíduos) e outros procedimentos da inferência clássica (tais como testes estatísticos de Razões de Verossimilhança) não podem ser utilizados*”.

A parte 3 (avaliação do modelo ajustado) sinalizada no início desta seção consiste em analisar a contribuição de cada preditor para o modelo multivariado e, a partir disso, decidir quais variáveis permanecerão na análise, enquanto que a parte 4 está relacionada à interpretação do modelo escolhido a partir dos testes de hipóteses e razão de chances.

No processo de especificação do modelo (parte 1.1 da recomendação de Hosmer e Lemeshow (2000)) é realizada análises bivariadas de relação entre a variável resposta e cada uma das candidatas à variável preditora.

A seleção das variáveis predictoras que tenham associação bivariada com a variável resposta Y (parte 1.2) é baseada no Teste de Wald executado a partir da função *regTermTest* da *library survey* disponível no software R. Esse teste é utilizado para analisar um a um os modelos com cada variável candidata à preditora. No Teste de Wald, a hipótese testada é de que todos os coeficientes associados a um termo de regressão específico são zero. Quando se observa no teste que o coeficiente estimado da variável candidata não apresenta significância estatística, opta-se pela sua exclusão.

⁶ Conforme Heeringa (2010) aponta, o desenvolvimento e a avaliação da abordagem Pseudo Máxima Verossimilhança são apresentados em Roberts, Rao e Kumar (1987), Morel (1989) e Skinner, Holt e Smith (1989). Nessa abordagem explora-se o método padrão para modelagem de regressão logística em todos os principais sistemas de software que suportam a análise de dados oriundos de amostras complexas.

A análise da contribuição de cada preditor para o modelo (parte 1.3) é realizada por meio do teste t com a incorporação do número de UPAs e estratos. Essa adaptação é útil para avaliar a significância dos parâmetros correspondentes às variáveis presentes no modelo com dados amostrais complexos. Conforme recomendado por Heeringa *et al.* (2010) para o caso de dados amostrais complexos, o presente trabalho faz uso do teste t conforme apresentado a seguir:

$$\begin{aligned} H_0: \beta_i &= 0 \\ H_1: \beta_i &\neq 0 \end{aligned}$$

Sendo β_i o parâmetro relacionado à variável i .

A estatística de teste sob H_0 é:

$$t^* = \frac{\widehat{\beta}_i - 0}{\sqrt{\widehat{V}(\widehat{\beta}_i)}} \sim t\text{-student} \left(u - L \right)$$

Onde:

$\widehat{V}(\widehat{\beta}_i)$ é a variância do estimador;

u é o número de Unidades Primárias de Amostragem (UPA);

L é o número de estratos.

Ao nível de significância α definido pelo pesquisador, se $P(t^* > |t|) \leq \alpha$, há evidências estatísticas para rejeitar H_0 .

Em relação à parte 1.4 dos apontamentos de Hosmer e Lemeshow (2000), destaca-se que foram utilizados apenas preditores categóricos no modelo construído a partir dos dados da PNAD Contínua (3º trimestre de 2016). Referente à parte 1.5, cabe ressaltar que foi utilizada uma interação (sexo*condição no domicílio) ao tomar como base os pontos levantados no resumo de literatura.

Para o presente modelo, são analisadas as seguintes candidatas a variáveis predictoras presentes na PNAD Contínua do 3º trimestre de 2016:

Tabela 1: Variáveis candidatas à inclusão no Modelo de Regressão Logística

Variável	Descrição
-	Região brasileira
V2005	Condição no domicílio
V2007	Sexo
V2009	Idade do morador na data de referência
V2010	Cor ou raça
VD3001	Nível de instrução mais elevado alcançado (pessoas de 5 anos ou mais de idade)
Região	Grande Região Brasileira

Fonte: Dicionário da PNAD Contínua – 3º trimestre de 2016.

A justificativa para o trabalho com essas variáveis via PNAD Contínua reside no seguinte fato: a questão central investigada é a não procura por ocupação em um cenário recente de recessão econômica. Apesar de a versão vigente da PNAD Contínua na data de elaboração deste trabalho ser uma pesquisa com menos variáveis disponíveis que o Censo Demográfico ou a PNAD, a PNAD Contínua é a única que consegue captar informações relevantes sobre o objeto de estudo no ano de 2016 com abrangência nacional.

Por fim, são tratados aspectos ligados à interpretação do modelo. Uma das formas possíveis de interpretá-lo é por meio da razão de chances. No presente trabalho, a razão de chances é elaborada como a razão entre a probabilidade de ocorrência da procura por ocupação e a probabilidade de não procurar ocupação, sendo comparada com a categoria de referência (*baseline*) da covariável analisada. Dessa forma, torna-se possível interpretar as chances de um NEET estar procurando ocupação com base nas comparações de grupos com características sociodemográficas distintas. Para analisar as razões de chance, é levado em consideração um Intervalo de Confiança de 95%. Outra forma utilizada para analisar os resultados do modelo é por meio das probabilidades preditas de procura por ocupação a partir de distintos perfis sociodemográficos disponíveis na análise.

Análise descritiva das variáveis candidatas à preditoras na modelagem

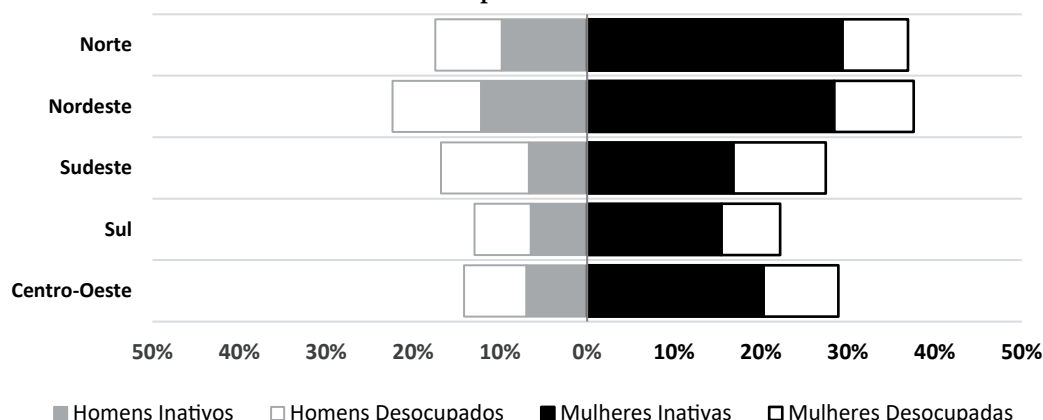
A presente seção tem como finalidade realizar uma análise prévia das variáveis candidatas à inclusão no modelo de regressão logística com dados disponíveis da PNAD Contínua do 3º trimestre de 2016. Antes de executar a análise, cabe apresentar alguns adendos relacionados à escolha de períodos e ao manuseio das variáveis:

A escolha desse período reside nos seguintes motivos:

- a) O ano de 2016 apresentou taxas de desocupação elevadas em um momento de recessão da economia brasileira. Há especial interesse em entender as características sociodemográficas dos NEETs associadas à não procura por ocupação nesse período.
- b) O 3º trimestre foi escolhido porque nessa mesma época do ano foram realizadas as PNADs dos anos anteriores. Na seção 4 e no anexo são apresentadas as regressões logísticas com dados de 2016 (para a PNAD Contínua, única pesquisa domiciliar disponível nesse ano) e 2014 (PNAD Contínua e PNAD).
- c) A escolaridade foi dividida em três faixas: sem instrução, fundamental incompleto e fundamental completo ou mais. Não foram apresentadas faixas de escolaridade iniciadas acima de fundamental completo, pois a análise contempla jovens entre 15 e 29 anos. Foi considerado na análise que jovens com 15 anos convencionalmente não poderiam ser considerados com ensino médio completo por motivo de idade. Assim, o maior grau de escolaridade possível a ser alcançado para jovens de 15 anos seria aquele contido na faixa de escolaridade “ensino fundamental completo ou mais”.

Para a análise das variáveis candidatas à inclusão no modelo de regressão logística, primeiro são analisadas as regiões brasileiras. Nota-se no gráfico 1 elevada diferenciação por sexo, principalmente para NEETs inativos (fora da força de trabalho). As regiões de menor rendimento domiciliar per capita (Norte e Nordeste) concentram as maiores Taxas NEET para homens e mulheres.

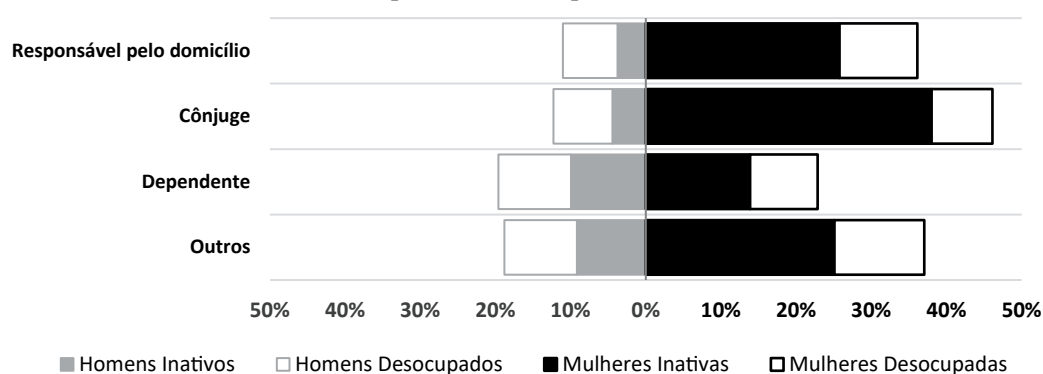
Gráfico 1: Taxa NEET no Brasil, por Região, no 3º trimestre de 2016 - Decomposta em desocupados e inativos



Fonte: Elaboração própria a partir da PNAD Contínua – 3º trimestre de 2016.

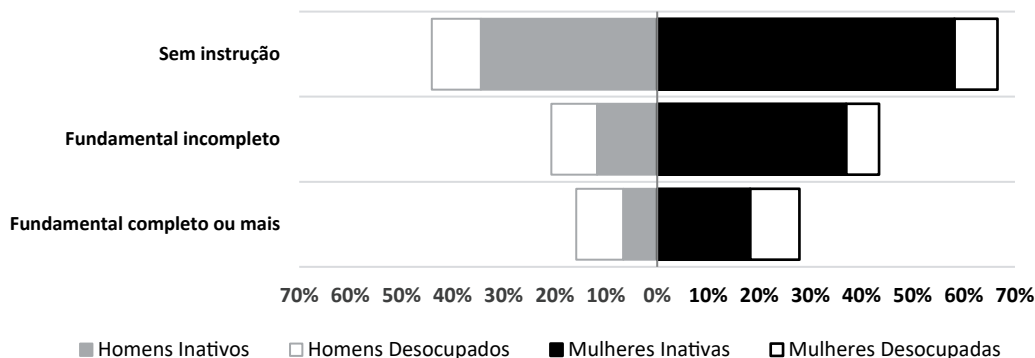
Ao analisar a condição no domicílio, nota-se que em nenhum momento a taxa NEET para homens ultrapassa 20% (conforme gráfico 2), enquanto que para mulheres essa taxa encontra-se sempre acima de 20%, chegando a alcançar mais de 40% para mulheres cônjuges. A menor diferenciação por sexo observada é no caso de jovens “dependentes”, aqueles associados à condição domiciliar de filhos, netos, bisnetos ou enteados.

Gráfico 2: Taxa NEET no Brasil, por Condição no Domicílio, no 3º trimestre de 2016 - Decomposta em desocupados e inativos



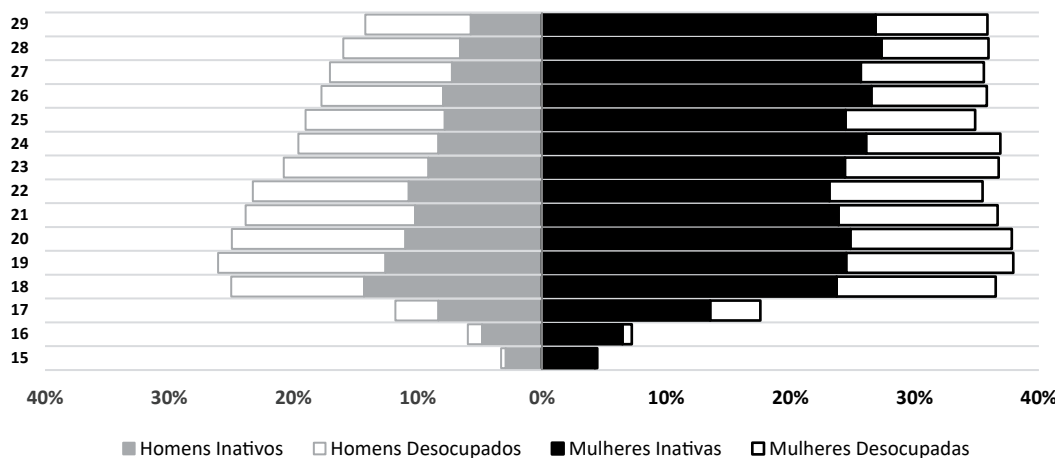
Fonte: Elaboração própria a partir da PNAD Contínua – 3º trimestre de 2016.

No caso da escolaridade, também se observa considerável diferenciação por sexo (conforme gráfico 3). Outro aspecto relevante é que a taxa NEET reduz conforme a escolaridade aumenta no caso das três faixas apresentadas: sem instrução, fundamental incompleto e fundamental completo ou mais. Faixas de escolaridade posteriores não foram incluídas, pois a faixa etária analisada começa com 15 anos. Dessa forma, “fundamental completo ou mais” foi tomada como a faixa de escolaridade máxima possível a ser alcançada para jovens de 15 anos.

Gráfico 3: Taxa NEET no Brasil, por Faixa de Escolaridade, no 3º trimestre de 2016 - Decomposta em desocupados e inativos

Fonte: Elaboração própria a partir da PNAD Contínua – 3º trimestre de 2016.

A taxa NEET em relação à idade, mostrada no gráfico 4, apresenta um marco importante: o fenômeno se mostra mais intenso a partir dos 18 anos, ou seja, idade relacionada ao término do ensino médio para jovens sem atraso escolar. A partir dos 20 anos, inicia-se uma tendência decrescente da taxa NEET para homens. No caso das mulheres, o fenômeno NEET é majoritariamente composto por aquelas fora da força de trabalho (inativas), principalmente a partir dos 18 anos. Também cabe destacar que a diferenciação por sexo se torna mais forte na faixa etária 25-29 anos, conforme pode ser observado no gráfico 4.

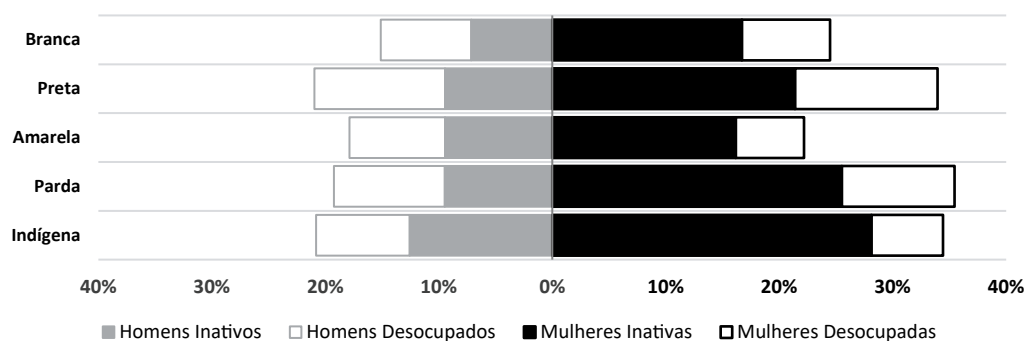
Gráfico 4: Taxa NEET no Brasil, por Idade, no 3º trimestre de 2016 - Decomposta em desocupados e inativos

Fonte: Elaboração própria a partir da PNAD Contínua – 3º trimestre de 2016.

Para a variável raça/cor, nota-se no gráfico 5 que a taxa NEET é menor para brancos e amarelos em relação aos pretos, pardos e indígenas. Também é observado que a diferenciação por sexo persiste para todos os grupos.

Tendo em vista as diferenciações observadas e o recorte da lei brasileira nº 12.711 de 29 de agosto de 2012 sobre ingresso em instituições federais de ensino universitário, técnico e médio para pretos/pardos/indígenas em virtude do histórico de maior exclusão social desses grupos em acesso a tais serviços, os grupos sociodemográficos posteriormente são divididos em brancos/amarelos e pretos/pardos/indígenas na análise da regressão.

Gráfico 5: Taxa NEET no Brasil, por Raça/Cor, no 3º trimestre de 2016 - Decomposta em desocupados e inativos



Fonte: Elaboração própria a partir da PNAD Contínua – 3º trimestre de 2016.

De acordo com os gráficos observados anteriormente, foi possível constatar que o fenômeno NEET é mais intenso para as seguintes características sociodemográficas: mulheres, cônjuges, sem instrução, faixa etária 18-29 anos e para o grupo de raça/cor preta/parda/indígena. Para mulheres o fenômeno mostra-se consideravelmente mais intenso entre aquelas fora da força de trabalho (inativas). Especificamente para a população que não procura ocupação, é possível observar os motivos associados a esse fato a partir da tabela 2.

Tabela 2: Motivos da não procura por ocupação para NEETs inativos no Brasil entre 15-29 anos no 3º trimestre de 2016

Motivo	Homens	%	Mulheres	%	Total	%
Tinha que cuidar dos afazeres domésticos, do(s) filho(s) ou de outro(s) parente(s)	60.473	2,9	2.961.365	57,3	3.021.838	41,7
Estava estudando (em curso de qualquer tipo ou por conta própria)	369.078	17,7	473.915	9,2	842.993	11,6
Por problema de saúde ou gravidez	385.843	18,5	517.402	10,0	903.245	12,5
Por ser muito jovem ou muito idoso para trabalhar	47.223	2,3	59.481	1,2	106.704	1,5
Por não querer trabalhar	470.130	22,6	348.506	6,7	818.637	11,3
Por outro motivo	751.559	36,1	807.496	15,6	1.559.055	21,5
Total	2.084.305	100,0	5.168.166	100,0	7.252.470	100,0

Fonte: Elaboração própria a partir da PNAD Contínua – 3º trimestre de 2016.

No caso das mulheres NEET, 57,3% daquelas que não procuram ocupação registram o motivo “tinha que cuidar dos afazeres domésticos, do(s) filho(s) ou de outro(s) parente(s)”, enquanto que apenas 2,9% dos homens declaram essa resposta. Um adendo especial é que, no caso de outros parentes, o cuidado pode abranger também pessoas de parentesco de maior idade, tal como mãe, pai, avô e avó. Contudo, considerando a estrutura etária atual do Brasil e os apontamentos da literatura, o presente trabalho se dedicou principalmente à questão do cuidado dos filhos. No entanto, há oportunidade de se explorar a questão dos cuidados de idosos e sua relação com a taxa NEET em trabalhos posteriores.

O segundo motivo mais forte verificado para mulheres que não procuram ocupação é “por outro motivo”, e representa 15,6% desse grupo. Para os homens NEET, o principal motivo é “por outro motivo” e representa 36,1%. Em seguida, são observados para homens os motivos “por não querer trabalhar” com 22,6% e por problema de saúde ou gravidez com 18,5%. Os mo-

tivos “por outro motivo” e “por não querer trabalhar” são especialmente críticos no sentido de que, dentre todas as respostas citadas, essas são as que, porventura, poderiam situar jovens em situação de desalento e exclusão social.

Características sociodemográficas dos NEETs relacionadas à procura por ocupação

O primeiro passo para a construção do modelo final é selecionar as variáveis preditoras com base no Teste de Wald para cada modelo univariado. As variáveis selecionadas para este teste são as seguintes:

- Condição no domicílio - organizada em 4 grupos: Cônjuge, Responsável pelo Domicílio, Dependentes (filhos, netos, bisnetos e enteados) e Outros.
- Escolaridade - organizada em 3 grupos: sem instrução, com instrução até fundamental incompleto e fundamental completo ou mais.
- Idade_Faixas - organizada em 3 grupos comumente utilizados na literatura: entre 15 e 17 anos completos, entre 18 e 24 anos completos e entre 25 e 29 anos completos.
- Raça/Cor - organizada em 2 grupos: brancos/amarelos e negros/pardos/indígenas⁷.
- Região: Centro-Oeste, Nordeste, Norte, Sudeste e Sul.
- Sexo: Feminino e Masculino.

As categorias de referência (baseline) são: Sexo Feminino, Condição no Domicílio Cônjuge, Faixa de Idade 15-17 anos, Escolaridade sem instrução e Região Nordeste. Das variáveis citadas anteriormente, apenas Raça/Cor apresentou p-valor maior que 5% no Teste de Wald (conforme tabela 3). Logo, ao não rejeitar H₀ (onde a hipótese é de que os coeficientes associados a um termo de regressão específico são zero), nota-se que não há evidências estatísticas apontando associação entre a variável raça/cor⁸ e a variável resposta (status da procura por ocupação do NEET) ao nível de significância de 5%.

Tabela 3: Teste de Wald dos Modelos Univariados – Base PNAD Contínua 3º trimestre de 2016

Modelo	Variável	Resultado - Teste de Wald
M1	Sexo	p= < 0,001 ***
M2	Cond_Domicilio	p= < 0,001 ***
M3	Escolaridade	p= < 0,001 ***
M4	Idade_Faixas	p= < 0,001 ***
M5	RegiaoBR	p= < 0,001 ***
M6	Cor	p= 0,14667
M7	Sexo*Cond_Domicilio	p= < 0,001 ***

Fonte: Análise baseada na PNAD Contínua – 3º trimestre de 2016.

Adiante é apresentada a tabela 4 com as estatísticas do modelo adotado. As variáveis foram incluídas no modelo com base nos seguintes pontos: disponibilidade na PNAD Contínua; referências obtidas no resumo de literatura; e resultado do Teste de Wald para os modelos univariados. O *baseline* é associado a mulheres residentes na Região Nordeste, com condição

⁷ A motivação para trabalhar com esses agrupamentos reside no recorte feito pela política de cotas raciais da lei nº 12.711, de 29 de agosto de 2012, cuja finalidade é permitir o acesso de grupos historicamente menos favorecidos ao ensino médio, técnico e superior.

⁸ Em relação a essa variável, também foi observada não significância estatística no estudo de Costa e Ulyssea (2014) quando as demais características observáveis são controladas.

domiciliar de cônjuge, sem instrução e com idade entre 15 e 17 anos. O evento em questão para o j-ésimo indivíduo é ser NEET que procura ocupação (desocupado).

Tabela 4: Estatísticas do Modelo – Base PNAD Contínua 3º trimestre de 2016

Nota- ção	Preditor	Catego- ria	Coefi- ciente	Erro Padrão estimado	Valor t	Pr (> t)	Razão de chances	IC 0,95	
								Lim. Infe- rior	Lim. Supe- rior
	Intercepto	-	-3,911	0,141	-27,801	< 0,001 ***	0,020	0,015	0,026
X1j	Sexo	Mascu- lino	2,282	0,172	13,262	< 0,001 ***	9,800	6,994	13,732
X2j	Cond_domi- cilio	Depen- dente	1,114	0,067	16,566	< 0,001 ***	3,048	2,671	3,477
X3j	Cond_domi- cilio	Resp_Do- micilio	0,633	0,082	7,677	< 0,001 ***	1,883	1,602	2,213
X4j	Cond_domi- cilio	Outros	0,885	0,094	9,455	< 0,001 ***	2,422	2,016	2,909
X5j	Escolaridade	Fundam. Incom- pleto	0,803	0,105	7,665	< 0,001 ***	2,233	1,818	2,742
X6j	Escolaridade	Fundam. Com- pleto ou mais	1,414	0,097	14,543	< 0,001 ***	4,114	3,400	4,978
X7j	Idade_Faixa	18-24	0,961	0,083	11,554	< 0,001 ***	2,613	2,220	3,076
X8j	Idade_Faixa	25-29	0,959	0,087	11,040	< 0,001 ***	2,608	2,200	3,092
X9j	Região BR	Norte	-0,211	0,054	-3,942	< 0,001 ***	0,810	0,729	0,899
X10j	Região BR	Centro- -Oeste	0,191	0,063	3,039	0,002 **	1,211	1,070	1,369
X11j	Região BR	Sul	0,213	0,056	3,784	< 0,001 ***	1,237	1,108	1,382
X12j	Região BR	Sudeste	0,508	0,046	10,955	< 0,001 ***	1,661	1,517	1,819
X1jX2j	Sexo*Cond_do- micilio	Mascu- lino e De- pendente	-1,774	0,179	-9,928	< 0,001 ***	0,170	0,119	0,241
X1jX3j	Sexo*Cond_do- micilio	Mascu- lino e Resp_Do- micilio	-0,584	0,213	-2,749	0,006 **	0,557	0,367	0,846
X1jX4j	Sexo*Cond_do- micilio	Mascu- lino e Outros	-1,412	0,208	-6,781	< 0,001 ***	0,244	0,162	0,367

Fonte: Análise baseada na PNAD Contínua – 3º trimestre de 2016.

Categorias de referência (baseline): Sexo Feminino, Condição no Domicílio Cônjuge, Faixa de Idade 15-17 anos, Escolaridade sem instrução e Região Nordeste.

Função de ligação: Logit

Nota: Tamanho da amostra: n = 35.103

Tamanho da população expandida: N = 11.711.396

A equação do modelo escolhido é apresentada adiante:

$$\ln(\pi_j/(1 - \pi_j)) = -3,911 + (2,282)X_{1j} + (1,114)X_{2j} + (0,633)X_{3j} + (0,885)X_{4j} \\ + (0,803)X_{5j} + (1,414)X_{6j} + (0,961)X_{7j} + (0,959)X_{8j} - (0,211)X_{9j} + (0,191)X_{10j} \\ + (0,213)X_{11j} + (0,508)X_{12j} - (1,771)X_{1j}X_{2j} - (0,584)X_{1j}X_{3j} - (1,412)X_{1j}X_{4j}$$

onde:

X_{1j} = Sexo; $X_{1j} = 1$ se masculino e $X_{1j} = 0$ caso contrário.

X_{2j} = Cond_domicilio; $X_{2j} = 1$ se for dependente e $X_{2j} = 0$ caso contrário.

X_{3j} = Cond_domicilio; $X_{3j} = 1$ se for responsável pelo domicílio e $X_{3j} = 0$ caso contrário.

X_{4j} = Cond_domicilio; $X_{4j} = 1$ se for outros e $X_{4j} = 0$ caso contrário.

X_{5j} = Escolaridade; $X_{5j} = 1$ se tiver fundamental incompleto e $X_{5j} = 0$ caso contrário.

X_{6j} = Escolaridade; $X_{6j} = 1$ se tiver fundamental completo ou mais e $X_{6j} = 0$ caso contrário.

X_{7j} = Idade_Faixas; $X_{7j} = 1$ se tiver entre 18 e 24 anos e $X_{7j} = 0$ caso contrário.

X_{8j} = Idade_Faixas; $X_{8j} = 1$ se tiver entre 25 e 29 anos e $X_{8j} = 0$ caso contrário.

X_{9j} = Região_BR; $X_{9j} = 1$ se for Norte e $X_{9j} = 0$ caso contrário.

X_{10j} = Região_BR; $X_{10j} = 1$ se for Centro-Oeste e $X_{10j} = 0$ caso contrário.

X_{11j} = Região_BR; $X_{11j} = 1$ se for Sul e $X_{11j} = 0$ caso contrário.

X_{12j} = Região_BR; $X_{12j} = 1$ se for Sudeste e $X_{12j} = 0$ caso contrário.

Sobre interações entre variáveis no modelo, cabe destacar que a escolha pela inclusão da interação entre sexo e condição no domicílio foi motivada pela relação existente entre NEETS fora da força de trabalho e dificuldades de conciliação trabalho/família apontadas no resumo de literatura. Pela tabela 4, outro aspecto que se pode destacar é a significância estatística a 5% de todos os parâmetros estimados. Contudo, cabe destacar certa instabilidade em relação à variável região quando analisados outros modelos ajustados (conforme tabela 5). Ainda sobre comparações com pesquisas de bases e anos diferentes, cabe destacar estabilidade nos sinais das variáveis (exceto região). Outro aspecto relevante diz respeito à variável situação censitária Rural/Úrbana disponível na PNAD 2014. Essa variável possui significância estatística ao nível de 5% e não estava disponível na PNAD Contínua utilizada neste trabalho. Para a situação urbana, o parâmetro positivo na tabela 5 aponta que a probabilidade de procura por ocupação aumenta para jovens residentes em área urbana. Em outras palavras, nota-se que a inatividade está mais associada, em média, às áreas rurais do que às urbanas. Por fim, um último aspecto relevante está relacionado ao sinal dos coeficientes para a PNAD Contínua do 3º trimestre de 2014 e para a PNAD de 2014. A leitura do fenômeno a partir das variáveis sociodemográficas pouco variou em intensidade e sinal em relação à PNAD Contínua do 3º trimestre de 2016, tendo em vista os coeficientes observados na tabela 5.

