

Documentos para Disseminação

Memória Institucional

26



POPULAÇÃO, TERRITÓRIO E ESTATÍSTICAS PÚBLICAS

20 anos do Programa
de Pós-Graduação da ENCE

Presidente da República
Jair Messias Bolsonaro

Ministro da Economia
Paulo Roberto Nunes Guedes

Secretário Especial de Fazenda
Waldery Rodrigues Junior

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE

Presidente
Susana Cordeiro Guerra

Diretora-Executiva
Marise Maria Ferreira

ÓRGÃOS ESPECÍFICOS SINGULARES

Diretoria de Pesquisas
Eduardo Luiz G. Rios Neto

Diretoria de Geociências
Claudio Stenner

Diretoria de Informática
Carlos Renato Pereira Cotovio

Centro de Documentação e Disseminação de Informações
Carmen Danielle Lins Mendes Macedo

Escola Nacional de Ciências Estatísticas
Maysa Sacramento de Magalhães

UNIDADE RESPONSÁVEL

Escola Nacional de Ciências Estatísticas

Ministério da Economia
Instituto Brasileiro de Geografia e Estatística - IBGE
Escola Nacional de Ciências Estatísticas

Documentos para Disseminação

Memória Institucional 26

População, Território e Estatísticas Públicas

20 anos do Programa
de Pós-Graduação da ENCE



Rio de Janeiro
2020

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 978-65-87201-34-4

© IBGE. 2020

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

Este volume foi organizado por César Marques, coordenador do Programa de Pós-Graduação da Escola Nacional de Ciências Estatísticas - ENCE entre 2016 e 2020.

Capa

Marcos Balster Fiore - Coordenação de *Marketing*/Centro de Documentação e Disseminação de Informações - CDDI

Ilustração

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

**Ficha catalográfica elaborada pela Gerência de Biblioteca e
Acervos Especiais do IBGE**

População, território e estatísticas públicas: 20 anos do Programa de Pós-Graduação da ENCE / Escola Nacional de Ciências Estatísticas ; [organizado por César Marques]. - Rio de Janeiro: IBGE, 2020.
227p. - (Documentos para disseminação. Memória institucional, ISSN 0103-6459; n. 26).

ISBN 978-65-87201-34-4

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

CDU: 314.1
SOC

Impresso no Brasil/*Printed in Brazil*

Sumário

- 5 **Apresentação**
- 7 **Introdução**
César Marques
- Seção 1
Produção de estatísticas públicas e condições de vida
- 25 **Tamanho dos municípios: estimativa populacional municipal comparada à enumeração censitária no Brasil em 2010**
Luiz André Ribeiro Zardo, Suzana Cavenaghi, Denise Britz do Nascimento Silva
- 47 **Economia Política da informação estatística: por uma linha de pesquisa sobre as conexões entre estatísticas e Políticas Públicas**
Paulo de Martino Jannuzzi, Luiz Marcelo Carvano, Leonardo Milhomem Rezende, Baiena Feijolo Souto
- 63 **Big Data na produção de estatísticas públicas: conceitos, oportunidades e desafios**
Rafael Bassegio Caumo
- 85 **Estimação de fluxos brutos das situações ocupacionais com pesquisas amostrais repetidas: uma aplicação na Região Metropolitana de Salvador**
Priscila Pagung de Aquino Lapa, Pedro Luis do Nascimento Silva, Maria Luíza Barcellos Zacharias
- 103 **Proposta de compatibilização de dados censitários para estudos longitudinais: o caso das áreas de ponderação do município do Rio de Janeiro nos Censos 2000 e 2010**
Vivian Alves da Costa Rangel Gomes

- 117 Indicadores de representatividade na pesquisa sobre o uso das Tecnologias de Informação e Comunicação nos domicílios brasileiros**
Mayra Pizzott Rodrigues dos Santos, Marcelo Trindade Pitta, Denise Britz do Nascimento Silva
- Seção 2
População e território
- 135 População, gênero e cor/raça na produção cinematográfica brasileira contemporânea**
Paula Alves de Almeida, José Eustáquio Diniz Alves, Denise Britz do Nascimento Silva
- 151 Breves considerações sobre a dinâmica dos espaços urbanos no mundo contemporâneo**
Letícia de Carvalho Giannella, Ana Carolina Chaves Católico, Annelize de Souza Pereira, Bruna de Castro Dias Bicalho, Cauan Braga da Silva Cardoso, Davi Bovolenta, Eloá Nascimento dos Santos, Raphael Henriques da Rocha, Ulisses Carlos Silva Ferreira
- 165 A contribuição das componentes demográficas nas projeções populacionais: um exercício para o Brasil incorporando hipóteses de imigração internacional**
Helena Nobre de Oliveira, César Augusto Marques da Silva, Antonio Tadeu Ribeiro de Oliveira
- 177 Morte e vida dos barnabés do IBGE**
André Bruno de Oliveira, Kaizô Iwakami Beltrão
- 195 Determinantes do tempo de aleitamento materno exclusivo no Brasil em 2006**
Letícia Batista de Paula Barros, Angelita Alves de Carvalho
- 211 Fatores que influenciam a medida socioeducativa aplicada ao adolescente autor de ato infracional, na comarca da capital do Rio de Janeiro**
Andrea Diniz da Silva, Pedro Luis do Nascimento Silva

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

As mudanças do mundo atual têm ocorrido cada vez mais rapidamente, as quais têm uma relação direta com o desenvolvimento tecnológico. A tecnologia tem permitido o acesso à informação em frações de segundo, a coleta, o armazenamento, a análise de bases de dados enormes, a disseminação de informação e de conhecimento, em diferentes meios e formas, e estes são apenas alguns dos papéis que a mesma tem impactado nossas vidas.

Neste contexto de mundo veloz, de tecnologia trazendo avanços, mas também riscos, impondo desafios e demandando soluções, o nosso Programa de Pós-Graduação *stricto sensu* em População, Território e Estatísticas Públicas completou 20 anos em 2018.

Os 20 anos do Programa, que se iniciou com o Mestrado, não só demonstra a consolidação do trabalho que vem sendo feito ao longo desses anos na formação de profissionais capacitados para atuar na área de População, Território e Estatísticas Públicas, mas também indica o reputado papel que o mesmo representa para o IBGE e para a sociedade.

Os cursos de Mestrado e Doutorado apresentam uma matriz curricular moderna, uma vez que ao longo desses anos vem sendo atualizada. Uma reformulação do Mestrado foi feita em 2005, quando era designado Programa de Pós-Graduação em Estudos Populacionais e Pesquisas Sociais, e outra em 2014, com a aprovação do Doutorado pela CAPES. O Programa foca em questões relevantes aos processos que tratam da dinâmica demográfica, territorial e da produção de estatísticas públicas, os quais são importantes para o entendimento da sociedade contemporânea e para a melhoria das condições de vida da população. A busca pelo conhecimento adequado, atualizado, nas questões da área de concentração do Programa permeia a formação que é transmitida aos nossos alunos.

O Programa atrai tanto estudantes do Brasil quanto de outros países latino-americanos e de língua portuguesa, como Colômbia, Costa Rica, Moçambique, Cabo Verde, entre outros. Egressos do Programa atuam em diversos setores da sociedade, em empresas e em instituições acadêmicas.

Várias pessoas contribuíram com o seu trabalho, esforço, empenho, trazendo os seus saberes, dando o seu tempo, em suma, colaborando para a construção do nosso Programa, para chegarmos aonde estamos hoje. Ressalta-se a importância dos docentes, bem como o papel primordial dos coordenadores da pós-graduação que estiveram à frente do Programa neste período de 20 anos, trabalhando sempre para o aperfeiçoamento contínuo do mesmo, colocando-o em uma posição de destaque no cenário nacional. Neste período de 20 anos tivemos oito coordenadores. Foram coordenadores do Programa, pela ordem cronológica, os professores: Victor Hugo Gouvêa, Jane Souto de Oliveira, César Ajara, Aída Verdugo Lazo, José Eustáquio Alves, Maysa S. de Magalhães, Suzana Cavenaghi e, atualmente, César Marques.

A vitalidade dos cursos fica bastante clara no presente livro, que reúne diversos estudos e importantes metodologias, realizados por docentes, discentes e egressos do Programa.

Este livro, que é composto por 12 capítulos distribuídos ao longo de duas seções refletindo as linhas de pesquisa do Programa, nos traz uma imensa satisfação pois demonstra que a ENCE tem vindo a cumprir, com o Mestrado e o Doutorado, não só uma missão de formação de recursos humanos de qualidade como também de promoção do saber para a sociedade.

Maysa Sacramento de Magalhães

Coordenadora-Geral da Escola Nacional de Ciências Estatísticas

Introdução

População, Território e Estatísticas Públicas*

César Marques**

O Programa de Pós-Graduação *stricto sensu* da Escola Nacional de Ciências Estatísticas (ENCE/IBGE) surgiu em 1998, com o Mestrado em Estudos Populacionais e Pesquisas Sociais. Criado como resultado de uma trajetória de 30 anos de discussões sobre as transformações da ENCE, como um reflexo do esforço do IBGE e da Escola pelo fortalecimento do ensino, pesquisa e extensão no nível da pós-graduação, o programa tinha como proposta a associação de pesquisadores das áreas de demografia, pesquisas sociais e estatística em torno de uma pós-graduação única.

Na CAPES, o Programa se inseriu na grande área de Ciências Sociais Aplicadas, na área de Planejamento Urbano e Regional e Demografia, subárea de Demografia. Logo nos primeiros anos de atividade, o programa incorporou também as discussões territoriais, dinamizando suas atividades para a compreensão dos fenômenos analisados em múltiplas dimensões.

Iniciavam-se assim atividades que buscavam fomentar as interfaces entre as linhas e projetos de pesquisa, a interdisciplinaridade, a composição de um corpo docente com maior qualificação, a integração entre o ensino e a pesquisa, o estreitamento dos vínculos com as áreas de produção do IBGE e o aumento da visibilidade do programa no meio acadêmico e entre centros de pesquisa e planejamento. Nesse âmbito, a obra organizada pelos docentes Alves, Bittencourt e Senra (2009) traz a memória e as trajetórias dos seus primeiros 10 anos, relacionando a criação do Mestrado tanto ao contexto da ENCE como ao do próprio IBGE. Na introdução da obra os autores indicam que o programa tinha um papel estratégico para a relação entre ENCE e IBGE, dado o grande potencial de desenvolvimento técnico na integração entre ensino, treinamento e pesquisa com a produção de informações, principalmente nas áreas de metodologia estatística, demografia e geociências.

* A presente introdução foi feita a partir da experiência acumulada como docente e coordenador do programa, refletindo sua perspectiva do mesmo. O autor agradece a compilação dos dados feitas por Joice de Souza Soares, secretária acadêmica do programa, e também pela revisão atenta e comentários da docente Ana Carolina Soares Bertho. O resultado, no entanto, é de inteira responsabilidade do autor.

** Professor permanente do programa de pós-graduação da ENCE/IBGE desde 2014. Foi coordenador do programa de fevereiro de 2016 a setembro de 2020.

Para Neide Patarra (2006) o programa do Mestrado propiciou a ampliação do quadro docente da escola e a formação e consolidação de linhas de pesquisa relacionadas à inserção do mesmo no IBGE, fazendo face aos desafios institucionais de ensino e pesquisa no âmbito da ciência e tecnologia do país.

Após quase duas décadas de atividades e o acúmulo de atuação e experiência de pesquisa nas áreas de demografia, condições de vida, dinâmica territorial, metodologia estatística para Censos e pesquisas amostrais e acerca de sistemas de informação estatística e geográfica, a partir de 2010 o programa passou por uma reformulação visando o início do doutorado, discutida amplamente pelo colegiado dos docentes e junto à direção do IBGE. Foi assim que, em 2014, quando o programa já possuía cerca de 250 dissertações defendidas, foi aprovada a proposta de um curso de Doutorado, expandindo as atividades da pós-graduação e de pesquisa da ENCE. O curso foi modificado, incluiu o nível de Doutorado e foi renomeado para População, Território e Estatísticas Públicas.

A reformulação do programa representou uma importante evolução do mestrado existente, permitindo a formação mais aprofundada e mais ampla de pesquisadores e profissionais interessados na temática, completando assim o ciclo acadêmico oferecido pela ENCE.

Com isso, o curso tem uma proposta interdisciplinar, tomando como ponto de partida que o estudo e as análises de questões populacionais e territoriais contemporâneas não podem prescindir de uma abordagem que combine conhecimentos da Demografia, Sociologia, Geografia, Estatística e Economia, principalmente.

No âmbito mais geral da formação em pós-graduação no Brasil, duas características marcam as especificidades do Programa. A primeira é o esforço de articulação entre diferentes cortes disciplinares e enfoques metodológicos na concepção e na prática da pesquisa social, sobretudo no que tange aos fenômenos territoriais e populacionais. A segunda é a vinculação direta ao IBGE, propiciando uma formação acadêmica integrada com a prática de produção e análise das estatísticas públicas. Essa vinculação institucional possibilita acesso e interação com parte do corpo de pesquisadores que acumula vasta experiência na produção de estatísticas públicas do IBGE e de outros órgãos, facilitando a troca de experiência dos alunos e uma formação que preza pela realização de pesquisas e levantamentos de elevado rigor metodológico.

Em 2018 a pós-graduação da ENCE completou 20 anos de existência, chegando à marca de 348 dissertações de mestrado e 2 teses de doutorado defendidas (ambas com antecipação de prazo). Assim, ao todo, a ENCE/IBGE teve, em seus 20 anos de cursos de pós-graduação *stricto sensu*, a expressiva produção de 350 trabalhos concluídos, entre dissertações e teses.

Como fruto das atividades que marcam as duas décadas do programa de pós-graduação *stricto sensu* na ENCE/IBGE, essa obra traz alguns dos trabalhos que foram oriundos das pesquisas aqui desenvolvidas.

Além disso, essa introdução compila algumas das perspectivas e questões centrais para as áreas de pesquisa do programa, mostrando as tendências e alguns dos desafios atuais e futuros, pensando especificamente na formação que propomos, em **População, Território e Estatísticas Públicas**. Por fim também trazemos nossos próprios dados, com informações sobre discentes, defesas e docentes que ilustram parte dos resultados dessas duas décadas de existência.

O Programa

Com a evolução e consolidação do Programa, que teve na criação do Doutorado um importante marco, ocorreu aumento de sua visibilidade e, conseqüente, de seu reconhecimento. A publicação de trabalhos de docentes do corpo permanente e discentes do programa em anais de congressos (nacionais e internacionais), em periódicos (nacionais e internacionais) e em livros, bem como a participação dos docentes em congressos nacionais e internacionais, tem-se ampliado.

Em termos gerais, a proposta do programa é propiciar uma formação profunda nas pesquisas sobre População, Território e Estatísticas Públicas, tendo como grandes especificidades

ser parte de uma instituição de estatística e geografia, o IBGE; e tratar do tema populacional na interface com dinâmicas territoriais e com as estatísticas públicas, como questões que são efetivamente indissociáveis. Para isso sua estrutura se dá a partir de duas linhas de pesquisa: População, Território e Condições de Vida (1), e Produção de Estatísticas Públicas (2).

O cotidiano das atividades de ensino, pesquisa e extensão do programa é marcado ao mesmo tempo por uma ampla diversidade temática e pela centralidade da dinâmica social presente e futura, que, de forma mais ou menos direta, perpassa as pesquisas aqui desenvolvidas. Projetos que tratam, por exemplo, do planejamento e análise de amostras complexas, modelagem e análise de dados e avaliação de estimativas populacionais em pequenas áreas coexistem e dialogam com projetos sobre mortalidade, preferências de fecundidade, saúde sexual e reprodutiva, mobilidade espacial da população, financeirização dos espaços urbanos e questões ambientais. São, assim, temas que dialogam com uma série de dinâmicas do mundo contemporâneo.

A partir dessa diversidade, o programa busca fomentar o desenvolvimento da área de Planejamento Regional e Urbano e Demografia. Assume-se que as visões sobre as mudanças populacionais ocorrem a partir de um debate amplo, compondo perspectivas sociais que moldam políticas públicas, considerando que a demografia não é destino (CAMARANO, 2014). Nesse campo, a proposta do programa é tratar dessas questões com a complexidade e densidade que lhes são inerentes, com a profundidade que se espera de um programa de pós-graduação. Para isso, promove-se a compreensão e o diálogo com referenciais teóricos clássicos e contemporâneos, a construção de indicadores e metodologias adequadas e robustas, que ofereçam diagnósticos, avaliações e reflexões com fundamentação científica.

Embora não tenhamos a intenção de fazer um inventário sobre tais questões no âmbito desse texto, é interessante pontuar que, epistemologicamente, um desafio que permanece nos estudos populacionais diz respeito a deixar de pensar a população apenas em termos de crescimento e sim nas suas relações e contradições.

Considerando a perspectiva da obra organizada por Hauser e Duncan (1972), a demografia seria definida pela aplicação e desenvolvimento de métodos para a mensuração dos fenômenos populacionais e pela interação com outras ciências, visando compreender tais transformações. Tais caminhos foram reconhecidos, de modo geral, como da análise demográfica e dos estudos populacionais. Contudo, vários pesquisadores apontaram as contradições dessa perspectiva, que tinha um potencial de conferir uma suposta neutralidade técnica ao campo de estudos. Neide Patarra, ex-docente do programa, tratou da questão conjuntamente com outros autores em Duarte, Montali, Oliveira e Patarra (1985) por exemplo. Contudo, a própria tradição latino-americana das pesquisas enfatizou o caráter interdisciplinar da área, construindo um pensamento sobre população que olhava seu objeto de maneira específica. Nesse âmbito, vale ressaltar os esforços de Carmen Miró, no CELADE, Giorgio Mortara, no próprio IBGE, dentre tantos outros.

A dificuldade de definição dos limites do alcance da demografia ainda existe. Ojima (2017) aponta que nesse caminho seria necessário chegar às fronteiras da ciência demográfica para então vislumbrar o que estaria para além dessa. Havendo segurança desse recorte teórico-metodológico (do objeto de análise e dos limites de abrangência), a interlocução com outras ciências seria mais produtiva.

O contexto desse debate, contudo, é ainda mais amplo, sendo inscrito na dimensão das crises das metanarrativas da modernidade. Essas, também para os estudos demográficos, implicaram em uma crise que não é puramente teórica, mas também epistemológica (CANALES, 2007), pois seriam necessárias novas formas de olhar e compreender o fenômeno demográfico, com revisão do alcance da própria categoria população, em seus usos, significados e domínios. No âmbito do processo de globalização e de constituição da sociedade na atual fase da “modernidade”, Canales argumenta a partir do marco teórico formulado por Ulrich Beck que seria importante problematizar a construção de saberes do tempo histórico presente, já que, nos termos que utiliza, “*a sociedad moderna no tiene formas de pensarse em términos de su globalización, pues las*

categorías y conceptos usados para su entendimiento están empapados de un nacionalismo metodológico” (CANALES, 2004, p. 50).

Soma-se a isso que, na América Latina, onde a modernização foi descontínua, incompleta, subordinada e híbrida, as ciências sociais foram construídas com a crítica de modelos e abordagens predominantemente europeus e norte-americanos (CANALES, 2004).

Esse cenário fomentou que, no campo dos estudos populacionais, ocorresse uma ampla diversificação de temáticas e áreas de investigação. Por outro lado, como não houve na região uma reflexão que articulasse um “metadiscurso” demográfico, a diversificação poderia ser interpretada como uma atomização e desarticulação.

Assim, seria preciso pensar no posicionamento dos estudos populacionais em relação às mudanças da sociedade contemporânea, considerando, por exemplo, as dinâmicas da sociedade informacional, da sociedade de risco e da reflexividade da modernidade. Essa direção encontra paralelos com a abordagem do texto clássico de Camargo (1980), que trata da importância de considerar a dinâmica demográfica no âmbito do processo histórico em que se inscreve.

Na sociedade contemporânea, seguindo a argumentação de Caneles (2007), poderia haver uma posição crítica em relação à demografia tradicional com problematização de seus três enfoques: da transição demográfica (onde o crescimento populacional é uma das etapas no processo de modernização das estruturas sociais), da relação população e desenvolvimento (que, genericamente, seria obstáculo ou impulso à modernização), ou do estruturalismo histórico (onde o crescimento demográfico é a base para a geração de um excedente populacional, sendo a consequência demográfica mais direta a manifestação de uma das principais contradições do capitalismo, que se expressa na exclusão contínua de contingentes demográficos). Embora antagônicos, tais enfoques são marcados por uma ideia de população moderna, ainda privilegiando questões de crescimento populacional e com a população de forma abstrata, onde a unidade que é representada anula a diversidade social e histórica, presente em cada grupo e cada indivíduo. Tem-se assim uma categoria – população -, que aparece como um conceito político e ideológico (LE BRAS, 2000), que não incorpora o pensamento de classes e da diversidade, mas sim a perspectiva da soma dos indivíduos iguais e indiferenciados (CANALES, 2007).

A proposta desse autor não é a de negar o potencial heurístico da população como uma abstração numérica, mas sim não se enganar pela ilusão que ela cria. Indo além, os estudos populacionais teriam como objeto a compreensão e análises das heterogeneidades e das iniquidades entre esses “sujeitos demográficos”. A população não seria um mero agregado de indivíduos iguais, mas sim uma articulação de sujeitos sociodemográficos, marcada por tensões, conflitos e contradições.

O desafio para a demografia seria deixar de pensar a população somente em termos de crescimento e passar a estudá-la nas relações e contradições diversas, entre indivíduos, gerações, gêneros, etnias e entre o ser humano e a natureza. Assim, temos uma mudança de um escopo de pensamento que compreende, além da preocupação com a dinâmica demográfica e seus componentes, uma abordagem sobre as estruturas demográficas, contemplando as relações, a diferenciação e as desigualdades (CANALES, 2007).

No contexto de tais discussões, amplas em conteúdo e contradições, globais em abrangência, o PPG da ENCE tem buscado contribuir para a produção técnica e acadêmica nacional ao pensar nessas questões levando em consideração toda a densidade que lhes é inerente. A desafiadora proposta de articulação entre População, Território e Estatísticas Públicas se configura como uma possibilidade de desenvolvimento dessas reflexões, permitindo avanços não apenas em uma demografia das totalidades, mas também das desigualdades, incorporando as relações territoriais e espaciais que são fundamentais. Para isso, é necessário o desenvolvimento e aplicação de modelos que permitam, em um contexto de Big Data, escassez de recursos e produção ampliada de estatísticas públicas, elaborar os melhores métodos e abordagens quantitativas.

É nesse sentido que o programa tem crescido e atuado, em cada uma das linhas de pesquisa e na interface entre as mesmas. A definição das linhas de pesquisa foi feita a partir da definição dos objetivos e da experiência acumulada no programa, refletindo a interdisciplinaridade

constante nos resultados das dissertações e, agora, teses. Todos os temas abordados no programa são classificados em uma das duas linhas de pesquisa. As linhas de pesquisa criadas foram denominadas “População, Território e Condições de Vida” e “Produção de Estatísticas Públicas”.

População, Território e Estatísticas Públicas

Tem como foco o estudo da dinâmica populacional e territorial e das condições de vida da população, englobando aspectos sociais, econômicos e ambientais. Inclui a utilização de indicadores e o emprego de métodos e modelos para análise da dinâmica populacional e territorial e das condições de vida, o monitoramento e a avaliação de políticas públicas. Para tal, explora as bases de dados existentes e realiza levantamentos de informações quantitativas e qualitativas, considerando aspectos conceituais e metodológicos, potencializando o uso das mesmas e observando seus limites.

Produção de Estatísticas Públicas

Engloba questões ligadas ao planejamento e à condução de levantamentos como censos, pesquisas amostrais e registros administrativos usados para a produção de estatísticas públicas que retratem o estado e a dinâmica populacional e territorial e as condições de vida da população. Inclui o desenvolvimento, a análise e a avaliação de métodos e modelos que contribuam para o avanço da produção de estatísticas públicas, da construção de indicadores e de sistemas de indicadores.

Teses, dissertações e integração entre linhas de pesquisa

A construção de um sistema estatístico com desenvolvimento de indicadores sociais e econômicos e que permita a implementação e o monitoramento de políticas públicas é central para a própria dinâmica social (JANNUZZI, 2018). A partir desse ponto de vista, pode-se dizer que as próprias conquistas sociais relacionadas a promoção da educação, redução da pobreza e avanços da proteção social foram ancoradas na produção de estatísticas públicas de qualidade.

Para tal, o desenvolvimento de pesquisas na interface entre produção de estatísticas públicas e análises de condições de vida, a partir das dinâmicas populacionais e territoriais, não é trivial. Essa abordagem demanda conhecimento de áreas com suas próprias tradições, metodologias, e, inclusive, epistemologias. Contudo, a busca por tal interação foi feita ao longo das duas décadas do programa e permanece.¹

Adicionalmente, a interação entre a produção do programa e a de estatísticas públicas, com o potencial de uso dos resultados de teses e dissertações no cotidiano das pesquisas e levantamentos do próprio IBGE, também é um desafio. Além desse potencial, a formação dos discentes propiciou importantes avanços para o Instituto, seja porque funcionários de carreira aprofundaram conhecimentos e buscaram nos cursos de pós-graduação da ENCE o caminho para a superação de desafios e aperfeiçoamento de processos que são parte das atividades cotidianas do IBGE, seja pelos egressos que ingressaram posteriormente no instituto. Em ambos os casos egressos da ENCE passaram a ocupar posições estratégicas, tanto no planejamento técnico como em posições gerenciais.

No programa de PPG da ENCE, principalmente em teses e dissertações, os discentes são estimulados à especialização em uma das linhas de pesquisa. Contudo, o diálogo entre as linhas também pode ser visto em algumas experiências concretas de pesquisa, o que constitui simultaneamente um potencial e um desafio. A seguir, nos exemplos de trabalhos de conclusão do programa que serão apresentados, temos, de modo mais latente ou mais explícito, dependendo do caso, tanto a integração entre as linhas como a interface com as atividades do IBGE.

¹ Até 2013 as linhas de pesquisa vigentes refletiam outra concepção de programa, no então Mestrado em Estudos Populacionais e Pesquisas Sociais. Naquele momento haviam 4 linhas de pesquisa: Dinâmica populacional, Condições de vida e Políticas Públicas; Metodologia Estatística para Censos, Pesquisas Amostrais e Registros Administrativos; Sistemas de Informação Estatística e Geográfica; e Sociedade, Economia e Território no Brasil Contemporâneo.

Esse é o caso de trabalhos que fomentaram o desenvolvimento da área de amostragem. Nessa, podemos citar as dissertações de Correa (2001), que trata da modelagem estatística em amostras complexas e produz indicadores de condições de vida que considerem a estrutura hierárquica das pesquisas sociais; de Rodrigues (2003) e Arantes (2012), sobre metodologias alternativas de amostragem para o Censo Demográfico; de Gonzalez (2005), propondo uma metodologia de amostragem para os registros administrativos da Previdência Social, de Lopes (2002), que analisa o desgaste do informante em pesquisas longitudinais, caracterizando seus determinantes na Pesquisa Mensal de Emprego (PME) e, por fim, de Freitas (2002), Rosseti (2013) e Junior (2015), que tratam das estimativas de emprego em pesquisas amostrais.

No âmbito das análises demográficas, podemos destacar trabalhos acerca das três variáveis básicas da demografia, tanto no âmbito das análises das condições de vida como focados no desenvolvimento de estimativas mais precisas. Assim, mortalidade (RAMOS, 2001; FIALHO, 2003; SOUZA, 2007; LÉON, 2015; SILVA, 2015; BARROSO JR., 2018), fecundidade (CORRÊA, 2004; LIMA, 2007; BARBOSA, 2008) e migração (OLIVEIRA, 2003; VIDAL, 2009; LIMA, 2017; OLIVEIRA, 2017) estiveram na agenda do programa ao longo dessas duas décadas. Ademais, a construção de projeções e estimativas populacionais, centrais para o entendimento das dinâmicas populacionais contemporâneas e futuras, também foram incorporadas em dissertações (BRITO, 2007; FRANÇA, 2012; ZARDO, 2013) e inclusive na primeira tese defendida no programa (BARROS, 2017).

Uma terceira área que ilustra as especificidades do programa traz abordagens sobre o próprio papel das estatísticas públicas no âmbito da sociedade moderna, como é o caso dos trabalhos de Santos (2004), Fonseca (2005), Santos (2006), Gama (2008) e Bello (2013).

O programa também apresentou dissertações que contemplaram as discussões territoriais e geocientíficas, de suma importância na contemporaneidade. Logo nos seus primeiros anos, a dissertação de Santos (2001) avaliou a qualidade de modelos digitais de terreno construídos por diferentes fontes de dados e Pereira (2003) debateu a delimitações de áreas para gestão ambiental a partir de bacias hidrográficas. Mais recentemente, Gonçalves (2014) utilizou ferramentas da geoestatística para análise da segregação socioespacial e Bastos (2016) caracterizou a vulnerabilidade à dengue a partir de indicadores dispostos em um Sistema de Informação Geográfica (SIG). Mais especificamente na discussão territorial, temas centrais para o mundo contemporâneo também foram incorporadas, como a reflexão sobre fronteiras (BRITO, 2006), políticas públicas e território (BERBAT, 2007), relação urbano-rural (PASSOS, 2005; GUEDES, 2017) e desigualdade socioespacial (OLIVEIRA, 2006; MELO, 2011).

Vale também ressaltar a atualização contínua das questões tratadas no programa em termos de desenvolvimentos metodológicos. Como exemplos, há dissertações que fomentam o uso de *data mining* (TORRES, 2003), analisam paradados de pesquisas censitárias para investigação de erros não amostrais (DUARTE, 2014), acompanham as inovações no IBGE, como a implementação da PNAD Contínua (SANTOS, 2017), que discorrem sobre metodologias para construção de índices de preços (SIMÕES, 2009; OLIVEIRA, 2018) e que inovam na mensuração da desigualdade (HECKSHER, 2017).

Por fim, observando as primeiras teses concluídas no programa, tanto o trabalho de Barros (2017) como o de Silva (2018) dialogam com temas que ganham espaço na agenda do IBGE. O primeiro, propondo novas metodologias para a construção de estimativas populacionais municipais nos períodos intercensitários, e o segundo, com o uso de registros administrativos e metodologias de pareamento para produção de estatísticas públicas.

Evidentemente tais exemplos cobrem apenas parte da produção temática do PPG, já que consideram apenas alguns dos resultados expressos diretamente em teses e dissertações. Há contribuições de relevo também em outros campos, dentre os quais podemos citar o da educação, dos debates de gênero, mercado de trabalho, desenvolvimento sustentável, epidemiologia, envelhecimento populacional, dentre outros. Contudo, essa breve exposição já traz exemplos de como os temas estiverem presentes, em diversos momentos da existência do programa.

Corpo docente

Um dos desafios inerentes a qualquer programa de pós-graduação está na contínua adequação e maturidade acadêmico-científica do corpo docente. Segundo a última avaliação feita pela CAPES (em 2017, relativa ao período 2013-2016)², os programas devem buscar ter um perfil de corpo docente adequado com a sua proposta, que busque aprimoramento contínuo, com dedicação equilibrada às atividades de pesquisa e formação e que demonstre capacidade de captar recursos para pesquisa. Nesse caso, é esperado que todos se dediquem de modo equilibrado às atividades do programa, evitando a concentração de atividades (publicações, orientações, aulas) em poucos docentes.

O Quadro 1 mostra a evolução recente do corpo docente credenciado no programa por ano de atuação, dimensionando o desafio corrente e futuro da estabilidade e da renovação.

Quadro 1 - Docentes com atuação no programa - 2012-2018

Docentes	2012	2013	2014	2015	2016	2017	2018
Denise Britz do Nascimento Silva							
José André de Moura Brito							
Julia Celia Mercedes Strauch							
Maria Salet Ferreira Novellino							
Maysa Sacramento de Magalhães							
Miguel Antonio Pinho Bruno							
Paulo de Martino Jannuzzi							
Pedro Luis do Nascimento Silva							
Alinne de Carvalho Veiga							
Ana Carolina Soares Bertho							
Angelita Alves de Carvalho							
Cesar Augusto Marques da Silva							
Fabio Giusti Azevedo de Britto							
Gustavo da Silva Ferreira							
Letícia de Carvalho Giannella							
Maurício Teixeira Leite Vasconcellos							
José Eustáquio Diniz Alves							
Suzana Marta Cavenaghi							
Sonoe Sugahara							
Cesar Ajara							
Moema de Poli Teixeira							
Aida Graciela Verdugo Lazo							
Marden Barbosa de Campos							
Maria Luiza Barcellos Zacharias							
Maria Luiza Guerra de Toledo							
Rogério dos Santos Seabra							

Fonte: Gerência de Pós-Graduação, ENCE/IBGE.

Em termos qualitativos, destacam-se as iniciativas de atuação integrada entre os docentes, que se reflete no alto número de coorientações, projetos de pesquisa em conjunto, coautoria de artigos; na atuação crescente em Sociedades Científicas e Técnicas correlatas às temáticas desenvolvidas no PPG; na incorporação de professores que atuam em demais diretorias do IBGE, na

² Os documentos de cada área da avaliação da CAPES, com os parâmetros da avaliação, estão disponíveis em: <http://avaliacaoquadrienal.capes.gov.br/documentos-de-area>.

regulamentação da participação de professores voluntários aposentados do IBGE no programa, que, dada a vasta experiência em suas áreas de expertise, continuam contribuindo e vinculados ao Programa), na regulamentação para o credenciamento de docentes e na busca por processos seletivos que permitam o ingresso de novos docentes.

Corpo discente

O corpo discente do programa é formado com a realização de processos seletivos anuais, que ofertam até 24 vagas para estudantes de Mestrado e até 10 vagas para o doutorado. Além de propiciar uma seleção rigorosa, baseada no mérito e qualificação, tal processo tem sido marcado também pela continuidade da diversidade do perfil dos ingressantes. Os ingressantes apresentam formação plural, sendo egressos de cursos em estatística, ciências sociais, geografia, engenharias e economia, dentre outros. Em termos regionais há uma significativa inserção de discentes do próprio Estado do Rio de Janeiro (um dos estados com maior número de egressos de cursos de graduação do país), além de estudantes da região Nordeste, do Sul e de outros Estados do Sudeste. Anualmente também há procura de candidatos estrangeiros no processo seletivo, principalmente de países da América Latina e pertencentes aos PALOP (Países Africanos de Língua Oficial Portuguesa). Em 2019 o programa contava com 3 estudantes estrangeiros cursando o doutorado. Em termos locais, muitos discentes já atuam profissionalmente em diversas instituições de pesquisa ou gestão e procuram o programa para a contínua qualificação. Dentre essas estão o próprio IBGE, IPEA, IPP, CETEM, Fundação CESGRANRIO, CEPEL, órgãos de ministérios, agências governamentais e Petrobras.

Como os Gráficos 1 e 2 mostram, nos processos seletivos a relação entre candidatos inscritos e candidatos aprovados têm sido, nos anos recentes, de aproximadamente 3 para 1. Em termos gerais, cerca de 80 inscrições são feitas (considerando mestrado e doutorado). Esse número pode ser avaliado como relativamente alto, dadas as especificidades e o nível dos cursos.

Considerando as pesquisas concluídas, até 2018 o programa titulóu 348 mestres e 2 doutores. Até 2020, as estimativas são de que 380 dissertações e 12 teses estejam concluídas. No doutorado, mesmo com um ciclo maior de formação (4 anos) e apenas 3 anos de existência completos em 2018, estima-se que o total acumulado de conclusões até 2025 possa superar as 50 teses. A evolução das conclusões está apresentada nos Gráficos 3 e 4, com a respectiva projeção para 2021.

Gráfico 1 - Candidatos inscritos e aprovados no mestrado da ENCE - 1998-2018

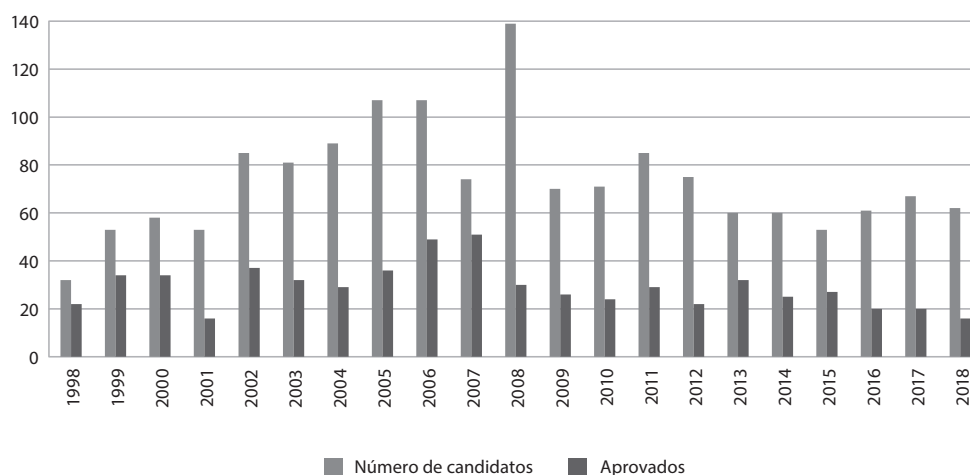
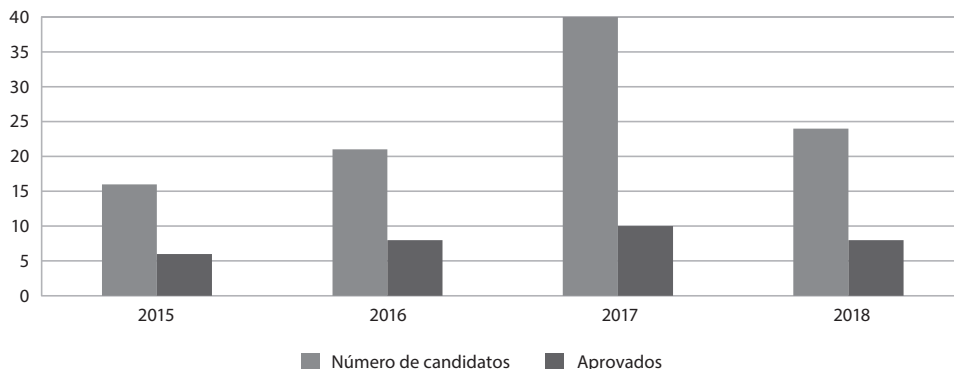


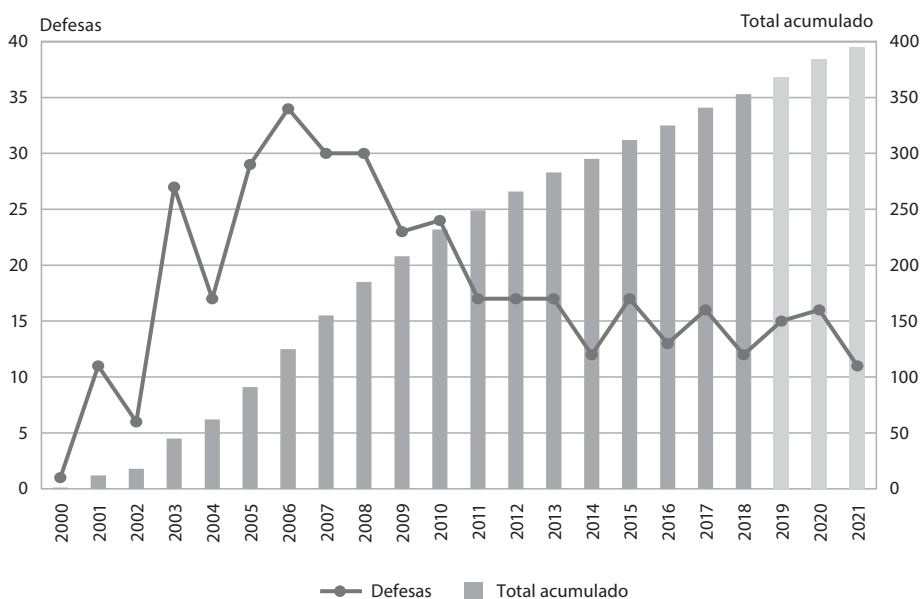
Gráfico 2 - Candidatos inscritos e aprovados no doutorado da ENCE - 2015-2018



Fonte: Gerência de Pós-Graduação, ENCE/IBGE.

Com a implementação do curso de doutorado, com teses já finalizadas e em andamento, nota-se que há uma capacidade de formação e orientação já bastante interessante, superando a dificuldade existente no início do curso do mestrado, quando havia um relativo baixo número de defesas (ALVES *et al.*, 2009). Após esse momento, entre 2003 e 2010, houve aumento do número de conclusões anuais devido a sobreposição de diferentes coortes, que concluíram o curso em um mesmo ano. Já em um terceiro momento, que perdura, o número de defesas de dissertação anuais foi estabilizado, considerando que não há mais sobreposição de coortes. Tal tendência também deve continuar dado o início do doutorado e o fomento a uma relação entre conclusões de mestrado e conclusões de doutorado próxima a 1 (como a CAPES preconiza). Adicionalmente, as atividades do curso de doutorado têm sido fundamentais para a ampliação da participação do corpo discente em eventos científicos da área, na realização de cursos de curta duração em outras instituições e nos chamados estágios “sanduíche”, nos quais uma parte do curso é realizada fora do país. Quatro estágios dessa modalidade estavam concluídos/em andamento até 2018, e outros 2 estavam em curso em 2019.

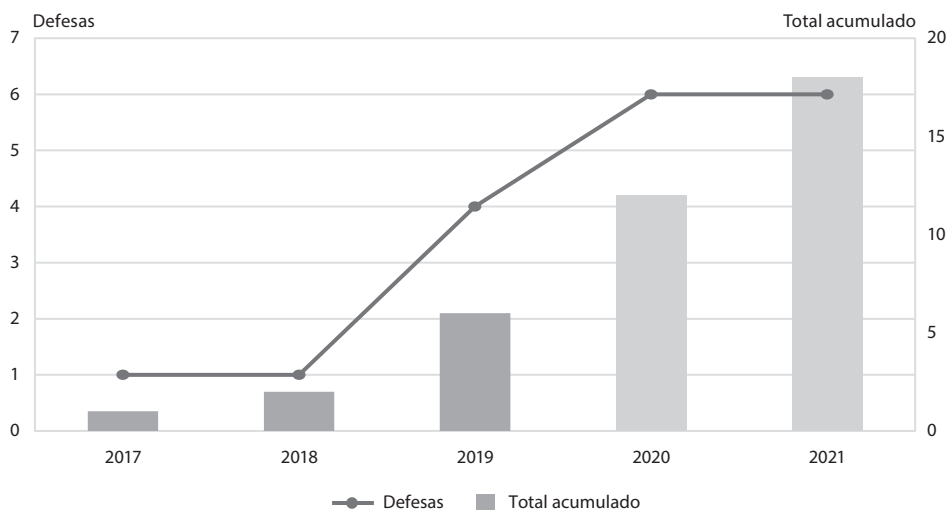
Gráfico 3 - Defesas de dissertações por ano e total acumulado de dissertações - 2000-2021



Fonte: Gerência de Pós-Graduação, ENCE/IBGE.

Nota: Os valores de 2019-2021 são projeções que consideram o número de alunos matriculados e o respectivo ano previsto de defesa.

Gráfico 4 - Defesas de teses por ano e total acumulado de teses - 2017-2021

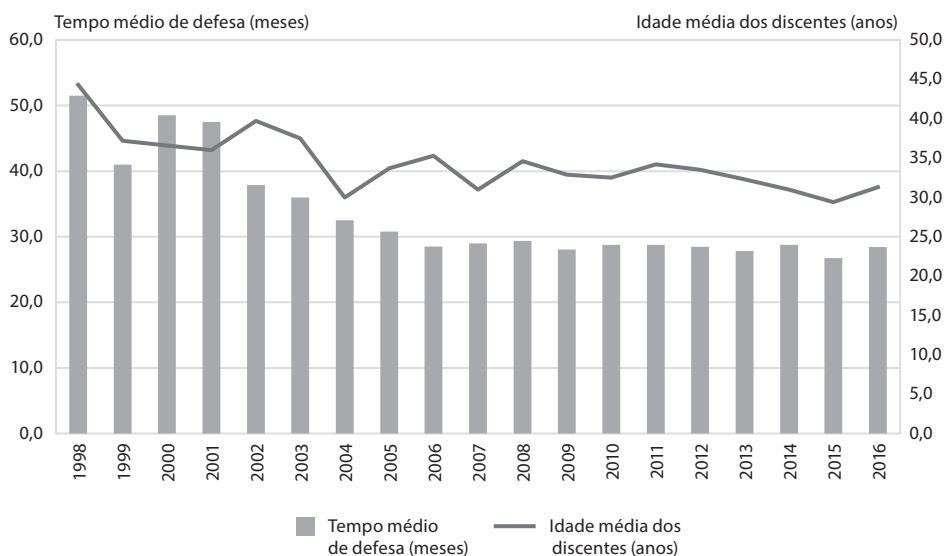


Fonte: Gerência de Pós-Graduação, ENCE/IBGE.

Nota: Os valores de 2019-2021 são projeções que consideram o número de alunos matriculados e o respectivo ano previsto de defesa.

Outro indicador das formações é o tempo médio para conclusão do curso. No mestrado, o tempo médio para a defesa da dissertação nos anos recentes variou entre 26 a 28 meses após o ingresso, sendo que não houve diferença notável entre as durações de discentes bolsistas e não bolsistas. Os dados estão dispostos no Gráfico 5. Além dessa informação, o gráfico também traz as idades médias dos alunos na conclusão do curso, mostrando que essas passaram por uma leve queda. Durante a primeira metade dos anos 2000, a média era superior a 40 anos, caindo para valores em torno de 35 anos posteriormente e, a partir de 2010, para 32 anos para o período 2010-2016. Já no doutorado há somente o tempo médio estimado para a conclusão da primeira coorte de ingressantes, que será em torno de 47 meses – portanto, abaixo do tempo esperado para a realização das defesas, que é de 48 meses.