Tabela 5: Comparação dos coeficientes estimados em cada modelo com base na PNAD Contínua (3º trimestre de 2016 e 3º trimestre de 2014) e na PNAD (2014)

Nota- ção	Preditor	Categoria	Coeficientes			
			PNADC 2016 3º trimestre	PNADC 2014 3º trimestre	PNAD 2014	PNAD 2014 c/ Rural- -Urbano
	Intercepto	-	-3,911	-4,991	-4,101	-4,681
X1j	Sexo	Masculino	2,282	1,930	2,430	2,365
X2j	Cond_domi- cilio	Dependente	1,114	1,162	1,111	1,088
X3j	Cond_domi- cilio	Resp_Domici- lio / Pessoa_ de_referencia	0,633	0,664	0,626	0,573
X4j	Cond_domi- cilio	Outros	0,885	0,885	0,784	0,748
X5j	Escolaridade	Fundam. In- completo	0,803	1,818	1,295	1,305
X6j	Escolaridade	Fundam. Com- pleto ou mais	1,414	2,390	1,719	1,688
X7j	Idade_Faixa	18-24	0,961	0,779	0,754	0,764
X8j	Idade_Faixa	25-29	0,959	0,752	0,731	0,730
X9j	Região BR	Norte	-0,211	-0,277	-0,134	-0,182
X10j	Região BR	Centro-Oeste	0,191	-0,119	-0,063	-0,171
X11j	Região BR	Sul	0,213	0,014	-0,053	-0,151
X12j	Região BR	Sudeste	0,508	0,260	0,288	0,168
X13j	S i t u a ç ã o Censitária	Urbana	NA	NA	NA	0,790
X1jX2j	Sexo*Cond_ domicilio	Masculino e Dependente	-1,774	-1,448	-2,113	-2,068
X1jX3j	Sexo*Cond_ domicilio	Masculino e Resp_Domici- lio / Pessoa_ de_referencia	-0,584	-0,540	-0,851	-0,765
X1jX4j	Sexo*Cond_ domicilio	Masculino e Outros	-1,412	-0,776	-1,680	-1,655

Coeficiente menor que 0

P-valor > 0,05

Fonte: Elaboração própria a partir da PNAD Contínua (3º trimestre de 2016 e 3º trimestre de 2014) e da PNAD (2014). Maiores detalhes dos modelos com dados de 2014 nas tabelas A.1, A.2 e A.3 do anexo).

Voltando especificamente ao modelo ajustado com dados do 3º trimestre de 2016 (momento de forte crise econômica no Brasil) por meio da PNAD Contínua (conforme tabela 4), são analisados adiante os principais resultados. Em relação à leitura da procura por ocupação dos NEETs de acordo com as características sociodemográficas, a análise é realizada por meio do sinal dos parâmetros estimados e da razão de chances. Uma conclusão sintética do modelo pode ser obtida a partir dos sinais dos parâmetros apresentados na tabela 4 (cuja estabilidade foi analisada na tabela 5). Nota-se que o log das chances de procura por ocupação apresenta os seguin-

tes comportamentos quando comparado com o perfil do *baseline* (Sexo Feminino, Condição no Domicílio Cônjuge, Faixa de Idade 15-17 anos, Escolaridade sem instrução e Região Nordeste):

- Aumenta se a pessoa for do sexo masculino;
- Aumenta se a pessoa não estiver na condição domiciliar de cônjuge;
- Aumenta conforme a faixa de escolaridade cresce;
- Aumenta para pessoas entre 18 e 29 anos;
- Aumenta se a pessoa residir nas Regiões Centro-Oeste, Sudeste ou Sul (localidades com maior rendimento domiciliar per capita). Aqui cabe cautela em relação à interpretação da variável Região, tendo em vista a instabilidade observada na tabela 5;
- Contudo, cabe destacar um aspecto importante relacionado à interação. O aumento da probabilidade de procura por ocupação para NEETs do sexo masculino, em relação ao baseline, é parcialmente reduzido no caso de homens na situação domiciliar de “dependentes” e “outros”. Por outro lado, o efeito da interação aponta que homens na situação de “responsável pelo domicílio” são aqueles relativamente mais propensos a procurar ocupação.

Ademais, cabe reforçar que a dinâmica sociodemográfica do fenômeno NEET associada à (não) procura por ocupação de parte dos NEETs apresenta um comportamento regular no período analisado associado às variáveis de sexo, condição no domicílio, escolaridade, idade e interação sexo*condição no domicílio (vide muitos parâmetros relativamente próximos e com mesmo sinal na tabela 5). Essa regularidade verificada nos resultados das regressões a partir de dados de 2014 (PNAD Contínua 3º trimestre e PNAD) e 2016 (PNAD Contínua 3º trimestre) aponta que a PNAD Contínua possui alto potencial de substituir a PNAD na análise com relativa estabilidade nos resultados obtidos. Contudo, a inclusão da variável rural/urbano na PNAD Contínua também seria de grande utilidade para refinamento da análise.

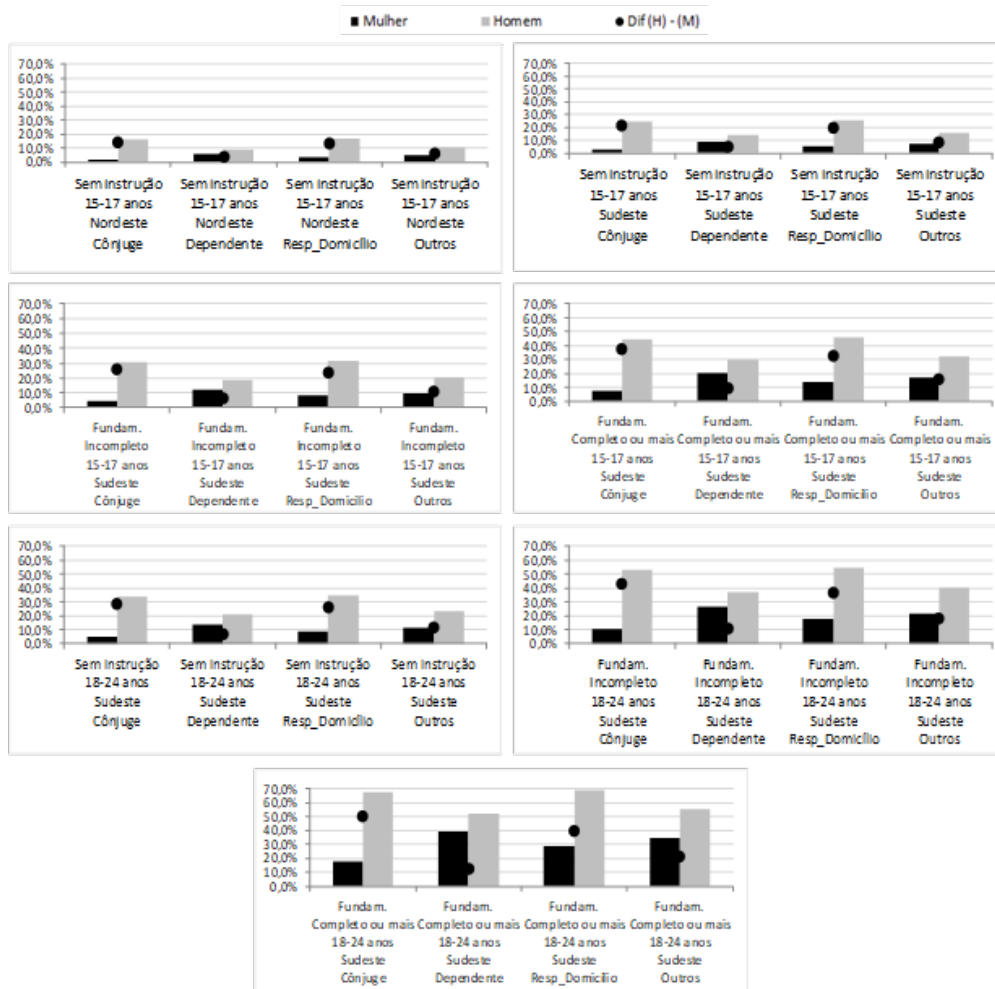
Adiante são apresentados, na ilustração 1, alguns resultados do modelo de regressão logística por meio de outra visualização: as probabilidades preditas de procura por ocupação dos NEETs para determinados perfis sociodemográficos. Em diferentes grupos, é possível notar que a diferença do fenômeno por sexo é consideravelmente menor para jovens na situação de dependente (grupo formado por filhos, netos, bisnetos e enteados). Ou seja, para aqueles que não estão em situações domiciliares características de uma transição realizada para a vida adulta, a diferenciação por sexo atua de forma menos intensa.

Alguns grupos sociodemográficos específicos foram selecionados para análise gráfica. Jovens NEET sem instrução entre 15 e 17 anos e moradores do Nordeste ou do Sudeste são comparados. A motivação dessa análise reside no fato de que a Região Nordeste apresenta aproximadamente metade do rendimento médio mensal domiciliar per capita do Sudeste. Nessa comparação observam-se os seguintes aspectos:

- As probabilidades preditas da procura por ocupação são maiores para jovens NEET entre 15 e 17 anos residentes na Região Sudeste. Em ambas as regiões, a diferenciação do fenômeno por sexo é menor para jovens dependentes (grupo formado por filhos, netos, bisnetos e enteados).
- Quando a escolaridade aumenta, a probabilidade predita da procura por ocupação dos NEETs aumenta para os diferentes perfis sociodemográficos apresentados. Isso é válido para o Nordeste e para o Sudeste quando são analisadas três faixas de escolaridade: sem instrução, fundamental incompleto e fundamental completo ou mais.
- Quando a faixa etária é alterada para 18-24 anos, as probabilidades preditas da procura por ocupação dos NEETs aumentam em relação ao observado na faixa 15-17 anos. Contudo, as probabilidades preditas não aumentam significativamente quando

se comparam as faixas 18-24 anos e 25-29 anos, então não foi construído gráfico considerando essa faixa etária, tendo em vista os resultados próximos da faixa 18-24 anos.

Ilustração 1: Probabilidades previstas da procura por ocupação dos NEETs em 2016 no Brasil fornecidas pelo Modelo de Regressão Logística e a diferença das probabilidades por sexo



Fonte: Elaboração própria a partir da PNAD Contínua do 3º trimestre de 2016.

Adiante é realizada uma análise adicional com o modelo mais completo possível a partir dos dados disponíveis. A base utilizada é a PNAD 2014 e foram analisadas inicialmente as variáveis por meio do Teste de Wald para cada um dos 12 modelos univariados (conforme tabela 6).

Tabela 6: Teste de Wald dos 12 Modelos Univariados analisados – Base PNAD 2014

Modelo	Variável	Resultado - Teste de Wald
M1	Sexo	p= < 0,001 ***
M2	Cond_Domicilio	p= < 0,001 ***
M3	Escolaridade	p= < 0,001 ***
M4	Idade_Faixas	p= < 0,001 ***
M5	RegiaoBR	p= < 0,001 ***
M6	Cor	p= 0,5231
M7	Rural_Urbano	p= < 0,001 ***
M8	Faixa_Renda	p= 0,017 *
M9	Tam_Familia	p= 0,023 *
M10	Filhos_Total	p= < 0,001 ***
M11	Tipo_Familia	p= < 0,001 ***
M12	Sexo*Cond_Domicilio	p= < 0,001 ***

Fonte: Análise baseada na PNAD 2014.

Ao não rejeitar H_0 (onde a hipótese é de que os coeficientes associados a um termo de regressão específico são zero), nota-se que não há evidências estatísticas apontando associação entre a variável raça/cor e a variável resposta (ser NEET que procura ocupação) ao nível de significância de 5%. Nos demais modelos univariados, foi possível rejeitar H_0 .

O modelo completo é apresentado adiante, na tabela 7:

Tabela 7: Estatísticas do Modelo completo – Base PNAD 2014

Nota- ção	Preditor	Categoria	Coefi- cientes	Erro		Pr (> t)	Razão de chan- ces	IC 0,95	
				Padrão esti- mado	Valor t			Lim. Infe- rior	Lim. Supe- rior
	Intercepto		-3,8141	0,2638	-14,457	< 0,001 ***	0,022	0,013	0,037
X1j	Sexo	b.Masculino	1,7636	0,2062	8,553	< 0,001 ***	5,833	3,894	8,739
X2j	Cond_Domicilio	b.Dependente	0,6914	0,0969	7,139	< 0,001 ***	1,997	1,651	2,414
X3j	Cond_Domicilio	c.Pessoa de referên- cia	0,4078	0,0916	4,450	< 0,001 ***	1,503	1,256	1,799
X4j	Cond_Domicilio	d.Outros	0,3828	0,0984	3,890	< 0,001 ***	1,466	1,209	1,778
X5j	Escolaridade	b.Fundamental_In- comp	1,3614	0,1300	10,474	< 0,001 ***	3,902	3,024	5,033
X6j	Escolaridade	c.Fundamental_ Comp_ou_mais	1,7578	0,1247	14,099	< 0,001 ***	5,799	4,542	7,405
X7j	Idade_Faixas	b.18-24	0,8115	0,0774	10,491	< 0,001 ***	2,251	1,935	2,620
X8j	Idade_Faixas	c.25-29	0,8693	0,0834	10,423	< 0,001 ***	2,385	2,026	2,809
X9j	RegiaoBR	b.Norte	-0,1212	0,0605	-2,001	0,045 *	0,886	0,787	0,998
X10j	RegiaoBR	c.Centro-Oeste	-0,0592	0,0766	-0,773	0,439	0,942	0,811	1,095
X11j	RegiaoBR	d.Sul	-0,0316	0,0777	-0,406	0,685	0,969	0,832	1,128
X12j	RegiaoBR	e.Sudeste	0,2623	0,0536	4,896	< 0,001 ***	1,300	1,170	1,444
X13j	Rural_Urbano	b.Urbano	0,8394	0,0703	11,950	< 0,001 ***	2,315	2,017	2,657
X14j	Faixa_Renda	b.Até ¼ salário mí- nimo	0,0971	0,1704	0,570	0,569	1,102	0,789	1,539
X15j	Faixa_Renda	c.Mais de ¼ até ½ sa- lário mínimo	-0,1487	0,1662	-0,894	0,371	0,862	0,622	1,194
X16j	Faixa_Renda	d.Mais de ½ até 1 sa- lário mínimo	-0,3039	0,1639	-1,854	0,064 .	0,738	0,535	1,017
X17j	Faixa_Renda	e.Mais de 1 até 2 sa- lários mínimos	-0,4063	0,1674	-2,428	0,015 **	0,666	0,480	0,925
X18j	Faixa_Renda	f.Mais de 2 até 3 sa- lários mínimos	-0,5767	0,2030	-2,841	0,005 **	0,562	0,377	0,836
X19j	Faixa_Renda	g.Mais de 3 até 5 sa- lários mínimos	-0,6591	0,2303	-2,862	0,004 **	0,517	0,329	0,812
X20j	Faixa_Renda	h.Mais de 5 salários mínimos	-0,9586	0,2636	-3,637	< 0,001 ***	0,383	0,229	0,643
X21j	Faixa_Renda	i.Sem declaração	-0,2697	0,2018	-1,336	0,182	0,764	0,514	1,134
X22j	Filhos_Total	Variável numérica	-0,5278	0,0772	-6,833	< 0,001 ***	0,590	0,507	0,686
X23j	Filhos_Total2	Variável numérica	0,0742	0,0195	3,815	< 0,001 ***	1,077	1,037	1,119
X24j	Tipo_Familia	b.Mãe com filhos menores de 14 anos e de 14 anos ou mais	-0,3198	0,1566	-2,042	0,041 *	0,726	0,534	0,987
X25j	Tipo_Familia	c.Mãe com todos os filhos de 14 anos ou mais	-0,2502	0,1163	-2,152	0,031 *	0,779	0,620	0,978
X26j	Tipo_Familia	d.Casal com todos os filhos menores de 14 anos	-0,5342	0,1046	-5,107	< 0,001 ***	0,586	0,478	0,720
X27j	Tipo_Familia	e.Casal com filhos menores de 14 anos e de 14 anos ou mais	-0,5082	0,1328	-3,827	< 0,001 ***	0,602	0,464	0,780

X28j	Tipo_Familia	f.Casal com todos os filhos de 14 anos ou mais	-0,3623	0,1181	-3,066	0,002 **	0,696	0,552	0,877
X29j	Tipo_Familia	g.Casal sem filhos	-0,2499	0,1285	-1,945	0,052 .	0,779	0,605	1,002
X30j	Tipo_Familia	h.Outros tipos de família	-0,2527	0,1305	-1,937	0,053 .	0,777	0,601	1,003
X1jX2j	Sexo:Cond_Domicilio	Masculino e Dependente	-1,5580	0,2096	-7,435	< 0,001 ***	0,211	0,140	0,317
X1jX3j	Sexo:Cond_Domicilio	Masculino e Pessoa de referência	-0,6312	0,2381	-2,651	0,008 **	0,532	0,334	0,848
X1jX4j	Sexo:Cond_Domicilio	Masculino e Outros	-1,2337	0,2301	-5,362	< 0,001 ***	0,291	0,186	0,457

Fonte: Análise baseada na PNAD 2014.

Categorias de referência (baseline): Sexo Feminino, Condição no Domicílio Cônjuge, Escolaridade sem instrução, Faixa de Idade 15-17 anos, Região Nordeste, Situação Censitária Rural, Faixa de Renda sem rendimento e Tipo de Família "Mãe com todos os filhos menores de 14 anos". Função de ligação: Logit | Nota: Tamanho da amostra: n = 17.807 | Tamanho da população expandida: N = 9.736.446

O modelo da tabela 7 apresenta algumas novidades (em relação ao modelo da tabela 4) que são listadas adiante:

- A partir da faixa de 1 a 2 salários mínimos per capita, notam-se coeficientes com significância estatística a 5%, sinal negativo e comportamento decrescente dos coeficientes estimados (o que atuaria na diminuição da probabilidade de procura por ocupação dos NEETs quando se observa um crescimento da faixa de rendimento domiciliar per capita). Contudo, cabe cautela quanto a essa observação, pois ela possui intervalos de confiança das razões de chance sobrepostos;
- Em relação ao número de filhos das mulheres moradoras no domicílio, observa-se coeficiente estimado da variável "Filhos_Total" com sinal negativo, porém sinal positivo para o coeficiente estimado da variável de filhos elevado ao quadrado "Filhos_Total²". Nesse sentido, a partir dos coeficientes estimados, pode-se notar que a quantidade de filhos contribui negativamente para a procura por ocupação da respectiva mulher NEET em famílias com até 7 filhos, em média. Contudo, a partir de 8 filhos, em média, o comportamento se inverte e essa variável passa a contribuir positivamente na procura por ocupação.
- Em relação ao tipo de família, notam-se coeficientes menores e negativos para famílias classificadas como "casais com filhos", contudo não foi possível extrair conclusões com significância estatística devido à sobreposição dos intervalos de confiança das razões de chance.

Ademais, é informado adiante o Pseudo R² de Nagelkerke para cada modelo estimado.

Tabela 8: Pseudo R² de Nagelkerke dos modelos apresentados

Bases utilizadas para construir um modelo de Regressão Logística	Pseudo R ² de Nagelkerke
PNADC 2016 3º trimestre - ajuste base (tabela 4)	0,1294
PNADC 2014 3º trimestre - ajuste base (tabela A.1)	0,1121
PNAD 2014 - ajuste base (tabela A.2)	0,0990
PNAD 2014 - ajuste base + variável situação censitária (tabela A.3)	0,1080
PNAD 2014 "completo" (tabela 7)	0,1219

Fonte: Elaboração própria.

Em suma, os resultados permanecem em linha com o que é visto na literatura. Em relação à variável raça/cor, a não significância estatística observada no presente artigo não é um evento inédito. Para os NEETs inativos (“nem-nem”), foi observada não significância estatística da variável raça/cor no estudo de Costa e Ulyssea (2014) quando as demais características observáveis são controladas. A análise de Costa e Ulyssea abrange as PNADs de 1992, 2002 e 2012, sendo a variável raça/cor instável e não apresentando significância estatística em 1992 e 2002.

Adicionalmente, uma das características mais marcantes associadas ao fenômeno NEET no Brasil é o fato de esse grupo ter uma forte composição de mulheres jovens com filhos e fora da força de trabalho (CAMARANO; KANSO, 2012; CABANAS *et al.*, 2014; COSTA; ULYSSEA, 2014). Carcillo *et al.* (2015) mostram que a Taxa NEET é sistematicamente mais alta para mulheres, sendo ainda mais intensa nos países emergentes do estudo. Costa e Ulyssea (2014) constataram que a propensão a estar nessa condição é fortemente associada ao fato de ser mulher jovem e ter filhos, ter baixa escolaridade e possuir baixa renda domiciliar. Contudo, no modelo “completo” analisado neste trabalho (conforme tabela 7), o efeito de ter filhos se inverte a partir do 8º filho, em média, no sentido de a mulher NEET sair da “inatividade” e passar a procurar ocupação. Esse dado suscita a hipótese de que no caso dessas famílias com 8 filhos ou mais, em média, irmãos mais velhos estejam auxiliando no cuidado dos mais novos enquanto a mãe NEET tenta buscar ocupação.

Conclusões

O presente estudo teve como objetivo analisar a população NEET e, especificamente, as respectivas características sociodemográficas e econômicas relacionadas à procura por ocupação. Para isso, a população NEET foi separada como objeto específico de estudo e segmentada em desocupados e inativos. Como estratégia empírica, foi construída uma regressão logística binária em que o evento analisado consistia em ser NEET desocupado (aquele que procura ocupação). Previamente à regressão foi levantada a literatura relacionada ao fenômeno, especificamente àquela que tangencia o tema e diz respeito à conciliação trabalho/família, uma vez que o fenômeno possui forte diferenciação por sexo e afeta, no Brasil, proporcionalmente mais mulheres, principalmente aquelas que se dedicam aos cuidados dos afazeres domésticos, dos filhos e de outros parentes.

Os resultados dialogam conforme os principais apontamentos da literatura de conciliação trabalho/família. O acesso a serviços estruturais como educação integral e creche liberam tempo de cuidados dos filhos e facilitam o ingresso na força de trabalho. A precariedade ou ausência de tais serviços, assim, tende a penalizar proporcionalmente mais o ingresso no mercado de trabalho de mulheres, principalmente aquelas com as seguintes características sociodemográficas apontadas pela regressão: cônjuge, baixa escolaridade, com filhos, residentes em área rural e situadas em regiões de menor renda. Baseado nesses resultados, pode-se observar que uma transição para a vida adulta em contexto de precariedade educacional e econômica gera um contexto socialmente mais propício para o crescimento da população NEET, principalmente aquela inativa (que não procura ocupação).

Assim, enquanto o fenômeno NEET impulsionado pela desocupação possui relação com a conjuntura econômica, vale reforçar que a questão dos NEETs fora da força de trabalho pode abranger uma série de adversidades que englobam questões ligadas à violência e evasão escolar (MONTEIRO; ARRUDA, 2011; PADILHA *et al.*, 2011); e também à dificuldade de acesso a serviços e direitos que favoreçam a conciliação entre trabalho e família, tendo geralmente maior impacto sobre as mulheres com filhos. Especificamente em relação aos direitos de licença-maternidade e à maior dificuldade de conciliação, as trabalhadoras informais teriam ainda mais dificuldades de exercer esse direito, tendo em vista a condição ocupacional não coberta pela legislação (GOLDANI, 2002; SORJ, 2004; SORJ *et al.*, 2007; BILAC, 2014).

Visto os pontos anteriores, é importante destacar que a minimização do fenômeno NEET não passa apenas por uma maior capacidade do mercado de trabalho absorver a população jovem em busca de ocupação. A dinâmica econômica certamente é parte importante na redução do fenômeno, principalmente em relação aos NEETs desocupados, contudo, também é preciso garantir um ambiente social propício para que os NEETs inativos ingressem no mercado de trabalho. Por ambiente propício, cabe destacar um ambiente econômico gerador de empregos e com elementos de bem-estar social que facilitem a conciliação trabalho/família tais como disponibilização de creches e educação integral.

Referências

- ALVES, J. E. D. *Crise no mercado de trabalho, bônus demográfico e desempoderamento feminino*. In: ITABORAI, N. R.; RICOLDI, A. M. (Org.). *Até onde caminhou a revolução de gênero no Brasil?* Belo Horizonte: Abep, p. 21-44. 2016.
- BILAC, E.D. *Trabalho e família: articulações possíveis*. *Tempo Social*, v. 26, n. 1, p. 129-145, 2014.
- BINDER D.A. *On the variances of asymptotically normal estimators from complex surveys*. *International Statistical Review*, 51, p.279-292. 1983.
- BRASIL. *Lei Nº 12.711, 29 de agosto de 2012. Dispõe sobre o ingresso nas universidades federais e nas instituições federais de ensino técnico de nível médio e dá outras providências*. 2012. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/lei/112711.htm>. Acesso em 06 de agosto de 2017.
- CABANAS, P.H.F.; MENEZES FILHO, N.A.; KOMATSU, B.K. *Crescimento da renda e as escolhas dos jovens entre os estudos e o mercado de trabalho*. Anais do 42º Encontro Nacional de Economia da ANPEC, Rio Grande do Norte, 2014.
- CAMARANO, A. A.; KANSO, S. *O que estão fazendo os jovens que não estudam, não trabalham e não procuram trabalho?* Boletim de Mercado de Trabalho – Conjuntura e Análise, Nota técnica, Rio de Janeiro, No. 53, nov, 2012.
- CARCILLO, S.; FERNÁNDEZ, R.; KÖNIGS, S.; MINEA, A. *NEET Youth in the Aftermath of the Crisis: Challenges and Policies*. OECD Social, Employment and Migration Working Papers, No. 164, OECD Publishing. Paris. 2015.
- COSTA, J.M.S.; ULYSSEA, G. *O fenômeno dos jovens nem-nem*. In: COURSEUIL, C.H.; BOTELHO, R.U. (Org.). *Desafios à Trajetória Profissional dos Jovens Brasileiros*, p. 115-137. Rio de Janeiro. Ipea, 2014.
- EUROFOUND. *NEETs – Young people not in employment, education or training: Characteristics, costs and policy responses in Europe*. Publications Office of the European Union. Luxembourg, 2012.
- GOLDANI, Ana Maria. *Família, gênero e políticas: famílias brasileiras nos anos 90 e seus desafios como fator de proteção*. *Revista Brasileira de Estudos de População*, v. 19, n. 1, p. 29-48, 2002.
- GORNICK, J. C.; MEYERS, M. K. *Families that work: policies for reconciling parenthood and employment*. New York: Russell Sage Foundation, 2003.
- HEERINGA, S. G.; WEST, B. T. & BERGLUND, P. A. *Applied Survey Data Analysis*. Boca Raton, FL: Chapman & Hall / CRC. 2010.
- HIRATA, H. & KERGOAT, D. *“Novas configurações da divisão sexual do trabalho”*. *Cadernos de Pesquisa – Fundação Carlos Chagas*, 37 (132): 595-609. 2007.
- HOSMER, D. W.; LEMESHOW, S. *Applied Logistic Regression*. Second Edition. New York: Wiley & Sons. 2000.
- IBGE. *Nota técnica - Principais diferenças metodológicas entre as pesquisas PME, PNAD e PNAD Contínua*. 2015(a). Disponível em: <ftp://ftp.ibge.gov.br/Trabalho_e_Rendimento/Pesquisa_Nacional_por_Amostra_de_Domicilios_continua/Nota_Tecnica/Nota_Tecnica_Diferencas_Metodologicas_das_pesquisas_PNAD_PME_e_PNAD_Continua.pdf>. Acesso em 21 de setembro de 2016.
- IBGE. *PNAD - Um Registro Histórico da Pesquisa Nacional por Amostra de Domicílios 1967-2015*. Documentos para disseminação. Memória Institucional nº 20. 2015(b). Disponível em: <<https://biblioteca.ibge.gov.br/visualizacao/livros/liv94878.pdf>>. Acesso em 28 de outubro de 2017.
- KOVROVA, I.; LYON, S. *NEET youth dynamics in Indonesia and Brazil: A cohort analysis*. UCW Programme and University of Rome “Tor Vergata”, Rome, 2013.