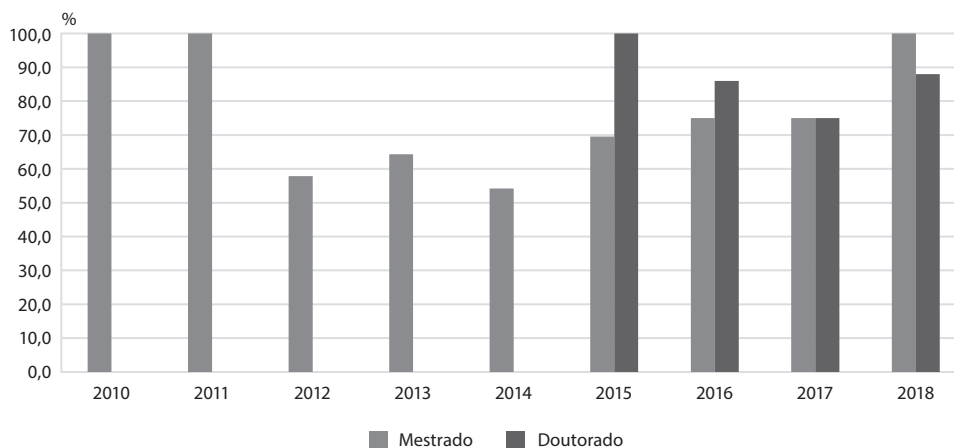
Gráfico 5 - Tempo médio para defesa de dissertação e idade dos discentes na conclusão do curso de Mestrado - 1998-2016



Fonte: Gerência de Pós-Graduação, ENCE/IBGE.

Por fim, uma última informação interessante sobre o corpo discente é a taxa de conclusão, medida pelo percentual de concluintes em relação ao total de ingressantes efetivos (que realizaram matrícula e efetivamente iniciaram o curso).

Gráfico 6 - Percentual de concluintes em relação ao ano de ingresso - 2010-2018



Fonte: Gerência de Pós-Graduação, ENCE/IBGE.

Nota: Para 2020 (Mestrado), o valor é estimado considerando o número de matrículas ativas no início de 2020, com perspectiva de defesa, em relação ao número de ingressos em 2018. Para o período 2016-2018 o mesmo é feito para o doutorado.

Tais taxas apresentam um dos principais desafios do programa e possuem margem significativa para avanços. Para alcançar o resultado ideal, com valores próximos a 100%, iniciativas internas estão em curso, como acompanhamento contínuo dos discentes, a maior regularidade nas avaliações de curso, o aprimoramento do processo seletivo e os seminários de apresentação dos projetos de dissertação, que permitem um acompanhamento do andamento das pesquisas e um momento de avaliação intermediário antes da defesa. Tais iniciativas visam fomentar o sucesso na conclusão dos cursos. A perspectiva é que, dentre os mestrandos, essa taxa chegue a 100% em relação aos ingressantes em 2018. Já para o doutorado as taxas anuais estimadas para os primeiros 4 anos foi de 85,7%.

Agenda futura e o PPG da ENCE

O Brasil passou por uma expansão significativa da pós-graduação durante as últimas décadas. O número de Programas de Pós-Graduação (PPGs) *stricto sensu* atingiu o impressionante crescimento de 6,9% ao ano entre 1998 e 2016, passando de 1.259 para 4.175 cursos. Segundo o Plano Nacional de Pós-Graduação, objetiva-se que até 2020 a pós-graduação brasileira tenha forte expansão, dobrando o número médio de doutores por mil habitantes entre 15 e 64 anos (passando de 1,4 em 2010 para 2,8 em 2020) e alcançando uma titulação anual de 19 mil doutores, 57 mil mestres e 6 mil mestres profissionais no mesmo ano (BRASIL, 2010).

Grande parte desse crescimento se deu em função da expansão da escolaridade e pelo estímulo a políticas públicas de educação superior. Assim, atualmente é a CAPES – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), fundação do Ministério da Educação (MEC), o órgão que define as estratégias de crescimento, acompanhamento e consolidação da pós-graduação *stricto sensu* em todo o Brasil. Nesse âmbito, o Sistema Nacional de Pós-Graduação (SNPG) é e continua sendo fundamental para o desenvolvimento da ciência e tecnologia, sendo que a avaliação da pós-graduação tem certa centralidade dentre as demais atividades que realiza. A avaliação se consolidou como parâmetro no estabelecimento dos referenciais de dis-

tribuição de bolsas de estudos, investimentos e recursos para o fomento à pesquisa, assim como na identificação das assimetrias entre as áreas do conhecimento e as regiões brasileiras, com o intuito de orientar ações de indução na criação e expansão de programas de pós-graduação. Dada sua importância, esse processo tem sido constantemente revisado e debatido, tanto na CAPES como na comunidade acadêmica.

Durante os últimos ciclos da avaliação (o último foi concluído em 2017, relativo ao período 2013-2016), a todos os programas, novos ou não, foi atribuída uma nota que variou entre 1 e 7. Embora a avaliação seja um processo bem sistematizado e relativamente complexo, estruturado a partir das informações preenchidas pelos programas de pós-graduação na Plataforma Sucupira, as mudanças são contínuas.

Em 2019 a CAPES publicou uma nova ficha de avaliação, avançou na criação de um Qualis de referência para periódicos (classificação única dos periódicos, visando avaliar a produção científica dos programas) e publicou uma série de relatórios técnicos, indicando que os programas deverão avançar nas suas atividades de autoavaliação, no acompanhamento dos egressos e na educação doutoral.

Especificamente para o programa da ENCE, que atualmente possui a nota 4 na avaliação, alguns desafios futuros são prementes. Estão tanto na melhoria da nota e na incorporação das novas tendências em relação às mudanças avaliativas (que traz uma série de indicadores e relatórios sobre as propostas dos programas, número de alunos, docentes, trabalhos de conclusão, publicações, dentre outros) como no protagonismo em relação à agenda científica da área e a interação com as demais diretorias do IBGE, no sentido da cooperação contínua para o desenvolvimento de pesquisas que permitam ao IBGE o cumprimento de sua missão institucional.

A pauta das iniciativas do programa perpassa, assim, o estímulo a publicações indexadas (com coautorias entre docentes e discentes egressos); a ampliação dos projetos de pesquisa com financiamento externos, diversificando o número de agências e buscando fontes alternativas para captação de recursos; a ampliação da divulgação do processo seletivo, com estímulo ao ingresso de alunos estrangeiros; o estímulo de realização de estágios pós-doutorais de forma rotativa, visando o treinamento continuado e expansão das redes de pesquisa de docentes; a integração entre os diversos níveis de ensino da ENCE e as demais áreas do IBGE e a consolidação de parcerias com instituições brasileiras e de fora do país. Em suma, trata-se de aumentar o impacto do programa, nas dimensões científicas, sociais e econômicas.

Referências

- ALVES, J. E. D.; BITTENCOURT, N. D.; SENRA, N. C. O Mestrado na Evolução Histórica da ENCE. In: ALVES, J. E. D.; BITTENCOURT, N. D.; SENRA, N. C.. (org.). *Mestrado em estudos populacionais e pesquisas sociais: memória dos primeiros 10 anos*. 1 ed. Rio de Janeiro: IBGE, 2009.
- ARANTES, S. B. *Planejamento de amostras domiciliares no Brasil explorando a malha setorial do Censo Demográfico 2010*. 2012. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2012.
- BARBOSA, A. M. *Panorama sociodemográfico da fecundidade de adolescentes e jovens: Brasil: 1970/2006*. 2008. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2008.
- BARBOSA E SILVA, C. J. *Comportamento da mortalidade infantil, e suas componentes etárias, nos municípios de Pernambuco*. 2015. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2015.
- BARROS, L. F. W. *Potencialidades e desafios na utilização de registros administrativos e de imagens noturnas de satélite para a realização de estimativas populacionais municipais intercensitárias no Brasil*. 2017. Tese (Doutorado em População, Território e Estatísticas Públicas) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2017.
- BARROSO JR., G. T. *A letalidade dos acidentes de trânsito nas rodoviárias Federais Brasileiras*. 2018. Dissertação (Mestrado em População, Território e Estatísticas Públicas) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2018.
- BASTOS, C. M. *Vulnerabilidade à dengue: um índice para o estado do Rio de Janeiro*. 2016. Dissertação (Mestrado em População, Território e Estatísticas Públicas) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2016.
- BELLO, L. A. *A mídia e os órgãos oficiais de estatísticas*. 2013. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2013.
- BERBAT, V. *A dimensão territorial nas políticas públicas: uma discussão sobre o Programa de Assentamentos Populares da Cidade do Rio de Janeiro*. 2007. (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2007.
- BRASIL. Ministério da Educação. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior Plano Nacional de Pós-Graduação - PNPG 2011-2020 / Coordenação de Pessoal de Nível Superior. - Brasília, DF: CAPES, 2010.
- BRITTO, L. P. G. de. *Avaliação da precisão de estimativas e projeções populacionais para pequenos domínios*. 2007. (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2007.
- CAMARANO, A. A. *Novo regime demográfico: uma nova relação entre população e desenvolvimento?* Rio de Janeiro: IPEA, 2014.
- CAMARGO, C. P. F. de. A Dinâmica Populacional como Processo Histórico-Social. In: SANTOS, LEVY & SZMRECSÁNYI (orgs.), *Dinâmica da População: teoria, métodos e técnicas de análise*. São Paulo: T. A. Queiroz, 1980.
- CANALES, A. I. Retos teóricos de la Demografía en la sociedad contemporánea. *Papeles de población*, Toluca, v. 10, n. 40, p. 47-69, jun. 2004.
- CANALES, A. La demografía latinoamericana en el marco de la postmodernidad. *Revista Latinoamericana de Población*, vol. 1, núm. 1, p. 17-33, jun. - dec. 2007.

CORREA, H. *A fecundidade na adolescência: a interpretação de um problema ou um problema de interpretação?* 2003. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2003.

CORREA, S. T. *Modelos lineares hierárquicos em pesquisas por amostragem relacionando o índice de massa corporal com as variáveis da pesquisa sobre padrões de vida do IBGE.* 2001. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2001.

DUARTE, J. C.; MONTALLI, L. T. OLIVEIRA, M. C. F. A. de; PATARRA, N. Alguns problemas teórico-metodológicos dos estudos de população na América Latina. *Textos NEPO 3*, UNICAMP, Campinas, 1985.

DUARTE, L. T. *Análise dos parados do Censo Demográfico 2010: investigação de fatores associados a erros não amostrais detectados na coleta das informações.* 2014. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2014.

FIALHO, E. O. M. *Propriedades estatísticas do estimador indireto da probabilidade de morte na infância, com base no método dos filhos sobreviventes: uma simulação Monte Carlo.* 2003. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2003.

FONSECA, S. M. *A notícia da estatística: A divulgação das estatísticas do IBGE na visão dos jornalistas.* 2005. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2005.

FRANÇA, V. O. *Avaliação de Métodos Estatísticos e Dasimétricos para Estimativa Populacional em Pequenas Áreas.* Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2012.

FREITAS, M. P. S. de. *Estratificação para a amostra de uma pesquisa domiciliar sobre mercado de trabalho.* 2002. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2002.

GAMA, W. P. G. *Privacidade e informação na era digital – a questão da confidencialidade em bases de dados de registros administrativos: o caso da Previdência Social Brasileira.* 2008. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2008.

GONÇALVES, A. P. V. *Uma aplicação da geoestatística ao estudo da segregação socioespacial na microrregião do Vale do Paraíba Fluminense: Um foco nos municípios de Resende e Volta Redonda.* 2014. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2014.

GONZALEZ, R. A. C. *Amostragem longitudinal em registros administrativos: Uma aplicação à Previdência Social.* 2005. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2005.

GUEDES, R. de S. *Racionalidade urbana e espaços rurais: a empresa agrícola UNIONE e a Vila Almeida em Juiz de Fora - MG.* 2017. Dissertação (Mestrado em População, Território e Estatísticas Públicas) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2017.

HAUSER, P. M. & DUNCAN, O. D. (eds.). *The study of population: an inventory and appraisal.* 7th Edition, Chicago: University of Chicago Press. 1972.

HECKSHER, M. D. *Preponderância dos ricos na desigualdade de renda no Brasil (1981-2016): aplicação da J- divergência a dados domiciliares e tributários.* 2014. Dissertação (Mestrado em População, Território e Estatísticas Públicas) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2017.

JANNUZZI, P. de M. A importância da informação estatística para as políticas sociais no Brasil: breve reflexão sobre a experiência do passado para considerar no presente. *Revista Brasileira de Estudos de População*, v. 35, n. 1, p. 1-10, 11 jun. 2018.

JÚNIOR, A. E. T. *Produção de pesos longitudinais para estimação e análise de dados da Pesquisa Mensal de Emprego do IBGE*. 2015. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2015.

LE BRAS, H. Peoples et populations. In: *L'Invention des populations: Biologie, idéologie et politique*. Paris, França. Editions Odile Jacob. pp. 9-54. 2000.

LÉON, D. A. D. *Aprimorando estimativas da mortalidade infantil para as unidades de federação no Brasil*. 2015. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2015.

LIMA, C. dos R. Uma análise das características dos beneficiários que trabalhavam nos anos de 2004 e 2014. 2017. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2017.

LIMA, I. S. de. *Fecundidade e Capital Humano: Uma análise dos dados da Pesquisa Mensal de Emprego na Região Metropolitana do Rio de Janeiro*. 2007. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2007.

LOPES, M. D. *Avaliação de desgastes de painéis em estudos longitudinais: uma aplicação na pesquisa mensal de emprego (PME/IBGE)*. 2002. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2002.

MELO, M. P. de. *Dinâmica do Território em Angra dos Reis: um foco na desigualdade socioespacial*. 2011. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2011.

OJIMA, R. Demografia e ciência: reflexões epistemológicas sobre a ciência das populações. *Novedades en población*, v. 13, n. 25, en. – jun. 2017.

OLIVEIRA, D. S. de. *Desigualdades Sócio-Espaciais e vulnerabilidade juvenil no contexto metropolitano: o caso do município de Nova Iguaçu*. 2006. (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2006.

OLIVEIRA, H. N. *Imigração internacional: uma alternativa para os impactos das mudanças demográficas no Brasil?* 2017. (Mestrado em População, Território e Estatísticas Públicas) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2017.

OLIVEIRA, K. F. de. *Dinâmica migratória em Sergipe dos anos 70 aos 90: uma análise a partir de alguns fatores estruturais*. 2003. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2003.

OLIVEIRA, L. da S. *Comparação de métodos de amostragem aplicáveis a estimação de índices de preços ao consumidor*. 2018. (Mestrado em População, Território e Estatísticas Públicas) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2018.

PASSOS, R. M. dos. *População em movimento, políticas públicas e pertencimento: Ruralidades Fluminense*. 2005. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2005.

PATARRA, N. L. Introdução. In: OLIVEIRA, J.M.S. de; AJARA, C.; PATARRA, N. L.; SILVA, P.L.N. (org.). *A ENCE aos 50 anos: um olhar sobre o Rio de Janeiro*. Rio de Janeiro: IBGE, 2006, v. , p. 11-25.

PEREIRA, C. J. de A. *O Geoprocessamento como ferramenta para estudos sociológicos: o caso da relação de influência entre urbanização e o perfil de capital cultural dos vestibulandos da UNESP*. 2003. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2003.

RAMOS, R. de S. *Diferenças socioespaciais dos riscos de mortalidade na infância*. 2001. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2001.

RODRIGUES, S. de C. *Análise da estrutura salarial revelada pela PPV incorporando peso e plano amostral*. 2003. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2003.

ROSSETI, E. S. *Análise de séries temporais da pesquisa mensal de emprego com incorporação dos efeitos do plano amostral*. 2013. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2013.

SANTOS, A. da R. *Estimação de indicadores educacionais por sexo e raça/cor através da Metodologia PROFLUXO*. 2006. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2006.

SANTOS, C. J. B. dos. *Avaliação da qualidade de modelos digitais do terreno construídos a partir de diferentes fontes de aquisição de dados*. 2001. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2001.

SANTOS, M. A. M. *Tipos Móveis: O olhar da imprensa sobre o censo do centenário*. 2004. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2004.

SANTOS, M. P. R. dos. *Indicadores para monitoramento de representatividade no caso da Pesquisa Nacional por Amostra de Domicílios Contínua*. 2017. Dissertação (Mestrado em População, Território e Estatísticas Públicas) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2017.

SILVA, A. D. *Proposta de Método de Pareamento para Integrar Dados sobre a Agropecuária*. 2018. Tese (Doutorado em População, Território e Estatísticas Públicas) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2018.

SIMÕES, P. F. M. 2009. *Modelagem longitudinal de dados de pesquisas de índices de preços ao consumidor, tratamento da não-resposta e índices com heterogeneidade controlada*. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2009.

SOUZA, V. dos R. de. *Análise espacial dos acidentes de trânsito com vítimas fatais na cidade do Rio de Janeiro em 2003*. 2007. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2007.

TORRES, A. C. P. *Uso de técnicas de data mining para a imputação de dados: Uma aplicação ao censo demográfico de 1991*. 2003. (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2003.

VIDAL, M. de O. *Emigrantes Brasileiros na Espanha: Fluxos, Políticas e Implicações Sociais*. 2009. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2009.

ZARDO, L. A. R. *Investigação de metodologias e avaliação de estimativas populacionais para pequenos domínios*. 2013. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2013.

Seção 1

Produção de estatísticas públicas e condições de vida

Tamanho dos municípios: estimativa populacional municipal comparada à enumeração censitária no Brasil em 2010*

Luiz André Ribeiro Zardo**

Suzana Cavenaghi***

Denise Britz do Nascimento Silva****

Resumo

Informações sobre o tamanho populacional para áreas subnacionais são frequentemente demandadas por diversos setores da sociedade, como subsídio para tomadas de decisão e repasse de recursos públicos. Diferentes abordagens metodológicas têm sido empregadas para estimar a população dos municípios e, em geral, dependem de indicadores sobre a dinâmica demográfica, que nem sempre estão disponíveis em nível de pequenos domínios. Incertezas sobre o tamanho populacional estão presentes tanto nos valores provenientes das projeções e estimativas, sujeitas a erros de estimação, quanto nas enumerações censitárias, sujeitas a erros de cobertura populacional. Para discutir sobre os desafios na estimação populacional para municípios no país, este estudo tem como objetivo avaliar os desvios entre estimativas populacionais e as respectivas populações enumeradas pelo Censo Demográfico 2010, buscando replicar a metodologia adotada pelo IBGE para obtenção das estimativas populacionais municipais para o ano de 2010. Os desafios encontrados e as alternativas utilizadas para análise são discutidos no texto.

Palavras-chave: Estimativas populacionais. Estimativas populacionais municipais. Erros de estimação. Erros de cobertura.

* A população oficial dos municípios brasileiros no ano de Censo Demográfico é decorrente dos dados coletados sem correção de cobertura. Para os anos intercensitários o IBGE utiliza as projeções e estimativas populacionais como fonte para o número oficial de habitantes nos municípios. Por escolha metodológica, nos anos de realização do Censo o IBGE não divulga as estimativas municipais, somente a população enumerada no censo. Motivados pelo fato que ambos os métodos estão sujeitos a erros, de diferentes tipos, e buscando discutir as estratégias adotadas pelo IBGE, foi desenvolvido um projeto de dissertação que buscava responder quão diferente seria a população municipal no ano censitário, caso essa fosse o resultado das estimativas populacionais. Este capítulo é decorrente da dissertação de mestrado defendida em 2013 por Luiz André Ribeiro Zardo, intitulada "Investigação de metodologias e avaliação de estimativas populacionais para pequenos domínios", sob orientação de Suzana Marta Cavenaghi e Denise Britz do Nascimento Silva, que resultou dessa inquietude.

** Docente no Departamento de Estatística da Universidade Federal de Mato Grosso (UFMT). Concluiu o mestrado em Estudos Populacionais e Pesquisas Sociais em 2013, e Doutor em População, Território e Estatísticas Públicas em 2019, pela ENCE.

*** Pesquisadora Independente. Professora e Pesquisadora do Programa de Pós-Graduação da ENCE de 2005 a 2019.

**** Pesquisadora da ENCE e Docente do seu programa de pós-graduação desde sua criação em 1998.

Introdução

As mudanças que ocorrem na sociedade em função do dinamismo populacional ao longo do tempo fazem com que seja necessária a obtenção de informações atualizadas sobre o contingente populacional de localidades administrativas de variados tamanhos, bem como a caracterização quanto à sua composição, de forma a sinalizar cenários específicos por demandas de serviços futuros para o bom planejamento e tomadas de decisões, baseadas nos melhores dados possíveis.

A principal fonte de informações demográficas com desagregação geográfica em pequenos domínios ainda é o censo demográfico, cuja periodicidade no Brasil é decenal. Sendo assim, faz-se necessário dispor de métodos adequados a cada contexto para estimar a população em períodos intercensitários, a fim de suprir as demandas por quantidades populacionais, desde aquelas vinculadas ao diagnóstico e desenho de políticas públicas em todas as áreas, ou para repasses de recursos públicos, até mesmo para servir como o denominador dos mais variados indicadores sócio-demográficos e econômicos, para os quais a base populacional é primordial.

O método conhecido como Componentes Demográficas (NACIONES UNIDAS, 2010) é o mais recomendado para se obter projeções populacionais, pois busca incorporar as tendências passadas e recentes das variações nas componentes demográficas, fecundidade, mortalidade e migração, as quais determinam o total populacional em um dado momento.

A aplicação do método das Componentes Demográficas para a obtenção de população em pequenos domínios pode não ser exequível e/ou mesmo recomendável em alguns casos, devido à ausência de estimativas acuradas e de qualidade dos indicadores demográficos em pequenos domínios, ou da impossibilidade da realização dos ajustes necessários adotados durante o processo de projeção. Nesse contexto, é oportuno considerar a adoção de procedimentos alternativos ao método das Componentes Demográficas, principalmente para a obtenção de estimativas populacionais em períodos intercensitários, como é realizado no Brasil há décadas. Nessa perspectiva, é importante conhecer bem o método empregado, seus pressupostos, bem como as fontes de dados que devem ser utilizadas na produção estimativas populacionais de forma que sejam coerentes com a realidade local e, sobretudo, realizar avaliações dos resultados obtidos periodicamente.

No entanto, o processo de avaliação não é tarefa fácil. Por um lado, é reconhecido que as estimativas populacionais não estão isentas de erros de estimação, por considerar suposições não precisas acerca do comportamento futuro das componentes da dinâmica demográfica ou por possíveis limitações técnicas em metodologias baseadas, por exemplo, em extrapolação matemática ou modelagem estatística. Por outro lado, em operações censitárias, também não se descarta a ocorrência de imprecisão quanto ao quantitativo populacional enumerado, pois essas operações dificilmente cobrem 100% da população e, ainda, tem cobertura variável ao longo do território. Pode ocorrer, por exemplo, erro de cobertura (enumeração) em decorrência de motivos operacionais ou por recusa em responder ao levantamento censitário. Assim, por mais cuidadosa, e metodologicamente correta, que sejam o planejamento e a implementação de uma operação censitária, dificilmente esta cobrirá de forma exaustiva e plena a população alvo.

No Brasil, para os períodos intercensitários, o Instituto Brasileiro de Geografia e Estatística - IBGE, responsável pelas estatísticas populacionais oficiais, tem empregado o método das Componentes Demográficas para projetar a população brasileira, por idade e sexo, em nível nacional, e mais recentemente em nível de Unidade da Federação. Entretanto, emprega o método de tendência de crescimento demográfico (*AiBi*) para obtenção de estimativas de total populacional em nível de municípios (IBGE, 2013). Nos anos censitários, os dados da enumeração do censo são considerados como as cifras oficiais das populações municipais, sem nenhuma correção de cobertura populacional.

Como essas estimativas populacionais em nível municipal são utilizadas para atender a diversos propósitos, sobretudo aquele relacionado ao cumprimento de leis vinculadas ao repasse de recursos em base *per capita*¹, elas são por vezes colocadas em xeque por alguns gestores mu-

¹ Como os Fundos de Participação dos Municípios, que são repasses de recursos financeiros da União aos municípios, a partir de previsão constitucional.

nicipais. Muitos gestores demandam modificação dos dados, visto que diferenças populacionais pequenas podem significar grandes montantes de recursos que o município deixa de receber.

Nesse contexto, este artigo tem o objetivo de analisar as diferenças que podem existir entre as estimativas provenientes da aplicação do método utilizado pelo IBGE (*AiBi*) para obtenção de totais populacionais municipais e aquelas provenientes da enumeração do censo demográfico, buscando a identificação de possíveis padrões nas diferenças. Mais especificamente, busca-se analisar as diferenças entre estimativas populacionais municipais calculadas para 2010 e suas respectivas cifras populacionais enumeradas pelo censo demográfico de 2010.

O propósito do estudo é, em princípio, bastante simples mas, visto que não existem estimativas populacionais para o ano de 2010, foi necessário obter as estimativas populacionais municipais para este ano. Isto foi realizado primeiramente tentando-se reproduzir o método utilizado pelo IBGE para o ano de 2009 que, como se verá, apresentou enorme desafio metodológico para alguns municípios. Foi então adotado um procedimento alternativo para estimar o total populacional municipal em 2010. Essas estimativas, não oficiais, foram comparadas com a população enumerada no Censo Demográfico 2010. As fontes de dados utilizadas foram os Censos Demográficos de 2000 e 2010, a Contagem da População de 2007 e as estimativas populacionais municipais de 2008 e 2009, todas realizadas pelo IBGE.

Estimativas populacionais para pequenos domínios: conceitos e métodos

Dado que neste artigo são utilizados alguns termos e conceitos nem sempre definidos de maneira unívoca na literatura, esta seção está dedicada a explicitá-los. Entre eles estão os associados a pequenos domínios e estimação para pequenas áreas; estimativas populacionais e projeções populacionais; erros de cobertura e de classificação, assim como apresenta um breve resumo dos métodos mais utilizados para cálculo de total populacional para áreas subnacionais.

Conceitos

Os conceitos de *estimativas populacionais* e de *projeções populacionais* não são tão claros. Em geral, entende-se como *estimativas populacionais* o quantitativo populacional obtido para algum momento específico do tempo (em geral, o ano corrente) para a população total ou para algum grupo específico desta população. Quanto aos procedimentos adotados, geralmente, para estimativas populacionais, utilizam-se técnicas com emprego de modelos matemáticos, econômicos e/ou demográficos (NACIONES UNIDAS, 2010).

O termo *projeção populacional*, por outro lado, pode usualmente ser compreendido como o resultado de cálculos que consideram as componentes da dinâmica populacional em conjunto – as taxas de fecundidade, de mortalidade e os movimentos migratórios – as quais a população de interesse tenha experimentado, bem como depende de suposições sobre o comportamento futuro dessas componentes (IBGE, 2004).

No entanto, as projeções populacionais para horizontes temporais futuros também podem ser obtidas pela aplicação de métodos de extrapolação populacional baseados em modelos matemáticos, bem como estimativas populacionais pontuais no tempo podem ser obtidas considerando as componentes da dinâmica populacional, quando informações que permitam mensurar as componentes demográficas estejam disponíveis (WALDVOGEL, 1997).

Neste estudo, para facilitar o entendimento, o termo projeções populacionais é entendido como aquele obtido para áreas maiores, por métodos demográficos para longos períodos de tempo. Já as estimativas populacionais são aquelas obtidas para áreas geográficas menores, por métodos matemáticos ou estatísticos, para anos intercensitários.

Outro termo que necessita de definição é o referente a pequenos domínios ou pequenas áreas. Existem situações para as quais se deseja obter estimativas para alguma característica de interesse, não apenas para uma população em estudo como um todo, como em nível nacional,

mas também para subgrupos populacionais, como no caso da obtenção de estimativas para níveis territoriais locais. Tais subgrupos populacionais são denominados *domínios* (SÄRNDAL *et al.*, 1991). Além de áreas geográficas, como regiões, estados, municípios e distritos, é importante mencionar que grupos sociodemográficos, tais como grupos específicos de idade, sexo e/ou raça, ou outras subpopulações de interesse, também podem ser definidos como domínios (RAO, 2003).

O termo *estimação para pequenas áreas* pode ser encontrado em diversos estudos e, em geral, refere-se a casos em que pequenos domínios de estudo são definidos por alguma delimitação territorial ou geográfica, independentemente do tamanho da área/território (SÄRNDAL *et al.*, 1991). No entanto, o termo *pequena área* pode ser compreendido como referência a um espaço geográfico que não necessariamente constitui um objeto de estudo, pois o foco de interesse, no caso de estimativas populacionais, pode ser a obtenção de cifras para uma localidade de grande extensão territorial, porém com baixa densidade populacional.

Assim, a fim de evitar confusão no entendimento do escopo deste estudo, utiliza-se o termo *pequeno domínio* como referência a uma subpopulação ou subgrupo populacional de menor magnitude, territorialmente definido, como unidade administrativa local.

Especificamente, o termo *estimação em pequeno domínio* aqui se relaciona à obtenção, por meio de métodos de projeções e estimativas populacionais, de quantitativos de população para os municípios brasileiros, que de fato constituem pequenos domínios, mas também outros de grande porte, visto que os municípios variam de poucos mil a milhões de habitantes.

Por fim, é importante definir o conceito de *erros* e *desvios*. Os procedimentos de projeções/estimativas populacionais, como mencionado, não estão livres de incertezas quanto ao valor da cifra populacional resultantes, pois são as taxas associadas às componentes demográficas experimentadas no passado, e que refletem a atual estrutura de uma população, que determinarão a composição de sua estrutura e seu tamanho no futuro. O desafio é elaborar suposições e cálculos associados às possíveis tendências futuras baseadas em experiências observadas no passado, que possam proporcionar estimativas mais prováveis para a população real.

O processo de recenseamento populacional também é complexo e gera erros. Além de competência técnica da instituição ou agência estatística responsável pela tarefa, depende de bom planejamento e emprego de diversos recursos (contratação e capacitação de pessoal, aquisição e implementação de recursos tecnológicos, etc.) para uma execução bem-sucedida. A Organização das Nações Unidas (ONU) tem revisado e publicado periodicamente princípios e recomendações associados a operações censitárias, visando assegurar boa qualidade tanto da cobertura territorial quanto de informações coletadas (CHACKIEL, 2010). Ainda assim, há o reconhecimento que uma operação censitária está sujeita a incertezas geralmente associadas a dois tipos de *erros*: de *cobertura* e de *classificação* ou *conteúdo* (NACIONES UNIDAS, 2010).

Erros de cobertura relacionam-se à captação da unidade recenseada, seja pela omissão ou duplicação de indivíduos ou de unidades domiciliares, e são de caráter quantitativo, impactando no total populacional. *Erros de classificação* ou *conteúdo* associam-se ao registro das informações referentes aos indivíduos ou domicílios, impactando na ausência de um, ou mais, quesitos que compõem as características da unidade recenseada.

De acordo com as Nações Unidas,

[...] Los errores de cobertura son los que se producen en el empadronamiento por la omisión o duplicación de personas o unidades de habitación. Errores de cobertura son, entre otros, unos mapas incompletos o incorrectos, el hecho de que los empadronadores no hayan visitado todas las unidades de las zonas asignadas, la duplicación en el recuento, la existencia de personas que por un motivo u otro no se dejan empadronar, el tratamiento equivocado de ciertas categorías de personas, como visitantes o extranjeros no residentes, y la pérdida o la destrucción de documentos censales después del empadronamiento. Los errores de contenido son los provocados por errores de comunicación y registro de las características detalladas de las personas, los hogares y las unidades de vivienda. Los errores de contenido se producen por equivocaciones del empadronador al formular las preguntas del censo, la incapacidad de responder o los mal entendidos por parte de los declarantes, el suministro deliberado de información incorrecta, los errores debidos a la cumplimentación del cuestionario por terceros, y los errores de codificación o asiento de datos, entre otras razones (NACIONES UNIDAS, 2010, p.91).

Apesar de usualmente se utilizar o termo *erros* para se referir às diferenças entre as estimativas populacionais calculadas por vários métodos, utiliza-se aqui a denominação de *desvios* populacionais (positivos ou negativos) e não erros entre a estimativa e a enumeração da população, visto que ambas as fontes podem conter erros e não será possível identificar qual o valor verdadeiro da população de qualquer município em questão no ano de 2010. Ou seja, assume-se explicitamente que a população enumerada nos censos demográficos geralmente não corresponde a 100% da população residente nos municípios.

Métodos para projeções e estimativas populacionais

O método das Componentes Demográficas (UNITED NATIONS, 2010) tem sido utilizado com o propósito de obter projeções populacionais em diversos países. Tal método tem sido considerado adequado, pois consiste em projetar a população, por sexo e idade, por meio de equação compensadora de população², incorporando as tendências das componentes demográficas: fecundidade, mortalidade e migração, as quais uma população tenha experimentado, ou esteja exposta, e que impactam na composição e na magnitude de uma população.

Uma questão substancialmente importante é a garantia de que as tendências das componentes demográficas estejam consistentes com as mudanças observadas no tamanho da população e sua distribuição por sexo e idade (UNITED NATIONS, 2008). Para tanto, em cada nova *rodada* ou *revisão* de projeções e estimativas de população, dois processos devem ser considerados: (a) inserção de informações novas relevantes sobre a dinâmica demográfica associada à população de interesse e; (b) a formulação de hipóteses, fundamentadas em metodologias adequadas, acerca das possíveis trajetórias associadas às componentes demográficas, fecundidade, mortalidade e migração (NACIONES UNIDAS, 2010). Essa é uma tarefa que demanda a maior quantidade de informações demográficas disponíveis, bem como o envolvimento de especialistas, de modo a garantir plausibilidade e coerência nos resultados (IBGE, 2007).

Métodos alternativos ao das componentes demográficas, com base em modelos matemáticos e/ou estatísticos (e/ou até combinados), têm sido desenvolvidos e experimentados por pesquisadores e por agências produtoras de estatísticas oficiais para suprir a demanda por informações populacionais para períodos intercensitários para domínios populacionais em níveis subnacionais, dentre outros domínios de interesse (CAVENAGHI, 2012).

No Brasil, fora do âmbito das estatísticas oficiais, diferentes abordagens têm sido empregadas para esse fim. Observam-se experiências tais como o emprego de modelos de extrapolação matemáticas (WALDVOGEL, 1997), modelagem estatística com base em variáveis sintomáticas auxiliares (JARDIM, 2003; PITTA, DINI e JANNUZZI, 2010) e projeções populacionais

² $P_{t+n} = P_t + B_{t,t+n} - D_{t,t+n} + I_{t,t+n} - E_{t,t+n}$. Tal que representam-se, n : o intervalo de tempo transcorrido entre dois instantes de tempo, geralmente entre duas operações de porte censitário; P_{t+n} : população no instante $t+n$; P_t : população no instante t ; $B_{t,t+n}$: nascimentos ocorridos entre t e $t+n$; $D_{t,t+n}$: óbitos ocorridos entre t e $t+n$; $I_{t,t+n}$: imigrantes que chegaram à localidade entre t e $t+n$; $E_{t,t+n}$: emigrantes que deixaram a localidade entre t e $t+n$.

com base em modelos adaptados do método das componentes demográficas (WALDVOGEL; 1997; JARDIM, 2003). Estudos também têm sido realizados na elaboração e avaliação de métodos de projeção e estimativas populacionais (BRITO, CAVENAGHI e JANNUZZI, 2010; BORGES, ERVATTI e SILVA, 2011; CORRÊA, MYRRHA e FÍGOLI, 2011), em um esforço de identificar as fragilidades e potencialidades dos métodos.

Experiências internacionais, como podem ser observadas em metodologias empregadas por instituições e agências oficiais de estatística, também abrigam uma diversidade de abordagens e métodos. Alguns países, como Estados Unidos (U.S. BUREAU CENSUS, 2012a, 2012b), Canadá (STATISTICS CANADA, 2009; 2010) e Austrália (ABS, 2009; 2012), utilizam o método das componentes demográficas, tanto para a obtenção de projeções ou estimativas populacionais em nível de nação, quanto em níveis subnacionais. No caso da Inglaterra e País de Gales (ONS, 2012), Brasil (IBGE, 2008; 2013) e alguns países da América Latina (GONZÁLEZ; TORRES, 2012), por exemplo, estimativas para domínios menores têm sido obtidas por meio de métodos de repartição, extrapolação matemática ou metodologia adaptada do método das componentes demográficas, a partir de projeções ou estimativas populacionais realizadas em nível de nação (ou algum domínio maior), cujas cifras populacionais foram obtidas pelo método das componentes demográficas.

Em projeções e estimativas para pequenos domínios populacionais, o método das componentes demográficas nem sempre pode ser empregado e/ou recomendado. Isso ocorre, sobretudo nos anos intercensitários, pela ausência de informações que permitam obter parâmetros referentes às componentes demográficas, principalmente sobre os movimentos migratórios, cuja mensuração é mais complexa em função de grande volatilidade, em geral, em função de maior e mais frequente mobilidade populacional entre localidades vizinhas. A ausência de registros com cobertura completa destes eventos inviabiliza a produção de estimativa de população em pequenos domínios pelo método das componentes demográficas (MESA; VALENZUELA, 2007).

A falta de informações de qualidade impõe dificuldades na construção de hipóteses futuras e/ou construção de cenários futuros e obtenção de projeções adequadas, coerentes e consistentes para pequenos domínios de interesse (BAY, 2012). Deve-se considerar que a formulação de hipóteses sobre o comportamento futuro das componentes demográficas está sujeita a base de dados e estatísticas oficiais que podem não ser precisas e/ou não estão disponíveis³ e detalhadas de forma oportuna e suficiente para tal propósito. Ou seja, a obtenção de insumos básicos para especificar padrões e tendências das componentes demográficas depende da disponibilidade⁴ de informações estatísticas de qualidade, a insuficiência destas, especialmente para pequenos domínios, torna inexecutável a elaboração de projeções populacionais pelo método das componentes demográficas e a obtenção de estimativas derivadas.

Barros e Cavenaghi (2016) apontam que a adoção de métodos que incorporem informações auxiliares sobre a dinâmica demográfica, objetivando melhor acurácia de estimativas populacionais para pequenos domínios por meio da utilização de registros administrativos, pode ser uma estratégia a ser adotada em alguns casos com devidos cuidados e restrições. No caso do Brasil, os autores (*idem*) atentam para a existência de situações em que para alguns municípios brasileiros pode ser baixa ou quase nula a correlação entre os registros administrativos e suas respectivas populações.

³ A disponibilização de informações estatísticas, por agências produtoras de estatísticas oficiais, requer etapas de ajustes e tratamentos de dados inerentes aos processos de captura, crítica e validação de dados.

⁴ Tendo em conta registros estatísticos com diferentes graus de cobertura entre níveis territoriais distintos, geralmente associados ao nível de desenvolvimento socioeconômico dentro e entre regiões e/ou localidades.

Metodologia e estratégias utilizadas na estimação oficial do total populacional para os municípios brasileiros

No Brasil, para anos em que ocorrem operações censitárias, a obtenção e a divulgação dos quantitativos populacionais oficiais são realizadas pelo IBGE, às vezes, por meio da enumeração e ajustes populacionais resultantes do Censo Demográfico (ou da Contagem de População no meio da década). Para anos *pós* ou *entre* operações censitárias, os totais populacionais oficiais são obtidos por métodos de projeções e estimativas populacionais.

Breve histórico

O IBGE tem empregado o método das Componentes Demográficas para projetar o total populacional para o Brasil, por sexo, idades simples e grupos quinquenais de idade (IBGE, 2013). Nas décadas de 1990 e 2000, o IBGE empregou, majoritariamente, o *método de tendência de crescimento demográfico* (*Apportionment Method - share-of-growth*) (SWANSON; TAYMAN, 2012), conhecido no Brasil como método *AiBi* (IBGE, 2008), para estimar populações para as grandes regiões brasileiras, as unidades da federação (UF) e os municípios. O processo de produção das estimativas populacionais subnacionais divulgadas pelo IBGE provê cifras apenas para totais populacionais, não dispondo de resultados desagregados por sexo, idade ou área de residência (rural ou urbana). No entanto, importante destacar que existem métodos disponíveis para este fim assim como algumas iniciativas no passado recente⁵.

Documentos do IBGE de períodos anteriores aos anos iniciais da década 1990 mencionam uso de técnicas distintas para a obtenção de estimativas de população. Documentos da década de 1960 (IBGE, 1965; 1967; 1969) apontam para o uso de projeções baseadas em taxas de crescimento geométrico observado no período intercensitário anterior, para estimar as populações nas UFs e municípios das capitais. Para os demais municípios, a diferença entre as estimativas obtidas para as UFs e suas respectivas capitais foram distribuídas proporcionalmente às quotas de participação de cada município segundo o censo demográfico anterior ao ano de referência das projeções. Documentos referentes às projeções populacionais para os anos de 1990 e 1991 (IBGE, 1990; 1991) relatam dificuldades encontradas na obtenção das estimativas para os municípios brasileiros naquela ocasião, em função do atraso na realização do censo que só ocorreu em 1991 e, conseqüentemente, da defasagem de informações em relação ao censo de 1980. As estimativas foram produzidas baseadas na participação relativa da população municipal no total de sua UF considerando o comportamento das mesmas no período de 1970-1980, resultando em algumas das estimativas não compatíveis com os dados do eleitorado inscrito para as eleições de 1988. Tal fato ocasionou a manifestação de algumas prefeituras sobre a possibilidade das estimativas populacionais produzidas estarem aquém da realidade. A solução adotada naquela ocasião foi ajustar as projeções considerando as proporções de eleitores e a das estimativas populacionais em 1990 e em 1991 (IBGE, 1991). Além desses dados de eleitorado, foram utilizadas também informações de domicílios levantados para a realização do censo de 1991 e do número médio de moradores por domicílio, obtidos pela Pesquisa Nacional por Amostra de Domicílios (PNAD) de 1989.

Na década de 2000, o IBGE produziu revisões de projeções populacionais para os anos de 2004 e 2008 e realizou uma Contagem de População⁶ em 2007. Para a elaboração, por exemplo, de estimativas populacionais para os anos de 2008 e 2009, o IBGE considerou: (i) as projeções⁷ de totais populacionais para o Brasil para os anos em questão (IBGE, 2008) e, (ii) informações

⁵ Estimativas para 2000 a 2013 efetuadas em estudo patrocinado pela Rede Interagencial de Informações para a Saúde – Ripsa e para 2014 e 2015 estimativas elaboradas pelo Ministério da Saúde/SVS/CGIAE, disponíveis em <<http://tabnet.datasus.gov.br/cgi/deftohtm.exe?novapop/cnv/popbr.def>>.

⁶ No Brasil, em 2007, existiam 5.564 municípios. A Contagem da População 2007 compreendeu os municípios com estimativa de população com até 170 mil habitantes e mais 21 municípios selecionados com população acima desta faixa. No entanto, 128 municípios, os maiores, tiveram seus respectivos totais populacionais estimados para o ano de 2007. Deve-se ressaltar que nos estados onde somente a capital tinha população estimada maior que 170 mil habitantes, nesta também houve a realização da contagem populacional (IBGE, 2007a).

⁷ Projeções populacionais realizadas pelo IBGE na ocasião da atualização da projeção da população do Brasil por sexo e idade 1980-2050, Revisão 2008 (IBGE, 2008).

referentes à tendência de crescimento populacional de cada domínio de interesse no período compreendido entre 2000 e 2007, tendo como base as populações oriundas do Censo Demográfico 2000 e da Contagem da População 2007 para o cálculo das estimativas por meio do método *AiBi* (IBGE, 2008; 2009).

A partir da revisão das projeções de população de 2013, o IBGE empregou o método das Componentes Demográficas para projetar tanto a população em nível nacional, para um horizonte temporal de projeção de 50 anos, quanto para as populações das UFs, individualmente, para um horizonte de projeção de 30 anos, mas manteve o método *AiBi* para estimar as populações municipais anualmente (IBGE, 2013, 2018).

O método *AiBi*

O método de tendência de crescimento demográfico, *AiBi* se baseia na hipótese de que o total populacional de um pequeno domínio (por exemplo, municípios) pode ser estimado por meio de uma função linear do total populacional de um domínio maior que o contém (como o total populacional do Brasil ou da respectiva Unidade da Federação). O total populacional do domínio maior deve ser conhecido por algum processo de projeção e/ou estimação de modo que a soma dos totais populacionais para os pequenos domínios resulte no total populacional do domínio maior no qual estão contidos (IBGE, 2008).

Em um momento t a estimativa populacional $P(t)$ de um domínio maior é subdividida em m pequenos domínios com população $P_i(t)$, $i = 1, \dots, m$ de modo que $\sum_{i=1}^m P_i(t) = P(t)$. Segundo o IBGE (2008), o método considera que a população de cada domínio menor i possa ser estimada por meio da função linear.

$$P_i(t) = a_i P(t) + b_i \quad ,$$

na qual a_i representa a proporcionalidade do incremento do domínio menor i em relação ao incremento do domínio maior, e b_i representa o coeficiente linear de correção. Os valores de a_i e b_i podem ser estimados pela resolução do sistema de equações:

$$P_i(t_0) = a_i P(t_0) + b_i \quad ,$$

$$P_i(t_1) = a_i P(t_1) + b_i \quad ,$$

representando, respectivamente, as populações estimadas do domínio menor i em dois momentos distintos t_0 e t_1 tais como as populações de dois censos consecutivos. Assim, os coeficientes a_i e b_i podem ser obtidos pelas expressões:

$$a_i = \frac{P_i(t_1) - P_i(t_0)}{P(t_1) - P(t_0)} \quad ,$$

$$b_i = P_i(t_0) - a_i P(t_0) \quad .$$

Medidas de desvios

A qualidade das cifras resultantes da aplicação de alguma metodologia para produção de estimativas populacionais pode ser avaliada por meio de medidas que buscam mensurar o desvio entre o quantitativo populacional estimado e o quantitativo populacional enumerado.