- MONTEIRO, V.B.; ARRUDA, E.F. *O impacto da violência urbana nos indicadores de evasão escolar na Região Metropolitana de Fortaleza*. Anais do I Circuito de Debates Acadêmicos. Área Temática 2º CODE:03 – Educação e Desenvolvimento. 2011.
- MOREL, G., *Logistic regression under complex survey designs*, Survey Methodology, 15, 202–223, 1989.
- OECD. *Youth not in employment, education or training (NEET indicator)*. 2016c. Disponível em: <<http://dx.doi.org/10.1787/72d1033a-en>>. Acesso em 27 de outubro de 2016.
- PADILHA, M.A.S.; HYPOLITO, A.M.; SOARES, M.C.; MEINCKE, S.M.K.; BUENO, M.E.N.; FEIJÓ, A.M.; SCHWARTZ, E. *Jovens Mães e Abandono Escolar*. Rev enferm UFPE on line, ago.;5(6). 2011.
- PESSOA, D.G.C.; NASCIMENTO SILVA, P.L. *Análise de Dados Amostrais Complexos*. São Paulo: Associação Brasileira de Estatística (ABE). 1998.
- ROBERTS, G.; RAO, J.N.K.; KUMAR, S., *Logistic regression analysis of sample survey data*, Biometrika, 74, 1–12, 1987.
- SIMÕES, A.; SANTOS, M.F.P.; VAZ, A.C. *Os Jovens que não estudam nem trabalham no Brasil: discussão conceitual, caracterização e evolução de 2001-2011*. Estudo Técnico nº 03/2013 – MDS/SAGI. 2013.
- SKINNER, C.J.; HOLT, D.; SMITH, T.M.F., *Analysis of Complex Surveys*, John Wiley & Sons, New York, 1989.
- SORJ, B. *Reconciling work and family: Issues and policies in Brazil*. Vol. 8. Geneva: International Labour Office, 2004.
- SORJ, B., FONTES, A.; MACHADO, D. C. *Políticas e práticas de conciliação entre família e trabalho no Brasil*, Cadernos de Pesquisa, São Paulo, n. 37, set-dez, 2007.
- SUSANLI, Z. B. *Understanding the NEET in Turkey*. Eurasian Journal of Economics and Finance, p. 42-57, 4(2), 2016.
- YATES, S.; PAYNE, M. *Not so NEET? A critique of the use of “NEET” in targeting interventions with young people*. Journal of Youth Studies, 9 (3), pp. 329-344, 2006.

Anexos

Tabela A.1: Modelo – Base PNAD Contínua do 3º trimestre de 2014

Nota- ção	Preditor	Categoria	Coeficien- tes	Erro Padrão estimado	Valor t	Pr (> t)	Razão de chan- ces	IC 0,95	
								Lim. Infe- rior	Lim. Supe- rior
	Intercepto	-	-4,991	0,175	-28,444	< 0,001 ***	0,007	0,005	0,010
X1j	Sexo	Masculino	1,930	0,162	11,906	< 0,001 ***	6,889	5,014	9,466
X2j	Cond_domi- cilio	Dependen- te	1,162	0,066	17,707	< 0,001 ***	3,195	2,810	3,634
X3j	Cond_domi- cilio	Resp_Do- micilio	0,664	0,096	6,916	< 0,001 ***	1,943	1,610	2,345
X4j	Cond_domi- cilio	Outros	0,885	0,097	9,156	< 0,001 ***	2,424	2,005	2,930
X5j	Escolaridade	F u n d a m . Incompleto	1,818	0,148	12,267	< 0,001 ***	6,158	4,606	8,233
X6j	Escolaridade	F u n d a m . Completo ou mais	2,390	0,142	16,847	< 0,001 ***	10,914	8,264	14,412
X7j	Idade_Faixa	18-24	0,779	0,082	9,473	< 0,001 ***	2,178	1,854	2,559
X8j	Idade_Faixa	25-29	0,752	0,087	8,682	< 0,001 ***	2,122	1,790	2,514
X9j	Região BR	Norte	-0,277	0,059	-4,718	< 0,001 ***	0,758	0,675	0,850
X10j	Região BR	C e n t r o - -Oeste	-0,119	0,070	-1,701	0,089	0,888	0,774	1,018
X11j	Região BR	Sul	0,014	0,059	0,233	0,816	1,014	0,903	1,138
X12j	Região BR	Sudeste	0,260	0,053	4,890	< 0,001 ***	1,296	1,168	1,439
X1jX2j	Sexo*Cond_ domicilio	Masculino e Depen- dente	-1,448	0,170	-8,542	< 0,001 ***	0,235	0,169	0,328
X1jX3j	Sexo*Cond_ domicilio	Masculino e Resp_Do- micilio	-0,540	0,206	-2,618	0,009 **	0,583	0,389	0,873
X1jX4j	Sexo*Cond_ domicilio	Masculino e Outros	-0,776	0,206	-3,767	< 0,001 ***	0,460	0,307	0,689

Fonte: Análise baseada na PNAD Contínua – 3º trimestre de 2014.

Categorias de referência (baseline): Sexo Feminino, Condição no Domicílio Cômjuge, Faixa de Idade 15-17 anos, Escolaridade sem instrução e Região Nordeste.

Função de ligação: Logit

Nota: Tamanho da amostra: n = 30.765

Tamanho da população expandida: N = 10.124.298

Tabela A.2: Modelo – Base PNAD de 2014

Nota- ção	Preditor	Categoria	Coefi- cientes	Erro		Valor t	Pr (> t)	Razão de chances	IC 0,95	
				Padrão estima- do					Lim. Infe- rior	Lim. Supe- rior
	Intercepto	-	-4,101	0,158	-25,883	< 0,001 ***	0,017	0,012	0,023	
X1j	Sexo	Masculino	2,430	0,191	12,754	< 0,001 ***	11,354	7,816	16,492	
X2j	Cond_domi- cilio	Dependente	1,111	0,061	18,081	< 0,001 ***	3,037	2,692	3,425	
X3j	Cond_domi- cilio	Pessoa_de_referen- cia	0,626	0,086	7,245	< 0,001 ***	1,869	1,578	2,214	
X4j	Cond_domi- cilio	Outros	0,784	0,087	8,986	< 0,001 ***	2,191	1,846	2,599	
X5j	Escolaridade	Fundam. Incom- pleto	1,295	0,128	10,095	< 0,001 ***	3,651	2,839	4,694	
X6j	Escolaridade	Fundam. Completo ou mais	1,719	0,124	13,877	< 0,001 ***	5,578	4,376	7,111	
X7j	Idade_Faixa	18-24	0,754	0,077	9,741	< 0,001 ***	2,126	1,826	2,474	
X8j	Idade_Faixa	25-29	0,731	0,082	8,899	< 0,001 ***	2,077	1,768	2,440	
X9j	Região BR	Norte	-0,134	0,060	-2,228	0,026 *	0,875	0,778	0,984	
X10j	Região BR	Centro-Oeste	-0,063	0,076	-0,824	0,410	0,939	0,808	1,091	
X11j	Região BR	Sul	-0,053	0,075	-0,705	0,481	0,948	0,818	1,099	
X12j	Região BR	Sudeste	0,288	0,051	5,630	< 0,001 ***	1,334	1,206	1,474	
X1jX2j	Sexo*Cond_do- micilio	Masculino e Depen- dente	-2,113	0,196	-10,759	< 0,001 ***	0,121	0,082	0,178	
X1jX3j	Sexo*Cond_do- micilio	Masculino e Pessoa_ de_referencia	-0,851	0,228	-3,737	< 0,001 ***	0,427	0,273	0,667	
X1jX4j	Sexo*Cond_do- micilio	Masculino e Outros	-1,680	0,217	-7,753	< 0,001 ***	0,186	0,122	0,285	

Fonte: Análise baseada na PNAD 2014.

Categorias de referência (baseline): Sexo Feminino, Condição no Domicílio Cônjuge, Faixa de Idade 15-17 anos, Escolaridade sem instrução e Região Nordeste.

Função de ligação: Logit

Nota: Tamanho da amostra: n = 17.807

Tamanho da população expandida: N = 9.736.446

Tabela A.3: Modelo – Base PNAD de 2014 com Situação Censitária

Nota- ção	Preditor	Categoria	Coefi- cientes	Erro			Razão de chances	IC 0,95	
				Padrão estima- do	Valor t	Pr (> t)		Lim. Infe- rior	Lim. Supe- rior
	Intercepto	-	-4,681	0,169	-27,779	< 0,001 ***	0,009	0,007	0,013
X1j	Sexo	Masculino	2,365	0,195	12,133	< 0,001 ***	10,641	7,263	15,592
X2j	Cond_domi- cilio	Dependente	1,088	0,062	17,688	< 0,001 ***	2,969	2,632	3,350
X3j	Cond_domi- cilio	Pessoa_de_referencia	0,573	0,086	6,633	< 0,001 ***	1,774	1,497	2,101
X4j	Cond_domi- cilio	Outros	0,748	0,088	8,530	< 0,001 ***	2,113	1,779	2,509
X5j	Escolaridade	Fundam. Incompleto	1,305	0,129	10,128	< 0,001 ***	3,689	2,866	4,750
X6j	Escolaridade	Fundam. Completo ou mais	1,688	0,124	13,568	< 0,001 ***	5,410	4,239	6,904
X7j	Idade_Faixa	18-24	0,764	0,078	9,852	< 0,001 ***	2,147	1,844	2,500
X8j	Idade_Faixa	25-29	0,730	0,082	8,887	< 0,001 ***	2,076	1,767	2,438
X9j	Região BR	Norte	-0,182	0,060	-3,040	0,002 **	0,834	0,741	0,937
X10j	Região BR	Centro-Oeste	-0,171	0,076	-2,264	0,024 *	0,843	0,727	0,977
X11j	Região BR	Sul	-0,151	0,076	-1,996	0,046 *	0,860	0,741	0,997
X12j	Região BR	Sudeste	0,168	0,051	3,276	0,0011 **	1,183	1,070	1,309
X13j	Situação Cen- sitária	Urbano	0,790	0,069	11,376	< 0,001 ***	2,203	1,922	2,524
X1jX2j	Sexo*Cond_ domicilio	Masculino e Depen- dente	-2,068	0,201	-10,298	< 0,001 ***	0,126	0,085	0,187
X1jX3j	Sexo*Cond_ domicilio	Masculino e Pessoa_ de_referencia	-0,765	0,232	-3,296	< 0,001 ***	0,465	0,295	0,733
X1jX4j	Sexo*Cond_ domicilio	Masculino e Outros	-1,655	0,221	-7,475	< 0,001 ***	0,191	0,124	0,295

Fonte: Análise baseada na PNAD 2014.

Categorias de referência (baseline): Sexo Feminino, Condição no Domicílio Cônjuge, Faixa de Idade 15-17 anos, Escolaridade sem instrução, Região Nordeste e Situação Censitária Rural.

Função de ligação: Logit

Nota: Tamanho da amostra: n = 17.807

Tamanho da população expandida: N = 9.736.446

Seção 5

Outras análises acerca de População, Território e Estatísticas Públicas

Fatores de risco associados à prática de maus-tratos contra pessoas idosas no município de São Paulo

Ana Carolina Soares Bertho*
Bruna Matter dos Santos**
Marília Miranda Forte Gomes***
Jair Lício Ferreira dos Santos****
Yeda Aparecida de Oliveira Duarte*****

Resumo

O objetivo da pesquisa é analisar as características associadas à ocorrência de maus-tratos contra pessoas no município de São Paulo em 2015. A fonte de dados utilizada foi a Pesquisa Saúde, Bem-Estar e Envelhecimento (SABE) de 2015, que entrevistou 1.224 pessoas com 60 anos ou mais residentes em São Paulo. Foi elaborado um modelo de regressão logística tendo como variável resposta a ocorrência de qualquer tipo de maus-tratos. O estudo evidenciou que pessoas com idade entre 60 e 69 anos, viúvas, sem plano de saúde, que receberam críticas daqueles que os cercavam, que estavam insatisfeitas com a afetividade que recebiam, com doenças cardiovasculares, com índice de bem-estar ruim e com depressão foram as que tiveram maiores chances de sofrerem maus-tratos. Na comparação com a SABE 2010, verificou-se um aumento de 10% para 12,6% na prevalência de maus-tratos contra pessoas idosas. O estudo contribui para a compreensão de características e fatores associados à ocorrência de maus-tratos contra pessoas com 60 anos ou mais de idade. Conhecer os aspectos que tornam determinados grupos populacionais mais vulneráveis é importante para a elaboração de políticas públicas eficientes.

Palavras-chave: pessoas idosas; maus-tratos aos idosos; vulnerabilidade em saúde.

* Professora do programa de pós-graduação da ENCE.

** Mestre pelo programa de pós-graduação da ENCE.

*** Universidade de Brasília (UnB), Brasil.

**** Universidade de São Paulo (USP), Brasil.

***** Universidade de São Paulo (USP), Brasil.

Introdução

Desde o início do século XXI o mundo está preocupado com o acelerado processo de envelhecimento populacional, que já se consolidou como um fenômeno em escala global, pois tem sido cada vez mais presente em diversos países, independentemente do estágio de desenvolvimento (CAMARANO; PASINATO, 2004). Caracterizado pela mudança no perfil etário da população, tal fenômeno tem ganhado cada vez mais atenção, pelas demandas por políticas públicas e transferências de recursos (CAMARANO; KANSO; FERNANDES, 2014), especialmente em países como o Brasil, onde o crescimento da população idosa tem ocorrido de maneira ainda mais acelerada.

No passado, a fecundidade, ou seja, o número médio de filhos por mulher, era alta. As mulheres tinham muitos filhos e a mortalidade em todas as idades era elevada, de forma que nem todos chegavam com vida à idade adulta. Aos poucos, a mortalidade foi caindo e o número de nascimentos por mulher diminuiu progressivamente nas últimas décadas. Com a entrada de menos crianças na população, aumenta a importância dos grupos populacionais mais velhos (adultos e pessoas idosas), o que torna a população mais envelhecida. Além disso, a queda das taxas de fecundidade associada à redução da mortalidade e à elevação da expectativa de vida faz com que o processo de envelhecimento se acentue ainda mais (KALACHE, 1987).

Os censos demográficos têm mostrado que, no Brasil, a proporção de pessoas com 60 anos ou mais passou de 6,1% em 1980, para 10,8% em 2010, chegando a 15,8% em 2022 (IBGE, 2023).

Um dos desafios que tem se mostrado presente quando se analisa o aumento da população idosa é a questão dos maus-tratos contra este grupo. Autores como Minayo e Souza (2005) apontam que ainda não há dados suficientes para demonstrar que, com o aumento da participação das pessoas idosas na população, aumentarão também os casos de violência contra esse grupo populacional. Apesar disso, Frazão et al. (2014) afirmam que tais eventos têm sido cada vez mais reconhecidos e registrados.

Para a Organização Mundial da Saúde (OMS), maus-tratos são conceituados como ações únicas ou repetidas, ou ainda, a ausência de ação que cause dano, sofrimento ou angústia e que ocorra dentro de um relacionamento de confiança (WHO, 2002). Segundo Minayo (2003):

“Violência” é um conceito referente aos processos, às relações sociais interpessoais, de grupos, de classes, de gênero, ou objetivadas em instituições, quando empregam diferentes formas, métodos e meios de aniquilamento de outrem, ou de sua coação direta ou indireta, causando-lhes danos físicos, mentais e morais. As violências contra idosos, também, frequentemente, são denominadas maus tratos e abusos, mas vou me omitir de fazer uma avaliação sobre as últimas duas noções, utilizando-as como sinônimo de violência. (MINAYO, 2003, p. 785)

Assim, sem desconsiderar que existem diferenças teóricas e históricas entre os termos, este estudo se alinha ao posicionamento de Minayo (2003), que considera a utilização de “maus-tratos” e “abusos” como sinônimos de violência. Segundo Pasinato, Camarano e Machado (2006), a violência se caracteriza como um problema social que parece fazer parte da história humana. É possível dizer até que a violência é como uma “doença social”, pois afeta toda a sociedade e pode tomar grandes proporções (BRASIL, 2006).

De acordo com Leite, Hildebrandt e Santos (2008), pessoas com 60 anos ou mais naturalmente possuem mais riscos à vitimização por maus-tratos em decorrência de sua fragilidade e dependência, principalmente pelas limitações funcionais. Questões culturais também podem ter influência na prática da violência contra as pessoas idosas, em particular, no ambiente doméstico, pois é comum a pessoa idosa ser desvalorizada, além de, muitas vezes, ser vista e tratada como um peso e/ou problema para a família.

Pasinato, Camarano e Machado (2006) apontam que os maus-tratos contra as pessoas idosas ocorrem, em geral, onde elas se encontram, ou seja, em suas casas, na comunidade e em instituições de longa permanência. As principais tipologias definidas na literatura para maus-tratos contra a pessoa idosa são: maus-tratos físicos, maus-tratos psicológicos, abuso financeiro

ou material, abuso sexual, negligência, abandono, autonegligência e negligência social difusa (SOUSA *et al.*, 2010).

Sendo assim, a violência contra a pessoa idosa se manifesta de diversas formas, ambientes e por distintos personagens (agressores). Além do próprio ambiente doméstico, dos meios de transporte públicos, instituições de longa permanência, políticas sociais também podem contribuir com um quadro de perpetuação de desigualdades socioeconômicas (PASINATO; CAMARANO; MACHADO, 2006), em que a pessoa idosa pode estar exposta a situações de violência. Até mesmo entre pessoas idosas hospitalizadas o risco de violência existe, afetando especialmente aquelas maior idade, com estresse percebido e com depressão leve a severa, como mostra o estudo transversal realizado por Antequera *et al.* (2021).

No Brasil, as famílias são designadas como a principal instituição responsável em atender às necessidades das pessoas idosas, que são amparadas pela Constituição Federal, Política Nacional do Idoso e o Estatuto do Idoso que “contribuem para prevalecer a responsabilidade do núcleo familiar sobre a proteção e o sustento de seus idosos como principal responsável pelo bem-estar do idoso” (OLIVEIRA *et al.*, 2012).

Entretanto, a violência intrafamiliar, vista como a mais frequente e mais complexa por envolver relações afetivas, ocorre no “silêncio” dos lares, o que a torna invisível à sociedade (BRASIL, 2006). Dessa forma, uma das grandes dificuldades em lidar com este tema consiste no registro desta forma de violência, pois além de envolver relações familiares (sejam elas consanguíneas ou não), muitas vezes o agressor é instintivamente protegido e defendido pela pessoa idosa que sofre agressão. É o que mostram Morilla e Manso (2021) em uma revisão integrativa englobando o período de 2010 a 2020. As autoras identificaram que, entre os agressores, se destacam os filhos, pessoas de 40 a 49 anos, que fazem uso de álcool e/ou drogas e que residem com as vítimas de violência. Com relação ao tipo de violência mais comum, de acordo com essas autoras, se sobressaem a violência psicológica e a negligência. Em estudo sobre o município de Campinas de 2009 a 2019, Lopes e Elboux (2021) também destacam que quase 93% dos casos de violência aconteceram na própria residência das vítimas. Não por acaso, a maior parte dos casos não chega aos serviços de saúde, pois ficam ocultos, seja dentro dos lares, seja pelos cuidadores da pessoa idosa ou nas instituições de longa permanência.

No Brasil, ainda que haja informações referentes às internações e aos óbitos de pessoas idosas por causas externas (incluindo quedas e acidentes), nem sempre é fácil relacionar tais eventos aos maus-tratos sofridos pelas pessoas idosas. E mesmo quando há violência física, nem sempre ela chega a gerar uma internação hospitalar, o que torna o evento invisível. E outras expressões de violência, tais como a psicológica e a financeira, não deixam marcas físicas que permitam a identificação do ponto de vista do atendimento médico.

Os maus-tratos contra a pessoa idosa são considerados problema de saúde pública, pois podem trazer diversos prejuízos às vítimas, com sequelas irreversíveis à saúde, o que onera gastos em saúde pública e prejudica o desenvolvimento econômico e social do país.

Diversos estudos mostram que existe um conjunto de características que podem aumentar ou diminuir o risco de uma pessoa idosa ser vítima de maus-tratos. Em geral, essas características podem ser agrupadas em grupos de fatores (sociodemográficos, econômicos e de saúde), por exemplo.

Cabe ressaltar que o risco de as pessoas idosas sofrerem maus-tratos foi potencializado em razão da pandemia de COVID-19. Santos *et al.* (2021) apontam, que já no primeiro ano da pandemia, diversos estudos de cunho teórico evidenciavam não apenas o agravamento do risco de violência, como também estratégias para mitigação. Alves *et al.* (2020) ressaltam que, muitas vezes, os idosos estavam isolados no mesmo ambiente dos seus agressores e distantes de outras pessoas (como cuidadores) que pudessem identificar os maus-tratos). Moraes *et al.* (2020) também reconheceram o agravamento do risco em razão do distanciamento social e elaboraram uma série de iniciativas de promoção à saúde e prevenção à violência contra idosos, em diversos níveis: macroestrutural (como redução da desigualdade social e da violência estrutural); comunitário (tais como atendimento 24 horas dos equipamentos de proteção aos idosos e manutenção

das consultas periódicas via videoconferência); relacional (buscando formas pacíficas para resolução de conflitos, por exemplo); e individual (como manter uma alimentação saudável, o contato com familiares e amigos por telefone e manter o uso de medicações de rotina). Kourti *et al.* (2023) apontam que a violência doméstica foi uma questão considerável imposta pela pandemia em todo o mundo. Se, por um lado, o confinamento das pessoas em seus domicílios aumentou o contato entre perpetradores e vítimas, resultando em aumento da violência, por outro, levou a uma queda no número de relatos (KOURTI *et al.*, 2023).

Ainda que o recorte temporal do presente artigo seja anterior ao da pandemia (pois a edição mais recente da Pesquisa SABE que abordou o tema da violência contra os idosos foi a de 2015), entende-se que o conhecimento dos fatores associados a esse fenômeno é essencial, uma vez que o planejamento de ações e programas eficazes de prevenção à violência se baseia na compreensão de tais fatores (PILLEMER *et al.*, 2016; PASINATO; CAMARANO; MACHADO, 2006). A pandemia agravou a ocorrência de maus-tratos em razão principalmente do distanciamento social e do conseqüente afastamento das redes de apoio, mas não alteraria o perfil dos idosos que sofreram violência. Nesse sentido, o objetivo geral desta pesquisa é analisar as características associadas à ocorrência de maus-tratos contra pessoas idosas no município de São Paulo em 2015.

Fonte de dados e métodos

Foram utilizados os dados da Pesquisa Saúde, Bem-Estar e Envelhecimento (SABE) de 2015. Essa Pesquisa foi iniciada em 2000, com o apoio da Organização Pan-Americana de Saúde (OPAS) e outras instituições conveniadas. A SABE se caracteriza como um estudo transversal de base populacional realizado em sete cidades da América Latina e Caribe, com o objetivo de conhecer aspectos relacionados às condições de vida e saúde das pessoas idosas não institucionalizadas nessas regiões, de forma que os resultados pudessem ser comparados entre as cidades (LEBRÃO; LAURENTI, 2005).

No Brasil, o Estudo SABE vem sendo desenvolvido com pessoas idosas residentes em áreas urbanas do município de São Paulo. A coleta de dados ocorreu em quatro momentos distintos: 2000, 2006, 2010 e 2015. Dessa forma, a pesquisa conta com quatro coortes de pessoas idosas. Além disso, sempre é adicionada uma nova coorte de pessoas entre 60 e 64 anos a fim de avaliar o padrão de envelhecimento entre as diferentes gerações de pessoas idosas. Os dados foram obtidos por intermédio de entrevistas domiciliares, cujo tamanho da amostra foi determinado a partir de delineamento amostral do tipo complexo (SILVA, 2003; LEBRÃO; DUARTE, 2008). Em 2015 foram entrevistadas 1.224 pessoas idosas.

Para análise estatística dos dados, em primeiro lugar verificou-se a relação entre cada variável explicativa selecionada e a variável resposta, por meio de modelos de regressão logística (simples), cuja aplicação é adequada quando a variável dependente é do tipo binária, onde há uma categoria de sucesso e outra de fracasso. Neste estudo, a categoria “sucesso” se refere às pessoas idosas que sofreram algum tipo de maus-tratos, isto é, que responderam “sim” para quaisquer das seguintes perguntas da seção Maus-tratos do Questionário do Estudo SABE em 2015:

- No último ano, alguma das pessoas que o rodeiam gritaram com o(a) Sr(a) sem razão?
- No último ano, alguma das pessoas que o rodeiam o(a) chamaram por algum nome ou apelido que o(a) Sr(a) não goste?
- Alguma das pessoas que o rodeiam o(a) ameaçaram por não fazer o que eles queriam que o(a) Sr(a) fizesse?
- Alguma das pessoas que o rodeiam bateram ou esbofetearam o(a) Sr(a)?
- Alguma das pessoas que o rodeiam sacudiram ou chacoalharam o(a) Sr(a)?

- No último ano, alguma das pessoas que o rodeiam usaram ou mexeram no seu dinheiro sem a autorização do(a) Sr(a)?
- Alguma das pessoas que o rodeiam roubaram seu dinheiro ou algum pertence importante para o(a) Sr(a)?

Uma vez que o processo de amostragem da Pesquisa SABE é do tipo complexo, o desenho amostral foi levado em consideração a fim de evitar viés nas estimativas obtidas. Para interpretação dos resultados, foram utilizadas as Razões de chance (RC) e identificadas como variáveis significativas, aquelas cujo p-valor foi menor que 0,05. O software para análise dos dados foi o STATA 14.0 que por meio do pacote *svy* conta com diversas funções que incorporam o plano amostral da pesquisa.

Tomando como ponto de partida as variáveis utilizadas em estudos anteriores relacionados à violência contra a pessoa idosa com dados da SABE, as variáveis selecionadas foram agrupadas em cinco blocos: variáveis sociodemográficas; rede de apoio familiar e social; econômicas; estilo de vida; e condições de saúde. Foi analisada a qualidade da informação dessas variáveis, a fim de identificar os níveis de *missing* de cada uma delas e aquelas cujos níveis foram superiores a 8% não foram consideradas, pois poderiam prejudicar a estimação dos coeficientes do modelo e as razões de chance levando em consideração o tamanho da amostra (GOMES, 2011). Quando o número de casos das categorias das variáveis foi muito pequeno, optou-se pelo agrupamento das categorias para uma melhor análise dos dados.

O método aplicado para estimação dos parâmetros do modelo foi de Pseudo Máxima Verossimilhança (MPV) que incorpora os pesos e o plano amostral para obtenção das estimativas dos parâmetros e as respectivas variâncias.

Em um primeiro momento, foram realizadas análises univariadas, sendo consideradas elegíveis para os modelos multivariados as variáveis com p-valor inferiores a 0,20. Nas modelagens multivariadas, as variáveis explicativas foram inseridas em blocos, começando pelos fatores sociodemográficos, seguidas pelas variáveis de rede de apoio familiar e social, fatores econômicos, estilo de vida e condições de saúde. A cada inserção, foram mantidas no modelo apenas variáveis cujo p-valor foi inferior a 0,05 dos blocos anteriores.

Após o cálculo das estimativas dos parâmetros foi realizada a inferência e analisada a qualidade do modelo a partir da aplicação de teste de bondade (ARCHER; LEMESHOW, 2006). No STATA, foi utilizado o comando *svylogitgof*, sendo mantidas somente as variáveis que contribuísem para um melhor ajuste e estimação dos coeficientes (BATTISTI, 2008). Foram calculadas as Razões de Chance (RC) a fim de estimar quantas vezes é mais provável ocorrer o evento de interesse (a pessoa idosa ter sido vítima de maus-tratos).

O Estudo Sabe foi aprovado pelo Comitê de Ética em Pesquisa da Faculdade de Saúde Pública da Universidade de São Paulo. Todos os participantes assinaram o Termo de Consentimento Livre e Esclarecido.

Resultados

A maior parte das pessoas entrevistadas eram mulheres (65,0%), se declararam brancas (53,7%), com idade entre 60 e 69 anos (55,2%) e estavam em uma união (49,9%). A variável referente à escolaridade apresentou um nível elevado de *missing* (10,9%), sendo excluída das demais análises.

Quanto à rede de apoio, 82,0% dos entrevistados moravam com outras pessoas. Os filhos foram os mais citados (43,1%) entre os que ajudaram quando a pessoa idosa precisou. Além disso, 63,6% afirmaram que as pessoas que as cercavam nunca faziam muitas exigências, cobranças ou críticas e 78,7% disseram nunca ter se sentido isoladas ou deixadas de lado. O Apgar de família mensura a funcionalidade familiar e é utilizado para avaliar a dinâmica dos relacionamentos

familiares no cuidado com uma pessoa dependente (SMILKSTEIN, 1978). Tal indicador mostrou que 84,8% dos entrevistados estavam satisfeitos com o apoio demonstrado pelos familiares.

A variável referente à receita mensal apresentou 11,4% de missing, sendo excluída das demais análises. Outras características foram usadas como proxy da situação econômica. Verificou-se que 53,8% dos entrevistados não possuíam plano de saúde. Quanto às condições de moradia e a situação de trabalho, 79,8% tinham residência própria e 68,6% não trabalhavam. Para 50,9% a renda que recebiam era suficiente para cobrir as despesas diárias.

Quanto ao estilo de vida, 70,1% relataram não ter feito uso de bebida alcoólica nos três últimos meses, 87,0% disseram nunca ter fumado ou somente haviam fumado no passado. Em relação às atividades físicas, 67,2% revelaram não ter feito nenhum exercício nos últimos três meses. A qualidade de vida da população idosa foi considerada boa, pois 84,0% obtiveram um bom resultado no Índice de Bem-Estar.