Em um processo de avaliação de estimativas, medidas de *precisão* podem indicar quanto as estimativas produzidas por um processo de estimação populacional estão próximas, ou desviam, dos quantitativos populacionais obtidos por enumeração censitária, ao passo que medidas de *viés*, as quais consideram a direção do desvio, buscam indicar a tendência da direção do desvio (SMITH; CODY, 2004).

Segundo os autores (SMITH; CODY, 2004), algumas medidas usualmente empregadas para avaliar a precisão e o viés em processos de projeções e estimativas populacionais, e que são utilizadas aqui, podem ser obtidas por:

- *Desvio total* (D_i), que indica a diferença total entre os quantitativos populacionais estimados (PE_i) e os quantitativos populacionais observados (PO_i) na localidade i , calculado por $D_i = PE_i - PO_i$;
- *Desvio Percentual* (DP_i), que indica o quanto a medida D representa, em percentual, em relação à população observada PO_i , obtida por $DP_i = (D_i / PO_i) * 100$, e;
- *Desvio Percentual em Módulo ou Absoluto* (DPA_i), que representa a Média do DP , permite verificar tendência de viés dos desvios, pois considera a direção (positiva ou negativa) dos desvios, dado por $DPA_i = DP_i$.

Como destacado por Brito (2007), a definição de uma margem aceitável para desvios em estimativas populacionais pode não ser uma tarefa simples, pois “[...] diversas características como tamanho da população, taxa de crescimento populacional, horizonte de projeção, entre outras, influenciam na definição de um percentual mínimo de erro” (*idem*, p.51).

No caso de avaliações de estimativas populacionais municipais para o Brasil, Brito, Cavenaghi e Jannuzzi (2010) indicam que “[...] métodos de projeções populacionais para pequenos domínios (municípios) considerados de boa precisão têm sido aqueles em que mais da metade dos municípios apresentam EP [erro percentual] entre -10% e 10%” (*idem*, p. 38). Assim, esses valores serão tomados em conta no momento da análise dos resultados neste estudo.

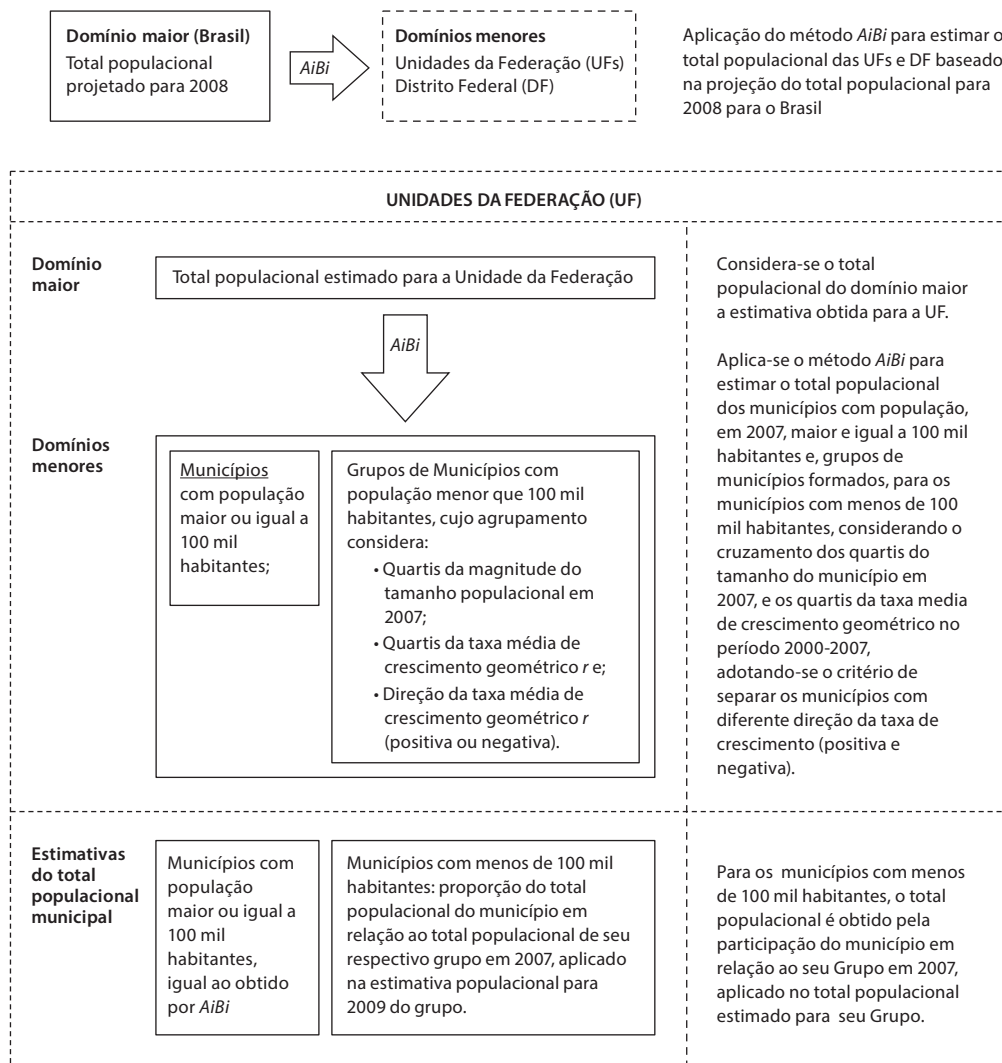
Obtenção das estimativas populacionais municipais para 2010

A Figura 1 apresenta, de forma sintética, um fluxograma dos procedimentos adotados pelo IBGE para a produção de estimativas do total populacional para os municípios brasileiros em 2008⁸ (e em 2009) conforme descrito no informe metodológico (IBGE, 2008). Vale destacar que apesar das projeções populacionais serem estimadas para um período longo de tempo para o país (e recentemente para UFs), e revisadas pelos menos duas vezes na década, as estimativas populacionais municipais somente são divulgadas anualmente pelo IBGE, com base nessas projeções populacionais para as áreas maiores.

Como há várias especificidades na aplicação do método *AiBi*, buscou-se reproduzir as estimativas calculadas pelo IBGE para o ano de 2009 para, então, se verificado consistências nos resultados, aplicar a metodologia para estimar o total populacional dos municípios em 2010.

⁸ Em 2013 houve mudanças nesse fluxo, baseado principalmente no fato que as estimativas populacionais para Unidades da Federação também foram obtidas pelo método das componentes demográficas.

Figura 1 - Fluxograma dos procedimentos adotados pelo IBGE para obtenção das estimativas dos totais populacionais em 1º de julho de 2008 para os municípios brasileiros



Fonte: Com base na metodologia das estimativas das populações residentes nos municípios brasileiros para 1º de julho de 2008 (IBGE, 2008).

Assim, considerando a metodologia utilizada pelo IBGE no fim década de 2000 para obter estimativas populacionais municipais pelo método *AiBi* (Figura 1), foram empregadas as seguintes informações e respectivas fontes de dados como insumos para a produção das estimativas populacionais municipais para 1º de julho de 2009:

A população total, segundo os municípios brasileiros para o ano de 2000, com data de referência 1º de agosto de 2000 (Fonte: IBGE, Censo Demográfico 2000);

- A população total, segundo os municípios brasileiros para o ano de 2007, com data de referência 1º de abril de 2007 (Fonte: IBGE, Contagem da População 2007) e;
- A projeção do total populacional para o Brasil, com data de referência 1º de julho de 2009 (Fonte: IBGE, Projeção da População do Brasil por Sexo e Idade 1980-2050, Revisão 2008).

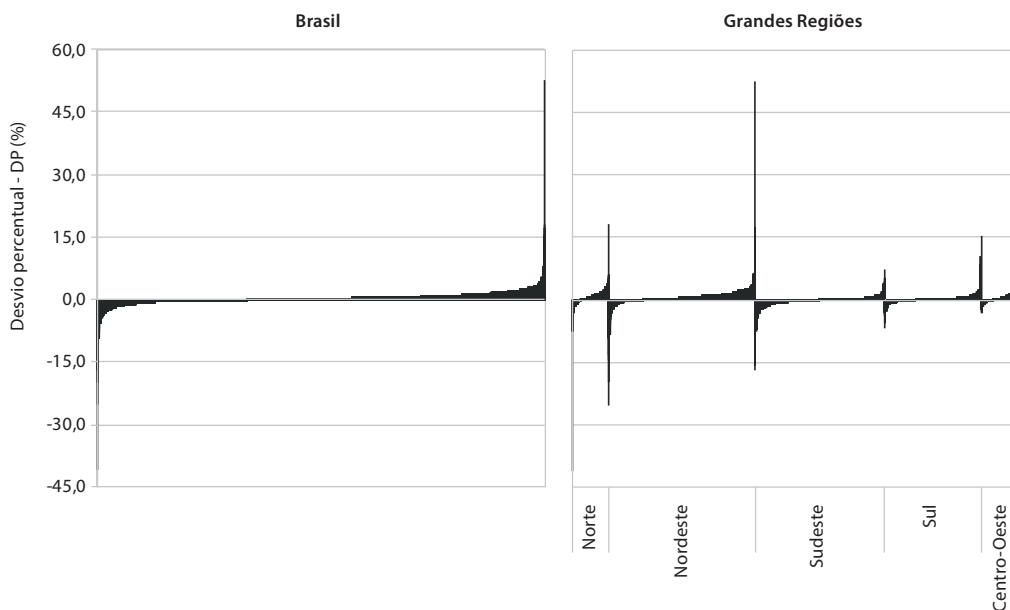
Mesmo seguindo todos os procedimentos descritos, a aplicação do método *AiBi* apresentou resultados em que as estimativas populacionais produzidas neste estudo se desviaram das

publicadas pelo IBGE para 2009. Na média, os desvios encontrados foram baixos, menor que 1% para as regiões Sudeste e Sul, que exibiram um DP médio de -0,29% e 0,23%, respectivamente. Verificou-se, ainda, que mais da metade das estimativas populacionais municipais aqui produzidas apresentaram desvios menores que 1% em valor absoluto (64,6% dos casos no Brasil), com exceção da região Norte, cuja proporção de casos com Desvio Percentual Absoluto (DPA) menor que 1% é 38,7%.

No entanto, foram obtidos alguns desvios altos espalhados em todo o país, que variaram em torno 40%, negativamente, a desvios positivos de cerca de 50%. Observou-se a ocorrência de desvios em todas as grandes regiões brasileiras com picos de desvio percentual (DP) de grande magnitude com valores de, aproximadamente, -40% e 18% na região Norte, -25% e 50% na região Nordeste, -17% e 7% na região Sudeste, -7% e 15% na região Sul e, -3% e 7% na região Centro Oeste, como apresentado no Gráfico 1.

Em geral, é esperado que na replicação do método (tendo-o como uma função linear), considerando os mesmos insumos (*inputs*), não sejam obtidos resultados que desviem amplamente dos resultados obtidos em uma primeira rodada. No entanto, deve-se ter em conta que o processo de elaboração de estimativas populacionais não constitui um procedimento puramente matemático, pois ajustes podem ser realizados durante o processo a partir de alguma informação auxiliar confiável e da *expertise* do demógrafo ou técnico responsável para a melhoria das estimativas. Contudo, ajustes que mudem sobremaneira as estimativas esperadas deveriam ser registrados no informe metodológico, visto que o método implica que a população calculada para um município inevitavelmente afeta a população estimada de outro município, já que o total da área maior é previamente definido em etapa anterior. Diante da ausência dessa documentação, foram consultados técnicos do IBGE, responsáveis pelas estimativas municipais, para buscar insumos que permitissem replicar os valores oficiais das estimativas municipais de 2009. Assim, foi adotada uma nova alternativa para estimação da população municipal em 2010.

Gráfico 1 - Desvio percentual das estimativas populacionais municipais produzidas para 2009, aplicando-se o método *AiBi* em relação às calculadas e publicadas pelo IBGE em 2009



Fonte: Desvios Calculados considerando estimativas obtidas com base no Censo Demográfico 2000 (COPIS/DPE/IBGE); Contagem da População 2007 (COPIS/DPE/IBGE); Projeção da População do Brasil por Sexo e Idade 1980-2050, Revisão 2008 (IBGE) e; Estimativas das populações residentes, em 1º de julho de 2009, segundo os municípios (IBGE).

A ausência de documentações precisas sobre os ajustes realizados pelos especialistas durante o processo de obtenção de estimativas, mesmo aplicando-se o método *AiBi* tal como descrito no informe metodológico (IBGE, 2008, 2009), impediu que se obtivesse valores mais próximos para os municípios com grandes desvios. Por esse motivo, buscou-se empregar um procedimento alternativo para estimar a população dos municípios em 2010. Assim, assumindo o pressuposto que foram utilizados os mesmos critérios técnicos pelos especialistas do IBGE na obtenção das estimativas para os anos de 2008 e 2009, com base na tendência de crescimento observada para esses anos, foram obtidas as estimativas para 2010.

Dessa forma, com base na taxa média de crescimento geométrico, considerou-se que a população em estudo tenha experimentado um crescimento geométrico no período tido como referência (UNITED NATIONS, 1969), tal que,

$$r = \left\{ \left[(P_t / P_{t_0})^{1/n} \right] - 1 \right\} * 100$$

em que, P_t representa a população no instante t , P_{t_0} representa a população no período inicial t_0 e n o número de anos, ou fração de ano, considerando o período de referência compreendido entre os instantes t_0 e t .

Dessa maneira, para estimar a população municipal em 2010, foram utilizados os seguintes insumos:

- A população total, segundo os municípios brasileiros para o ano de 2008, com data de referência 1º de julho de 2008 (Fonte: Estimativas das populações residentes, em 1º de julho de 2008, segundo os municípios) e;
- A população total, segundo os municípios brasileiros para o ano de 2009, com data de referência 1º de julho de 2009 (Fonte: Estimativas das populações residentes, em 1º de julho de 2009, segundo os municípios);

Pelo desenvolvimento algébrico da expressão de r , pode-se obter a população em 2010 pela expressão:

$$E_{2010} = P_{2009} (1 + r_{08-09})^{\left[(2010-2009) + \frac{1}{12} \right]}$$

em que, E_{2010} representa a estimativa para o quantitativo populacional municipal em 1º de agosto de 2010; P_{2009} representa a população municipal estimada em 1º de julho 2009 e r_{08-09} representa a taxa de crescimento populacional no período de referência compreendido entre 2008 e 2009 nas datas de 1º de julho.

Análise das diferenças nas estimativas populacionais para municípios brasileiros, ano de referência 2010

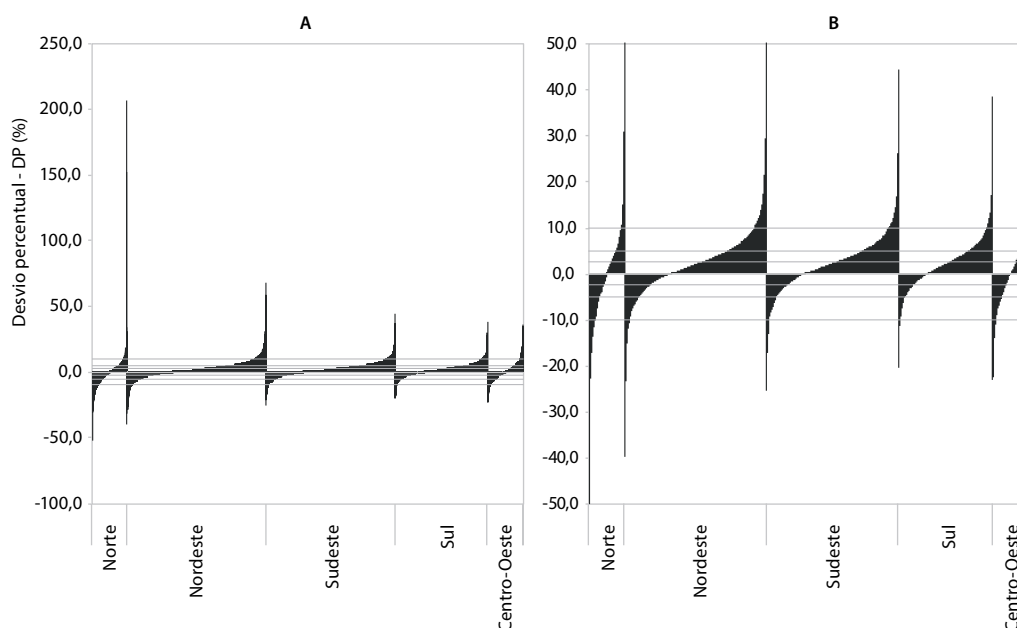
Nesta seção, apresenta-se uma análise dos desvios entre as estimativas populacionais produzidas para os municípios brasileiros para o ano de 2010, como descrito anteriormente, e as respectivas cifras populacionais enumeradas pelo Censo Demográfico de 2010.

Como resultado do processo de estimação populacional adotado neste trabalho, obteve-se um total populacional para o Brasil, pela soma das estimativas calculadas para os municípios, de 193.541.238 habitantes, cerca de 1,5% maior do que o total populacional enumerado pelo censo demográfico, considerando que neste já foi incorporado uma imputação de domicílios e população (IBGE, 2011). Para as regiões Nordeste, Sudeste, Sul e Centro-Oeste, obteve-se total populacional estimado 2%, 1,68%, 2,10% e 0,41% maior que a população enumerada pelo censo demográfico, respectivamente. A região Norte apresentou população estimada 1,66% menor do que a enumerada.

Quanto à magnitude dos desvios, em média, os municípios apresentaram desvios em torno de 5,13%, em que, na maioria dos casos, a mediana dos desvios foi menor do que 5,0%, com exceção para a região Norte, cujos desvios foram, em média, maiores na maioria das Unidades da Federação.

Embora tenha havido desvios de grande magnitude como, por exemplo, de aproximadamente 200% na região Norte, grande parte dos municípios (90% dos casos) apresentaram desvios percentuais até 10%, considerando os desvios percentuais em valor absoluto. Observa-se ainda, no Gráfico 2 (b), que este padrão é semelhante em todas as regiões apresentadas.

Gráfico 2 - Desvio percentual das estimativas dos totais populacionais municipais calculadas para 2010 e totais populacionais municipais publicados pelo Censo Demográfico 2010, segundo as Grandes Regiões



Fonte: Obtidas com base nas Estimativas das populações residentes, em 1º de julho de 2008, segundo os municípios – IBGE (IBGE, 2008); Estimativas das populações residentes, em 1º de julho de 2009, segundo os municípios – IBGE (IBGE, 2009); Censo Demográfico 2010.

Nota: O Gráfico B difere-se do A quanto à escala adotada, o primeiro apresentando ênfase na amplitude que varia de -50 a 50 pontos percentuais.

Ao observar as estatísticas dos desvios considerando o tamanho populacional (número de habitantes) dos municípios (Tabela 1), há indícios de que a magnitude do desvio se relaciona com o tamanho do município. As médias dos Desvios Percentuais Absolutos – DPAs apresentaram tendência de redução de acordo com o aumento dos tamanhos dos municípios.

A maioria dos municípios apresentou desvios positivos, em torno de 60% dos casos, exceto municípios cujas populações encontram-se na faixa de até 10 mil habitantes, com desvios positivos em torno de 70%. Verifica-se ainda que, em municípios de menor tamanho populacional, o percentual de desvios positivos é maior, tendendo a reduzir com o aumento do tamanho populacional do município.

Considerando a direção da taxa média de crescimento geométrico dos municípios, verificou-se (Tabela 1) que as magnitudes do tamanho do desvio tenderam a aumentar quando as taxas de crescimento populacional se distanciam de zero.

Quando analisados os desvios entre as estimativas populacionais municipais calculadas e a população obtida a partir da enumeração censitária sem a inclusão da população obtida por meio de imputação, os resultados indicam semelhança no comportamento dos desvios quanto à forma, havendo maiores desvios onde ocorreu maior imputação de dados populacionais.

Tabela 1 - Distribuição absoluta e relativa de municípios e estatísticas de desvio entre as estimativas dos totais populacionais municipais e totais populacionais municipais, segundo faixas de tamanho de população municipal e faixas de taxa média de crescimento geométrico - 2010

Categorias	Distribuição dos municípios (1)		Diferença entre a população estimada e a enumerada				
	Total	Percentual (%)	Média do desvio percentual absoluto	Média do desvio percentual	Percentual de desvios positivos (%)	Percentual de desvio percentual absoluto (%)	
						Menor que 5	Maior que 10
Faixas de tamanho populacional							
Total	5 565	100,0	5,1	1,9	65,5	62,7	11,8
Até 2 000 hab.	117	2,1	6,1	3,2	71,8	51,3	16,2
De 2 001 a 5 000 hab.	1 184	21,3	5,6	3,1	72,1	56,4	14,3
De 5 001 a 10 000 hab.	1 212	21,8	5,4	2,4	69,4	58,8	12,7
De 10 001 a 20 000 hab.	1 401	25,2	5,5	1,8	64,5	60,8	13,2
De 20 001 a 50 000 hab.	1 043	18,7	4,4	0,4	58,0	71,2	8,4
De 50 001 a 100 000 hab.	325	5,8	3,5	0,2	59,7	78,5	5,5
De 100 001 a 500 000 hab.	245	4,4	4,4	1,5	58,0	69,8	9,0
De 500 001 a 1 000 000 hab.	23	0,4	3,7	2,2	65,2	65,2	0,0
Mais de 1 000 000 hab.	15	0,3	3,7	2,9	60,0	80,0	6,7
Taxa média de crescimento geométrico							
Menor que -5,00%	14	0,3	17,7	(-) 17,1	14,3	14,3	57,1
De -4,99% a -2,50%	34	0,6	9,3	(-) 7,6	5,9	32,4	29,4
De -2,49% a -1,00%	259	4,7	5,5	(-) 2,9	37,1	62,2	14,3
De -0,99% a 0,00%	1 008	18,1	4,4	0,6	57,8	69,4	8,9
De 0,01% a 1,00%	2 439	43,8	4,3	2,0	70,5	69,4	7,6
De 1,01% a 2,50%	1 569	28,2	5,9	3,0	69,5	52,9	15,2
De 2,51% a 5,00%	226	4,1	10,2	5,4	63,3	38,9	35,0
Mais que 5,00%	16	0,3	14,5	10,9	68,8	25,0	56,3

Fonte: Obtidas com base nas Estimativas das populações residentes, em 1º de julho de 2008, segundo os municípios - IBGE (IBGE, 2008); Estimativas das populações residentes, em 1º de julho de 2009, segundo os municípios - IBGE (IBGE, 2009); Censo Demográfico 2010.

Nota: Dados para 2010 comparados aos dados observados no Censo Demográfico 2010.

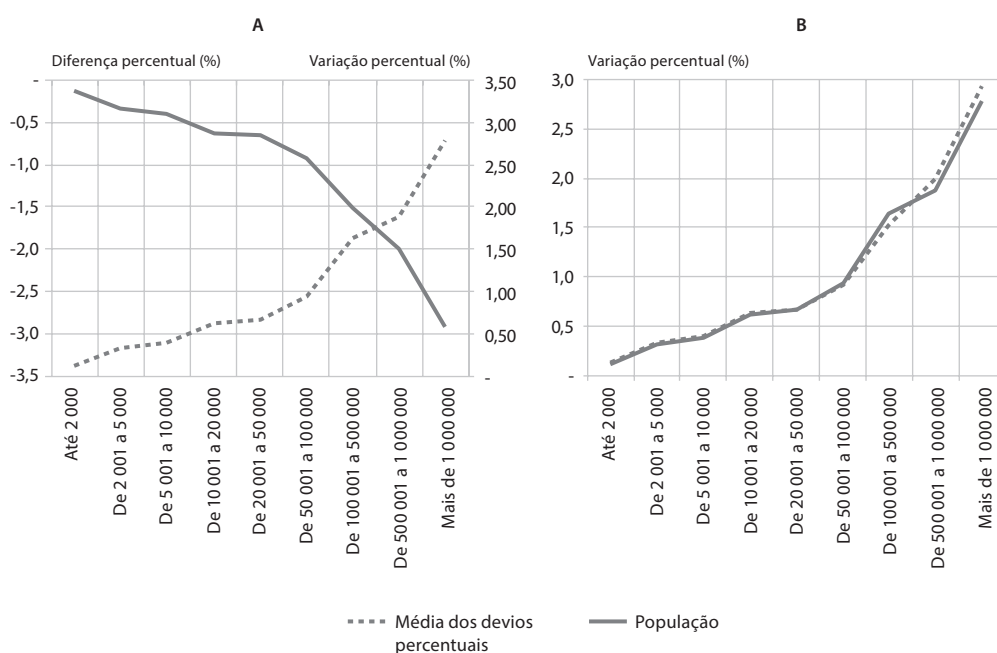
(1) Distribuição segundo dados enumerados no Censo Demográfico 2010.

Verificaram-se indícios da tendência de maiores desvios, em média, para os municípios onde ocorreram as maiores variações no percentual entre a população enumerada e a população enumerada menos imputada. Pode-se observar que há uma tendência de redução no valor das diferenças das médias dos DP para as faixas de tamanho populacional municipal onde houve maior percentual de população imputada (Gráfico 3 (a)).

Os resultados do Gráfico 3 (b) indicam que as diferenças dos desvios percentuais das estimativas calculadas em relação à população enumerada e dos desvios percentuais calculados em relação à população enumerada sem imputação variam, em média, de forma semelhante à variação do percentual da população imputada em relação à população enumerada, de acordo com cada faixa de tamanho populacional municipal. O aumento na amplitude da diferença entre a população enumerada e a população enumerada menos a imputada foi semelhante à amplitude da diferença, em valor absoluto, entre a média dos DP da população enumerada e as

médias dos DP da população enumerada menos a imputada. Por exemplo, para uma variação positiva de aproximadamente 1,0% da população imputada na faixa de tamanho populacional municipal entre 50 mil a 100 mil habitantes, verificou-se a diferença de aproximadamente 1,0 ponto percentual entre as médias dos DPs da população enumerada e a média dos DPs da população sem imputação. Semelhantemente, uma variação positiva de cerca de 3% na população na faixa populacional municipal com mais de 1 milhão de habitantes, verificou-se a diferença de aproximadamente 3 pontos percentuais na diferença entre as médias dos DPs.

Gráfico 3 - Diferença percentual das médias dos desvios percentuais e variação percentual populacional (A), e diferença percentual absoluta das médias dos desvios percentuais, e variação percentual populacional (B), baseadas nas diferenças entre população enumerada e população enumerada menos população imputada no Censo Demográfico, segundo as faixas de tamanho populacional - Brasil - 2010



Fonte: Obtidas com base nas Estimativas das populações residentes, em 1º de julho de 2008, segundo os municípios – IBGE (IBGE, 2008); Estimativas das populações residentes, em 1º de julho de 2009, segundo os municípios – IBGE (IBGE, 2009); Censo Demográfico 2010.

Na ausência de informações detalhadas sobre os ajustes populacionais realizados pelos especialistas do IBGE durante o processo de obtenção de estimativas para 2008 e 2009, o método aqui aplicado admitiu como suposto que quaisquer ajustes de população para os municípios brasileiros (por exemplo, os provenientes de alguma modificação de limites territoriais municipais, etc.) tenham sido incorporados no cálculo das estimativas populacionais municipais para 2008 e 2009, as quais foram utilizadas como insumo neste estudo. E ainda, que quaisquer mudanças ocorridas nos limites municipais entre 2009 e 2010 não seriam consideradas.

Importante registrar que no Brasil, em alguns casos ao longo do tempo, houve definição de quantitativo populacional municipal por meio de decisão judicial para efeito, por exemplo, de distribuição do Fundo de Participação dos Municípios (Lei nº 8.443, de 16 de julho de 1992). Mais recentemente, o IBGE tem informado sobre os valores populacionais obtidos por processos de obtenção de projeções/estimativas de população e aqueles referentes a decisões judiciais⁹. Dos dados aqui utilizados, apenas para o município Uruará/PA, em 2009, houve informações sobre o valor populacional obtido por decisão judicial, sem explicações dos motivos.

⁹ Vide arquivos dispondo das estimativas da população residente no Brasil e unidades da federação com data de referência em 1º de julho em: <<https://www.ibge.gov.br/estatisticas-novoportal/sociais/populacao/9103-estimativas-de-populacao.html?=&t=downloads>> e <<https://sidra.ibge.gov.br/tabela/6579#>>

Há uma dificuldade de identificar explicações precisas para as diferenças encontradas entre as quantidades populacionais municipais estimadas e aquelas provenientes da enumeração censitária. Por exemplo, como explicar um DP de cerca de 200% encontrado para o município de Jacareacanga/PA? Se for observado para este município, por exemplo, os valores populacionais enumerados pelo Censo Demográfico de 2000 (cerca de 24 mil habitantes) e pela Contagem de População de 2007, aproximadamente 37 mil habitantes. Pela tendência de crescimento observada no período, era de se esperar, pelo método adotado pelo IBGE (*idem*), um incremento populacional no município para os anos seguintes, cujas estimativas para totais populacionais foram de aproximadamente 39 mil e 41 mil habitantes em 2008 e 2009, respectivamente. O Censo Demográfico 2010 apresentou, entretanto, enumeração populacional para Jacareacanga/PA em torno de 14 mil habitantes, supostamente um decréscimo populacional de cerca de 62% em relação à Contagem de População de 2007, e de 65% em relação à estimativa de população para 2009.

Nesse caso de Jacareacanga/PA, o que teria determinado essa enorme diferença? Incertezas associadas aos processos de enumeração censitária ou aos processos de obtenção de estimativas populacionais para o domínio de interesse? A resposta para essas questões não é simples nem imediata, mas há uma grande probabilidade que, neste caso, ter sido a enumeração de algum dos censos demográficos que apresentou problema: ou em 2000 e 2007 houve uma expressiva sobre-enumeração populacional neste município ou em 2010 houve uma substancial subenumeração.

Esse único caso não permite generalizar que todos os demais desvios são devidos a enumerações e não devido a erros nas estimativas. No entanto, o caso deixa patente que existem problemas de enumeração censitária. Portanto, uma supervisão de população esperada deve ser realizada com muito detalhe durante o levantamento de dados censitários para minimizar sub ou sobre enumerações. A decisão judicial¹⁰ sobre esse caso, após a realização do Censo Demográfico de 2010, mantém a população de Jacareacanga/PA com o valor estimado de 41.487 habitantes, mas não estão disponíveis informações mais detalhadas sobre os motivos que levaram a esta decisão. No entanto, o valor estimado da população do município durante a década permaneceu com o valor estimado a partir do Censo Demográfico de 2010, visto que a alteração implicaria “retirar” população de outros municípios, dado que, no método *AiBi*, a soma das áreas menores devem compor a população da área maior, neste caso o estado do Pará, que teve população estimada para a década de 2010 pelo método das Componentes Demográficas.

Considerações finais

O principal objetivo deste estudo foi realizar uma análise de estimativas para os totais populacionais dos municípios brasileiros quanto às possíveis diferenças, ou desvios, quando comparadas com os totais populacionais obtidos por meio de operação censitária.

Os métodos de obtenção de cifras populacionais, enumerações censitárias e procedimentos de projeções/estimativas populacionais, como visto, estão sujeitos a algum tipo de erro. Experiências em produção de projeções e estimativas populacionais para pequenos domínios, vivenciadas por pesquisadores e/ou especialistas de agências produtoras de estatísticas, apontam para a possibilidade de adoção de diferentes métodos para esse fim. A escolha do método mais adequado, em diferentes casos, é uma tarefa que requer além de conhecimento das diferenças técnicas e suas especificidades, um saber acerca da área a ser projetada/estimada (WALDVOGEL, 1997), e da disponibilidade de informações confiáveis, detalhadas e atualizadas que subsidiem a construção de projeções e estimativas populacionais para os domínios de interesse. Deve-se considerar que a acurácia da informação populacional proveniente de métodos de projeção/estimativas populacionais não depende apenas da forma como são operacionalizadas as funções

¹⁰ Desde 2012 é informado que “Por determinação judicial e para efeito de distribuição do Fundo de Participação dos Municípios, a população do Município de Jacareacanga-PA é de 41.487 habitantes. Processo Judicial nº 798-41.2011.4.01.3902, Seção Judiciária de Itaituba-PA.” (Online).

dos cálculos matemáticos, estatísticos ou demográficos utilizados no processo, pois associa-se fortemente à qualidade e quantidade de insumos, e de ajustes necessários.

Enumerações populacionais, de outra forma, não estão isentas de incertezas quanto ao quantitativo populacional levantado em uma operação censitária em função de *erros* associados a *falhas no processo* de coleta de dados, tais como: omissão ou duplicidade de unidades de pesquisa, domicílios fechados, não resposta e/ou imprecisão na mensuração de um ou mais quesitos, respostas errôneas ou equivocadas (intencionais ou não), codificação de quesitos, etc. Tais falhas implicam em impacto, tanto na quantificação do total populacional para diversos domínios de interesse, quanto na caracterização e estrutura das respectivas populações alvo.

Os resultados encontrados neste estudo mostraram o padrão das diferenças verificadas entre as estimativas populacionais municipais calculadas e os resultados do Censo Demográfico 2010 com os valores resultantes da população imputada para domicílios fechados. De forma geral, em média, houve um desvio absoluto de aproximadamente 5% entre as estimativas calculadas e suas respectivas cifras populacionais enumeradas pelo Censo Demográfico 2010, mesmo após imputação de domicílios fechados, quando a estimativa de subenumeração do Censo Demográfico de 2010, se comparada com o esperado pela Revisão das Projeções Populacionais de 2008, seria de 2,4% (NACIONES UNIDAS, 2013). Verificou-se que a maioria dos municípios (cerca de 89%) apresentaram desvios percentuais, em valor absoluto, de até 10%. Cerca de 63% dos municípios apresentaram desvios absolutos menores do que 5%. O processo de obtenção das estimativas resultou em valores populacionais maiores do que a cifra populacional enumerada no censo demográfico em aproximadamente 65% dos casos.

Há indícios de que a magnitude do desvio se relaciona com o tamanho populacional do município. Em média, os desvios percentuais, em valor absoluto, apresentaram tendência de redução com o aumento populacional dos municípios, e municípios que apresentaram menores taxas médias de crescimento geométrico apresentaram, em média, menores desvios.

Os resultados encontrados evidenciam padrões semelhantes quanto ao comportamento dos desvios calculados com base na população sem imputação populacional, porém com diferença no nível que se altera de acordo com a variação da magnitude da população imputada.

Os resultados também foram semelhantes quando comparados em outros níveis geográficos, como os de grandes regiões, em que os desvios percentuais comparados, tanto em relação à população enumerada quanto para a enumerada sem imputação populacional, se assemelham na forma, variando a amplitude de acordo com o tamanho populacional municipal.

Não é difícil imaginar que qualquer outro método que aqui fosse aplicado certamente produziria estimativas com algum desvio, de maior ou menor magnitude, em relação à população enumerada pelo censo demográfico. Isto é fato, pois dependeriam, também, de avaliação, escolhas e ajustes que se fazem necessários no processo, em função das especificidades (insumos, suposições e limitações) inerentes a cada metodologia.

Nesse sentido, a avaliação de desvios, utilizando-se vários métodos diferentes, é uma importante ferramenta no processo de elaboração de projeções e estimativas populacionais, pois permite ao analista realizar ajustes na aplicação da metodologia ou, ainda, avaliar duas ou mais técnicas na escolha de qual seria a mais apropriada para a produção de projeções/estimativas populacionais mais plausíveis à realidade.

Para finalizar, considerando o ambiente de incertezas que cercam a obtenção de estimativas populacionais e aquelas envolvidas em um levantamento censitário, é importante que avaliações periódicas sejam realizadas, assim como ocorra o empenho máximo de esforços possíveis, seguindo as recomendações e experiências internacionais e utilizando ferramentas tecnológicas e metodologias avançadas, para obter melhor enumeração nos levantamentos censitários. Importa sempre estar cientes, em todos os casos, que o objetivo é obter valores com os menores erros possíveis e conhecer as margens de erros, desvios ou diferenças nas cifras populacionais. Em última instância, deve-se primar pela utilização desses dados, sujeitos a erros de medida ou de estimação, de maneira adequada, nos mais variados programas e políticas públicas sabendo-se de suas imprecisões.

Referências

AUSTRALIAN BUREAU OF STATISTIC. *ABS corporate plan*. Belconnen. 2012. Disponível em: [http://www.ausstats.abs.gov.au/Ausstats/subscriber.nsf/0/C581FC0D399660D5CA257A4B001468A2/\\$File/abs_corp_plan_web1.pdf](http://www.ausstats.abs.gov.au/Ausstats/subscriber.nsf/0/C581FC0D399660D5CA257A4B001468A2/$File/abs_corp_plan_web1.pdf). Acesso em: jun. 2013.

_____. *Population estimates: concepts, sources and methods*. Belconnen. 2009. Disponível em: [https://www.ausstats.abs.gov.au/ausstats/subscriber.nsf/0/7B92336BE30E0D7DCA2575D2001B0013/\\$File/3228055001_2009.pdf](https://www.ausstats.abs.gov.au/ausstats/subscriber.nsf/0/7B92336BE30E0D7DCA2575D2001B0013/$File/3228055001_2009.pdf). Acesso em: jun. 2013.

BAY, G. Reflexiones sobre las estimaciones y proyecciones de población en América Latina: innovaciones metodológicas y dificultades para implementarlas. In: CAVENAGHI, S. M. (org.). *Estimaciones y Proyecciones de Población em América Latina: desafios de uma agenda pendiente*. 1.ed. Rio de Janeiro: ALAP, 2012. v. 1. 262p. Disponível em: http://www.alapop.org/alap/Serie-E-Investigaciones/N2/Serie_e-InvestigacionesN2.pdf.

BARROS, L. F. W.; CAVENAGHI, S. M. *Potencialidades e desafios no uso de registros administrativos nas estimativas populacionais municipais: uma análise exploratória para o Brasil no período 1990-2010*. VII Congresso da Associação Latino Americana de PopulaçãoXX Encontro Nacional de Estudos Populacionais. 17 a 22 de outubro de 2016. Foz do Iguaçu, Paraná. Disponível em: <http://www.abep.org.br/publicacoes/index.php/anais/article/view/2730/2629>, Acesso em: mar. 2019.

BORGES, A. S. *et al. Projeções Populacionais no Brasil: Subsídios para o seu aprimoramento*. XV Encontro Nacional de Estudos Populacionais. ABEP. 18 a 22 de setembro de 2006. Caxambú, Minas Gerais. Disponível em: <http://www.abep.org.br/publicacoes/index.php/anais/article/view/1516>. Acesso em: mar. 2012.

BORGES, G. M.; ERVATTI, L. R.; SILVA, L. G. C. Desafios para o IBGE nas estimativas populacionais dos municípios brasileiros: aplicação de distintas metodologias. In: INTERNATIONAL SEMINAR ON POPULATION ESTIMATES AND PROJECTIONS: Methodologies, innovations and estimation of target population applied to public policies. ALAP. 09 a 11 de novembro de 2011. CIC/IBGE, Rio de Janeiro.

BRITO, L. P. G. *Avaliação da precisão de estimativas e projeções populacionais para pequenos domínios: Rio de Janeiro 1980-2007*. Rio de Janeiro, 2007. 162 f. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) - Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2007.

BRITO, L. P. G.; CAVENAGHI, S.; JANNUZZI, P. M. *Estimativas e projeções populacionais para pequenos domínios: uma avaliação da precisão para municípios do Rio de Janeiro em 2000 e 2007*. REBEP. Rio de Janeiro, V. 27, n.1, p. 35-57, jan/jul. 2010.

CAVENAGHI, S. M. El seminario Estimaciones y Proyecciones de Población - metodologias, innovaciones y estimación de grupos objetivo de políticas públicas: informe detallado y comentado. In: CAVENAGHI, S. M. (org.). *Estimaciones y Proyecciones de Población em América Latina: desafios de uma agenda pendiente*. 1.ed. Rio de Janeiro: ALAP, 2012. v. 1. 262p. Disponível em: http://www.alapop.org/alap/Serie-E-Investigaciones/N2/Serie_e-InvestigacionesN2.pdf.

CEPAL/CELADE. Recomendaciones para los censos de la década de 2010 en América Latina. *Serie Manuales*. n. 72. CEPAL/CELADE/Nações Unidas. Santiago de Chile. 2011. Disponível em: http://repositorio.cepal.org/bitstream/handle/11362/5511/1/S1100525_es.pdf. Acesso em: abr. 2012.

CHACKIEL, J. Evaluación post-empadronamiento de la cobertura en los censos de población. *Notas de población*. n. 91. CEPAL. 2010. Disponível em: http://repositorio.cepal.org/bitstream/handle/11362/37691/1/S1000525_es.pdf. Acesso em: abr. 2012.

CORRÊA, C. S.; MYRRHA, L. J. D.; FÍGOLI, M. Métodos AiBi e Logístico para projeções de pequenas áreas: uma aplicação para a microrregião de Angicos - RN. In: INTERNATIONAL SEMINAR ON POPULATION ESTIMATES AND PROJECTIONS: Methodologies, innovations

and estimation of target population applied to public policies. ALAP. 09 a 11 de novembro de 2011. CIC/IBGE, Rio de Janeiro.

GONZÁLEZ, L.M; TORRES, E. Estimaciones de Población en Áreas Menores en América Latina: revisión de métodos utilizados. In: CAVENAGHI, S. (org.). ASOCIACIÓN LATINO-AMERICANA DE POBLACIÓN. *Estimaciones y proyecciones de población en América Latina: desafíos de una agenda pendiente*. Rio de Janeiro: ALAP, 2012, p.71-103. (Serie e-Investigaciones; n. 2).

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). *As projeções de população*. Rio de Janeiro, Fundação IBGE. 1968

_____. *Banco de Metadados*. Rio de Janeiro, Fundação IBGE. Disponível em: <http://www.metadados.ibge.gov.br/detalhePesquisa.aspx?cod=CD>. Acesso em: maio 2013.

_____. *Base de informações do Censo Demográfico 2010: Resultados do Universo por setor censitário*. Rio de Janeiro, Fundação IBGE. 2011.

_____. *Contagem da População 2007*. Rio de Janeiro, Fundação IBGE. 2007a.

_____. *Estimativa de população, para os municípios brasileiros, em 1º de julho de 1965*. Rio de Janeiro, Fundação IBGE. 1965.

_____. *Estimativa de população, para os municípios brasileiros, em 1º de julho de 1967*. Rio de Janeiro, Fundação IBGE. 1967.

_____. *Estimativa de população, para os municípios brasileiros, em 1º de julho de 1969*. Rio de Janeiro, Fundação IBGE. 1969.

_____. *Estimativa de população, para os municípios brasileiros, em 1º de julho de 1990, segundo as unidades da federação e municípios*. Rio de Janeiro, Fundação IBGE. 1990.

_____. *Estimativa de população, para os municípios brasileiros, em 1º de julho de 1991, segundo as unidades da federação e municípios*. Rio de Janeiro, Fundação IBGE. 1991.

_____. *Metodologia das Estimativas da População residente nos municípios brasileiros para 1º de julho de 2011*. 2. ed. Rio de Janeiro: IBGE, 2011. Disponível em: <https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=2101695>. Acesso em: mar. 2012.

_____. *Metodologia das Estimativas das Populações de 128 municípios e Distrito Federal para 2007*. Rio de Janeiro: IBGE, 2007b. Disponível em: http://ftp.ibge.gov.br/Contagem_da_Populacao/Contagem_da_Populacao_2007/Notas_Tecnicas. Acesso em: mar. 2012.

_____. *Pesquisa Mensal de Emprego Reponderação das estimativas, 2009a*. Disponível em: http://ftp.ibge.gov.br/Trabalho_e_Rendimento/Pesquisa_Mensal_de_Emprego/Notas_Tecnicas/notatecnicareponderacao2009.pdf. Acesso em: maio 2012.

_____. *Projeções da população: Brasil e unidades da federação: revisão 2013*. IBGE, Coordenação de População e Indicadores Sociais. 2. ed. Rio de Janeiro: IBGE, 2013.

_____. *Projeções da população: Brasil e unidades da federação: revisão 2018*. IBGE, Coordenação de População e Indicadores Sociais. 2. ed. Rio de Janeiro: IBGE, 2018.

_____. *Projeção da população do Brasil por sexo e idade 1980-2050: revisão 2008*. Rio de Janeiro: IBGE, 2008. Disponível em: http://www.ibge.gov.br/home/estatistica/populacao/projecao_da_populacao/2008/projecao.pdf. Acesso em: mar. 2012.

_____. *Projeções da população: Brasil e Unidades da Federação (Relatório Metodológico n. 40)*. Rio de Janeiro, 2013. Disponível em: <https://biblioteca.ibge.gov.br/visualizacao/livros/liv86444.pdf>.

JARDIM, M. L. T. *Metodologias de estimativas e projeções populacionais para o Rio Grande do Sul e seus municípios*. Série Documentos FEE n. 56, outubro de 2003. Disponível em: http://www.fee.rs.gov.br/sitefee/download/documentos/documentos_fee_56.pdf.

MESA, M. J. R; VALENZUELA, M. F. *Teoría y métodos demográficos para elaboración de estimaciones y proyecciones de población* (Insumos para la Planificación). Bogotá D.C., Mayo 2007.

NACIONES UNIDAS. *Gestión orientada a asegurar la calidad de los datos en los institutos nacionales de estadística*. In: SEGUNDA REUNIÓN DE LA CONFERENCIA ESTADÍSTICA DE LAS AMÉRICAS DE LA COMISIÓN ECONÓMICA PARA AMÉRICA LATINA Y EL CARIBE. SANTIAGO DE CHILE. 2003. Disponível em: <http://www.eclac.org/deype/ceacepal/documentos/lcl1889e.pdf>. Acesso em: jun. 2013.

_____. *Manual de organización estadística*. El funcionamiento la organización de una oficina de estadística. Estudios e métodos. Serie F n° 88. Tercera edición. Nueva York. 2004. Disponível em: http://repositorio.cepal.org/bitstream/handle/11362/3976/1/S2004021_es.pdf. Acesso em: jun. 2013.

_____. *Principios y recomendaciones para los censos de población y habitación*. *Informes estadísticos*. Series M n° 67. Rev2. Nueva York. 2010.