Com relação às condições de saúde, 48,4% das pessoas idosas autoavaliaram seu estado de saúde como muito bom ou bom e 82,8% tiveram um bom resultado nos testes de avaliação cognitiva. Do total de entrevistados, 92,3% se consideraram bem nutridos e 69,5% afirmaram não ter sofrido quedas no ano anterior à Pesquisa.

A maioria das pessoas idosas (96,6%) afirmaram ser capazes de andar por conta própria e 67,9% relataram não ter quaisquer dificuldades funcionais. Quanto às atividades instrumentais, 77,5% afirmaram ter pelo menos uma dificuldade. Além disso, 41,7% disseram necessitar de um cuidador para auxiliá-los em suas atividades diárias e 16,8% sofriam de depressão.

Quanto à presença de Doenças Crônicas Não Transmissíveis (DCNTs), 51,5% possuíam de uma a duas DCNTs, sendo as mais referidas hipertensão arterial (66,3%), doenças articulares (35,6%) e diabetes (28,6%).

A partir dessa análise descritiva geral da amostra, as variáveis de escolaridade e receita mensal, que apresentaram elevada proporção de missing, foram excluídas. No total, 20 variáveis foram selecionadas para a modelagem multivariada: sociodemográficas (sexo, idade e estado civil); rede de apoio familiar e social (auxílio, críticas, isolamento, Apgar de família); econômicas (plano de saúde e suficiência financeira); estilo de vida (índice de bem-estar); condições de saúde (auto-avaliação de saúde, estado nutricional, quedas nos últimos 12 meses, número de cuidadores, depressão, quantidade de DCNTs, diabetes, doença articular, doença do coração e doença psiquiátrica).

Os coeficientes estimados (Tabela 1) mostram que, com exceção da idade, os efeitos das variáveis significativas foram positivos, indicando que as chances de uma pessoa idosa ter sofrido maus-tratos foram maiores do que as chances das pessoas que estavam nas categorias de referência.

Tabela 1 - Coeficientes estimados para a prática de maus-tratos contra pessoas idosas, município de São Paulo, 2015.

Variáveis explicativas	Maus-tratos		
	Coeficiente	Erro padrão	p-valor
Intercepto	-3,070	0,012	0,000*
Idade			
60 a 69 anos		Referência	
70 a 79 anos	-0,471	0,170	0,086
80 anos ou mais	-1,110	0,135	0,008*
Situação conjugal			
Solteiro	0,762	0,897	0,070
Unido		Referência	
Separado	0,219	0,442	0,537
Viúvo	0,764	0,536	0,003*
Críticas			
Nunca		Referência	
Alguma vez	0,858	0,536	0,000*
Afetividade			
Nunca ou raramente	1,320	1,257	0,000*
Alguma vez		Referência	
Plano de saúde			
Sim		Referência	
Não	0,579	0,348	0,003*
Índice de Bem-Estar			
Bom		Referência	
Ruim	0,027	0,308	0,928
Depressão			
Sim	-0,025	0,291	0,933
Não		Referência	
Doença do coração			
Sim	0,478	0,370	0,039*
Não		Referência	
Índice de Bem-Estar e Depressão			
Índice Bom e Idoso sem depressão		Referência	
Índice Ruim e Idoso com depressão	1,177	1,456	0,008*

Fonte: Estudo SABE (2015).

Nota: * p-valor < 0,05.

Na Tabela 2 são apresentadas as razões de chance do modelo. As pessoas com 80 anos ou mais possuíam menos chances de sofrer maus-tratos, quando comparadas àquelas com 60 anos ou mais. Além disso, os viúvos possuíam 1,1 vez mais chances para violência do que os unidos. Possuir plano de saúde também é um fator de proteção, pois pessoas sem plano tiveram 78,5% mais chance de ter sofrido violência.

Aqueles que alguma vez receberam críticas tiveram 1,3 vez mais chance de ter sofrido maus-tratos. Os que nunca ou raramente se sentiram satisfeitos quanto à forma como a família reagia às emoções deles apresentaram 2,7 vezes mais chance do que aqueles que alguma vez se sentiram satisfeitos. As pessoas com depressão e índice de bem-estar classificado como ruim

tiveram 2,2 vezes mais chances para maus-tratos, quando comparadas àquelas sem depressão e com bem-estar classificado como bom. Além disso, as chances para maus-tratos foram 61,3% maiores em pessoas idosas com doenças cardiovasculares.

Tabela 2 - Modelo multivariado dos fatores de risco para maus-tratos contra pessoas idosas, município de São Paulo, 2015.

Variáveis explicativas	Maus-tratos		
	Razão de chances	IC (95%)	p-valor
Intercepto	0,046	0,02 - 0,07	0,000
Idade			
60 a 69 anos	1,000	-	-
70 a 79 anos	0,624	0,36 - 1,06	0,086
80 anos ou mais	0,330	0,14 - 0,74	0,008*
Situação conjugal			
Solteiro	2,144	0,94 - 4,88	0,070
Unido	1,000	-	-
Separado	1,245	0,63 - 2,49	0,537
Viúvo	2,146	1,33 - 3,54	0,003*
Críticas			
Nunca	1,000	-	-
Alguma vez	2,357	1,50 - 3,69	0,000*
Afetividade			
Nunca ou raramente	3,743	1,92 - 7,26	0,000*
Alguma vez	1,000	-	-
Plano de saúde			
Sim	1,000	-	-
Não	1,785	1,21 - 2,62	0,003*
Índice de Bem-Estar			
Bom	1,000	-	-
Ruim	1,027	0,56 - 1,85	0,928
Depressão			
Sim	0,975	0,54 - 1,76	0,933
Não	1,000	-	-
Doença do coração			
Sim	1,613	1,02 - 2,53	0,039*
Não	1,000	-	-
Índice de Bem-Estar e Depressão			
Índice Bom e Idoso sem depressão	1,000	-	-
Índice Ruim e Idoso com depressão	3,245	1,36 - 7,88	0,008*

Fonte: Estudo SABE (2015).

Nota: * p-valor < 0,05.

Em 2015 houve um pequeno aumento na prevalência de maus-tratos contra pessoas idosas quando comparada à SABE em 2010, passando de 10% para 12,6% na pesquisa mais recente.

Discussão

A prevalência de maus-tratos contra pessoas idosas obtida neste estudo (12,6%) mostra que houve um pequeno aumento quando comparada aos resultados obtidos por Machado (2014) com dados de 2010 do Estudo SABE, que identificou que 10,1% das pessoas entrevistadas tinham sofrido violência.

Diversos estudos apontam maior incidência de maus-tratos em mulheres idosas (DUQUE *et al.*, 2012; BRASIL, 2014; MINAYO, 2003; LOPES *et al.*, 2018; ARAÚJO *et al.*, 2007; CUNHA *et al.*, 2021). Nesta pesquisa, a variável sexo não foi significativa, o que também foi observado em uma pesquisa realizada em Niterói (APRATTO JÚNIOR, 2010). Quanto à idade, pessoas entre 60 e 69 anos apresentaram maior risco, resultado já destacado na literatura (MINAYO, 2003). Aguiar *et al.* (2015) também identificaram maior risco para violência em pessoas idosas mais “jovens” em Aracaju (SE) e afirmaram que essa situação pode estar associada ao fato de que as pessoas nessa faixa etária, em geral, possuem mais autonomia funcional, conhecimento em relação aos seus direitos e à possibilidade de buscá-los, o que colabora para um aumento do número de denúncias. As autoras destacam ainda que as pessoas com idade avançada poderiam ter mais dificuldades em fazer denúncias, em virtude das limitações físicas e de um estado de saúde mais debilitado que as manteriam restritas ao ambiente doméstico. Para esses autores, a maior vitimização de pessoas idosas “mais jovens” pode estar relacionada à maior autonomia e compreensão sobre os próprios direitos, o que aumentaria o número de denúncias. Entretanto, tal resultado diverge de estudos que apontaram maior risco de sofrer violência para pessoas acima de 70, de 75 ou de 80 anos (DINIZ; SANTO; RIBEIRO, 2021; SOUSA *et al.*, 2010; BRASIL, 2014; SANCHES, 2006).

Em relação à situação conjugal, as chances de a pessoa ter sofrido violência foram menores quando ela era casada ou possuía companheiro (a), o que é recorrente na literatura (BOLSONI *et al.*, 2016; PINTO; BARHAM; ALBUQUERQUE, 2013). Estudo feito por Sanches (2006) com pessoas idosas que participaram do Estudo SABE em 2000 mostra que tanto homens quanto mulheres idosas, em que a situação conjugal era solteiro, divorciado ou viúvo, estavam mais propensos a sofrer maus-tratos, o que corrobora com os achados deste estudo que indicaram que pessoas idosas solteiras e viúvas apresentaram mais chances de ter sido vítimas de violência, sendo as chances 156% e 82% maiores, respectivamente, quando comparadas às que tinham companheiro (a). Dessa forma, o estado conjugal unido (casado ou com companheiro) indica ser um fator de proteção contra abusos em pessoas idosas.

Quanto às condições financeiras, as pessoas idosas que não tinham plano de saúde apresentaram risco 78% maior para violência, o que corrobora com estudos que apontam que pessoas com 60 anos ou mais com melhores condições econômicas teriam mais condições de se cuidar (BRASIL, 2014). A associação entre pobreza e maior vitimização de pessoas idosas também já foi apontada no “*World report on violence and health*” (WOLF; DAICHMANN; BENNETT, 2002). Sanches (2006) identificou que entre as pessoas idosas cuja renda era suficiente para cobrir as despesas, 5,6% tinha sinais de ter sofrido violência, enquanto entre aquelas cuja renda era insuficiente, a prevalência foi de 12,1%. Segundo Araújo *et al.* (2007), quanto maior for a precariedade social, maior é o risco de a pessoa idosa sofrer alguma forma de abuso. Dessa forma, ambos os fatores identificados são indicativos de que pessoas idosas em condições socioeconômicas piores estão em situação de maior vulnerabilidade à violência.

Pessoas idosas com doenças do coração apresentaram 61,3% mais chance de ter sofrido violência. Nas três primeiras ondas do Estudo SABE verificou-se um aumento na prevalência das doenças cardiovasculares, o que traz preocupações em relação ao futuro (MASSA, 2018). As condições de saúde são apontadas como fator de risco para maus-tratos, pois aumentam a dependência, o que poderia elevar os níveis de estresse dos cuidadores e familiares, ocasionando ações de violência (SANCHES, 2006; BOLSONI *et al.*, 2016; PILLEMER *et al.*, 2016; FRAZÃO *et al.*, 2014).

A interação entre as variáveis índice de bem-estar e depressão, relacionada à violência, aponta a importância tanto da qualidade de vida, como da saúde mental da pessoa idosa. A relação entre violência e depressão já havia sido identificada em rodadas anteriores da SABE (MACHADO, 2014; SANCHES, 2006). O índice de bem-estar mede a qualidade de vida das pessoas idosas, abrangendo aspectos como níveis de alegria, disposição, tranquilidade, energia, entre outros. Uma avaliação negativa dele pode ter associação com maus-tratos psicológicos, uma vez que esse tipo de abuso muitas vezes ocorre de maneira sutil, por meio de ameaças, coações, deterioração da imagem da pessoa idosa, apelidos negativos, isolamento, restrição da liberdade e comumente coloca a pessoa idosa em situações de humilhação e medo (BRASIL, 2014; SOUSA *et al.*, 2010; BOLSONI *et al.*, 2016).

Antunes *et al.* (2018) em estudo sobre desigualdades sociais quanto à autoavaliação do estado de saúde de pessoas idosas participantes do Estudo SABE em 2010, verificaram que aquelas menos escolarizadas e com menor renda foram as que tiveram as piores percepções quanto ao estado de saúde, evidenciando que os estratos socioeconômicos mais carentes estão em maior situação de vulnerabilidade à saúde. A maior vulnerabilidade de mulheres de baixa renda e baixa escolaridade também foi um dos achados de Morilla e Manso (2021), ao analisarem as publicações sobre violência contra idosos de 2010 a 2020. Tais resultados corroboram com os achados deste estudo que mostram que uma pior percepção do estado de saúde também torna essas pessoas mais vulneráveis à violência, assim como aquelas que estão em condições socioeconômicas piores, conforme indicado pelos fatores econômicos associados.

Estudo feito por Dyer *et al.* (2000) mostra que tanto a depressão como a demência são doenças mais comuns em pacientes idosos que foram negligenciados. O estudo se caracteriza como o primeiro que utilizou dados primários e concluiu que a prevalência de depressão é maior em casos de pessoas idosas que sofreram negligência ou outras formas de abuso. Os autores destacam que a depressão é uma condição médica reversível associada a abusos ou negligências e pode ser uma das principais razões pelas quais pacientes negligenciados, especialmente aqueles que sofrem de autonegligência, recusam tratamento médico ou assistência em casa.

Uma dificuldade está em não saber se o abuso antecedeu ou ocorreu após o diagnóstico de depressão, o que também é válido para outros fatores associados, como por exemplo, a autoavaliação da saúde. Park (2019), ao analisar a relação entre tipos de maus-tratos contra pessoas idosas e depressão na Coreia do Sul aponta evidências de que a depressão é uma consequência dos abusos, e contribui na diminuição da qualidade de vida das pessoas idosas, podendo levá-los até mesmo a pensamentos suicidas.

Como agravante, de acordo com Kosberg (1988 *apud* PAIXÃO JÚNIOR; REICHNHEIN, 2006), a subnotificação de denúncias de maus-tratos sofridos pelas pessoas idosas, muitas vezes, está associada a um quadro de solidão, pois a pessoa com 60 anos ou mais que está isolada e sozinha, terá dificuldades em relatar suas adversidades ou buscar ajuda. A dificuldade da pessoa idosa em relatar que sofreu maus-tratos é um fator que pede por mais atenção dos profissionais de saúde, pois ações como ansiedade na presença do cuidador ou familiar, lesões no corpo, desidratação ou ainda a não aceitação em conversar sobre o assunto podem ser sinais que a pessoa idosa sofre violência (BRASIL, 2006).

Conclusão

O estudo identificou os fatores associados à ocorrência de maus-tratos contra pessoas idosas em São Paulo em 2015. Dadas as características do Estudo SABE, não seria possível afirmar que o evento (ou eventos) violento ocorreu antes ou depois dos fatores que foram identificados. O modelo mostrou que as pessoas que apresentaram maiores chances de sofrerem maus-tratos tinham idade entre 60 e 69 anos, viúvas, sem plano de saúde, que receberam críticas daqueles que os cercavam, que estavam insatisfeitas com a afetividade que recebiam, com doenças cardiovasculares, com índice de bem-estar ruim e com depressão. Os resultados corroboraram com a

literatura, ainda que alguns aspectos não sejam unanimidade, como a faixa etária mais afetada (alguns estudos apontam idosos mais jovens; outros, apontam que a violência seja frequente entre os de idade mais elevada).

Algumas limitações deste estudo concernem à seletividade da amostra, uma vez que somente pessoas idosas que estavam disponíveis e que não se encontravam num estado de saúde muito debilitado, responderam ao Estudo SABE.

Diante do cenário de crescimento da população idosa, pois as taxas de fecundidade permanecem em queda, é possível que algumas características associadas à ocorrência de maus-tratos mudem, pois com mais pessoas idosas na população, pode ser que ondas futuras do Estudo SABE contem com uma amostra maior de pessoas, inclusive as de 80 anos ou mais, uma vez que a expectativa de vida da população tem aumentado.

A identificação das características associadas à ocorrência de maus-tratos contra as pessoas idosas é apenas o primeiro passo para compreensão deste fenômeno que pode tomar grandes proporções. A violência é um problema social que pode gerar sequelas irreversíveis que atinge não somente as vítimas, mas a sociedade como um todo. Segundo Block (1995) *apud* Souza *et al.* (2004), a violência pode ser vista como “produto e produtora das condições de vida social”. Esse fenômeno faz parte de uma série de determinantes sociais que influencia e (é influenciada) pelas condições de saúde da população (BUSS; PELLEGRINI FILHO, 2007).

Os resultados obtidos revelam o quanto esse tema é complexo, uma vez que os fatores associados se encontram vinculados a diversas esferas da realidade, indicando também que o enfrentamento deste problema é uma tarefa igualmente difícil, o que torna fundamental a ampliação de investigações nesta área e a promoção de políticas públicas que atuem com o objetivo de melhorar a qualidade de vida e saúde das pessoas idosas e prevenir a ocorrência de maus-tratos.

Há um consenso na literatura de que a pandemia de COVID-19 agravou os quadros de maus-tratos contra idosos, deixando-os não só mais expostos, mas também dificultando o registro dos casos de violências. Nesse contexto, torna-se ainda mais importante a existência de estudos como este, que evidenciem quais são as características e fatores mais associadas à violência, permitindo a definição de melhores estratégias de proteção para essas pessoas.

Referências

- AGUIAR, M. P. C. de; LEITE, H. A.; DIAS, I. M.; MATTOS, M. C. T. de; LIMA, W. R. *Violência contra idosos: descrição de casos no Município de Aracaju, Sergipe, Brasil*. Esc. Anna Nery; 19(2):343-349, 2015.
- ALVES, R.M.; COSTA, V.C.G.S.F.; OLIVEIRA, T.M.; ARAÚJO, M.O.; ARAÚJO, M.P.D. *Violência contra a população idosa durante a pandemia da COVID-19*. Saúde Coletiva (Barueri), v. 10, n. 59, p. 4314-4325, 2020.
- ANTEQUERA, I. G.; LOPES, M.C.B.T.; BATISTA R.E.A.; CAMPANHARO, C.R.V.; COSTA, P.C.P.; OKUNO, M.F.P. *Rastreamento de violência contra pessoas idosas: associação com estresse percebido e sintomas depressivos em idosos hospitalizados*. Escola Anna Nery, v. 25, p. e20200167, 2020.
- ANTUNES, J.L.F.; CHIAVEGATTO FILHO, A.D.P.; DUARTE, Y.A.O.; LEBRÃO, M.L. *Desigualdades sociais na autoavaliação de saúde dos idosos da cidade de São Paulo*. Rev. Bras. Epidemiol., 21(SPPL2), 2018.
- APRATTO JÚNIOR, P. C. *A violência doméstica contra os idosos nas áreas de abrangência do Programa de Saúde da Família de Niterói (RJ, Brasil)*. Ciên. Saúde. Colet. 15(6):2983-2995, 2010.
- ARAÚJO, C. L. O.; MORITA, A. B. P. S.; SOUZA, L. P. de; FARIA, A. C. S.; MENDES, A.C.A.; OLIVEIRA, E. M. C.; SANTOS, L. H. N. dos; LEAL, L. M. R.; GONÇALVES, V. N. *Maus tratos em idosos*. Janus, Lorena, v.4, n.5, p. 73-86, 2007.
- ARCHER, K. J.; LEMESHOW, S. *Goodness-of-fit Test for a Logistic Regression Model Fitted using Survey Sample Data*. The Stata Journal, 6(1), 97-105, 2006.
- BATTISTI, I. D. E. *Análise de dados epidemiológicos incorporando planos amostrais complexos [tese]*. Universidade Federal do Rio Grande do Sul. Porto Alegre, 2008.
- BOLSONI, C. C.; COELHO, E. B. S.; GIEHL, M. W. C.; D'ORSI, E. *Prevalência de violência contra idosos e fatores associados, estudo de base populacional em Florianópolis, SC*. Rev. Bras. Geriatr. Gerontol., Rio de Janeiro, v. 19, n. 4, p. 671-682, 2016.
- BRASIL. Secretaria de Direitos Humanos da Presidência da República. *Manual de Enfrentamento à Violência contra a Pessoa Idosa. É possível prevenir. É necessário superar*. Brasília, DF, 2014.
- BRASIL. Ministério da Saúde. Secretaria de Atenção à Saúde. Departamento de Atenção Básica. *Envelhecimento e saúde da pessoa idosa*. Brasília: Ministério da Saúde, 2006.
- BUSS, P. M.; PELLEGRINI FILHO, A. *A saúde e seus determinantes sociais*. Physis, Rio de Janeiro, v. 17, n. 1, p. 77-93, 2007.
- CAMARANO, A. A.; KANSO, S.; FERNANDES, D. *A população brasileira e seus movimentos ao longo do século XX*. In: CAMARANO, A. A. Novo regime demográfico: uma nova relação entre população e desenvolvimento? Rio de Janeiro: Ipea, 2014.
- CAMARANO, A. A.; PASINATO, M. T. *O envelhecimento populacional na agenda das políticas públicas*. In: CAMARANO, A. A. (Org.). Os novos idosos brasileiros: muito além dos 60? IPEA, Rio de Janeiro, setembro de 2004.
- CUNHA, R.I.M.; OLIVEIRA, L.V.A.; LIMA, K.C.; MENDES, T.C.O. *Perfil epidemiológico das denúncias de violência contra a pessoa idosa no Rio Grande do Norte, Brasil (2018-2019)*. Rev. Bras. Geriatr. Gerontol. 24(6), 2021.
- DINIZ, C.X.; SANTO, F.H.E.; RIBEIRO, M.Z.S.. *Análise do risco direto e indireto de violência intrafamiliar contra pessoas idosas*. Rev. Bras. Geriatr. Gerontol. 24(6), 2021.
- DUQUE, A. M.; LEAL, M. C. C.; MARQUES, A. P. DE O., ESKINAZI, F. M. V.; DUQUE, A. M. *Violência contra idosos no ambiente doméstico: prevalência e fatores associados (Recife/PE)*. Ciênc. Saúde Coletiva, Rio de Janeiro, v. 17, n. 8, p. 2199-2208, 2012.

- DYER, C.B., PAVLIK, V.N., MURPHY, K.P.; HYMAN, D.J. *The High Prevalence of Depression and Dementia in Elder Abuse or Neglect*. Journal of the American Geriatrics Society, 48: 205-208. 2000.
- FRAZÃO, S. L.; SILVA, M. S.; NORTON, P.; MAGALHÃES, T. *Domestic violence against elderly with disability*. J Forensic Leg Med. 2014.
- GOMES, M. M. F. *Passado e presente: uma análise dos determinantes da mortalidade entre idosos com base nos dados da SABE 2000-2006*. Tese (Doutorado em Demografia) – Centro de Desenvolvimento e Planejamento Regional, Universidade Federal de Minas Gerais, Belo Horizonte, 2011.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). *Censo Demográfico 2022: população por idade e sexo : pessoas de 60 anos ou mais de idade : resultados do universo: Brasil, Grandes Regiões e Unidades da Federação*. Rio de Janeiro: IBGE, 2023. Disponível em: <https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=2102038> . Acesso em: 31 out. 2023.
- KALACHE, A. *Envelhecimento populacional no Brasil: uma realidade nova*. Cad. Saúde Pública, Rio de Janeiro, v. 3, n. 3, p. 217-220, 1987.
- KOURTI, A.; STAVRIDOU, A.; PANAGOULI, E.; PSALTOPOULOU, T.; SPILIOPOULOU, C.; TSOLIA, M.; SERGENTANIS, T.N.; TSITSIKA, A. *Domestic Violence During the COVID-19 Pandemic: A Systematic Review*. Trauma, Violence, & Abuse 24(2): 719-745, 2023.
- LEBRÃO, M. L.; DUARTE, Y. A. de O. *Desafios de um estudo longitudinal: O Projeto SABE*. Saúde Coletiva. 2008.
- LEBRÃO, M. L.; LAURENTI, R. *Saúde, bem-estar e envelhecimento: o estudo SABE no município de São Paulo*. Rev. Bras. Epidemiol. 8(2):127-41, 2005.
- LEITE, T. M.; HILDEBRANDT, L.; SANTOS, A. *Maus-tratos a idosos no domicílio: concepção de familiares*. Revista Brasileira de Geriatria e Gerontologia, 11, 2008.
- LOPES, E.D.S.; D' ELBOUX, M.J. *Violência contra a pessoa idosa no município de Campinas, São Paulo, nos últimos 11 anos: uma análise temporal*. Revista Brasileira de Geriatria e Gerontologia, v. 24, n. 6, p. e200320, 2021.
- LOPES, E. D. S.; FERREIRA, A. G.; PIRES, C. G.; MORAES, M. C. S. de. ELBOUX, M. J. D. *Maus-tratos a idosos no Brasil: uma revisão integrativa*. Rev. bras. geriatri. gerontol., Rio de Janeiro, v. 21, n. 5, p. 628-638, 2018.
- MACHADO, D. R. *Violência contra idosos e qualidade de vida relacionada à saúde: estudo populacional no município de São Paulo*. São Paulo: Escola de Enfermagem da Universidade de São Paulo, 2014.
- MASSA, K. *Análise longitudinal multinível da morbidade por doenças cardiovasculares em idosos do município de São Paulo [tese]*. São Paulo: Faculdade de Saúde Pública da USP; 2018.
- MINAYO, M. C. S. *Violência contra idosos: relevância para um velho problema*. Cad. Saúde Pública, Rio de Janeiro, v. 19, n. 3, p. 783-791, 2003.
- MINAYO, M. C. S.; SOUZA, E. R. *Violência contra idosos: é possível prevenir*. In: BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. *Impacto da violência na saúde dos brasileiros*. Brasília, 2005.
- MORAES, C.L.; MARQUES, E.S.; RIBEIRO, A.P.; SOUZA, E.R. *Violência contra idosos durante a pandemia de COVID-19 no Brasil: contribuições para seu enfrentamento*. Ciência & Saúde Coletiva, v. 25, p. 4177-4184, 2020.
- MORILLA, J.L.; MANSO, M.E.G. *A violência contra a mulher idosa no Brasil e os fatores relacionados ao tema: uma revisão integrativa*. VITALLE-Revista de Ciências da Saúde, v. 33, n. 2, p. 66-82, 2021.

- OLIVEIRA, M. L. C. de; GOMES, A. C. G.; AMARAL, C. P. M.; SANTOS, L. B. dos. *Características dos idosos vítimas de violência doméstica no Distrito Federal*. Rev. Bras. Geriatr. Gerontol., Rio de Janeiro, v. 15, n. 3, p. 555-566, 2012.
- PAIXÃO JÚNIOR, C. M.; REICHENHEIM, M. E. *Uma revisão sobre instrumentos de rastreamento de violência doméstica contra o idoso*. Cad. Saúde Pública, Rio de Janeiro, v. 22, n. 6, p. 1137-1149, 2006.
- PARK, E. O. *Most prevalent type of elder abuse and its correlation with elder depression*. Acta paul. enferm., São Paulo, v. 32, n. 1, p. 95-100, 2019.
- PASINATO, M. T.; CAMARANO, A. A.; MACHADO, L. *Idosos vítimas de maus-tratos domésticos: estudo exploratório das informações levantadas nos serviços de denúncia*. Texto para discussão nº 1200. Rio de Janeiro: IPEA; 2006.
- PILLEMER, K.; FINKELHOR, D.; BURNES, D.; RIFFIN, C.; LACHS, M. *Elder Abuse: Global Situation, Risk Factors, and Prevention Strategies*. The Gerontologist. 56. S194-S205. 2016.
- PINTO, F. N. F. R.; BARHAM, E. J.; ALBUQUERQUE, P. P. de. *Idosos vítimas de violência: fatores sociodemográficos e subsídios para futuras intervenções*. Estudos e Pesquisas em Psicologia., 13(3):1159-81, 2013.
- SANCHES, A. *Violência doméstica contra idosos no município de São Paulo: estudo SABE, 2000*. Dissertação (Mestrado em Epidemiologia) - Faculdade de Saúde Pública, Universidade de São Paulo, São Paulo, 2006.
- SANTOS, M. A. B. SÁ, G.G.; BRITO, A.A.; NOLÊTO, J.S.; OLIVEIRA, R.K. *Fatores associados à violência contra o idoso: uma revisão sistemática da literatura*. Ciência & Saúde Coletiva, v. 25, n. 6, p. 2153-2175, 2020.
- SILVA, N. N. Amostragem. In: LEBRÃO, M.L.; DUARTE, Y.A.O, organizadores. *O projeto SABE no município de São Paulo: uma abordagem inicial*. Brasília (DF): Organização Pan-Americana de Saúde; 2003.
- SMILKSTEIN, G. *The family APGAR: a proposal for a family function test and its use by physicians*. The Journal of Family Practice, v. 6, n. 6, p. 1231-1239, 1978.
- SOUSA, D. J.; WHITE, H. J.; SOARES, L. M.; NICOLOSI, G. T.; CINTRA, F. A.; D'ELBOUX, M. J. *Maus-tratos contra idosos: atualização dos estudos brasileiros*. Rev. Bras. Geriatr. Gerontol., Rio de Janeiro, v. 13, n. 2, p. 321-328, 2010.
- SOUZA, A. S.; MEIRA, E. C.; NERI, I. G.; SILVA, J. A.; GONÇALVES, L. H. T. *Fatores de risco de maus-tratos ao idoso na relação idoso/cuidador em convivência intrafamiliar*. Textos sobre Envelhecimento, v.7, n.2, Rio de Janeiro, 2004.
- WOLF, R.; DAICHMANN, L.; BENNETT, G. *Abuse of the elderly*. In: KRUG, E.G et al. (Eds.) *World report on violence and health*. Geneva: World Health Organization, p. 123-45, 2002.
- WHO. INPEA. *Missing voices: views of older persons on elder abuse*. Geneva, World Health Organization, 2002.