_____. *Principales cambios en las boletas de los censos latinoamericanos de las décadas de 1990, 2000 y 2010*. *Serie Manuales*. Santiago de Chile, Novembro, 2013.

OFFICE FOR NATIONAL STATISTICS. *Methodology Note on production of super outputs area population estimates*. April 2013.

_____. *Quality and methodology information*. March 2012.

PITTA, M. T.; DINI, N. P.; JANNUZZI, P. M. *Modelos para produção de estimativas demográficas para bairros e domínios intraurbanos: a experiência para o município de São Paulo*. Trabalho apresentado no XVII Encontro Nacional de Estudos Populacionais, ABEP, realizado em Caxambú – MG – Brasil, de 20 a 24 de setembro de 2010.

RAO, J. N. K. *Small area estimation*. Hoboken, N.J.: J. Wiley, c2003. xxiii, 313 p. (Wiley series in survey methodology).

SCHWARTZMAN, S. *Legitimidade, Controvérsias e Traduções em Estatísticas Públicas*. 1996. Disponível em: <http://www.schwartzman.org.br/simon/estpub.htm>. Acesso em: jun. 2013.

SANTOS, T. F. *Projeções da população de Pernambuco, desagregadas por microrregiões, até o ano 2010: aplicação de métodos alternativos*. Recife: UFMG/Cedeplar, Dissertação de Mestrado, 1989. 181p.

SÄRNDAL, C.-E.; SWENSSON, B; WRETMAN, J.H. *Model assisted survey sampling*. New York: Springer-Verlag, c1992. xv, 694 p. (Springer series in statistics).

SMITH, S. K; CODY, S. An evaluation of population estimates in Florida: April 1, 2000. *Population Research and Policy Review*. 23, 1-24, 2004.

STATISTICS CANADA. *2006 Census Technical Report: Coverage*. Social Survey Methods Division. Ottawa. 2010.

SWANSON, D. A.; TAYMAN, J. *Subnational population estimates*. Springer Netherlands. 2012.

_____. *Population and Family Estimation Methods at Statistics Canada*. Demography Division. Ottawa. 2012.

U.S. CENSUS BUREAU. *Methodology for the United States and County Population Estimates by Age, Sex, Race, and Hispanic Origin (Vintage 2012)*: April 1, 2010 to July 1, 2012a.

_____. *Methodology for the Porto Rico Commonwealth and Its Municipios Estimates by Age, Sex (Vintage 2012)*: April 1, 2010 to July 1, 2012b.

UNITED NATIONS. *Dicionário Demográfico Multilíngüe*. (Português - projeto da tradução da segunda edição). Demopaedia, 1969. Disponível em: <http://pt-ii.demopaedia.org/wiki/70#702>. Acesso em: fev. 2013.

_____. *Principles and Recommendations for Population and Housing Censuses Revision 2*. Statistical Papers. Series M n° 67.Rev2. New York. 2008.

_____. *World Population Prospects: The 2004 Revision*, vol. III, Analytical Report (United Nations publication, Sales No. E.05.XIII.7) 2005. Disponível em: http://www.un.org/esa/population/publications/WPP2004/WPP2004_Vol3_Final/Chapter6.pdf. Acesso em: abr. 2012.

_____. *World Population Prospects: The 2010 Revision. Volume II: Demographic Profiles*. United Nations publication. 2011. Disponível em: http://esa.un.org/unpd/wpp/Documentation/pdf/WPP2010_Volume-II_Demographic-profiles.pdf. Acesso em: abr. 2012.

WALDVOGEL, B. *Técnicas de projeção populacional para o planejamento regional*. Belo Horizonte: UFMG/Cedeplar, 1997 (Estudos Cedeplar, 1).

Economia Política da informação estatística: por uma linha de pesquisa sobre as conexões entre estatísticas e Políticas Públicas*

Paulo de Martino Jannuzzi**

Luiz Marcelo Carvano***

Leonardo Milhomem Rezende****

Baiena Feijolo*****

Resumo

O objetivo do trabalho é fazer uma reflexão sobre o contexto atual do Sistema Estatístico e as Políticas Públicas no Brasil Contemporâneo. A tese subjacente a esse texto, e para a qual vários estudos ainda são necessários na comunidade epistêmica dos Estudos Populacionais e da Ciência Política, é de que as Estatísticas Sociais foram funcionais e muito pertinentes para a especificação, implantação e legitimação de uma série de políticas públicas nos últimos 30 anos. Sem a produção regular e consistente das Estatísticas Públicas, a efetividade social das Políticas Públicas seria certamente menor, assim como a capacidade de aprimorá-las. O texto inicia com uma revisão histórica sobre Estado de Bem-Estar Social e as Estatísticas, passando para uma exposição sobre a instrumentalidade das mesmas para o ciclo de Políticas Públicas. A guisa de conclusão, apresenta-se alguns dos desafios institucionais da produção de informação no contexto brasileiro atual. Um desses desafios, com reflexos para os programas de Pós Graduação da Escola é, certamente, o fortalecimento de uma linha de pesquisa em Economia Política da Informação Estatística.

Palavras-chave: Sistema Estatístico. Estado de Bem-Estar. Economia Política da Informação.

* A elaboração desse texto se deve a duas motivações centrais. De um lado, ele é uma resposta às preocupações com o contexto crescente de deslegitimação e de dificuldades de recursos orçamentários, humanos e tecnológicos que os Sistemas Estatísticos vêm passando no esteio da desacreditação do Estado de Bem-Estar Social que se passou a conviver nos últimos 10 anos. De outro, em boa medida como consequência dessa preocupação, este texto procura incentivar a elaboração de estudos sobre Sistema Estatístico e Políticas Públicas ou ainda, sobre Economia Política da Informação Estatística. No marco de vinte anos do Programa de Pós-Graduação em População, Território e Estatísticas Públicas parece oportuno retomar questões e temáticas que, presentes em dissertações de alunos e pesquisas de professores, não ganharam o espaço que podem e devem ter, pela origem de seus docentes, natureza do programa e sua vinculação institucional com o IBGE.

** Professor da Escola Nacional de Ciências Estatísticas do Instituto Brasileiro de Geografia e Estatística (ENCE/IBGE), Pesquisador PQ/CNPq e Pesquisador em Estágio Pós-Doutoral na EBAPE/FGV.

*** Sociólogo pela UFRJ, bolsista da Fundação Ford na ENCE (2000-2002) e mestre em Gestão da Informação pela Universidade Nova de Lisboa/IMS

**** Auxiliar de Assuntos Educacionais do Ministério da Educação (MEC) e Mestre em Gestão de Políticas e Sistemas Educacionais pela Universidade de Brasília (UnB).

***** Coordenadora de Projetos da FGV Projetos, Mestre em Estudos Populacionais e Pesquisas Sociais pela ENCE/IBGE.

Introdução

Estatísticas são, como a etimologia da palavra revela, assuntos – fatos e dados – acerca do Estado. Se foram fundamentais para constituição do Estado Moderno, no dimensionamento da população dos países e de seus exércitos, na extensão de suas fronteiras e da cobertura da base tributária, as estatísticas cumprem, para o Estado Contemporâneo, desígnios mais abrangentes, civilizatórios e certamente mais complexos, seja nas decisões dos agentes econômicos, seja nas atividades de formulação, monitoramento e avaliação de políticas públicas.

Fato é que, ao longo do Século XX, a ampliação do escopo e a regularidade das Estatísticas Públicas acompanharam *pari-passu* o fortalecimento do Estado e das políticas públicas pelo mundo afora. Nesse período, parece ter imperado uma lógica virtuosa e auto-reforçadora entre Estatísticas e Políticas Públicas, em que demandas mais sofisticadas de diagnósticos e desenho de políticas e programas requereram e viabilizavam pesquisas mais detalhadas que, com seus indicadores derivados e mais específicos, conduziram a mudanças incrementais e inovações nas políticas e programas. Ações mais efetivas do Estado, por sua vez, levavam a novas necessidades de informação, retomando nova etapa do ciclo.

No caso do Brasil esse processo é mais recente, particularmente mais intenso após a Constituição Federal de 1988, que ampliou consideravelmente os direitos sociais para sociedade e os deveres do Estado em provê-los por meio de políticas e programas públicos. Se é verdade que o IBGE foi criado muito tempo antes, há mais de 80 anos, como resposta à necessidade de fornecer insumos para estruturação de um projeto pioneiro- mas marcadamente autoritário- de Nação, com foco sobretudo na dimensão econômico-territorial, é somente nos últimos 30 anos que o instituto desenvolveu mais amplamente sua capacidade de responder às demandas mais abrangentes de políticas públicas da sociedade brasileira. Sem negar sua relevância enquanto instituição governamental pela qualidade de seus produtos e pesquisas nesse largo período, seu reconhecimento como órgão de Estado se concretiza à medida que começa a se vertebrar no país um conjunto mais amplo de políticas públicas, notadamente as sociais, para o qual o instituto é chamado, por diferentes atores, a produzir informações multitemáticas sobre a realidade social, ambiental e geocientífica, para além da econômica. Se antes a modernização do Sistema Estatístico brasileiro era movido pela replicação, sobretudo a partir dos anos 1970, da agenda internacional no campo, reverberada pelos institutos de estatística dos países desenvolvidos, pela Divisão Estatística das Nações Unidas e pelos organismos multilaterais (Organização Internacional do Trabalho, Fundo de População das Nações Unidas etc), nos anos seguintes à promulgação da Constituição, é esta que vai definir pautas e planos de trabalho mais específicos.

Nos últimos anos o contexto mudou de forma radical. Como bem analisa Blyth (2017), a internalização da dívida dos grandes bancos internacionais no orçamento público dos países desenvolvidos, após a crise no setor em 2007/2008, com as consequências macroeconômicas geradas desde então, acabou reacendendo a sempre presente – embora nem sempre vitoriosa- crítica liberal ao Estado de Bem-Estar Social. Em um mundo pós-“Queda do Muro de Berlin”, sem contraponto ao modelo capitalista, tal crítica parece ter amplificado sua influência, com desdobramentos mais intensos e decisivos que em outros momentos análogos- como na década de 1980, de proeminência das ideias do Consenso de Washington- sobre a legitimidade e financiamento das Políticas Sociais e, por conseguinte, das Estatísticas Públicas. Se antes o custo de produção das Estatísticas se legitimava- e se internalizava - pelo reconhecimento de sua utilidade instrumental no diagnóstico, desenho e avaliação das Políticas Públicas, no presente, pesquisas e instituições passaram a conviver com dificuldades de recursos orçamentários, humanos e tecnológicos e um novo ambiente de negócios. Em vários países iniciam-se esforços, até então inéditos, para se demonstrar que a produção da informação estatística é altamente custo-efetiva, isto é, que os custos empregados na realização de pesquisas amostrais, de censos demográficos, na manutenção de registros administrativos e sistemas de informação geográfica são plenamente justificáveis em função da utilidade instrumental para agentes públicos e privados em uma grande variedade de situações e aplicações (UNECE, 2017).

Refletir sobre esse ambiente contextual da Produção Estatística no momento atual é o objetivo deste texto. Trata-se menos de apresentar e defender uma tese específica – ainda que ela, de fato exista –, mas ensaiar contribuições para análise situacional das instituições públicas dedicadas à produção de estatísticas e pesquisas e para incentivar a elaboração de estudos sobre Sistema Estatístico e Políticas Públicas ou ainda, sobre Economia Política da Informação Estatística. No marco de vinte anos do Programa de Pós-Graduação em População, Território e Estatísticas Públicas parece oportuno retomar questões e temáticas que, presentes em dissertações de alunos e pesquisas de professores, não ganharam o espaço que podem e devem ter, pela origem de seus docentes, natureza do programa e sua vinculação institucional com o IBGE.

Em que pesem a relevância, ineditismo e mérito de pesquisas desenvolvidas ou orientadas nessas temáticas no âmbito do programa – particularmente pelos ex-professores Nelson Senra, Jane Souto e Neide Patarra- e os registros institucionais que o Centro de Documentação e Disseminação de Informações vem publicando nas últimas décadas¹, as Estatísticas Públicas ainda não constituíram objeto de reflexão sistemática como outros campos da Demografia e das Ciências Sociais Aplicadas. Se não é óbvio, e não parece ser- o que revela a pertinência desse texto- as Estatísticas cumprem e cumpriram papel muito relevante na defesa da prioridade política, nos esforços de implementação de programas e na obtenção de resultados das políticas sociais no Brasil.

A tese subjacente a esse texto, e para a qual vários estudos ainda são necessários, é de que as Estatísticas Sociais foram funcionais e muito pertinentes para a especificação, implantação e legitimação de uma série de políticas públicas nos últimos 30 anos. É fato que sem a determinação constitucional, compromisso político e a implantação de políticas sociais específicas o país não teria logrado a ampliação de atendimento na educação infantil, a universalização no ensino fundamental, a mitigação da fome e insegurança alimentar, a inserção laboral mais qualificada para segmentos expressivos da força de trabalho, o acesso à Proteção Social por larga maioria da população idosa ou vulnerável, para citar alguns dos resultados da agenda das políticas públicas “contratadas” desde a Constituição de 1988. Mas sem a produção regular e consistente das Estatísticas Públicas, a efetividade social das Políticas Públicas seria certamente menor, assim como a capacidade de defende-las e aprimorá-las.

O texto está estruturado em duas seções, além dessa introdução e da conclusão. Nessas seções discute-se as conexões entre a formação do Estado de Bem-Estar Social, o fortalecimento das Políticas Públicas e a estruturação dos Sistemas Estatísticos ao longo do Século XX. À guisa de conclusão, apresenta-se alguns dos desafios institucionais da produção de informação no contexto brasileiro atual.

A constituição do Estado de Bem-Estar ao longo do século XX

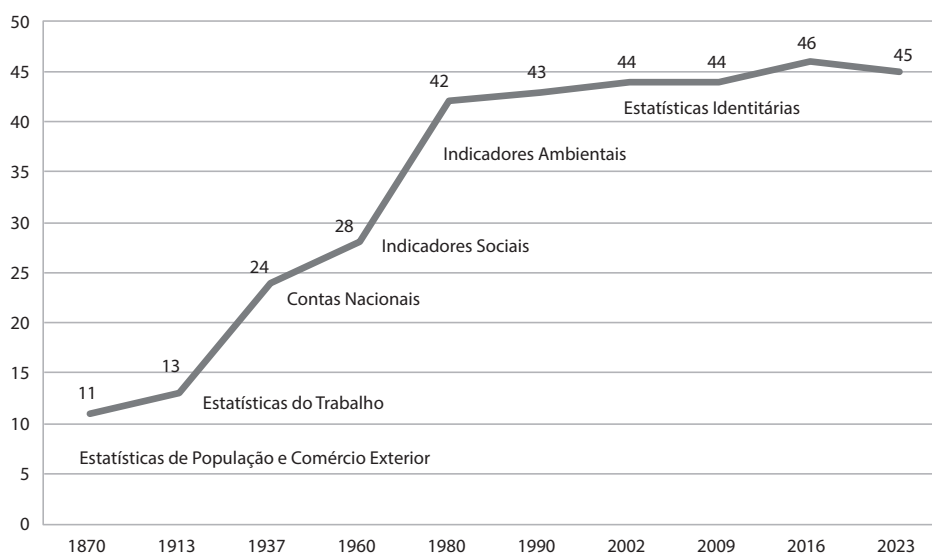
Até final do século XIX, as estatísticas de população, mortalidade e de comércio exterior eram suficientes para a gestão do então Estado Regulador e suas funções de administração da justiça, manutenção da burocracia de arrecadação tributária, de provimento de segurança interna e defesa nacional. Como mostra Tanzi (2005), para um Estado com essas funções básicas, bastava uma carga fiscal não muito superior a 10% do Produto Interno Bruto (PIB). Nesse contexto, Estatística relevante era a que servia para dimensionar o contingente disponível para recrutamento nos exércitos, quantificar os sobreviventes das recorrentes epidemias e pestes que assolavam a Europa, e que permitia contabilizar os superávits e déficits de balança comercial entre países, variável crítica na deflagração de várias guerras nos séculos passados.

Ao longo do século XX, nos países hoje desenvolvidos, novas demandas coletivas e responsabilidades públicas foram sendo incorporadas ao Estado. Para que tais demandas pudessem ser atendidas através de políticas, programas, regulamentação ou isenções, houve necessidade de

¹ Vide, entre outros, Oliveira (2003), Oliveira e Simões (2005), Senra (2005), Bello (2013), Senra (2017) e Jannuzzi (2017).

ampliação significativa de recursos tributários. Assim, no período entre as Guerras Mundiais, a carga fiscal média dos Estados europeus já estava acima de 20% do PIB, vindo a superar os 30% do PIB no pós-guerra. Ao longo dos “Anos Dourados do Capitalismo”, entre 1950 e meados dos anos 1970, ampliou-se para mais de 40%, como revelam dados compilados pelo autor. A cada momento de ampliação do escopo e escala de Políticas Públicas, novas demandas de Estatísticas foram se agregando, das informações sobre mercado de trabalho nos anos 1920 aos indicadores ambientais nos anos 1990 e estatísticas de reconhecimento identitário nos anos 2000.

Gráfico 1 - Estruturação do Estado de Bem-Estar (expresso pela Carga Fiscal) e a ampliação do escopo das Estatísticas Públicas



Fonte: Elaboração própria a partir de Tanzi (2005), Mauro et al (2013), IMF Fiscal Monitor (Tabela A5) e Jannuzzi (2017).

Nota: A Carga Fiscal ilustrada representa o percentual de tributos sobre o Produto Interno Bruto dos 11 países mais desenvolvidos da OCDE.

Vale observar que a carga fiscal – somatória dos tributos arrecadados pelas esferas de governo, assim como as isenções fiscais, como parcela do PIB- é um indicador com várias limitações para retratar a estruturação do Estado de Bem-Estar e o fortalecimento das Políticas Públicas. É ainda mais imperfeito para ilustrar a conformação dos diferentes modelos de Estado Social que lograram se constituir no período. O uso desse indicador nesse texto se justifica por sua simplicidade argumentativa entre necessidades de financiamento do Estado e as demandas por informação estatística, além do fato de ter uma série histórica mais longa e representativa de um amplo conjunto de países, frente a outras medidas possíveis e consensualmente mais válidas, como o percentual de gasto de políticas sociais pelo PIB ou gasto público por habitante, entre outras².

A estruturação do Estado de Bem-Estar no século passado foi resultado de vários processos e determinantes históricos, alguns comuns entre os países hoje desenvolvidos, outras mais particulares, como assevera Esping-Andersen (1991). Ampliação do sufrágio eleitoral e, associadamente, a democratização política foram certamente os mecanismos fundamentais e mais comuns para esse processo nos diferentes países. Afinal, quanto mais a sociedade – e seus diversos segmentos como trabalhadores urbanos, camponeses, mulheres, analfabetos- pôde participar das escolhas políticas e influenciar as agendas de governo, mais o Estado foi instado

² CEPAL, OCDE e Banco Mundial produzem alguns destes indicadores mais específicos para dimensionamento das Políticas Públicas e caracterização do papel maior ou menor dos Estados Nacionais. Mas as séries históricas são mais restritas e sujeitas a maiores problemas de comparabilidade entre países.

a atender as demandas coletivas de proteção social – seguro, assistência e previdência pública, por exemplo – e de promoção social – educação básica e qualificação profissional, entre outros³.

Naturalmente, essa incorporação de demandas coletivas adquiriu contornos particulares em cada país, conforme emergência e combatividade do movimento sindical e maior ou menor sucesso de partidos socialistas, social-democratas e comunistas na conquista de cadeiras nos parlamentos, na formação de maiorias e na liderança de governos na Europa. A devastação das Guerras Mundiais no território europeu e as agruras a que a sociedade sofreu – os chefes de famílias pobres foram às trincheiras e campos de batalha e classes médias que perderam empregos ou pequeno negócio- foram também determinantes do modelo de Estado Social – mais solidário e com políticas universais - que lá se estabeleceu, como bem coloca Sousa (2017). Não menos importante, como lembra Blyth (2017), a presença de vizinhos como a então União Soviética e outros países do leste europeu, com modelos de provisão estatal de uma ampla gama de serviços coletivos, também foi outro componente importante a pressionar a constituição das soluções sociais-democratas de Estado na Europa.

A urbanização e envelhecimento populacional foram outros fatores determinantes na constituição do Estado Social no século XX (KERTERNESKY 2016). A pobreza do campo havia chegado à cidade desde a primeira Revolução Industrial, requerendo ao longo do tempo, soluções coletivas ao que antes era resolvido no âmbito familiar, comunal e assistencial-religioso. Era necessária a estruturação de serviços públicos para fazer frente ao aumento de idosos que haviam perdido a capacidade laboral, aos trabalhadores acidentados ou que haviam perdido o emprego e os meios de subsistência familiar, às crianças órfãs de pais vitimados pelas pestes, em contexto de famílias cada vez menores, de redes mais restritas de sociabilidade e de limitação de oferta de serviços assistenciais pelas ordens religiosas.

A necessidade de reprodução do próprio Capital foi outro fator impulsionador do Estado Social. De um lado, a construção de infraestrutura de energia e transportes requeriam investimentos vultosos e riscos expressivos para agentes privados; de outro, era necessário dispor de trabalhadores com alguma formação educacional básica, em escala maior do que o provido pelas entidades religiosas-assistenciais. Como bem observou a Profa. Maria Conceição Tavares, a intervenção do Estado sempre foi absolutamente relevante, seja para dar lugar a um processo de industrialização e acumulação capitalista, seja em situações de guerra ou reconstrução (TAVARES, 1988).

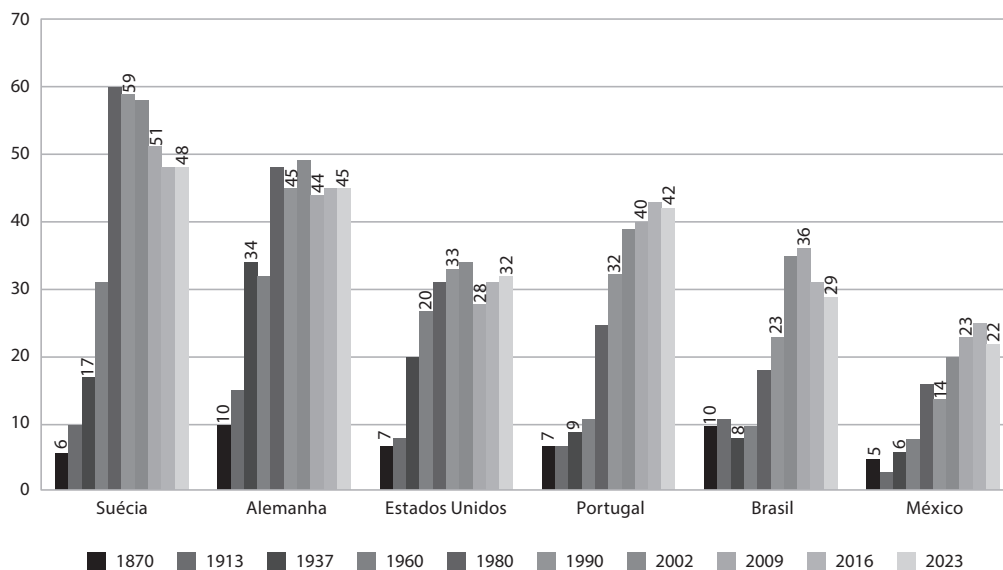
Com tantas trajetórias históricas possíveis, cada país hoje desenvolvido acabou trilhando seu próprio percurso civilizatório e constituindo um tipo específico de Estado de Bem-Estar, mais residual ou mais solidário, mais comprometido com a igualdade de oportunidades ou mais na igualdade de resultados, baseado sob princípios meritocráticos ou sob os de equidade e coesão social. Como compilado por Moreno *et al* (2014), a partir da proposta clássica de Esping-Andersen (1991), os Estados de Bem-Estar instituídos nos países centrais conformaram quatro modelos de Estado de Bem-Estar: Estado Anglo-Saxão Liberal (característico dos EUA), Mediterrâneo Familista (modelo da Espanha, Portugal e Itália), Continental Conservador (como Alemanha e França) e Nórdico Socialdemocrata (como a Suécia e outros países escandinavos).