Cidades inteligentes no Brasil: sistema de mensuração com foco em bem-estar social

Renata Curi de Moura Estevão Nagatomi*

Daniel Takata Gomes**

Julia Celia Mercedes Strauch***

Resumo

Nos últimos anos, novos paradigmas socioeconômicos, digitais e em rede, advindos das tecnologias de informação e comunicação (TIC) aliadas ao rápido processo de urbanização global, vêm gerando impactos significativos na qualidade de vida e bem-estar das populações e, em muitos casos, de forma não planejada, devido às limitações de recursos e informações dos governos locais. Neste contexto surge a concepção de Smart Cities ou Cidades Inteligentes, reunindo planejamento e gestão com TIC, para tratar dos novos desafios enfrentados pelas urbes com uma abordagem holística. Este trabalho apresenta uma metodologia para construir um sistema de indicadores perpassando aspectos econômicos, de mobilidade, do ambiente, das pessoas, de vida e de governança para mensurar a aplicação desta distinção para as cidades brasileiras com mais de 100.000 habitantes em 2010. Para isso são consideradas as diretrizes internacionais da norma ISO 37120:2018, empregados dados oficiais disponíveis e utilizada a técnica estatística multivariada exploratória. Pela análise fatorial os 78 indicadores foram reduzidos a 12 fatores que representem estatisticamente um novo conjunto de indicadores sintéticos. Para comparar as cidades brasileiras, foi aplicado o método Ward obtendo-se cinco agrupamentos por similaridade de fator. Os resultados evidenciam que existem cidades que participam de ambos os extremos em fatores distintos e que a maioria das capitais do Norte e Nordeste se encontram no mesmo grupo.

Palavras-chave: Indicadores socioeconômicos e ambientais; Estatísticas Públicas; cidades médias; Smart Cities (Cidades Inteligentes).

* Doutora pelo programa de pós-graduação da ENCE.

** Professor do curso de graduação da ENCE.

*** Professora do programa de pós-graduação da ENCE.

Introdução

Atualmente há uma demanda econômico-social mundial, globalizada, pressionando por mudanças devido ao novo paradigma socioeconômico, digital e em rede, oriundo do advento da *web 2.0* e do uso de dispositivos eletrônicos cada vez mais poderosos. Isto decorre da evolução da comunicação via satélite, da expansão da rede por fibra óptica e a comunicação sem fios, que aumentam a capacidade de transmissão de dados. Aliado a esta evolução e das técnicas de engenharia de sistemas ocorre um boom de desenvolvimento de programas computacionais para os mais diversos fins. Isto tem se refletido nos centros urbanos levando a uma demanda socioeconômica por infraestrutura e serviços digitais a serem disponibilizados nas cidades.

Há uma externalização da vida “em rede” (na *web*), na qual relações sociais são virtualizadas, trabalhos são compartilhados, comércio e serviços ganham uma nova perspectiva. A agilidade que a conectividade promove alterações na forma de consumo, amplia formas de comércio e de serviços e economiza tempo. A tecnologia afeta a forma de viver, de trabalhar, de consumir, de governar nas cidades. Ter acesso à informação torna-se essencial na melhoria da qualidade de vida e bem-estar das pessoas. Entretanto surge uma questão: Será que somente internet ou equipamentos tecnológicos bastam para que haja melhoria na qualidade de vida das cidades?

Ao longo deste primeiro quarto de século, a fim de corresponder à demanda por melhoria da qualidade de vida nas cidades, diversas iniciativas político-econômicas vêm sendo desenvolvidas e aplicadas pelo mundo. Associado a essas iniciativas, diferentes nomenclaturas e definições buscam retratar estas condições vividas ou almejadas para as cidades, dependendo dos interesses acadêmicos e daqueles que guiam a sua aplicação em áreas urbanas (MEDEIROS DA SILVA; BAGGIO, 2023). Dentre essas denominações encontram-se cidades digitais e *Smart Cities*. A denominação de cidade digital é associada à aplicação de tecnologias da informação e comunicação (TIC) para a melhora da gestão do município e o atendimento ao cidadão.

Por outro lado, *Smart Cities* ou Cidades Inteligentes, apesar de ainda não possuírem uma definição consolidada, está associada a uma abordagem holística dos desafios enfrentados pelas cidades, abrangendo aspectos econômicos, sociais e ambientais. Na literatura são encontrados alguns trabalhos (MATTERN, 2017; VALLIANATOS, 2015, DIAS, 2024) que abordam o perigo de medir o sucesso/crescimento econômico de uma cidade com as inovações tecnológicas, desconsiderando a complexa ecologia de produção e circulação de informações, que envolvem aspectos humanos, institucionais e culturais, e que ao final moldam o uso dessas informações nas cidades para a gestão e a definição de políticas públicas. Dias (2024), por exemplo, discute como as inovações tecnológicas podem ampliar desigualdades sociais ao priorizar o crescimento e a competitividade em detrimento das realidades sociais e econômicas das populações urbanas. Valinattos (2015) reflete em como aplicar a análise computacional aos desafios sociais e físicos de uma cidade considerando o urbanismo digital na era de big data. Mattern (2017) argumenta que a cidade não deve ser reduzida a simples processos de otimização tecnológica, pois envolve interações humanas e históricas imprevisíveis e ao limitar o planejamento urbano a algoritmos e a modelos técnicos, corre-se o risco de ignorar as dimensões sociais, culturais e históricas que tornam as cidades espaços dinâmicos e vivos.

Todavia, o conceito de Cidade Inteligente é uma realidade tangível e para uma cidade alcançar esse patamar é necessário um caminho de transformação através das tecnologias de informação e comunicação, orientando a inclusão digital dos cidadãos para a sua construção (OLIVEIRA E GRIN; 2023). Neste trabalho faz-se uso da proposição de Caragliu *et al.* (2009, p.50) que entende que uma cidade é inteligente quando aplica investimentos em capital humano e social, e em infraestrutura de comunicação e propicia crescimento econômico ambientalmente sustentável e melhoria na qualidade de vida, por meio de governança participativa e emprego consciente dos recursos naturais, como definição de referência.

Destarte este trabalho apresenta uma metodologia para construir um sistema de indicadores para mensurar a distinção de Cidades Inteligentes para as cidades brasileiras com mais

de cem mil habitantes no Censo 2010, perpassando aspectos econômicos, de mobilidade, do ambiente, das pessoas, de vida e de governança. Esse sistema de indicadores tem por premissas observar os Princípios Fundamentais das Estatísticas Oficiais (UNITED NATIONS STATISTICS, 2013; IBGE, 2014) - relevância, imparcialidade e igualdade de acesso; ética e padrões profissionais; responsabilidade e transparência; prevenção do mau uso dos dados; eficiência e uso de padrões internacionais – e priorizar a mensuração do bem-estar social, diferenciando-o de uma perspectiva comercial e mercantilista voltada para o uso de *ranking* ou classificação entre cidades como forma promoção. A partir deste sistema de indicadores é aplicado uma análise de *cluster* para identificar grupos de cidades homogêneas. Assim este trabalho oferece um sistema de indicadores por município proporcionando aos gestores locais um instrumento para desenvolver políticas públicas voltadas para o bem-estar social baseada em dados no âmbito da concepção de Cidades Inteligentes.

Delimitação do escopo geográfico da análise

Para a delimitação da área de estudo buscou-se criar um panorama nacional que não se restringisse às capitais e regiões metropolitanas, partindo da hipótese de que cidades de porte médio poderiam obter resultados mais promissores do que grandes capitais. Neste intuito, se estabeleceu como linha de corte os municípios brasileiros que, em 2010, tivessem mais de cem mil habitantes.

Na data de referência do Censo Demográfico de 2010, o Brasil apresentava uma população de 190.732.694 distribuídos em 5.560 Municípios. Destes, somente 283 possuíam mais de 100 mil habitantes, o que o equivale 54,76% da população brasileira nesta data. A escolha do ano de 2010 se deu por prever que dados do Censo Demográfico seriam necessários para o cômputo dos indicadores.

A distribuição espacial dos municípios selecionados bem como a síntese por região geográfica e unidade da federação reflete o grande contingente populacional mais próximo ao litoral Atlântico (Figura 1).

Figura 1 - Municípios brasileiros com mais de 100.000 habitantes em 2010 por região geográfica, por unidade da federação e totais populacionais.



Região Geográfica /UF	Municípios Selecionados	População em 2010	Região Geográfica /UF	Municípios Selecionados	População em 2010
N	20	7.151.971	NE	58	21.192.355
AC	1	336.038	AL	2	1.146.754
AM	2	1.904.047	BA	16	5.555.631
AP	2	499.466	CE	8	3.775.909
PA	10	3.204.154	MA	9	2.163.196
RO	2	545.137	PB	4	1.329.712
RR	1	284.313	PE	12	4.263.232
TO	2	378.816	PI	2	959.935
S	48	12.817.625	RN	3	1.266.010
PR	18	5.151.333	SE	2	731.976
RS	18	5.000.962	SE	139	55.463.812
SC	12	2.665.330	ES	9	2.157.689
CO	18	7.810.914	MG	29	8.508.413
DF	1	2.570.160	RJ	26	13.973.702
GO	9	9.294.015	SP	75	30.824.008
MS	4	1.188.326			
MT	4	1.112.269			
Total				283	104.436.677

Fonte: IBGE/Censo Demográfico, 2010.

Dimensões de Cidades Inteligentes

Da concepção de Cidades Inteligentes depreendem-se termos como “Dimensões de Inteligência” ou apenas “Dimensões” de Cidades Inteligentes. Estes termos são utilizados pelos autores quando se referem a aspectos concretos de uma cidade inteligente. Albino *et al.* (2015, p.8) apresenta uma discussão existente entre pesquisadores no que se refere a classificar em dimensões os diversos aspectos de Cidades Inteligentes.

Alguns pesquisadores acreditam que as cidades devem ser vistas de uma maneira orgânica, na qual a soma das partes não corresponde à plenitude do todo. No entanto, muitos pesquisadores, com a intenção de dar clareza ao que constitui uma cidade inteligente tendem a separar esta concepção em muitas características e dimensões, justificando esta decisão com a complexidade de gerir a concepção de Cidade Inteligente de uma forma holística (ALBINO *et al.*, 2015). Desta forma, neste trabalho, concebe-se para fins operacionais e visando um melhor entendimento da realidade o uso de dimensões.

Para isto é utilizado como referência os trabalhos de Giffinger e Gudrun (2010) que definem seis dimensões relacionadas a Cidades Inteligentes: economia, mobilidade, ambiente, pessoas, vida e governança. Essa escolha é referendada pela comparação com as dimensões propostas pelos autores Mahizhnan (1999), Giffinger *et al.* (2007), Eger (2009), Caragliu *et al.* (2009), Thuzar (2011), Nam e Pardo (2011), Barrionuevo *et al.* (2012), Kourtit e Nijkamp (2012) e Chourabi *et al.* (2012), conforme apresentado de forma sintética no Quadro 1.

Quadro 1 - Dimensões chave de uma cidade inteligente.

Autor(es)	Dimensões					
	Economia	Ambiente	Mobilidade	Governança	Cidadão	Vida
Mahizhnan (1999)	X		X		X	X
Giffinger <i>et al.</i> (2007)	X	X	X	X	X	X
Eger (2009)	X					X
Caragliu <i>et al.</i> (2009)	X	X	X	X	X	
Thuzar (2011)	X	X		X		X
Nam e Pardo (2011)	X	X	X	X	X	X
Barrionuevo <i>et al.</i> (2012)	X	X		X	X	X
Kourtit e Nijkamp (2012)	X		X		X	X
Chourabi <i>et al.</i> (2012)	X	X	X	X	X	X

Fonte: Adaptado de Albino *et al.* (2015, p. 10)

Mais recentemente vem sendo desenvolvida a família de normas ISO 37120: “Cidades e comunidades sustentáveis: indicadores para serviços urbanos e qualidade de vida”, da Organização Internacional de Padronização (*International Organization for Standardization* - ISO). Esta família de normas seleciona um conjunto de indicadores que possam servir de base de comparação internacional entre as cidades no que se refere à distinção de Cidades Inteligentes e Cidades Sustentáveis. As normas são desenvolvidas pelo Comitê Técnico TC-268 da ISO, responsável pela metodologia de padronização dos indicadores.

Em 2014, foi lançada a primeira versão da norma, traduzida em 2017 pela Associação Brasileira de Normas Técnicas (ABNT). Com a denominação ABNT NBR ISO 37120, a versão “Desenvolvimento sustentável em comunidades – indicadores para serviços urbanos e qualidade de vida” (ABNT, 2017), é a versão traduzida existente. Em 2018, a norma foi reformulada e atualizada, estando vigente uma nova versão da norma, a ISO 37120:2018. Ela apresenta mudanças tanto nos indicadores propostos quanto nas dimensões ou temáticas abordadas em relação a versão

de 2014. Estão sendo desenvolvidos dois desdobramentos nesta família de normas, que visam diferenciar indicadores específicos para Cidades Inteligentes e Cidades Sustentáveis. Na versão vigente, a norma propõe o uso de dezenove áreas temáticas (dimensões) e cento e vinte oito indicadores. As dimensões são: Economia, Educação, Energia, Ambiente e Mudanças Climáticas, Finanças, Governança, Saúde, Habitação, População e Condições Sociais, Recreação, Segurança, Resíduos Sólidos, Esporte e Cultura, Telecomunicações, Transporte, Agricultura urbana/local e Segurança alimentar, Planejamento Urbano, Esgoto, Água.

Propostas de mensuração para Cidades Inteligentes

No contexto internacional, existem iniciativas distintas de mensurar e classificar ou ainda avaliar como nível de maturidade (estágio de evolução) de Cidades Inteligentes. Giffinger *et al.* (2007) elaborou um *ranking* para cidades europeias de médio porte com objetivo específico de permitir uma avaliação com um olhar para o desenvolvimento futuro baseado na combinação de circunstâncias locais, atividades praticadas por políticos, negócios e os habitantes. Lombardi *et al.* (2012), por outro lado, correlaciona as seis dimensões de Cidades Inteligentes; a chamada tripla hélice que movimenta a sociedade: governo, indústria e academia. O *ranking* foi elaborado através de um processo analítico de rede empregando um modelo e agrupamentos que possam medir o desempenho de Cidades Inteligentes. Selada e Silva (2013) propõe um índice de cidades definido a partir de um indicador composto, resultante da média ponderada das pontuações atribuídas às cinco dimensões de análise, traduzidas como inovação, sustentabilidade, inclusão, governança e conectividade.

Há que se considerar ainda as reflexões e indicadores apontados no Projeto ASCIMER – *Assessment Methodology for Smart City Projects: Application to the Mediterranean Region*, desenvolvido pela Universidade Politécnica de Madrid (UPM, 2016), para concessão de financiamento pelo Banco Europeu de Investimento. Esse projeto utiliza-se da visão de especialistas da região para pontuar os principais desafios da área de abrangência (Mediterrâneo), a fim de dar peso às variáveis e aos indicadores.

No Brasil, as pesquisas, trabalhos e iniciativas relativas às Cidades Inteligentes são recentes, datando do último decênio. O governo federal, por meio do Ministério das Comunicações, entre 2011 e 2016, dispunha de um programa intitulado “Cidades Digitais”. Este programa tinha por objetivos: a) constituir redes digitais locais de comunicação nos municípios brasileiros; b) promover a produção e oferta de conteúdos e serviços digitais; e c) facilitar a apropriação de tecnologias da informação e da comunicação pela gestão pública local e pela população, de maneira coordenada e integrada entre esferas dos poderes públicos e da sociedade (BRASIL, 2011). Em 2016, tendo à frente um novo governo, esta política é reformulada. Sendo parte do projeto “Brasil Inteligente”, é lançada nova fase do Programa, agora intitulado “Cidades Inteligentes”. Essa fase ainda promove majoritariamente políticas relacionadas a aspectos de cidades digitais, tal como infraestrutura digital para as cidades (BRASIL, 2016), sendo esta apenas uma das dimensões de Cidades Inteligentes.

No âmbito de governo local, a iniciativa intitulada “Brasil 2030: Cidades Inteligentes e humanas” é promovida pela Rede Brasileira de Cidades Inteligentes e Humanas (RBCIH). Através de uma associação de prefeitos de cidades de médio e grande porte, essa iniciativa almeja incentivar o desenvolvimento das cidades, mediante a criação de um ranking da instituição e de um Selo correspondente ao nível da certificação anual em que a cidade se encontre (RBCIH, 2016). Esta entidade procura, com a participação da academia, governo, sociedade e iniciativa privada, estruturar indicadores para Cidades Inteligentes e humanas de forma a oferecer aos gestores municipais de mecanismos de autoavaliação e planejamento.

A iniciativa privada no Brasil, por intermédio da empresa *Urban Systems*, colaboradora da RBCIH, em conjunto com a *Sator*, promotora de eventos, desenvolveram um *ranking* de Cidades Inteligentes. Este *ranking* provê avaliações anuais das cidades brasileiras, sendo 2015 o primeiro

ano de avaliação e 2021, o último. Na documentação de referência do *Ranking Connected Smart Cities*, disponível no site do evento de premiação, são explicitados os indicadores utilizados e que, de ano para ano, esses indicadores são alterados de forma a melhorar a metodologia empregada.

A academia no Brasil também atua em pesquisas relacionadas a *ranking* de Cidades Inteligentes. Felden e Teixeira (2017) questionam a falta de correlação entre os indicadores que tratam da Saúde no *ranking* de Cidades Inteligentes da *Urban Systems* e os resultados da pesquisa Vigilância de Fatores de Risco e Proteção para Doenças Crônicas por Inquérito Telefônico (VIGITEL). Afonso *et al.* (2013) elaboram o BR-SCMM: Modelo Brasileiro de Maturidade para Cidades Inteligentes. Este *ranking* é considerado bem estruturado, mas com abrangência temática parcial. Este modelo de maturidade para cidades trata somente os aspectos relacionados à Água, Saúde e Educação como dimensões de inteligência das cidades, a fim de compor um conjunto de índices sintéticos disponíveis sobre os respectivos temas. Cunha *et al.* (2016) desenvolveram uma pesquisa abrangente, com viés empresarial e tecnológico. A metodologia utilizada efetua uma avaliação dos estágios de evolução (maturidade) de Cidades Inteligentes, voltada para questões tecnológicas e de infraestrutura. Destaca-se ainda que nesse trabalho houve uma pesquisa sobre as expectativas dos cidadãos de diferentes perfis acerca do que seria uma cidade inteligente e as principais demandas segundo o ponto de vista dos cidadãos. Outro trabalho relevante é um estudo capitaneado pelo Observatório das Metrôpoles que propôs o Índice de Bem-Estar Urbano (IBEU) (RIBEIRO; RIBEIRO, 2013, p. 7). Os indicadores propostos para a construção do IBEU contemplam parcialmente as dimensões analisadas no contexto de Cidades Inteligentes avaliando a dimensão urbana do bem-estar usufruído pelos cidadãos brasileiros promovidos pelo mercado, via consumo, e pelos serviços sociais prestados pelo Estado.

Um sistema de indicadores não deveria ser utilizado fundamentalmente para qualificar em *ranking* de Cidades Inteligentes, pois vai muito além de ser um instrumento de atratividade para as cidades com os melhores resultados. Ele contribui com os gestores locais ao oferecer instrumentos de autoavaliação e planejamento, bem como informações para os gestores definirem estratégias para as cidades com base em seus resultados.

Metodologia

Para o desenvolvimento do trabalho foram selecionados indicadores para cada dimensão, a seguir foram analisados através de técnicas de estatística descritiva e de análise multivariada, conforme ilustrado na Figura 2.

Figura 2 – Técnicas utilizadas para avaliação das Cidades Inteligentes no Brasil.



A seleção de indicadores utilizados neste trabalho foi norteada pelas dimensões e indicadores da versão vigente da norma ISO 37120:2018: *Sustainable cities and communities – Indicator for city services and quality of life* que abrange indicadores tanto para Cidades Inteligentes quanto Cidades Sustentáveis. Para cada uma das dimensões são propostos indicadores, os quais parte foi implementada diretamente ou de forma adaptada, conforme disponibilidade de dados relativos ao tema (Tabela 1). Foram, portanto, utilizados 78 indicadores neste trabalho pertencentes a 16 das 18 dimensões propostas da ISO 37120:2018. O detalhamento metodológico pode ser conferido em Nagatomi (2019).

Na seleção dos indicadores foi observada a disponibilidade de dados, requisito necessário e primordial para esta pesquisa devido à abrangência nacional, e nível de desagregação adequado. Um desafio nessa atividade foi cobrir a ampla diversidade de temas, com dados de bases oficiais armazenados de forma sistematizada. Aliado a este fato foi difícil encontrar fontes de dados apropriadas considerando critérios de qualidades para os indicadores relativos à sua relevância, acurácia, oportunidade e pontualidade, acessibilidade e transparência, comparabilidade e coerência.

As fontes de dados selecionadas foram bases de dados referentes a pesquisas, cadastros e mapeamento do IBGE e outras agências públicas. Dentre as fontes de dados produzidas pelo IBGE foram utilizadas: Censo Demográfico (2010); Pesquisa de Informações Básicas Municipais, MUNIC (2017, 2016, 2014, 2013); Pesquisa de Serviços de hospedagem (PSH, 2016); Produto Interno Bruto dos Municípios, PIB dos Municípios (2015); Cadastro Central de Empresas, CEMPRE (2016); Estimativas da População (2015, 2016); Base Cartográfica Contínua do Brasil ao Milionésimo, BCIM (2015); Malha Municipal para fins estatísticos (2010, 2018).

Tabela 1 – Temáticas de Cidades Inteligentes e sustentáveis, segundo ISO37120:2018, e quantidade de indicadores propostos pela norma e selecionados para este trabalho.

Área temática	Número de indicadores	
	Proposta ISO	Selecionados
Agricultura urbana/local e Segurança Alimentar	4	0
Água	7	3
Ambiente e Mudanças Climáticas	9	5
Economia	11	9
Educação	6	6
Energia	9	1
Esgoto	4	4
Esporte e Cultura	3	3
Finanças	6	0
Governança	4	4
Habitação	10	6
Planejamento Urbano	7	5
População e condições sociais	9	9
Recreação	2	0
Resíduos sólidos	10	5
Saúde	6	6
Segurança	10	4
Telecomunicação	2	2
Transporte	9	6
Total	128	78

Fonte: Elaborado a partir de ISO 37120 (2018).

Além destas fontes de dados do IBGE foram utilizados dados: i) da Secretaria Nacional de Saneamento (SNS): Sistema Nacional de Informações sobre Saneamento, SNIS Série Histórica (2016); ii) do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep): Censo Escolar, Sinopse Estatística (2017) e Censo da Educação Superior, Sinopse Estatística (2017); iii) do Tribunal Superior Eleitoral (TSE): Estatísticas Eleitorais (2016), Estatísticas do Eleitorado (2018); iv) do Departamento de Informática do Sistema Único de Saúde (DATASUS): TABNET – Informações de Saúde, utilizando o Sistema de Indicadores de Mortalidade, SIM (2016), o Sistema de Internações Hospitalares, SIH (2017), o Cadastro Nacional de Estabelecimentos de Saúde, CNES, (2017) e Estatísticas demográficas calculadas com base no Censo Demográfico (2010); v) da Secretaria Nacional de Segurança Pública (Senasp): Sistema Nacional de Informações de Segurança Pública, Prisionais, de Rastreabilidade de Armas e Munições, de Material Genético, de Digitais e de Drogas, SINESP (2016); vi) da Secretaria Nacional de Proteção e Defesa Civil (SEDEC): Sistema Integrado de Informações sobre Desastres, S2ID (2017).

A estatística descritiva das 78 variáveis foi realizada como o objetivo de conhecer os indicadores e seus desempenhos nas cidades selecionadas. A análise fatorial exploratória utilizada é empregada com o objetivo de reduzir o número de indicadores selecionados para a avaliação de Cidades Inteligentes construindo assim fatores socialmente interpretáveis e utilizados como mensuração para a comparação entre as cidades participantes. Neste trabalho os escores são utilizados para classificar as cidades, dando interpretação prática aos fatores com base nos dados observados por município. Para comparar valores dos escores por município em uma mesma escala, estes foram normalizados pela faixa de variação aplicando a seguinte transformação, atribuindo-se zero ao valor mínimo e um ao valor máximo:

$$y = \frac{x - \text{mín}}{\text{máx} - \text{mín}}$$

Onde:

y é o escore normalizado,

x é o valor do escore original,

min e máx são os valores de máximo e mínimo do escore do fator analisado.

A análise de *cluster* visou identificar os grupos de Cidades Inteligentes distribuídos geograficamente no Brasil. Para isto foi utilizado o método de Ward, que efetua um procedimento de agrupamento hierárquico. Neste, a medida de similaridade usada combina dois agrupamentos que apresentarem menor aumento na soma global de quadrados dentro dos grupos. Esse método tende a resultar em agrupamentos de tamanhos aproximadamente iguais devido à minimização de variação interna dos índices sintéticos municipais obtidos na análise fatorial (HAIR *et al.*; 2005).

Todos os processos de análise estatística foram realizados com o uso do programa R (R versão 3.5.3: *Great Truth*) e operando a versão 1.1.463 – © 2009-2018 do programa RStudio, Inc. As representações cartográficas foram realizadas no Arcgis 10.1. sob licença acadêmica da ENCE.

Resultados e discussões

Os 78 indicadores apresentaram na análise estatística descritiva distribuições que retratam alguns comportamentos próximos à normalidade, mas também distribuições assimétricas, bimodais, exponenciais e irregulares. Sendo assim, foi descartado o uso de técnica baseada em normal multivariada. Para indicadores ordinais, os resultados foram apresentados através das frequências das classes de valores. Não foi detectado um comportamento típico entre estes indicadores.

Tendo em vista que alguns indicadores são variáveis numéricas e outros ordinais foi utilizado no cálculo das correlações a função de cálculo “hetcor” (*Heterogeneous Correlation Matrix*) do pacote Polycor do R. Esta função calcula uma matriz de correlação heterogênea aplicando a Correlação de Pearson entre variáveis numéricas, Correlação Poliserial entre variáveis numéricas e ordinais, e Correlação Policórica entre as variáveis ordinais.

Nesta análise três variáveis apresentaram uma fraca correlação positivo e negativo abaixo 0,20, sendo, portanto, retiradas do conjunto de variáveis. Foram elas: Percentual de mulheres eleitas em relação ao número total de eleitos no município, Número de mortes relacionadas a desastres naturais por 100.000 habitantes e Participação em ações de capacitação relacionadas ao meio ambiente.

O teste de esfericidade de Bartlett foi aplicado a fim de verificar se as variáveis não são correlacionadas na população. A hipótese nula foi rejeitada, com a estatística de teste qui-quadrado no valor de 6546,681; grande em relação ao p-valor 1.224654e-304, chegando-se à conclusão de que a análise fatorial é adequada como forma de sintetizar as variáveis observadas em fatores. Aplicando-se o teste KMO (Kaiser-Meyer-Olkin) e empregando-se os critérios de Kaiser (1974), no qual valores acima de 0,5 podem ser considerados razoáveis, obteve-se o valor de 0,68 para o conjunto das 75 variáveis, o que representa um ajuste razoável do modelo.

Para a determinação do maior número de fatores que sintetizem as variáveis observadas utilizou-se a análise do *Scree Plot*, considerando como parâmetros a “regra do cotovelo” combinada com a linha de corte dos autovalores acima de 1. Para a determinação do número final de fatores foi avaliado o diagrama de Análise de fatores, a Carga Fatorial e a Variância Explicada Acumulada. Neste processo foi testada a remoção de variáveis para o aumento da variância explicada e, tendo em vista que o resultado não aprimorava significativamente o modelo, optou-se por preservar um maior número de variáveis a fim de refinar a interpretação do modelo e manter maior relação aos indicadores inicialmente propostos.

Assim, as 75 variáveis observadas foram agrupadas em 12 fatores, obtendo-se uma variância explicada acumulada que chegou a 0,59. Os fatores e as variáveis observadas que foram agrupadas pelo método de análise fatorial são apresentados no Quadro 2. Os fatores são ordenados neste quadro de modo decrescente, da maior para a menor variância explicada pelo fator. A sequência de apresentação das variáveis no contexto de cada fator também obedece, de maior para menor, a carga fatorial da variável, a contribuição da variável para o fator. A interpretação de cada fator está descrita também no Quadro 2.

Quadro 2 - Fatores, suas interpretações e variáveis agrupadas por fator, em ordem da carga fatorial das variáveis.

<p>Fator 3: Emprego, Trabalho e Condições de vida</p> <p>Taxa de desemprego de jovens, na área urbana do município em relação a PEA</p> <p>Taxa de desemprego da cidade</p> <p>Taxa de desemprego de jovens, no município, em relação a PEA</p> <p>Percentual de pessoas empregadas em tempo integral, em relação a população da cidade</p> <p>Percentual de crianças vivendo abaixo da linha nacional de pobreza (baixa renda)</p> <p>Percentual de pessoas empregadas em tempo integral, em relação a população Economicamente Ativa da cidade</p> <p>Percentual da população vivendo abaixo da linha nacional de pobreza (baixa renda)</p> <p>Relação empregos por habitação</p> <p>Percentual de pessoas com baixíssima renda</p> <p>Número de empresas do município por 100.000 habitantes</p> <p>Número de homicídios dolosos por 100.000 habitantes</p> <p>Taxa de suicídios por 100.000 habitantes</p> <p>Taxa estudante/professor na educação primária</p> <p>Percentual de pessoas que moram em domicílios cujo entorno não possui lixo acumulado nos logradouros</p> <p>Mortalidade infantil abaixo de cinco anos por 1.000 nascidos vivos</p> <p>Índice de Gini da renda domiciliar per capita do Município</p>
<p>Fator 11: Vida urbana, densidade e infraestrutura</p> <p>Densidade de pessoas em área construída (Habitantes por quilômetro quadrado)</p> <p>Percentual da população urbana com coleta regular de resíduos sólidos domésticos</p> <p>Percentual da população da cidade com serviço elétrico autorizado</p> <p>Percentual de domicílios com telefone celular</p> <p>Pessoas por domicílio (média)</p>
<p>Fator 1: Características do domicílio e entorno</p> <p>Percentual da população urbana vivendo em habitações inadequadas</p> <p>Percentual da população urbana atendida por serviço de coleta de esgoto</p> <p>Percentual da população urbana com serviço de abastecimento de água potável</p> <p>Percentual da população urbana atendida por serviço de coleta de esgoto. Índice de atendimento urbano de esgoto</p> <p>Percentual da população urbana com serviço de abastecimento de água potável. Índice de atendimento urbano de água</p> <p>Percentual de pessoas que moram em domicílios cujo entorno não possui esgoto a céu aberto</p> <p>Percentual de pessoas que moram em domicílios cujo entorno possui arborização</p> <p>Percentual de tamanho das áreas de assentamentos informais em relação a área urbana total</p> <p>Grau de diversidade de Instalações Esportivas Municipais</p> <p>Percentual de áreas designadas para a proteção ambiental</p>
<p>Fator 2: Saúde, mobilidade e equipamentos culturais</p> <p>Número de leitos hospitalares por 100.000 habitantes</p> <p>Número de profissionais da saúde por 100.000 habitantes</p> <p>Número de internações hospitalares por 100.000 habitantes</p> <p>Percentual de pessoas que trabalham fora do domicílio de residência e retornam do trabalho diariamente no período de até 1 hora</p> <p>Número de médicos por 100.000 habitantes</p> <p>Número de óbitos por causas externas por local de ocorrência por 100.000 habitantes</p> <p>Grau de diversidade de Equipamentos Culturais existentes</p>

Fator 12: Características socioeconômicas

- Percentual de pessoas que moram em domicílios cujo entorno possui rampa para cadeirante
- Percentual da população que nasceu fora do país
- Número de pessoas com ensino superior por 100.000 habitantes
- Renda Familiar Média (Reais)
- Percentual de domicílios com acesso à internet
- Consumo diário doméstico de água per capita (litros/dia)
- PIB per capita do município (Reais)

Fator 9: Influência social no ambiente urbano

- Número de funcionários locais eleitos para o cargo por 100.000 habitantes
- Número total de residências
- Grau de diversidade de Grupos Artísticos existentes
- Densidade demográfica da unidade territorial (Habitantes por quilômetro quadrado)
- Grau de abrangência do Plano Municipal de Transporte
- Grau de realização de iniciativas para melhoria do setor
- Grau de hierarquia política do órgão gestor de Transporte
- Grau de amplitude de tipos de serviços regulares de transporte de passageiros
- Grau de participação feminina em cargos de destaque na administração municipal

Fator 4: Receptividade local

- Percentual da população de novos imigrantes
- Percentual da população vivendo em residências a preços acessíveis
- Percentual de domicílios alugados em relação ao total de domicílios existentes na área urbana do município
- Taxa de desocupação de domicílios
- Percentual da população de 10 anos ou mais não naturais do município em relação ao total de pessoas do município
- Número de leitos de hospedagem por 100.000 habitantes
- Percentual do esgotamento sanitário coletado que recebe tratamento centralizado
- Número de mortes relacionadas a fogo por 100.000 habitantes

Fator 7: Participação no ensino e Trabalho

- Percentual da população em idade escolar matriculada em escola
- Percentual de mulheres em idade escolar matriculada em escolas
- Taxa de trabalho infantil

Fator 8: Aproveitamento escolar

- Percentual de estudantes completando o ensino médio
- Percentual de estudantes completando o ensino fundamental

Fator 6: Resíduos sólidos e estratégias ambientais

- Percentual da população urbana com coleta de regular de resíduos sólidos domésticos, Taxa de cobertura do serviço
- Total coletado de resíduos sólidos municipais per capita (Kg por habitante dia)
- Grau de diversidade de estratégias locais adotadas e implementadas para prevenção, redução e gestão de riscos e desastres alinhadas às estratégias nacionais
- Percentual de resíduos sólidos urbanos que são reciclados

Fator 10: Conscientização Cidadã

- Percentual de eleitores registrados participantes na última eleição em relação ao total de eleitores registrados
- Número de mortes relacionadas a acidentes de trânsito por 100.000 habitantes
- Número de mortes relacionadas a fogo por 100.000 habitantes

Fator 5: Responsabilidade ambiental

- Grau de hierarquia política do órgão gestor de Ambiente
- Grau de diversidade de Impactos ambientais e/ou processos/ações de maior ocorrência nos últimos 24 meses

Os escores correspondem às realizações das variáveis latentes segundo o modelo da análise fatorial implementado, e são utilizados para classificar as cidades. Dezoito municípios que obtiveram os máximos e mínimos escores foram reunidos e listados na Tabela 2.

Tabela 2 – Municípios participantes dos extremos de máximo e mínimo dos escores fatoriais para classificação de Cidades Inteligentes.

Município (UF)	F. 3	F. 11	F. 1	F. 2	F. 12	F. 9	F. 4	F. 7	F. 8	F. 6	F. 10	F. 5
Parintins (AM)	0,248	0,420	0,570	0,554	0,360	0,124	0,000	0,013	0,818	0,723	0,605	0,626
Boa Vista (RR)	0,083	0,854	0,815	0,904	0,182	0,000	0,400	0,122	0,474	0,786	0,722	0,612
Marituba (PA)	0,351	1,000	0,000	0,462	0,315	0,155	0,154	0,657	0,441	0,674	0,517	0,677
São José de Ribamar (MA)	0,332	0,000	0,732	0,216	0,224	0,230	0,686	0,910	0,086	0,800	0,783	1,000
Itapipoca (CE)	0,260	0,349	0,707	0,574	0,355	0,128	0,259	0,396	0,801	0,475	0,612	0,000
Sobral (CE)	0,249	0,718	0,841	0,985	0,000	0,194	0,418	0,432	0,710	0,382	0,569	0,380
N. Sra. do Socorro (SE)	0,000	0,948	0,699	0,483	0,310	0,068	0,393	0,404	0,343	0,527	0,823	0,498
Montes Claros (MG)	0,089	0,833	1,000	0,909	0,130	0,035	0,332	0,611	0,357	0,739	0,836	0,643
Maricá (RJ)	0,421	0,897	0,306	0,447	0,345	0,151	0,464	0,819	0,233	0,695	0,000	0,496
Rio das Ostras (RJ)	0,499	0,816	0,398	0,316	0,480	0,230	1,000	1,000	0,570	0,375	0,435	0,215
Barretos (SP)	0,341	0,887	0,877	1,000	0,293	0,045	0,308	0,435	0,579	0,552	0,021	0,589
Santana de Parnaíba (SP)	0,427	0,876	0,540	0,468	0,546	0,089	0,361	0,521	0,480	1,000	0,723	0,628
São Paulo (SP)	1,000	0,655	0,182	0,000	0,879	1,000	0,079	0,274	1,000	0,208	0,411	0,121
Balneário Camboriú (SC)	0,469	0,879	0,633	0,511	1,000	0,090	0,760	0,144	0,586	0,054	0,814	0,267
Blumenau (SC)	0,728	0,844	0,546	0,543	0,383	0,241	0,252	0,137	0,449	0,534	1,000	0,135
São Leopoldo (RS)	0,327	0,903	0,899	0,553	0,216	0,095	0,305	0,346	0,000	0,703	0,473	0,548
Rio Verde (GO)	0,430	0,860	0,718	0,640	0,221	0,146	0,520	0,000	0,570	0,493	0,176	0,548
Trindade (GO)	0,578	0,875	0,524	0,465	0,281	0,181	0,373	0,221	0,626	0,000	0,454	0,607

Pela análise de agrupamento pelo método hierárquico de Ward, quanto maior o valor do índice de Silhouette melhor é o agrupamento. Desta forma o melhor resultado apontou dois grupos com um índice de Silhouette de 0,09, onde o Grupo 1 e o Grupo 2 teriam 169 municípios e 114 municípios, respectivamente. Entretanto, nesta mesma análise foi considerado um índice de Silhouette mais flexível de 0,075 com cinco grupos. Assim, 17,7 % de municípios estão no Grupo 1; 17,3% no Grupo 2; 24,7% no Grupo 3; 14,1% no Grupo 4 e 26,1% no Grupo 5, conforme sistematizado no Quadro 3 por região e Unidade da Federação.

Quadro 3 - Resultado do agrupamento por Unidade da Federação e regiões.

Região	UF	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5	Total	Grupo da Capital
Norte	AC	0	1	0	0	0	1	G2
	AM	1	1	0	0	0	2	G2
	AP	0	2	0	0	0	2	G2
	PA	8	2	0	0	0	10	G2
	RO	2	0	0	0	0	2	G1
	RR	0	1	0	0	0	1	G2
	TO	1	1	0	0	0	2	G2
	Total	12	8	0	0	0	20	
	Percentual	60	40	0	0	0	100,0	
Nordeste	AL	0	2	0	0	0	2	G2
	BA	1	12	0	3	0	16	G2
	CE	1	4	3	0	0	8	G2
	MA	5	2	2	0	0	9	G2
	PB	0	4	0	0	0	4	G2
	PE	3	3	6	0	0	12	G3
	PI	1	1	0	0	0	2	G2
	RN	0	1	2	0	0	3	G2
	SE	0	2	0	0	0	2	G3
	Total	11	31	13	3	0	58	
Percentual	19,0	53,4	22,4	5,2	0,0	100,0		
Sudeste	ES	1	1	2	5	0	9	G2
	MG	3	2	9	5	10	29	G3
	RJ	7	1	1	15	2	26	G1
	SP	5	1	5	5	59	75	G1
	Total	16	5	17	30	71	139	
	Percentual	11,5	3,6	12,2	21,6	51,1	100,0	
Sul	PR	1	2	13	0	2	18	G3
	SC	1	1	10	0	0	12	G3
	RS	0	0	12	6	0	18	G3
	Total	2	3	35	6	2	48	
	Percentual	4,2	6,3	72,9	12,5	4,2	100,0	
Centro-Oeste	DF	0	0	0	0	1	1	G5
	GO	6	0	3	0	0	9	G1
	MS	2	0	1	1	0	4	G3
	MT	1	2	1	0	0	4	G2
	Total	9	2	5	1	1	18	
Percentual	50,0	11,1	27,8	5,6	5,6	100,0		
Brasil	Total	50	49	70	40	74	283	
	Percentual	17,7	17,3	24,7	14,1	26,1	100,0	

Na Região Norte, os Grupos 3 e 4 e não tiveram representações e na Região Nordeste o Grupo 5 também não teve representação. Por outro lado, na Região Sudeste o Grupo 5 apresentou 51% dos municípios. Cabe ainda destacar que a maioria das capitais da Região Norte e Nor-

mente não são suficientes para avaliar a melhoria da qualidade de vida nas cidades. Todavia, cabe ressaltar ainda que o panorama explicitado através da análise exploratória dos indicadores da base de dados elaborada revela grandes discrepâncias nacionais entre as cidades selecionadas para este trabalho.

Penalizado pela questão temporal, o trabalho se baseou em dados desatualizados, ao se considerar estar próximo ao fim da década, contudo, potencialmente interessante por estar por vir uma nova rodada de dados atualizados. A dependência de indicadores baseados em informações somente disponíveis no Censo Demográfico pode ser avaliada negativamente quanto à usabilidade do modelo para o acompanhamento de políticas ao longo de um governo. Contudo, a produção de indicadores frente a uma demanda crescente de informações mais frequentes e desagregadas se encontra aquém do desejável.

Finalizando, o emprego da técnica estatística Análise Fatorial Exploratória se mostrou eficiente na redução de 78 variáveis observadas em 12 fatores que representam estatisticamente as informações originais através de uma visão global. O novo conjunto de indicadores sintéticos criados na análise fatorial são passíveis de serem utilizados pelos tomadores de decisão, gestores locais, bem como pelas associações sociais para o pleito por melhorias em suas cidades ou municípios.

Considerações finais

Assumindo que uma cidade pode ser chamada como inteligente quando seus governantes, de forma participativa, aplicam recursos em capital humano e social e em infraestrutura de TIC, visando crescimento econômico ambientalmente sustentável e melhoria na qualidade de vida, sendo assim, a tomada de decisão racional e voltada para os temas mais importantes e urgentes deve ser fundamentada em fatos e informações. Estes devem ser obtidos a partir de fontes confiáveis e seguir uma metodologia de cálculo documentada e replicável. Especialmente para a administração das cidades, entende-se que seja tecnicamente recomendável e um grande desafio.

Destarte, este trabalho propõe um método de cálculo replicável que sintetiza os dados presentes nas variáveis observadas, segundo uma metodologia estatística, a fim de subsidiar a avaliação e comparação da realidade das cidades brasileiras em relação às dimensões de Cidades Inteligentes, incluindo o monitoramento dos resultados ao longo do tempo.

Todavia cabe ressaltar que uma cidade é um espaço complexo de comunicação e armazenamento de saberes, que possui uma função histórica, social, econômica e ambiental. Desta forma, não foi o objetivo desse trabalho reduzir as cidades a simples processos de otimização tecnológica mensurando seu desenvolvimento por indicadores de desempenho, típicos do mundo dos negócios, como se as cidades fossem apenas conjuntos de dados a serem processados por meio de algoritmos, mas sim apresentar uma metodologia que retrate a situação das cidades brasileiras.

Dessa forma foi considerada a necessidade de uma avaliação sistemática de indicadores que possibilitem comparar os resultados de cidades e municípios do Brasil e monitorar, ao longo do tempo, de modo integrado. Essa síntese pode subsidiar políticas públicas com dados confiáveis e metodologia transparente, por meio de métricas que dão suporte à decisão baseada em dados. Todavia se faz necessário ainda considerar uma profunda reflexão sobre a tecnologia e os valores sociais, políticos e culturais que definem as cidades inteligentes.

Ressaltam-se ainda as diversas dificuldades em encontrar fontes dados minimamente adequadas para uso, considerando as dimensões de avaliação da qualidade propostas: relevância, acurácia, oportunidade e pontualidade, acessibilidade e transparência, comparabilidade e coerência. Assim devem ser consideradas as limitações quanto a temporalidade das bases, pois os dados disponíveis pertencem a datas, períodos e anos de apuração distintos. A variedade na temporalidade não foi o ideal, mas a possível.

Referências

ABNT. ABNT NBR ISO 37120:2017 – *Desenvolvimento sustentável de comunidades: Indicadores para serviços urbanos e qualidade de vida*. 1 ed. Rio de Janeiro: ABNT, 2017.

AFONSO, R. A.; da S., W. M.; TOMAS, G. H. R. P.; GAMA, K.; OLIVEIRA, A.; ALVARO, A.; GARCIA, V.C. BR-SCMM: *Modelo Brasileiro de Maturidade para Cidades Inteligentes*. In: Simpósio Brasileiro de Sistemas de Informação, IX. 2013. Anais... v. 4, p. 511–516, 2013.

ALBINO, V.; BERARDI, U.; DANGELICO, R. M. *Smart cities: Definitions, dimensions, performance, and initiatives*. Journal of Urban Technology, v. 22, n. 1, p. 1–19, 2015.

BARRIONUEVO, J.M.; BERRONE, P.; RICART, J.E. *Smart Cities, Sustainable Progress*. IESE Insight n. 14, p. 50–57. 2012.

BRASIL. Ministério das Comunicações. *Portaria nº 376, de 19 de agosto de 2011*. Diário Oficial [da] República Federativa do Brasil, Brasília, DF, 22 ago. 2011. nº 161, Seção 1, pág. 76.

BRASIL. Presidência da República. Casa Civil. *Decreto nº 8.776, de 11 maio de 2016*. Diário Oficial [da] República Federativa do Brasil, Brasília, DF, 12 mai. 2016. pág. 20.

CARAGLIU, A.; BO, C. DEL; NIJKAMP, P. *Smart cities in Europe*. Journal of Urban Technology, v. 18, n. 2, p. 65–82. 2009.

CHOURABI, H.; TAEWOO, N.; WALKER, S.; GIL-GARCIA, J.R.; MELLOULI, S.; NAHON, K.; PARDO, T.A.; SCHOLL, H.J. *Understanding Smart Cities: An Integrative Framework*. In: HICSS, 45. Anais... Hawaii Conference, 2289–2297, 2012.

CUNHA, M. A.; PRZEYBILOVSCZ, E.; MACAYA, J. F. M.; BURGOS, F. *Smart Cities: transformação digital das cidades*. São Paulo: Programa de Gestão Pública e Cidadania – PGPC, 2016, 161p.

DIAS, Reinaldo. *Cidades Inteligentes e desigualdade social: uma análise crítica das dinâmicas econômicas e sociais*. EcoDebate. Setembro de 2024. Disponível em: <https://www.ecodebate.com.br/2024/09/04/cidades-inteligentes-e-desigualdade-social-uma-analise-critica-das-dinamicas-economicas-e-sociais/>. Acessado em 09/11/2024.

EGER, J.M. *Smart Growth, Smart Cities, and the Crisis at the Pump a Worldwide Phenomenon*. I-Ways, v. 32, n. 1, p. 47–53, 2009.

FELDEN, É. P. G.; TEIXEIRA, C. S. *Cidades Inteligentes e saúde nas capitais brasileiras*. In: Conferência ANPROTEC, 2017. Anais... 2017

GIFFINGER, R.; FERTNER, C.; KRAMAR, H.; KALASEK, R.; PICHLER-MILANOVIC, N.; MEIJERS, E. *Smart cities: Ranking of European medium-sized cities*. Viena; Centre of Regional Science. 2007. Disponível em: <www.smart-cities.eu>.

GIFFINGER, R.; GUDRUN, H. *Smart Cities Ranking: an effective instrument for the positioning of Cities?* ACE Architecture, City and Environment, n. 12, p. 7–25, fev. 2010.

HAIR, J. F. Jr.; ANDERSON, R. E.; TATHAM, R. L. e BLACK, W. C. *Multivariate Data Analysis with Readings*, 6ª ed., Prentice Hall, Englewood Cliffs, 2006.

IBGE. *Princípios Fundamentais das Estatísticas Oficiais: orientações para divulgações de resultados pelo IBGE*, 2014, 5p. Disponível em: https://www.ibge.gov.br/np_download/novoportal/codigos_e_principios/Principiosfundamentais_orientacoes_divulgacoes.pdf. Acessado em 10/07/2024.

ISO - International Organization for Standardization. *Sustainable cities and communities – Indicators for city services and quality of life*. ISO37120:2018(E). Suíça, 2018. Disponível em: <<https://www.iso.org/obp/ui/#iso:std:iso:37120:ed-2:v1:en>>. Preview acessado em: 22 ago. 2018.

- JOHNSON, R. A.; WICHERN, D. W. *Applied Multivariate Statistical Analysis*. New Jersey: Prentice Hall, 2007. 773p.
- KAISER, H.F. *An index of factor simplicity*. *Psychometrika*, v. 39, n. 1. P. 31-36. 1974.
- KOURTIT, K.; NIJKAMP, P. *Smart Cities in the Innovation Age*. *Innovation: The European Journal of Social Science Research*, v. 25, n. 2, p. 93–95, 2012.
- LOMBARDI, P.; GIORDANO, S.; FAROUH, H.; YOUSEF, W. *Modelling the smart city performance*. *Innovation: The European Journal of Social Science Research*, v. 25, n. 2, p. 137–149, 2012. DOI: 10.1080/13511610.2012.660325
- MAHIZHNAN, A. *Smart Cities: The Singapore Case*. *Cities*, v. 16, n. 1, p. 13–18, 1999.
- MEDEIROS DA SILVA, J. P.; BAGGIO, D. K. *Cidades inteligentes: uma revisão sobre conceitos, indicadores e estudos de caso*. *Contribuciones a Las Ciencias Sociales, São José dos Pinhais*, v.16, n.11, p. 25740-25759, 2023.
- MATTERN, Shannon. *A City Is Not a Computer*. *Places Journal*. Fevereiro de 2017. Disponível em: https://placesjournal.org/article/a-city-is-not-a-computer/?gad_source=1&gclid=Cj0KCQjwj4BhCaARIsAIfc7YbsX0g1whVgPokWyEZB54eb4sXPh4Pp_af5JTt7hVgr48QBAgEJxwEaAnfBEALw_wcB&cn-reloaded=1&cn-reloaded=1. Acessado em: 09/11/2024.
- NAGATOMI, R. C. de M. E. *Cidades Inteligentes no Brasil: Sistema de mensuração com foco em bem-estar social, a partir de estatísticas públicas*. Dissertação (Mestrado em População, Território e Estatísticas Públicas) – Escola Nacional de Ciências Estatísticas, p. 268. 2019.
- NAM, T.; PARDO, T.A. *Conceptualizing Smart City with Dimensions of Technology, People, and Institutions*. In: *Conference on Digital Government Research*, 12., 2011. College Park, MD. Anais... College Park, MD, 2011.
- OLIVEIRA, Ângela Maria de; GRIN, E. J. *As cidades inteligentes e o desafio da inclusão digital*. *Revista Rua, Campinas – SP, Volume 29 – Número 2*, | p. 433-458, novembro 2023
- RBCIH – Rede Brasileira de Cidades Inteligentes e Humanas. *Brasil 2030: Cidades Inteligentes e Humanas*. [2016].
- RIBEIRO, L. C. DE Q.; RIBEIRO, M. G. *Índice de Bem-estar Urbano*. Rio de Janeiro, RJ: Observatório das Metrôpoles. Letra Capital, 2013.
- SELADA, C.; SILVA, C. *As Cidades Inteligentes na Agenda Europeia: Oportunidades para Portugal*. In: *Conferência de PRU, II; ENPLAN, VIII; Workshop APDR: “Europa 2020: retórica, discursos, política e prática”, XVIII*. 2013. Anais... n. 1, p. 31, 2013.
- THUZAR, M. *Urbanization in Southeast Asia: Developing Smart Cities for the Future?* *Regional Outlook*, 2011, p. 96–100.
- UPM - UNIVERSIDADE POLITÉCNICA DE MADRID e EUROPEAN INVESTMENT BANK: INSTITUTE. *ASCIMER - ASSESSMENT METHODOLOGY FOR SMART CITY PROJECTS: Application to the Mediterranean Region*. Madrid. [2016]. Disponível em: <www.eiburs-ascimer.transyt-projects.com>. Acesso em: 20 jun. 2018.
- UNITED NATIONS STATISTICS. *Fundamental Principles of Official Statistics*. 2013, 2p. Disponível em: <https://unstats.un.org/unsd/dnss/gp/FP-Rev2013-E.pdf>. Acessado em 10/07/2024.
- URBAN SYSTEMS. *Ranking Connected Smart Cities*. p. 57, 2017. Disponível em: <https://materiais.connectedsmartcities.com.br/csc18_ranking-connected-smart-cities_lp>. Acessado em: 20 jun. 2018.
- VALLIANATOS, Mark. *Uncovering the Early History of ‘Big Data’ and ‘Smart City’ in Los Angeles, Boom California*. Junho de 2015. Disponível em: <https://boomcalifornia.org/2015/06/16/uncovering-the-early-history-of-big-data-and-the-smart-city-in-la/>. Acessado em 09/11/2024.

Estimating conditioning factors to proficiency via a Bivariate Multilevel Model using MCMC

Natália Raquel de Souza Pires*
Pedro Luis do Nascimento Silva**
James Brown***

Abstract

In educational research, Multilevel Models are widely considered as the best approach to estimate factors associated with assessment test scores. Such models are commonly applied to analyse test scores from one subject of interest, where the most common subjects are Mathematics, Language or Science. In Brazil, a national assessment test (SAEB - Sistema de Avaliação da Educação Básica) is applied every two years to students in public education enrolled in Years 5 and 9 in primary schools, and Year 3 of secondary schools. SAEB provides both Portuguese and Mathematics proficiency scores for each student, besides a rich data set with information about their socioeconomic characteristics and the education environment they share with peers in classrooms and schools. We analyse test scores for Year 9 students from SAEB 2017 estimating a Bivariate 3-level Model using Markov Chain Monte Carlo (MCMC) methods for students from the municipality of Recife, in Pernambuco State, Brazil. The approach we propose here can be easily applied to other municipalities or domains in Brazil to estimate locally relevant conditioning factors to students' proficiency, respecting the cultural and social diversity present in the country. Our results could be useful to subsidize decision making in public policies and to encourage the use of quantitative approaches to investigate education issues in the country.

* Doutora pelo programa de pós-graduação da ENCE.

** Professor aposentado do programa de pós-graduação da ENCE.

*** University of Technology Sydney, Australia.

Introduction

Modelling proficiency test scores is a widely used approach in Social Statistics, and it is especially useful to analyse school effectiveness and social and educational influences in student proficiency (Goldstein (2003), Raudenbush and Willms (1995), Leckie and Goldstein (2009)). Although the use of prior assessment scores is extremely important and recommended in such analyses, the unavailability of such information should not inhibit the study of educational issues from a statistical point of view.

In Brazil, there is a National Assessment Test called SAEB (Sistema de Avaliação da Educação Básica) which is applied every two years to students in public education enrolled in Years 5 and 9 in primary schools, and Year 3 of secondary schools. SAEB provides both Portuguese and Mathematics proficiency scores for each student, together with rich data about their socio-economic characteristics and the educational environment they share with peers in classrooms and schools. By design, SAEB does not support use of prior assessment scores for two reasons: first, tests are conducted every two years, and the gap between years included in the test is either four years (Years 5 to 9 in primary) or three years (from Years 9 in primary to 3 in secondary); second, even if we would wait 4 years to find the same students in year 9 who have taken the test in year 5, SAEB public data sets do not enable linking their test scores because no unique identifiers are provided.

Since its first application in 1995, results from SAEB revealed that the proportion of children achieving required levels of proficiency are small, thus highlighting the challenges faced by the public education system in Brazil. In 2017, the calculation of one of the SDG 4 (Sustainable Development Goal 4) indicators associated with Quality of Education (United Nations (2015)) for Year 9 students showed that only 34% of them had achieved appropriate proficiency levels of literacy skills and only 15% had achieved the required proficiency in numeracy skills.

Poor educational results in Brazil are the topic of discussion in Politics (Ribeiro (1984)), Pedagogy (Freire (2011)), Sociology (Ioschpe (2016)) and more recently, in Social Statistics (Ferrão et al. (2018); Barbetta et al. (2018); Pereira et al. (2020)). However, different from the rest of the world, Brazil does not have a large literature of quantitative approaches applied in this field. A recent article (Bof (2016)) pointed to the need "to determine the conditioning factors associated with proficiency scores to change them for improving the quality of teaching and the process of learning in public schools". According to the authors, "this information could support formulation of public policies related with quality of education" (Bof (2016)).

Conditioning factors for student proficiency are widely studied in the international literature. A meta-meta-analysis focused on student achievement research identified eight hundred meta-analyses, which covered 52.637 studies, identifying about 150 thousand effect sizes that influence proficiency. "The majority of the findings of the meta-analysis were derived from studies conducted in English-speaking, highly developed countries" (Hattie (2012)). There are no references from studies conducted in Brazil within these meta-analyses, supporting the need for research in this area in the country.

Multilevel Models are generally considered in educational research as the best approach to estimate factors associated with assessment test scores. This class of regression models is widely used to describe and explain school, classroom and individual effects on student proficiency scores obtained from standardized assessment tests. Statistical methods to fit such models have been extended over the past three decades, and a wealth of approaches and software tools are now available to facilitate such endeavours.

In this paper we use Multilevel Models to estimate conditioning factors impacting proficiency in Brazil using SAEB data. Our model considers the 3-level grouping structure of data, where each student is enrolled in one classroom and one school. We innovate by taking both proficiency scores in Portuguese Language (simply Portuguese from now on) and Mathematics as responses in a bivariate model. This model enabled us to identify conditioning factors rele-

vant for student proficiency in these two subjects. We refer to the proposed model as the BMM (Bivariate Multilevel Model), and we used Markov Chain Monte Carlo (MCMC) methods to fit it.

We analysed proficiency scores for Year 9 pupils from Recife. As we explain in Section 2, the prevailing cultural and socioeconomic diversity in Brazil makes analyses at a national level difficult, so we chose one municipality to illustrate our approach.

Section 3 describes SAEB details and the data preparation for analysis. Section 4 presents the model and methods used for data analysis. The main results are provided in Section 5. Section 6 explores the implications of our results towards formulation of public education policy.

SAEB Data

Public and official data on proficiency for Brazilian students are available for several years. The main source for such data is the so-called SAEB test. This nationwide standardised test is applied every two years to students enrolled in Years 5 and 9 of primary schools, and Year 3 of secondary schools. INEP (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira) is the government agency responsible for applying the SAEB test and for making its data available to the public.

All public schools in Brazil are required to participate in the SAEB, and sampled private schools are also invited to participate. Within public schools, all students in the target Years are required to take the test, while in private schools' samples of students are selected. In 2017, participating students were required to take tests in Mathematics and Portuguese, and to fill in a Socioeconomic Questionnaire (SEQ). The enrolled student non-response rate is around 20% (varying per Year). Item non-response rates vary per item in the socioeconomic questionnaire and by test score. A SEQ is also applied to teachers and school principals.

In the 2017 edition, more than 2.3 million Year 9 students enrolled in more than 70 thousand public schools took part in SAEB. Besides students, SAEB also collects data on teachers and school principals, comprising a total of 3.1 million observations in its data base for Year 9 students. The student SEQ had 57 questions covering topics such as social background, cultural habits, parent's education, and student learning habits. Teachers and school principals' questionnaires had 125 and 111 questions, respectively, about their formal education, continued professional education, pedagogic practices, and financial resources available to use in the school. One questionnaire (74 questions) was also applied regarding participating schools' infrastructure.

SAEB provides a rich data source to explore relationships between school, teacher and student conditioning factors that may affect learning outcomes. Given the cultural diversity existing in Brazil, it is not recommended to analyse data from all its 5,570 municipalities together. Each municipality adopts its own independent strategies to incorporate the National Curriculum Basis (Brasil, 1996), including how to hire teachers and principals, which pedagogic practices to use in classrooms and how to select students to the schools. Therefore, we focused on a single municipality – Recife – and developed a modelling approach which could be replicated for other areas.

Recife is the city with the highest HDI (Human Development Index) in the Northeast region of Brazil and the municipality with the seventh largest GDP (Gross Domestic Product) in the country. Despite good HDI results, in 2017 only 32% of its Year 9 students had achieved the minimum proficiency score for reading and writing in Portuguese, and only 12% had achieved the minimum proficiency score for maths-related skills, similar numbers to Brazil's overall indices. This helped explain why we chose this municipality for our analysis, in addition to the fact that the lead author is from Recife and knows the local situation well.

Tidying Process

SAEB 2017 data are available for public download from INEP website¹. The microdata available for download includes the student, teacher, principal and school infrastructure datasets, besides corresponding data dictionaries, technical notes, and summarized results.

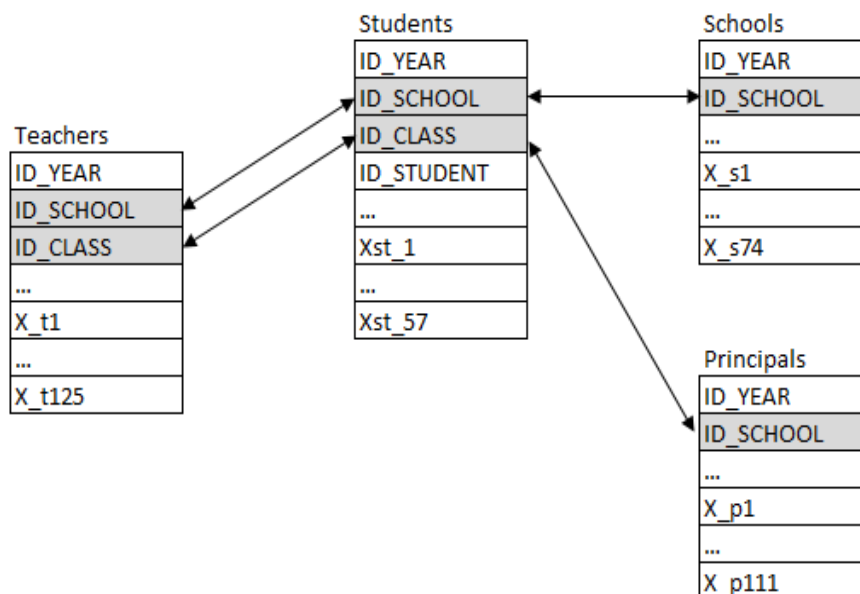
The first step comprised importing datasets into R, attributing labels to codes representing the responses, and classifying variables into relevant types. The information required for this step is contained in data dictionaries provided by INEP.

Then Year 9 students belonging to public schools in Recife who answered both tests and the SEQ were selected. Year 9 students were chosen because they can better reflect the effects of conditioning factors on proficiency, given that they spent 4 years more at school than Year 5 students. They are also the main target population of PISA (<https://www.oecd.org/en/about/programmes/pisa.html>) and the results of these two tests are commonly used to study potential conditioning factors to proficiency after completing primary education.

After selection of our target population (Year 9 Students from Public Schools in Recife, in 2017), and the two test scores for the response, we selected a total of potential explanatory 24 variables to use in the analysis, including variables from student, teacher, principal and school questionnaires – see the full list in the table in the Appendix. The choice of potential explanatory variables considered the quality of the information available from SAEB: some variables were rejected after reviewing the SEQ due to poor question design such as double-meaning and overlapping categories. Some variables had categories collapsed to summarize information. The final dataset resulting from the tidying process has data for 11,769 students enrolled in 372 classrooms from 127 schools. We lost a total of 232 students, from four schools, because these did not have the covariates due to confidentiality protection.

The last step in the tidying process was the ‘join’ task, where we combined the four datasets, namely student, teacher, principal, and school datasets into one. Figure 1 shows the key linking variables used in our study, where student records contain variables measured at this level together with those for the classroom, teacher, principal, and school levels.

Figure 1 - Diagram with key variables used to join datasets from SAEB.



¹ <http://portal.inep.gov.br/web/guest/microdados>

Model and Methods

Given the structure of SAEB data, where scores from both Mathematics and Portuguese tests are available, a natural choice of model to estimate conditioning factors for learning is the so called Bivariate Multilevel Model (BMM). This approach was used, within the education literature, by Yang *et al.* (2002) who explored the joint effects of studying subjects within a school structure, and Fielding and Yang (2005) where the authors used the grades rather than subject scores within a complex structure to look at teacher effects as well as pupil effects in a cross-classified structure for multiple subjects.

Here the BMM corresponds to a regression model where everyone in the target population has two outcomes, which can be correlated, within a strict hierarchy. The test scores for students within the same class and school are not independent, and we can use different covariates associated to each level in the hierarchy to explain the outcomes. We also have the potential to un-pick the unmeasured classroom or teaching group effects and the teacher effects, as the students are taught both subjects in the same group but potentially by different teachers.

The Bivariate Multilevel Model can be written as:

$$\begin{pmatrix} y_{1ijk} \\ y_{2ijk} \end{pmatrix} = \begin{pmatrix} \alpha_0 + \sum_{q=1}^Q \alpha_q x_{qijk} \\ \beta_0 + \sum_{q=1}^Q \beta_q x_{qijk} \end{pmatrix} + \begin{pmatrix} b_{1k} \\ b_{2k} \end{pmatrix} + \begin{pmatrix} \mu_{1jk} \\ \mu_{2jk} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1ijk} \\ \varepsilon_{2ijk} \end{pmatrix} \quad (1)$$

where $\begin{pmatrix} y_{1ijk} \\ y_{2ijk} \end{pmatrix}$ is the bivariate outcome for student $i=1, \dots, n_{jk}$ in classroom $j=1, \dots, n_k$, in school $k=1, \dots, K$; x_{qijk} is the q^{th} exploratory variable used in the model with associated model parameters α_q and β_q ; $\begin{pmatrix} b_{1k} \\ b_{2k} \end{pmatrix} \sim N[0; \Sigma_b]$, $\begin{pmatrix} \mu_{1jk} \\ \mu_{2jk} \end{pmatrix} \sim N[0; \Sigma_\mu]$ and $\begin{pmatrix} \varepsilon_{1ijk} \\ \varepsilon_{2ijk} \end{pmatrix} \sim N[0; \Sigma_\epsilon]$ are the random effects at school, classroom and student level, and the covariance matrices allow for heteroscedasticity of variances across the two subjects within a level as well as a non-zero covariance between random effects within a level. As usual, random effects are uncorrelated across levels.

There are diverse ways to estimate parameters from a BMM model, via both frequentist and Bayesian approaches. Each one of these approaches has pros and cons when applied to real problems. When working within the full generalised linear mixed modelling framework, frequentist approaches, such as REML (Restrict Estimation via Maximum Likelihood), use approximate likelihood methods and they are faster to find solutions for parameter estimates (Goldstein (1986)). Bayesian approaches are also approximations, "but the accuracy of the approximation increases the longer the analysis is run for, being exact at the limit" (Hadfield *et al.* (2010)).

In this article, the outcome is assumed normally distributed; therefore, the standard IGLS (Iterative Generalized Least Squares) algorithm would perform well to estimate the model parameters with asymptotic efficiency (Goldstein (1986)). However, while it can efficiently estimate the scale and variance components individually, the estimates are not constrained to obey the properties of a complex covariance matrix. Therefore, we used a Bayesian approach and the MCMC method to estimate both the fixed and random effects parameters in the Bivariate 3-level Model described above. "MCMC measures of confidence are exact, up to Monte Carlo error, and provide an easy and intuitive way of obtaining measures of confidence on derived statistics such as ratios of variances, correlations, and predictions" (Hadfield *et al.* (2010)).

The idea behind the Bayesian approach is to use prior information $\Pr(\theta)$ about the parameter values θ together with data likelihood $D = \Pr(y|\theta)$ to approximate a posterior distribution of parameters $\Pr(\theta|y)$. Therefore, if the distribution of the parameters is of interest, the Bayesian approach affirms that the distribution of the parameters given the prior information and data is:

$$\Pr(\theta|y) \cong \Pr(y|\theta) \times \Pr(\theta) \quad (2)$$

The Multivariate Normal distribution was used to define the prior distributions for the bivariate fixed effects, namely α_q and β_q , with a large parametric space to allow any estimate measure (negative or positive) to associated parameters. The multivariate Inverse-Wishart distribution was used as prior for the variance matrices of the normally distributed random effects, namely Σ_b , Σ_μ , and Σ_ϵ , because it allows specifying covariance matrices with different structures, depending on the model.

Here we used the package MCMCglmm in R to estimate the model parameters (Hadfield et al. (2010)).

Residual Analysis

There is no established approach to analyse residuals when a BMM is fitted via MCMC. We compared estimated and realized observations in each information level (Ntzoufras (2011)). Considering the model in Equation (1), we can write:

$$\begin{pmatrix} r_{1ijk} \\ r_{2ijk} \end{pmatrix} = \begin{pmatrix} y_{1ijk} \\ y_{2ijk} \end{pmatrix} - \begin{pmatrix} E(y_{1ijk}|D_1) \\ E(y_{2ijk}|D_2) \end{pmatrix} \quad (3)$$

where r_{qijk} is the residual associated with the observation y_{qijk} and D_q represents the data likelihood. Also, $E(y_{qijk}|D_q) = E_{\theta|D_q}(y_{qijk}|\theta)$, with θ representing all the model parameters. Since we have a large sample, we assume that the expected value of y_{qijk} conditioned on the estimated parameters is approximated by a linear function of the posterior expectation of random and fixed estimated parameters (Ntzoufras (2011))

$$E_{\theta|D_q} \left(y_{qijk} | \theta \right) \cong \bar{\theta}^T (x_{qijk} + b_j + u_{jk})$$

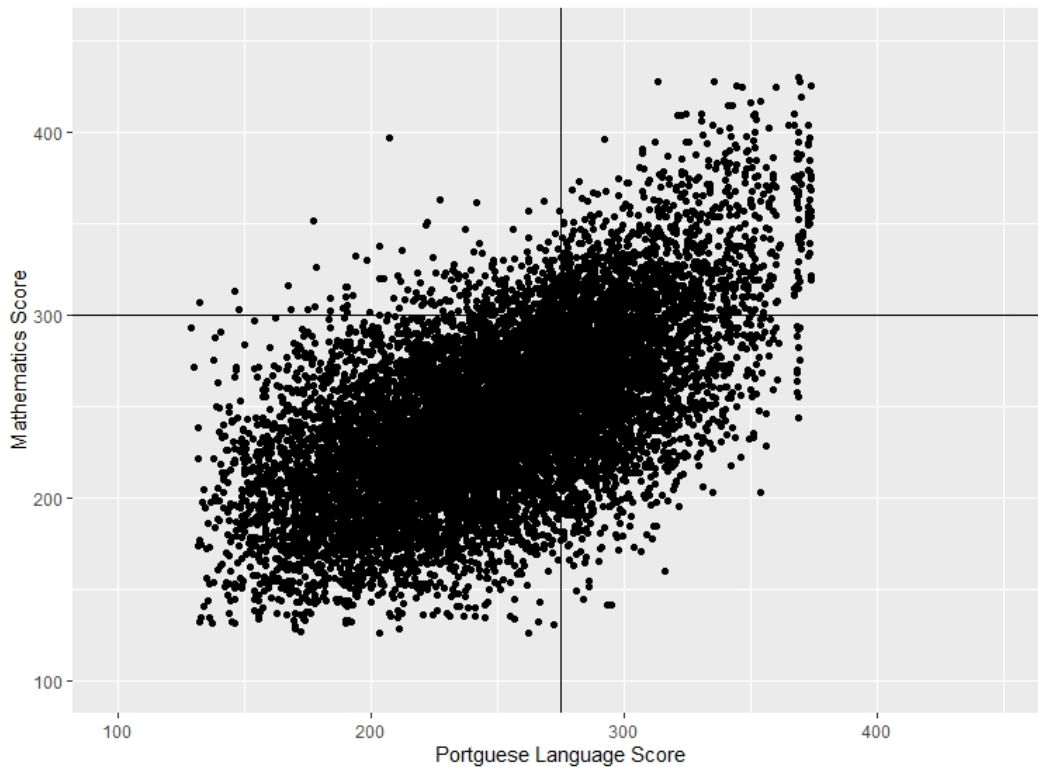
We adopted $\bar{\theta}^T$ as the mode of the posterior distribution of the parameters. Together with a convergence analysis, the residual analysis helped us evaluate the quality of our estimates.

To the best of our knowledge, this is the first time an analysis of Brazilian educational data is approached using a BMM fitted via MCMC. Some papers have already explored SAEB data looking for conditioning effects using multilevel regression. However, they did not explore the bivariate structure present in the data. At the international level, researchers from Italy (where a national evaluation like SAEB is conducted) applied a BMM to educational data (Masci et al (2017)) but they used the BMM structure only to estimate the random effects (using a software called ASReml). The fixed effects were estimated via Ordinary Least Squares (OLS) using LASSO algorithm. We used the open-source R software (Team (2013)) to analyse SAEB data, and we estimated fixed and random effects parameters together using MCMC.

Analysing SAEB data

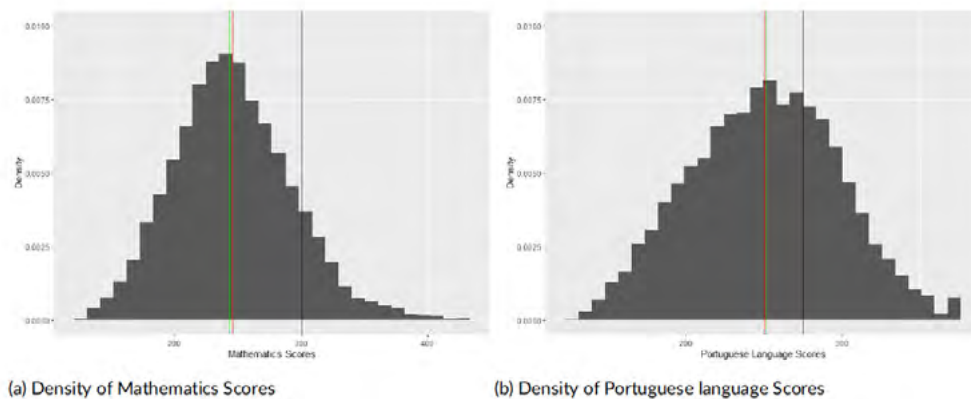
We used SAEB data from Recife municipality to obtain conditioning factors for proficiency scores. Figure 2 shows the dispersion diagram for the selected student assessment scores in Mathematics and Portuguese in 2017. The black lines represent the relevant minimum target scores for these subjects used to evaluate SDG indicator related to quality of learning. Considering the high correlation between Mathematics and Portuguese scores, observed in Figure 2, the use of BMM seems a plausible alternative.

Figure 2 - Dispersion of scores in Mathematics and Portuguese for Year 9 students from Recife public schools - SAEB 2017



Individual score histograms for both tests are shown in Figure 3. The histograms suggest that both scores have symmetric distributions. Only 32.3% of the 11,769 students had Portuguese scores achieving skills required to understand or write a text in Portuguese and 12.7% to realize simple operations in Mathematics (INEP (2020)). The metric used to define "skills required" is based on the average proficiency of OECD countries in assessment tests and is commonly used in Brazil by Education officials, researchers and organizations that work with educational data (Bof (2016)). Given the poor results observed, it is imperative to identify ways to tackle the problem and find out which factors affect learning to help improve learning outcomes in Recife and elsewhere.

Figure 3 - Density of Mathematics and Portuguese scores for Year 9 students from Recife public schools in 2017. The green line represents the median, the red line the mean and the black line the minimum standard score in each subject.



In a multilevel approach, we can use the "Null Model" to identify how much of the total variance is explained by each level of information. Table 1 shows the Variance Partition Coefficients (VPC) for schools, classrooms and students for univariate multilevel null models fitted via MCMC separately to Portuguese and Mathematics scores. Together, school and classroom account for 21.3% of the variability in Mathematics scores and 15.1% in Portuguese scores, respectively. We assumed Normal distributions for each response variable, and we ran 50.000 iterations for each univariate model, where every 10th estimate was stored after discarding the first 10.000 (burn-in) iterations in the MCMC algorithm. The prior distributions for fixed (means) and random effects covariances were, respectively, $N[0;1E+10]$ and *Inverse Wishart*[$V=1;v=0$].

Small VPC associated to classroom level, within school, in Table 1 can be explained by little diversity (in terms of economic, cultural, and social position) of peers across classrooms within a school. In addition, each school has just a few classrooms (73.7% of schools have less than 4 classrooms), which impacts the calculation of variance components in a multilevel model. However, since we want to use explanatory variables at classroom level in the model, and we have an interest in both the teacher and peer group effects of a common teaching group, for now we maintained the three levels in our model.

Table 1 - Estimated random effects variances in null univariate models for Mathematics and Portuguese and their respective VPC - Recife 2017

Level of Information	Mathematics		Portuguese L.	
	Variance	VPC	Variance	VPC
School ($\hat{\sigma}_b$)	404.0	18.0%	251.9	11.1%
Classroom ($\hat{\sigma}_\mu$)	73.9	3.3%	91.5	4.0%
Student ($\hat{\sigma}_\epsilon$)	1,763.0	78.7%	1,932.1	84.9%
Total	2,240.9		2,275.4	

Source: Own elaboration

Bivariate Null Model

The same set of prior parameters used to estimate univariate null models via MCMC was used to fit a bivariate null model. The estimated covariance matrices are shown in Table 2.

Table 2. Estimated covariance matrices for the bivariate null model - Recife 2017

School	Classroom	Student
$\Sigma_b = \begin{pmatrix} 255.6 & 285.5 \\ \square & 396.0 \end{pmatrix}$	$\Sigma_\mu = \begin{pmatrix} 87.2 & 76.4 \\ \square & 70.3 \end{pmatrix}$	$\Sigma_\epsilon = \begin{pmatrix} 1939.7 & 1023.1 \\ \square & 1771.8 \end{pmatrix}$

Source: Own elaboration

Random effects impact Mathematics and Portuguese scores in similar ways, in both school $\rho_b = 0.90$ and classroom levels $\rho_\mu = 0.98$. These high correlations between student scores at levels of information support our choice for adopting a bivariate modelling approach, where we formally consider these correlations in the covariance matrix estimation process. Such a high correlation is also indicative of the influence of common factors across subjects such as peer group, common teacher, and school environment. The correlation at student level is also substantial $\rho_\epsilon = 0.55$, yet small if we compare with that of the other levels.

An analysis considering all covariance matrices indicates that, even if general results for random effects of schools and classrooms have strong correlations (if a school has a positive im-

fact in Mathematics, the impact is also positive in Portuguese), a random student can still have different scores in Mathematics and Portuguese at an individual analysis because of the not so high correlation of proficiency scores at the student level.

High correlation at classroom level leads to the question about which factors at this level could jointly affect students' scores. Considering classroom is a common physical space that a group of students share during the whole year at the school, affected by social interactions and the teachers' influence, one possible explanation for the high correlation at this level is that it is common to observe only one teacher teaching both subjects for the same classroom in public schools since teachers do not need to have specific qualifications to teach classes of Mathematics or Portuguese in Brazil. Also, there are just few Year 9 classrooms per school.

An exploratory analysis of SAEB showed, however, that each teacher (identified by a unique ID) has only one register of questionnaire per classroom, contradicting our initial hypothesis - the dataset contains teachers who teach in more than one classroom but none who teaches Mathematics and Portuguese in the same classroom. Therefore, the extremely high correlation in the classroom level effects for the two subjects in the presence of different teachers implies little evidence in the null model of a strong teacher effect operating in either subject.

Peer effects within a common teaching environment is another possible explanation for high correlation at classroom level (Hattie (2012)) and it is plausible in the null model, since the main criterion to select students for public schools in Brazil is proximity to place of residence, which implies similar economic and social conditions for pupils in the same class. To test this last hypothesis, we added different variables associated with students' socioeconomic characteristics in our BMM and explored if this would weaken correlations. We also explored the impact of teacher covariates in search of a conditional teacher effect.

Bivariate Multilevel Model

We selected 24 covariates for potential inclusion in our model, based in previous meta-analysis in the same field (Hattie (2012)) – see the full list in the Appendix.

An exploratory analysis of the distribution of covariates showed a moderate percentage of item nonresponse in student variables, which could lead to about a 35% reduction in the database to be used for modelling if we were to discard incomplete student records. To address this problem, we decided to create a new category, called NR (representing the item nonresponse), for all covariates considered in our analysis.

We used a "forward-backward" selection procedure to test the inclusion of covariates in the model. We started with a Null Model, where only the random effects were estimated and then we included, step by step, blocks of covariates from each level in the hierarchy (forward).

We followed a sequence to include blocks of covariates: first the covariates at school level, then the covariates at classroom level and finally the covariates at student level. At each step, we checked the Credible Intervals (CI) for the coefficients for each category of each covariate to decide if it should remain in the model (backward). Then we fitted the model again with only the selected covariates and compared both models, the full and the selected model, via DIC (Deviance Information Criterion). The model with smaller DIC was selected for the next step, where we included the next block of covariates until we tested all pre-selected covariates in the model (forward-backward).

With the inclusion of groups of covariates in the model, and so increasing the number of parameters to estimate, we faced convergence problems, including high autocorrelation estimates and heteroscedasticity along the chain. These convergence problems were noted analysing trace plots for school, classroom, and residual estimates, which presented different trends along the chain, for the model selected after the forward-backward procedure.

We adopted different strategies to address this issue, including changing the starting values of the algorithm and increasing the number of iterations, but they were not sufficient. Then, considering the very high correlation observed at the classroom level, we tested the use of

a single parameter for effects at this level, which resolved the convergence problems. Therefore, instead of a covariance matrix associated with classroom random effects, we have a common estimated variance associated with both proficiency scores at the classroom level. We still have a covariance matrix for random effects at school and student levels and they are specified in Table 3 with the estimates for the new Null model.

Table 3. Estimated covariance matrices in bivariate null model with univariate effect for classroom level - Recife 2017

School	Classroom	Student
$\Sigma_b = \begin{pmatrix} 264.6 & 315.1 \\ \square & 413.2 \end{pmatrix}$	$\sigma_\mu = 79.2$	$\Sigma_\epsilon = \begin{pmatrix} 1936 & 1021 \\ \square & 1760 \end{pmatrix}$

Source: Own elaboration

Compared with the results of the null model specified in Table 2, we see the additional variance that was attributed to the classroom level is now attributed to the school level, and not to the subject level. It validates our choice to take the bivariate effect in classrooms out of the model, since we do not increase the subject level residual variance but slightly adjust the correlation structure from classroom and school. The estimates of random effects, including residuals, for the conditional model are presented in Table 4. If we compare the results with Table 3, we can see that a good part of the random effects was explained when we included fixed effects in the model, as expected.

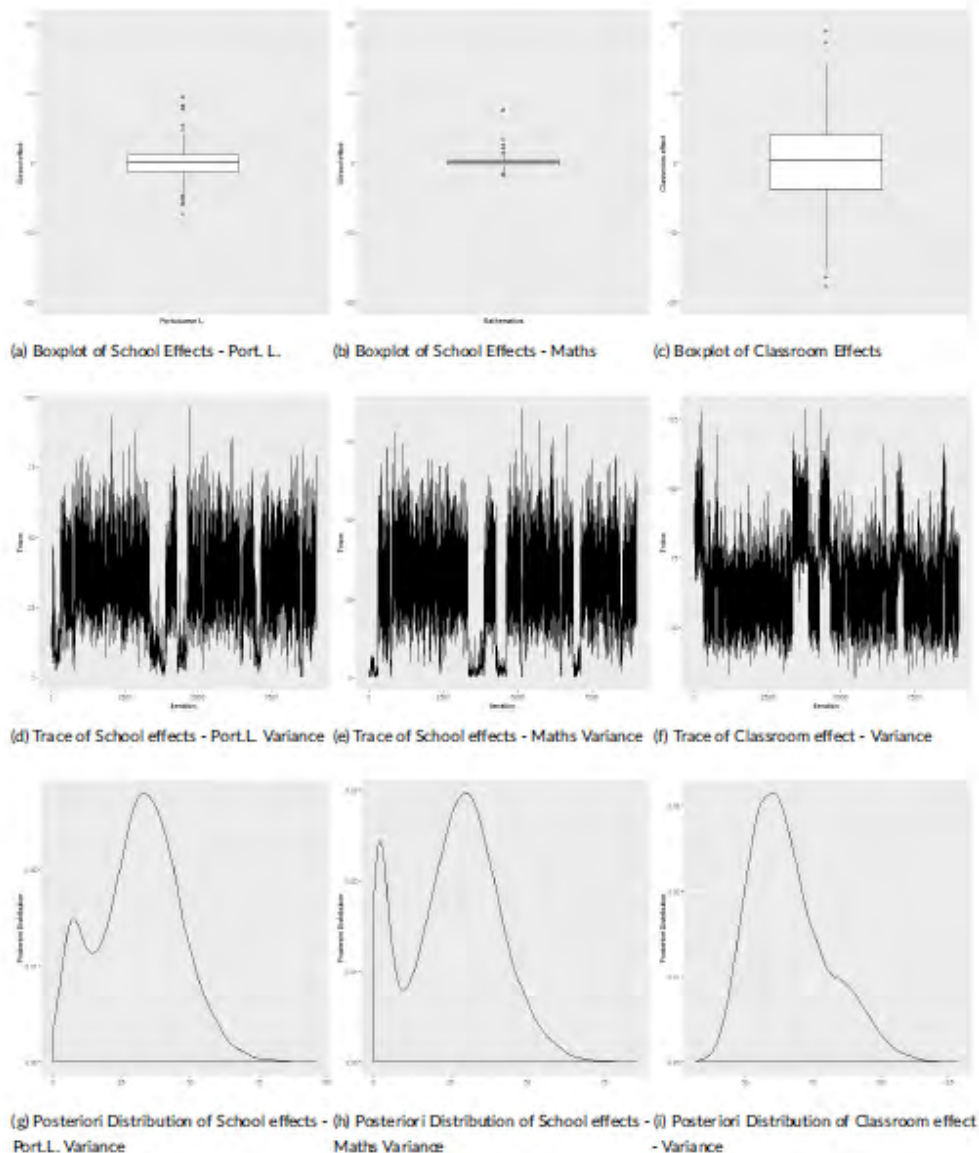
Table 4. Estimated covariance matrices in BMM with univariate effect in classroom level - Recife 2017

School	Classroom	Student
$\Sigma_b = \begin{pmatrix} 264.6 & 315.1 \\ \square & 413.2 \end{pmatrix}$	$\sigma_\mu = 79.2$	$\Sigma_\epsilon = \begin{pmatrix} 1936 & 1021 \\ \square & 1760 \end{pmatrix}$

Source: Own elaboration

The distribution of school and classroom estimated effects, followed by the convergence plots, are in Figure 4. Notice that, after the inclusion of significant variables at school level, the school effect residual is almost insignificant, meaning our explanatory variables are effective at explaining variation due to schools in the model. The variables used to explain the classroom effects, however, were not so efficient. The trace plots for the final model show convergence of the estimated random effects, although the chains had some variability during the MCMC process.

Figure 4. Plots of Random Effects - Recife, 2017



The conditioning factors associated with students' achievement in Portuguese and Mathematics are described in Tables 5 and 6. Twelve of the twenty-four variables were retained to explain the variability of SAEB student scores in the municipality of Recife. This is larger than the average of eight explanatory variables identified in previous studies (Hattie (2012)).

At the school level, only three variables were significant from the credibility interval (CI) point of view. The variables which did not include zero in the CI were the Mechanism of Student Selection, the School Socioeconomic Level, and the Participation Rate of students in the test. The first variable is a proxy for a prior attainment measure at the student level and it represents the strongest effect in our model, both in Mathematics $CI_{95\%}=[22.99; 54.06]$ and Portuguese ($CI_{95\%}=[50.50; 80.04]$) for individuals that accessed schools by selection tests. There are only four schools in Recife municipality using this type of student selection mechanism, and between them, we have two schools associated with the Federal Educational System and two military schools, which are partially funded and administered by the Federal Educational System.

Table 5. School and classroom conditioning factors to proficiency - Recife 2017

Conditioning Factor	Mathematics		CI		Portuguese		CI	
	N	$\hat{\beta}_{ikj}$	L-95%	U-95%	$\hat{\beta}_{ikj}$	L-95%	U-95%	
β_0		128.94	106.63	151.93	143.31	119.58	166.65	
Participation Rate	11,769	0.72	0.48	0.96	0.60	0.35	0.84	
Socioeconomic Level (Group 4 or less)								
Group 5	115	11.97	7.66	33.72	19.94	2.13	41.53	
Group 6	87	45.86	26.07	67.40	41.56	20.14	63.44	
NR	270	4.49	-8.14	17.36	3.7	-10.19	17.13	
Student selection Process (Geographic/ Others)								
Test	322	65.63	50.50	80.04	38.87	22.99	54.06	
Type of Contract of Maths' Teacher (Others)								
State/Fix Contract	5,902	6.12	3.29	8.91	6.12	3.29	8.91	
NR	1,423	0.97	3.44	4.82	0.97	-3.44	4.82	

Source: Own elaboration

Table 6. Student level conditioning factors to proficiency - Recife 2017

Conditioning Factor	Mathematics		CI		Portuguese		CI	
	N	$\hat{\beta}_{ikj}$	L-95%	U-95%	$\hat{\beta}_{ikj}$	L-95%	U-95%	
Gender (Female)								
Male	5,523	14.47	12.97	15.98	-4.62	-6.25	-3.01	
NR	486	-5.93	-10.04	-1.78	-17.08	-21.39	-12.83	
Race (African (Black)/ Indigenous/Others)								
Europe (White)/Asia (Yellow)	2,713	2.47	0.7	4.22	5.19	3.37	7.04	
NR	340	-4.88	-10.15	0.39	-4.04	-9.53	1.35	
Leisure time (Less than 2 hours)								
More than 2 hours	7,182	13.91	12.37	15.47	15.76	14.08	17.36	
NR	310	15.01	7.7	22.26	14.61	7.06	22.46	
If Work (Yes)								
No	10,394	9.04	6.46	11.75	14.32	11.57	17.11	
NR	379	1.46	-5.16	7.81	9.92	3.02	16.86	
Year that accessed school (After Year 1)								
At Year 1	912	4.58	-0.77	9.87	11.19	5.56	16.66	
Before Year 1	9,996	11.80	6.94	16.46	16.77	11.81	21.75	
NR	572	10.43	4.10	16.78	12.47	6.04	19.0	
If does Mathematics Homework (Never/Almost Never)								
Sometimes	4,417	2.64	0.02	5.17	-1.82	-4.43	0.87	
Always/Almost Always	5,504	16.01	13.34	18.56	2.82	-0.07	5.44	
NR	392	-1.47	-7.69	5.04	-0.45	-7.12	6.44	
If does Portuguese Homework (Never/Almost Never)								
Sometimes	4,294	-3.0	-6.0	0.08	-0.32	-3.52	2.92	
Always/Almost Always	6,116	-2.47	-5.69	0.53	8.51	5.20	11.79	
NR	429	-1.41	-7.66	4.98	-3.35	-10.18	3.28	
Mother' Education (Don't know/Until Year 9)								
High School	3,796	9.96	8.3	11.53	11.59	9.84	13.27	
Diploma	1,104	8.72	6.12	11.52	7.70	4.82	10.6	
NR	228	2.68	-3.56	9.31	5.44	-1.72	12.01	

Source: Own elaboration

More than 70% of students enrolled in the four schools that practice student selection via an entry test achieved the minimum required scores in Mathematics and Portuguese, while the average proportions for Recife are 35% for Portuguese and 15% for Mathematics. A further investigation of SAEB data showed that teachers working in these schools have higher wages and qualifications, as shown in Table 7 for Mathematics Teachers. The concentration of more qualified teachers, with better wages, in these schools can explain why covariates (at the clas-

sroom level) were not important in our model. Considering that public schools aim to reduce inequalities in Brazil, the concentration of the “best teachers” in schools which apply tests for selecting “better students” does not seem a good practice.

Table 7. Distribution of variables at Mathematics teacher level

Mathematics Teacher' variables	Students Selection	
	Tests	Others
Income		
Less than 5 m.w.	150 (46.6%)	8,343 (72.8%)
Between 5 and 7 m.w.	-	1,276 (11.1%)
More than 7 m.w.	115 (35.7%)	349(3.1%)
NR	57(17.7%)	1,479(13%)
Qualification (Diploma)		
Area that teaches	265 (82.3%)	6,718 (58.7%)
Other areas	-	3,035 (26.5%)
Don't have diploma	-	172 (1.5%)
NR	57 (17.7%)	1,522(13.2%)

Source: Own elaboration

The covariate School Socioeconomic Level (SSL) is the second strongest effect in our model. It is also a proxy for the ESCS indicator from PISA. The ESCS indicator represents a combination of three variables: Index of Student home possessions, Highest parental education, and Highest occupational status of parents. The SSL indicator uses the same idea of the ESCS indicator but is calculated at the school level and is an average of the socioeconomic level of the school's students. The SSL indicator has eight categories but, in Recife municipality, the two highest levels do not occur. The level 5 represents Students that, in general, said that their homes have two or more bathrooms, three or more bedrooms, two or three TVs and goods such as computers, washing machines, car and freezer. They do not have a housekeeper; the family average income is between 2.5- and 7-times minimum wage (R\$937,00 in 2017); and their parents completed High school or University. Level 6 has housekeeper, their average income is between 7- and 20-times minimum wage and the parents have completed University. The other levels represent lower SSL.

Previous studies identified associations between peer effects and learning outcomes (Thrupp et al. (2002)). One of those affirms that "the effects of peers appear to be greater for low-ability students than for high-ability students" (Zimmer and Toma (2000)), which supports the importance of the heterogeneity of students' backgrounds for superior results in general student achievement. Once more, public policies should be addressed to spread students with different economic backgrounds in schools, which could be made by establishing quotas for certain groups, ensuring the desirable diversity at public schools – which is also an aim of the SDG 4 (Nations (2015)).

The last covariate at school level was the Participation rate. As expected, this factor is positively correlated with the student scores, which means that schools with higher participation rates in SAEB displayed better results in the tests. This hypothesis was tested in the model because we have a high rate of early abandonment of education for students at Year 9, where 7,6% do not go to Secondary School and 23,1% of 15-17 years old are enrolled in years below their ideal grade. These indicators were calculated from the Annual National Household Sample Survey 2019, from IBGE².

² <https://agenciadenoticias.ibge.gov.br/agencia-noticias/2012-agencia-de-noticias/noticias/25883-abandono-escolar-e-oito-vezes-maior-entre-jovens-de-familias-mais-pobres>

The fact that this factor was significant in our model indicates that schools with a low participation rate can be the same schools which have a high abandonment rate at the end of the year, which can be associated with poor learning environment. Unfortunately, it is not possible to evaluate this last hypothesis using SAEB data and we cannot say that this is exactly the reason behind the positive association between participation rate and student test scores.

At the Classroom level, besides the size of the classrooms (number of students in each class), Teacher's covariates were tested but only the factor "Type of contract" used to hire Maths teachers at schools was retained. As with the random effects, the fixed effects at the classroom level are univariate. This means that the same effect equally impacts in Portuguese and Mathematics scores.

A considerable percentage of teachers (43%) working in Recife's public schools did not have a firm contract. These teachers are temporary workers, and they work less hours and receive smaller wages than civil servants, as analysis of SAEB data indicates. This practice often induces them to accumulate jobs, dedicating less time to study and help students from one specific classroom or school. Once again, public policies can help to solve the "type of contract" problem in the form of a policy that would require teacher's exclusive dedication to one school, or in the form of a career plan for teachers who want to dedicate themselves to working in public schools. The NR categories for school and classroom variables are like the corresponding reference categories.

Turning to Student level factors, all the pre-selected covariates were important to explain the variance of student scores. Considering we did not use any individual economic factor at student level, the demographic variable "Race" was tested as a proxy for economic capacities in Brazil, especially because there is evidence that Black and Indigenous people have the worse wages and jobs (IBGE (2017)). Students who declare to be Black or Indigenous have lower expected proficiency scores for both subjects, consistent with their economic backgrounds.

Gender commonly appears as an important variable in proficiency studies, and the fact that girls perform better in Portuguese and worse in Mathematics than boys is not new in the literature (Parker *et al.* (2018)). The unsurprising result from our fitted model that boys are doing better in Mathematics, and girls are doing better in Portuguese, however, highlights a gender discrepancy problem that exists in Brazil in general form. There are already policy programs to address this issue and give equal opportunities to all genders in schools, companies, political parties etc. but the fact that children are still affected by this social problem is an indicative that it is a cultural pattern and so must be dealt with as such. Policy programs that involve parents in discussions about this problem can help to decrease the discrepancy we found here.

A curious significant factor in our model was the leisure time by children. According with the results, children who spent more hours with leisure (watching TV, using internet, or playing electronic games) have better achievement scores than the ones who spent less time with leisure. Although this seems contradictory, it can be explained by the direct relationship that it has with the access to internet and the possession of electronic games in Brazil. According to data from NIC.br (Brazilian Network Information Center), only 61% of urban households in Brazil had access to internet in 2017. The same survey presented data about possession of TV (97%), electronic games (19%) or tablets (18%)³. These indicators decrease according to economic classes and regions and provide evidence that families who have access to these facilities are the ones that have more economic capacities, which can explain why the leisure time factor has a positive association with achievement scores in Recife.

The indicator "If the student works" is another associated with economic background that was significant in our model. Three hundred seventy-nine (379) students in Year 9 answered that they were working in 2017 and presented lower average proficiency scores than students who were not working at that time. These children are 14-17 years old, living in an urban centre and still had not completed mandatory primary education. They should not be working, and efforts

³ https://data.cetic.br/cetic/explore?idPesquisa=TIC_DOM

should be made to ensure that they conclude their primary education before starting work. Some public policies such as "Bolsa Família" put "attending school" as a condition for the payment of financial aid to poor families, aiming to address this problem, but it was not enough to make students dedicate full time to their education. Our analyses shows that we must advance in proper programs to solve this socioeconomic issue.

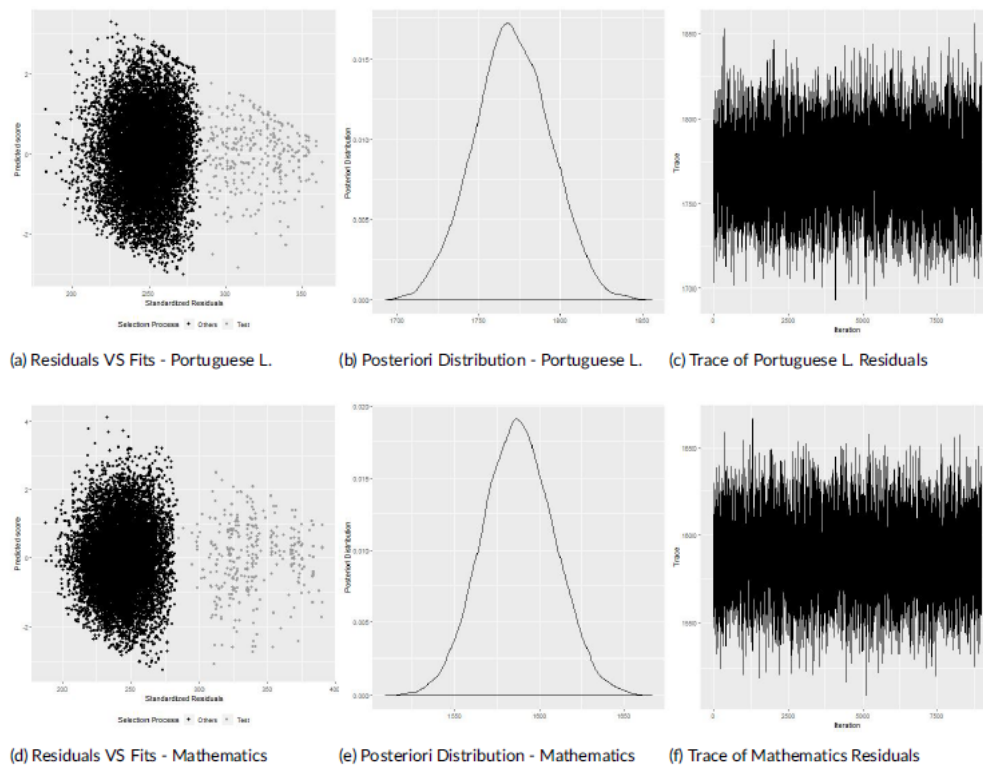
The last few covariates, "Year of access to school", "If student does homework" and "Mother's education" are well known as important conditioning factors - see (Hattie (2012)) - to explain the variability of students achievement scores and they presented expected associations with the proficiency scores, reinforcing that students who access schools early, do homework and have a mother with good education do better.

Different from what happened with school and classroom fixed effects, NR categories had significant effects for some variables at the student level. Students who did not answer the questions about Gender and Race, for example, had negative effects in scores from both subjects. This can be related with the fact they do not accept the INEP's categorisation for these questions, or it could mean they did not understand the questions. On the other hand, NR category had positive effect in variables regarding to Work, Leisure time and Access to School. The first two can be related with a lack of explanation about the concepts of "Work" and "Leisure" in the SEQ. Other questions had more evident problems in their formulation and categories, as we already discussed, but the significant NR in the covariates tested in the model can indicate that these questions also have problems in their formulation. The significant NR category in "Access to School", in its turn, can be explained by the lack of memory of students, since the access to school usually occurs when they are very young. The remaining student covariates have NR effects like the reference categories.

Although we estimated 61 parameters in the final model, there remains a large residual variance attributed to the student level, as we can see in Table 4. The absence of other variables related with socioemotional capabilities, cultural and social environment and latent characteristics associated with good practices of studying (as motivation and focus), as well as a direct control for pupil level prior attainment, is an explanation for the persistent large variance of student level residuals. We suggest that questions on these topics should be included on SAEB SEQ and enabling the possibility of linking datasets in the public-use files to support better control for pupil-level prior attainment.

Figure 5 shows that our residuals, calculated using Equation 3, appear to be normally distributed around zero, both for Portuguese and Mathematics. We can clearly see two groups of observations in Graphs a) and d) representing students in categories of the variable "Student's selection process", as expected.

Figure 5 - Residual graphs for BMM - Recife, 2017



Concluding remarks

We aimed to identify conditioning factors of learning in Brazil using a BMM fitted to SAEB data. We chose the municipality of Recife to illustrate the use of MCMC approach to fit the model and found twelve covariates from school, classroom and student levels associated with student achievement in Portuguese and Mathematics. The fitted model revealed the impact of the hierarchical structure of SAEB data in the estimation process, as well as the correlation between test results in Portuguese and Mathematics.

Our results confirm associations already found in other studies between different conditioning factors and student proficiency scores, such as the Age that students entered school, Mother's education, positive impacts of Homework and full-time dedication to studies. Some effects, such as the ones associated with gender and race of students are highly associated with socioeconomic conditions in Brazil, explaining their significance in the model. From the point of view of public policies, covariates associated with students are difficult to change but affirmative policies related to decreasing the effects of socioeconomic differences can help.

The conditioning factors associated with policy decisions popped-up in the model at Tables 5 and 6. The first one indicates that higher Participation rates are associated with increased student proficiency. In other words, if a school has a low participation rate in SAEB, its students had smaller proficiency scores. Despite the causes of the low participation of some schools, a public policy that encourages students to participate in SAEB tests would help to improve the quality of model estimates, since they could decrease the non-response rate.

The Student Selection Mechanism factor is also important, since it revealed that public schools that use tests to select students with better prior learning outcomes present a sizeable increase in average proficiency scores. These schools are also the schools with better funding from the Brazilian government. Considering that public schools should reduce inequalities, to apply tests to select "better students" does not seem a good practice to adopt. The diversity of

social and financial conditions of peers can help in improving learning outcomes and should be encouraged by public policies.

Conversely, we found little evidence that the teachers have much direct impact on student performance, either through direct inclusion of teacher level covariates in the model or in the unmeasured classroom level random effect. The only significant factor in our model, the Type of contract used to hire teachers, is also a policy problem, since there is not a proper law to establish which qualification or hours of work a public-school teacher must follow.

The evidence that some of the conditioning factors tested are strongly associated with proficiency outcomes can be used to support decisions to prioritize actions from governments (Municipal, State and Federal) to improve the quality of education. Considering that the approach we proposed here can be used to evaluate SAEB results in other municipalities or domains besides Recife and adding the fact that few quantitative analyses regarding factors affecting quality of public education in Brazil are available, this article contributes to the development of this research area and also to public policy planning for improving quality of education in Brazil.

References

- Barbetta, P. A., Andrade, D. F., and Tavares, H. R. (2018). *Estudo de fatores associados através de regressão quantílica hierárquica*. Estudos em Avaliação Educacional, 29(71):320–349.
- Bof, A. M. (2016). *A aprendizagem dos alunos e os desafios do PNE*. Série PNE em Movimento.
- Brasil (1996). *Lei no 9.394, de 20 de dezembro de 1996*.
- Ferrão, M. E., de Freitas Barros, G. T., Bof, A. M., and de Oliveira, A. S. (2018). *Estudo longitudinal sobre eficácia educacional no brasil: comparação entre resultados contextualizados e valor acrescentado*. Dados-Revista de Ciências Sociais, 61(4):265–300.
- Fielding, A. and Yang, M. (2005). *Generalized linear mixed models for ordered responses in complex multilevel structures: effects beneath the school or college in education*. Journal of the Royal Statistical Society: Series A (Statistics in Society), 168(1):159–183.
- Freire, P. (2011). *Pedagogia do Oprimido*. Rio de Janeiro: Paz e terra.
- Goldstein, H. (2003). *Multilevel modelling of educational data*. In Methodology and epistemology of multilevel analysis, pages 25–42. Springer.
- Hadfield, J.D. et al. (2010). *MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package*. Journal of Statistical Software, 33(2):1–22.
- Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. Routledge.
- IBGE (2017). *Síntese de indicadores sociais: uma análise das condições de vida da população brasileira, 2017*. No. 32. IBGE, Coordenação de População and Indicadores Sociais.
- INEP (2020). *Escalas de proficiência do SAEB*.
- Ioschpe, G. (2016). *A ignorância custa um mundo: o valor da educação no desenvolvimento do Brasil*. Objetiva.
- Leckie, G. and Goldstein, H. (2009). *The limitations of using school league tables to inform school choice*. Journal of the Royal Statistical Society: Series A (Statistics in Society), 172(4):835–851.
- Masci, C., Ieva, F., Agasisti, T., and Paganoni, A. M. (2017). *Bivariate multilevel models for the analysis of mathematics and reading pupils' achievements*. Journal of Applied Statistics, 44(7):1296–1317.
- Ntzoufras, I. (2011). *Bayesian modeling using WinBUGS*, volume 698. John Wiley & Sons.
- Parker, P. D., Van Zanden, B., and Parker, R. B. (2018). *Girls get smart, boys get smug: Historical changes in gender differences in math, literacy, and academic social comparison and achievement*. Learning and Instruction, 54:125–137.
- Pereira, J. B., Nobre, W. S., Silva, I. F., Schmidt, A. M., et al. (2020). *Spatial confounding in hurdle multilevel beta models: the case of the brazilian mathematical Olympics for public schools*. Journal of the Royal Statistical Society Series A, 183(3):1051–1073.
- Raudenbush, S.W. and Willms, J. (1995). *The estimation of school effects*. Journal of educational and behavioral statistics, 20(4):307–335.
- Ribeiro, D. (1984). *Nossa escola é uma calamidade*. Salamandra Editora.
- Team, R. C. (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Thrupp, M., Lauder, H., and Robinson, T. (2002). *School composition and peer effects*. International journal of educational research, 37(5):483–504.

United Nations (2015). *Transforming our world: the 2030 agenda for sustainable development*. Division for Sustainable Development Goals: New York, NY, USA.

Yang, M., Goldstein, H., Browne, W., and Woodhouse, G. (2002). *Multivariate multilevel analyses of examination results*. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 165(1):137–153.

Zimmer, R. W. and Toma, E. F. (2000). *Peer effects in private and public schools across countries*. *Journal of Policy Analysis and Management: The Journal of the Association for Public Policy Analysis and Management*, 19(1):75–92..

Appendix

List of potential explanatory variables considered in the modelling of proficiency scores

Level	Variable	Categories
school	Administrative Dependence	Federal; State; Municipal
	Socioeconomic Level	Level 1 to 7 (Low to High)
	Student Selection Mechanism	Test; Geographic; Arrival order; Draw; Others
	Participation Rate	72 - 100%
	Indicator of Education(Principal's responses)	Does not have diploma; Diploma of bachelor; Other diplomas
	Wage range (Principal's responses)	Less than 5 minimum wages; Between 5 and 7 m.w; More than 7 m.w
	Years working with education (Principal's responses)	Less than 6 years; Between 6 and 10 years; More than 10 years
classroom	How are principals appointed (Principal's responses)	Indication; Public competition; Others
	Estimated Class size	14-60
	Indicator of Education(Portuguese Teachers' responses)	Does not have diploma; Diploma in Portuguese L. ; Other diplomas
	Wage range (Portuguese Teachers' responses)	Less than 5 minimum wages; Between 5 and 7 m.w; More than 7 m.w
	Type of contract (Portuguese Teachers' responses)	State; Fixed Term; Temporary
	Indicator of Education(Mathematics Teachers' responses)	Does not have diploma; Diploma in Mathematics; Other diplomas
	Wage range (Mathematics Teachers' responses)	Less than 5 minimum wages; Between 5 and 7 m.w; More than 7 m.w
students	Type of contract (Mathematics Teachers' responses)	State; Fixed Term; Temporary
	Gender	Male; Female
	Race/Color	White;Yellow; Black; Indigenous; Pardo
	Indicator for working	Yes; No
	Indicator for Mathematics homework	Always;Almost always; Sometimes; Never;Almost never
	Indicator for Portuguese L. homework	Always;Almost always; Sometimes; Never;Almost never
	Mother 6 Education	Until Year 9; High School; Diploma; Don't know
	Leisure Time	More than 2h; Less than 2h
Indicator for Year that student entered school	Before Year 1; At Year 1; After Year 1	
Indicator for Repetition	None; Once; Twice or more	

Equipe técnica

Escola Nacional de Ciências Estatísticas

Jorge Abrahão de Castro

Colaboradores

Revisão dos capítulos

Alinne de Carvalho Veiga (IBGE/ENCE)
Amanda Mendonça (UERJ)
Ana Carolina Radd Lima
Ana Carolina Soares Berto (IBGE/ENCE)
Andrea Diniz da Silva (IBGE/GRI)
Angelita Alves de Carvalho (IBGE/ENCE)
Barbara Cobo Soares (IBGE/ENCE)
Bianca Walsh (ENCE/IBGE)
Carlos Henrique Leite Corseuil (IPEA)
Cayo de Oliveira Franco (IBGE/DGC)
César Augusto Marques da Silva (IBGE/ENCE)
Cimar Azeredo Pereira (IBGE)
Cristiano Ferraz (UFPE)
Danielle Carusi Machado (UFF)
Denise Britz do Nascimento Silva (IBGE) (aposentada)
Denise Lobato Gentil (UFRJ)
Denise Guichard Freire (IBGE/DPE)
Eduardo Campana Barbosa (UFV)
Fernanda Siqueira Malta (IBGE/DPE)
Fernando Henrique Ferreira de Oliveira
Gustavo da Silva Ferreira (IBGE/ENCE)
Igor Cavallini Johansen (UNICAMP)
Joice de Souza Soares (IBGE/ENCE)
Júlia Celia Mercedes Strauch (IBGE) (aposentada)
Juliana Vasconcelos de Souza Barros (UFMG)
Luana Junqueira Dias Myrrha (UFNR)
Marcos Dantas Hecksher (IPEA)
Miguel Antonio Pinho Bruno (IBGE/ENCE)
Octávio Costa de Oliveira (IBGE/COAGRO)

Paula Alves de Almeida
Rafael Chaves Vasconcelos Barreto
Raphael Villela Almeida
Ricardo de Sampaio Dagnino (UFRGS)
Romay Conde Garcia (IBGE/ENCE)
Silvana Nunes de Queiroz (URCA)
Wagner Lopes Soares (IBGE/DPE)
Wasmália Bivar (IBGE) (aposentada)
Zelia Magalhaes Bianchini (IBGE) (aposentada)

Projeto Editorial

Centro de Documentação e Disseminação de Informações

Coordenação-Geral

José Daniel Castro da Silva

Coordenação de Produção Editorial e Gráfica

Mauro Emílio Araújo

Gerência de Sistematização de Conteúdos Informacionais

José Marcio Batista Rangel

Gerência de Editoração

Estruturação e diagramação textual

Luiz Paulo do Nascimento

Programação visual da publicação

Luiz Carlos Chagas Teixeira

Gerência de Gráfica

Newton Malta de Souza Marques

Gerência de Impressão, Acabamento e Logística

Edmilson Ramos Raya

Se o assunto é **Brasil**,
procure o **IBGE**.



www.ibge.gov.br 0800 721 8181

POPULAÇÃO, TERRITÓRIO E ESTATÍSTICAS PÚBLICAS

25 anos da Pós-Graduação

stricto sensu da ENCE

Em comemoração aos 25 anos da Pós-Graduação em População, Território e Estatísticas Públicas da Escola Nacional de Ciências Estatísticas - ENCE, o presente livro reúne, em cinco seções, uma coletânea de artigos que discorrem sobre a exitosa jornada desse Programa, bem como sobre os seus reflexos nas trajetórias acadêmica e profissional dos alunos egressos dos Cursos de Mestrado e Doutorado, e apresenta resultados de alguns estudos realizados, mais recentemente, no âmbito de suas linhas de pesquisa.

A seção de abertura realça importantes conquistas do Programa, com destaque para o alcance da nota 5 que lhe foi atribuída pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES em 2023, e traz evidências sobre a sua relevância, tanto na formação de profissionais com alta qualificação quanto na inserção destes no mercado de trabalho.

A segunda seção aponta os desafios enfrentados no processo de produção de estatísticas públicas e as suas diversas condicionantes; discute a natureza do conceito de “população de difícil acesso”, com uma proposta de extensão de tal conceito com vistas à melhor compreensão das dificuldades enfrentadas na coleta de informações em contextos complexos; e mostra os limites e as possibilidades de utilização dos dados da histórica Plataforma Nilo Peçanha - PNP para pesquisa sobre a atuação da Rede Federal de Educação Profissional, Científica e Tecnológica - RFEPC.

A terceira seção abarca temas essenciais para o desenvolvimento sustentável e a formulação de políticas públicas e oferece contribuições para a estimação de indicadores e o monitoramento de metas globais, com foco nos ODS 6.1 e 11.7.1, além de estudos sobre a expansão, a interiorização e os desequilíbrios entre a oferta e a demanda da educação superior no País.

A quarta seção se dedica a análises setoriais e de gênero no mercado de trabalho brasileiro em variados períodos, com base em estatísticas institucionais provenientes da Pesquisa Nacional por Amostra de Domicílios - PNAD, encerrada em 2015, e da PNAD Contínua, cuja série teve início em 2012.

A última seção contempla outros estudos também abarcados pelas linhas de pesquisa do Programa, entre os quais se destaca uma proposta metodológica para construção de um sistema de indicadores para mensurar a aplicação do conceito de cidades inteligentes aliado ao bem-estar social, considerando-se para tal as municipalidades brasileiras com mais de 100 000 habitantes no Censo Demográfico 2010.

Este livro espelha o esforço contínuo de reflexão sobre o desenvolvimento do Programa de Pós-Graduação *stricto sensu* da ENCE. Espera-se que as análises e contribuições ora formuladas inspirem novas pesquisas, políticas públicas mais eficazes e uma valorização ainda maior das estatísticas como ferramenta fundamental para a transformação social.