Cada um desses modelos reflete um determinado nível de carga fiscal que, segundo dados do Fundo Monetário Internacional, situavam-se entre 35% a 55 % do PIB em 2016 (MAURO *et al* 2013), conforme amplitude setorial e cobertura populacional das políticas públicas e, naturalmente, conforme nível de coesão, desigualdade e bem estar social desejado pela sociedade. Na realidade, mais do que a carga fiscal, os modelos de Estado de Bem-Estar se diferenciam pelo conjunto de políticas públicas ofertadas para atender demandas coletivas da sociedade, da

³ Na definição de Castro (2011), as políticas de Proteção Social são aquelas vinculadas à seguridade social- previdência, saúde, assistência social e seguro-desemprego, “destinadas a reduzir e mitigar os riscos e vulnerabilidades a que qualquer indivíduo está exposto numa sociedade de mercado, tal como o de não poder prover o seu sustento e de sua família por meio do trabalho, seja por velhice, morte, doença ou desemprego” (p. 71). As políticas de Promoção Social constituem-se naquelas “que, a partir de determinados instrumentos, pretendem garantir aos cidadãos oportunidades e resultados mais amplos e mais equânimes de acesso aos recursos e benefícios conquistados pela sociedade em seu percurso histórico[compreendendo] ações de formação e desenvolvimento do cidadão – educação, cultura, trabalho e qualificação profissional – [e] democratização do acesso a ativos – como as políticas voltadas à agricultura familiar, à economia solidária, à habitação, à mobilidade urbana”. (CASTRO 2011: 72)

educação, saúde à assistência social. Melhor ainda, tais modelos diferenciam-se pelo grau de “desmercantilização” e “desfamiliarização” provido pelas políticas de proteção e promoção social, isto é, da garantia com que indivíduos, em qualquer momento de seu ciclo de vida, qualquer que seja sua origem social e situação socioeconômica, possam dispor de serviços, políticas e meios, providos pelo Estado, para assegurar uma sobrevivência digna sem depender primariamente do emprego exercido ou do acolhimento familiar.

Gráfico 2 - Estimativas e projeções de carga fiscal em países selecionados - 1870-2023



Fontes: Dados 1870-2002: Vito Tanzi. *The Economic Role of the State in the 21st Century*. Cato Journal, 2005 e Fiscal Prudence and Profligacy database em: <http://www.imf.org/external/datamapper/rev@FPP/USA/FRA/JPN/GBR/SWE/ESP/ITA/ZAF/IND>. Dados 2009-2016: IMF Fiscal Monitor Tabela A13 em: <http://www.imf.org/~media/Files/Publications/fiscal-monitor/2018/April/pdf/fm1801.ashx?la=en>

Ainda assim, com todas as limitações desse indicador, é sintomático o fato de que, com exceção dos EUA, todos os demais países mencionados acima apresentassem carga fiscal acima de 40% do PIB em 2016 (Gráfico 2). Nos EUA, a carga fiscal estacionou nos 35% do PIB nos anos 1970, abortando o projeto de constituição de políticas sociais mais abrangentes iniciado com o Plano de Combate à Pobreza *Great Society* no Governo Lyndon Johnson. A retomada de tal iniciativa nas décadas seguintes- como a tentativa de constituição de um sistema de provimento de saúde pública mais acessível à população por governos democratas- Clinton e Obama- tampouco foi bem sucedida⁴.

Esse mesmo patamar de carga fiscal dos EUA- cerca de um terço do PIB - é o que o Brasil se encontra desde os anos 2000, depois da ampliação de escopo e escala de Políticas Públicas que a Constituição acabou ensejando e que governos sucessivos vieram a efetivar nas décadas seguintes (CASTRO 2011). Tal como ocorrido em Portugal, a redemocratização foi um fator chave para início da estruturação do Estado Social no Brasil, mas com evolução e desdobramento diferente nas últimas décadas em cada país. Segundo dados estimados por Mauro et al (2013), desde a Revolução dos Cravos, em meados dos anos 1970, a carga fiscal do Estado português ampliou-se pelas quatro décadas seguintes, saindo de 25% nos anos 1980, 32% em 1990 e alcançando o padrão mínimo das principais democracias europeias- 40%- na década de 2010⁵. No caso brasileiro, a am-

⁴ Para Tanzi (2005), até mesmo os EUA estariam com carga fiscal excessiva, já que para ele “Government should now scale down its operation... In our 2000 book, Schuknecht and I speculated that no country needs to spend more than 30 percent of its GDP for public-sector activities. This level ought to be enough to finance all the legitimate interventions by the state.(TANZI 2005: 637)

⁵ A Espanha teve movimento semelhante após o enfraquecimento do regime franquista, mas não tão virtuoso como o de Portugal.

pliação da carga fiscal inicia-se mais tardiamente e estabiliza-se mais precocemente, em patamar mais baixo. Nos anos 1980 a carga fiscal estimada era da ordem de 18%, passando, logo alguns anos após a promulgação da Constituição de 1988, a cerca de um quarto do PIB. Em 2002, a carga fiscal já havia atingido o patamar de 33% e, desde então, passou a crescer bem mais lentamente do que no passado, estabilizando-se em torno de 35%. Não fosse o crescimento elevado do PIB no país – mais de 50% em valores reais entre 2000 e 2014 – não teria sido possível a expressiva ampliação de escala e escopo de Políticas Sociais nesse período (JANNUZZI, SOUZA 2016).

Estimativas do FMI para 2016 apontam que a carga fiscal no Brasil teria caído para 31 % do PIB, dado ainda mais preocupante se considerar-se que o produto brasileiro caiu 3,6% nesse ano. Mas ainda mais impactante é a projeção do organismo para 2023, incorporando os efeitos da Ementa Constitucional 95 – Teto dos Gastos Federais- quando a carga fiscal cairia para 29% do PIB. Considerando os efeitos restritivos do Teto para os próximos 20 anos, tal como estimados por Rossi e Dweck (2016), é possível que o Brasil chegue em 2030 com uma carga fiscal próxima ao que o México tem apresentado nos últimos 20 anos (de 20% do PIB) ou próximo daquela que o Brasil apresentava nos anos 1980. Ou seja, o Brasil não estaria evoluindo para a constituição de um padrão de oferta de políticas e serviços públicos com alguma semelhança ao de Portugal, Mediterrâneo-Familista, como sequer ao modelo americano de Estado Social Liberal. Frente às iniquidades já existentes no país e à continuidade do crescimento e do envelhecimento populacional no país, se estados e municípios não ampliarem significativamente a receita de tributos, o país estará condenado a uma regressão ao passado, uma mexicanização da Política Social brasileira, com todas as implicações que isso significa em termos de retomada da fome, indigência e trabalho infantil, de resiliência da pobreza e agudização da desigualdade e da violência, para citar os desdobramentos mais evidentes.

No contexto de uma gestão economicista de recursos públicos, em que as decisões políticas são subordinadas à restrição fiscal- ao contrário da experiência histórica dos países que lograram constituir Estados de Bem-Estar robustos, em que a determinação política pautou a gestão econômico-fiscal – as disputas de orçamento entre atividades meio – como a produção de Estatísticas Públicas – e atividades-fim serão cada vez mais frequentes. Trata-se de um ambiente muito diferente do que se passou ao longo dos últimos 100 anos, da sinergia Estado e Estatística.

Não deveria ser assim, pois a busca de maior eficiência requer informações com maior precisão sobre tamanho e características de públicos-alvo, localização dos mesmos etc. Como se procurou argumentar nesse texto, a efetividade – e eficiência- das Políticas Públicas depende da qualidade das Estatísticas disponíveis.

Estatísticas Públicas para atender às Políticas Públicas

Como se procurou adiantar na **Introdução**, as Estatísticas Públicas não são autônomas, mas refletem o estágio de estruturação do Estado de Bem-Estar e das políticas, programas e outras iniciativas implementadas para sua efetivação (ATKINSON, 2005). De um lado, as Políticas Públicas pautam as demandas de informação para sua formulação e acompanhamento; de outro, o Sistema Estatístico contribui para aprimorar a ação governamental, pelo desvelamento de realidades socioeconômicas nos diagnósticos em diferentes escalas e pela avaliação dos efeitos, defeitos ou não efeitos das políticas e programas.

Estatísticas podem ser de grande utilidade para os diversos agentes e organizações envolvidas na definição das prioridades sociais e na alocação de recursos do orçamento público (Figura 1). Se bem empregados, as estatísticas e indicadores sociais podem enriquecer a interpretação empírica da realidade social e orientar de forma mais competente a análise, a formulação e a implementação de políticas. Na defesa e na negociação das prioridades públicas, os indicadores podem contribuir no apontamento da magnitude das carências a atender nas diversas áreas de intervenção. Ajudam na elaboração de diagnósticos sobre públicos-alvo, priorização da ação em determinadas regiões, auxiliam o monitoramento dos efeitos e a efetividade – ou não – das

políticas na mitigação da problemática originária ou no atendimento da demanda originária da política. Estatísticas Públicas podem ajudar na formação de consensos sobre quais demandas devem ser priorizadas, auxiliar no diálogo com a sociedade e prover meios técnicos para justificar as decisões políticas.

As Estatísticas também respondem à complexidade da base econômica, às demandas dos agentes econômicos e ao conjunto de preocupações sociais a cada momento histórico, verbalizado pela academia, movimento social ou movimento cultural. Índice de preços, Variação do Produto Interno Bruto (PIB) e outras estatísticas conjunturais são exemplos de informações cruciais para gestão cotidiana dos agentes econômicos, e sua produção por uma instituição técnica de Estado garante regularidade, consistência metodológica e acesso franco e universal a todos interessados, ao contrário de muitos indicadores econômicos mais específicos, produzidos por consultorias especializadas. Estatísticas de gênero, de raça/cor e sobre iniquidades sociais de grupos vulneráveis são, por outro lado, instrumentais para movimentos sociais e estudos acadêmicos nas Ciências Sociais Aplicadas.

Mas diferentemente de uma consultoria especializada- que pode cobrar por seus produtos- ou um grupo de pesquisa- que pode ter financiamento de pesquisa episódica de campo- instituições encarregadas de produção das Estatísticas Públicas precisam legitimar o uso dos seus recursos a partir de marcos referenciais estáveis que as Políticas Públicas proporcionam. Em países democráticos as demandas da sociedade, de seus representantes e atores acabam influenciando as agendas de Políticas Públicas, o que, por sua vez, gera demanda por informação mais específica. Assim, sem negar a importância de outros vetores influenciadores do escopo das Estatísticas Públicas são as demandas de Políticas de Estado que tendem a ser determinantes estruturais para a regularidade e especificidade com que determinados temas são investigados pelos órgãos técnicos encarregados pelos levantamentos, frente a tantas possíveis demandas legítimas.

Figura 1 - Papel das estatísticas sociais no ciclo de políticas e programas



Fonte: Elaboração própria a partir de Jannuzzi (2017).

Foi isso que se mostrou na seção anterior. Qualquer que seja a natureza do Estado Social, mesmos para países com pautas mais restritivas de Políticas Públicas- como em países assentados em sistemas de bem-estar típicos do Estado Liberal Anglo-Saxão - o escopo de estatísticas públicas é muito mais amplo que no passado, refletindo o processo histórico de produção de informação para tratar de agendas públicas cada vez mais adensadas. Países com Sistemas de Proteção Social tão diferentes como os EUA, Alemanha, Portugal e Suécia precisaram, pelo menos até os anos 1970, recolher parcelas crescentes do PIB para financiamento de funções estatais mais alargadas. Em todos eles, censos, pesquisas amostrais e registros administrativos continuaram a ser implantados para responder às demandas mais complexas de gestão pública

e da estrutura econômica, seja para promoção ativa de políticas públicas, seja para regulação do mercado provedor de serviços.

De fato, da década de 1920 ao final dos anos 1990, sucederam-se diferentes ênfases da produção das estatísticas e indicadores econômicos, sociais e ambientais, decorrência direta dos determinantes históricos já descritos. De um lado, pela ação de sindicatos e de movimentos sociais (feminista, negro, minorias etc) e, sobretudo, pela institucionalização dos mecanismos de representação democrática – voto universal, representação política, partidos políticos- o Estado passou a ser pressionado a atender demandas coletivas cada vez mais abrangentes. Aos pleitos pela proteção aos riscos de desemprego, acidentes de trabalho e da perda da capacidade laboral da população idosa, foram se agregando às demandas por acesso à educação básica, aos serviços de saúde, aos serviços socioassistenciais e, inclusive, em vários países, à moradia subsidiada. De outro lado, recursos públicos foram demandados para atender as necessidades de financiamento da infraestrutura física para viabilização do crescimento industrial e, mais recentemente, para lidar com as externalidades ambientais decorrentes desse crescimento e da exploração dos recursos naturais.

Assim como a crise de 1929 gerou um ambiente propício para que, nas décadas seguintes, se concebessem os sistemas de contas nacionais e pesquisas para acompanhar a conjuntura econômica e o desemprego, a persistência da pobreza e desigualdade, mesmo em situações de ampliação da produção econômica nas décadas seguintes à Segunda Guerra, acabou evidenciando a necessidade de dispor, a partir dos anos 1960, de pesquisas para acompanhar a mudança social (NEUFVILE 1975). É nesse contexto que se empreendeu um imenso esforço conceitual e metodológico para o desenvolvimento de instrumentos de mensuração do bem-estar e da mudança social nas agências estatísticas dos diversos países, registrado na literatura como o ‘Movimento de Indicadores Sociais’ CARLEY, 1985; SIMÕES, et al 2017).

Em meados dos anos 1980, como discutido em Jannuzzi (2017), um novo ‘Movimento dos Indicadores Sociais’ começou a se espalhar pelo mundo, mobilizados pelas discussões suscitadas pela introdução do Relatório do Desenvolvimento Humano e, em especial, nos anos 1990, pelas várias Cúpulas Sociais Mundiais promovidas pelas agências temáticas das Nações Unidas, entre elas a Conferências Mundiais sobre Meio Ambiente (1992) e Desenvolvimento Social (1995). Criou-se uma ampla mobilização para promoção de políticas sociais nos países não desenvolvidos e de renda média, com menor ou maior escopo setorial e grau de cobertura populacional, que viriam a ser acompanhadas por pressão pela produção de novas estatísticas e relatórios, procurando destacar as iniquidades sociais de grupos sociodemográficos específicos: as Estatísticas Identitárias de negros, mulheres, jovens, indígenas etc. É nesse contexto que se estabelece, em 2000, a Declaração dos Objetivos de Desenvolvimento do Milênio (ODM), uma agenda certamente minimalista em relação à agenda social discutida na década anterior, mas ainda assim impactante na produção de pesquisas e indicadores sociais.

Em um contexto de redemocratização, sindicatos e movimentos sociais passaram a fazer uso crescente de estudos e indicadores para fundamentar demandas por políticas. Estatísticas que permitam reconhecimento identitário – dos negros, indígenas, quilombolas, da população em situação de rua- passaram a ser cada vez mais demandadas e produzidas nos últimos anos. A mídia e outras organizações sociais vislumbraram o potencial jornalístico ou de pressão por transparência do gasto governamental (SAVIGNANO; JANNUZZI 2003). A profissionalização da Administração Pública, nas três esferas de governo, foi mais um vetor impactante para legitimação das Estatísticas no ciclo de formulação e avaliação de políticas de programas sociais.

Desde então, o aumento da desigualdade social na Europa e EUA, os dilemas entre crescimento econômico e uso dos recursos naturais, assim como a crise internacional de 2008, acabaram ensejando nova onda do ‘Movimento’, o terceiro, com maior repercussão mundial e maior amplitude temática. Foi nesse contexto que apareceu o Relatório da Comissão Stiglitz-Sen-Fitoussi, que propunha revisão das medidas de desenvolvimento e bem-estar (STIGLITZ *et al* 2014). Posteriormente, em 2012, fortalecem-se as discussões que culminaram com a proposição da agenda dos Objetivos de Desenvolvimento Sustentável (ODS), com desafios consideráveis na produção de indicadores

de monitoramento (KOEHLER 2016). Novas pesquisas e sistemáticas de registros de informação seriam necessárias para produzir indicadores que, de fato, pudessem acompanhar as mudanças sistêmicas da “boa” política pública preconizada nessa agenda, qual seja, a que simultaneamente geraria prosperidade econômica, inclusão social e sustentabilidade ambiental.

Um rápido exame das duas centenas de indicadores estabelecidos – de um conjunto cinco vezes maior, quando do início das discussões sobre a agenda ODS entre 2010-2014- permite constatar que ainda há muitas lacunas a serem sanadas nos sistemas estatísticos para responder plenamente às demandas nacional e internacional de acompanhamento. O esforço de monitoramento da agenda ODM deixou um legado importante, de cerca de um painel mínimo de cerca de sessenta indicadores internacionais comparáveis, mas com periodicidade muito diferentes de atualização entre os países. O fato é que não há como produzir indicadores sociais, econômicos e ambientais que respondam à Agenda 2030 sem investimentos em recursos humanos, tecnológicos e de pesquisas⁶.

Nesse contexto, ao contrário do apontado ao final da última seção, Estatísticas Públicas, mais regulares, mais desagregadas e com escopo ampliado seriam, pois, cada vez mais necessárias em todas as fases dos ciclos das políticas públicas, principalmente as Sociais. É praticamente impossível atualmente na administração pública a construção de políticas, programas ou até mesmo projetos sociais - sérios - que não levem em conta indicadores sociais e, por consequência, Estatísticas Públicas capazes de os originar. Cada vez mais são necessárias informações periódicas, desagregadas, confiáveis e sensíveis sobre as políticas públicas, seja para a definição de agenda e planejamento como - os diversos - indicadores constantes dos PPAs, seja para acompanhar a execução e realização de monitoramento, seja para avaliar sua efetividade ou seja para *accountability*.

À guisa de conclusão: os rumos incertos da Produção Estatística

Até há pouco tempo, o paradigma institucional dos Sistemas Estatísticos se constituiu aos moldes das necessidades do Estado Social emergente. Ao longo do século XX, pelas características da oferta e provisão de acesso às políticas e serviços públicos- e também pelo porte populacional de cada país-, cada modelo de Estado acabou moldando seu sistema de Estatísticas de forma particular, ancorado mais em Censos Demográficos e Pesquisas Amostrais Nacionais, como no caso dos EUA e países ibéricos – Estado Liberal e Famílistas –, ou em Cadastros Públicos e Registros Administrativos, como no caso da Alemanha e Suécia – Estado Corporativo e Socialdemocrata. A necessidade de provisão universal de uma ampla gama de serviços foi determinante, por exemplo, nos países nórdicos, para estruturação de múltiplos registros de programas, com informações individualizadas de cada cidadão. Certamente contou para isso o porte populacional relativamente modesto desses países e o nível de estruturação do serviço público presente. Com advento da Informática, barateamento dos equipamentos e sofisticação dos aplicativos desde os anos 1970, tornou-se possível a integração dos registros setoriais- antes armazenadas em papel ou computadores isolados- e a produção regular de Estatísticas deles derivadas.

Em países em que a oferta de políticas é mais restrita ou não é universal, em que o acesso aos serviços públicos é pautado por critérios de seletividade (pelo nível de pobreza, origem étnica etc), os cadastros e registros de programas tem cobertura populacional menor e específica. Nesses casos, Estatísticas de ampla representatividade da população e de seus grupos sociode-

⁶ Há metas para as quais não se dispõe ainda de medidas regulares e amplamente disponíveis pelos países, o que particularmente afeta os indicadores do “Planeta”. Para as dimensões de “Paz” e “Parceria” ainda há muito por se trabalhar em termos conceituais, para depois, se chegar a medidas mais “válidas” e factíveis de serem construídas. Em outras situações, embora existam indicadores, como é o caso de boa parte das dimensões de “Pessoas”, estes não podem ser desagregados para todos os segmentos desejados por raça/cor, gênero ou região. Nos vários indicadores de “Prosperidade” volta-se ao tema da necessidade de definições consensuais para além dos princípios e práticas usados no Sistema de Contas Nacionais.

mográficos requerem a realização de Censos Demográficos e levantamento amostrais temáticos com alguma regularidade. Esse tem sido o modelo brasileiro, muito bem sucedido em termos comparativos internacionais (DARGENT et al. 2018). Para além da produção do IBGE, a criação de registros e cadastros públicos nos Ministérios -e o investimento na melhoria da qualidade dos dados declarados e disponibilizados- deixou um legado informacional bastante relevante para múltiplos propósitos no país.

É preciso ressaltar que se não fosse esse portfólio variado de levantamentos estatísticos disponíveis no Brasil, a natureza pública e regularidade da produção das mesmas pouco se conheceria sobre a mudança social ao longo do século XX no Brasil e, em especial, após a Constituição de 1988. Não houvesse o IBGE e outras instituições no país, a produção de conhecimento empírico acerca da realidade brasileira seria muito menor e muito menos aprofundada, assim como também seria incomensuravelmente mais baixa a efetividade da ação pública no equacionamento e mitigação dos grandes problemas nacionais.

Contudo, o contexto ambiental parece ter mudado muito rapidamente nas duas primeiras décadas desse século. Novos provedores de informação, com custos mais baixos de produção e, supostamente, com maior tempestividade na divulgação, passam a disputar os recursos públicos, oferecendo dados e pesquisas embalados em pacotes de alta tecnologia- Big Data- mas com sérias limitações de representatividade populacional e territorial, regularidade temporal e, sobretudo, qualidade informacional. Se estes eram atributos fundamentais às Estatísticas Públicas no passado, aparentemente deixaram de ter importância decisiva no presente. Para isso contribui o fato de que a informação econômica produzida a partir da compilação de dados de associações patronais, de consultorias econômicas e de rumores veiculados na mídia especializada passa a ter proeminência nas decisões públicas, pela fatídica subordinação que a Política passou a ter da Economia em tempos de Globalização Financeira (SANTOS 2002; BRUNO,CAFFE 2017).

Em círculos ligados à Gestão Pública, há certo encantamento com volume, baixo custo e tempestividade das informações que trafegam nas mídias digitais, nos registros de vendas e de transações individuais e com os resultados e factoides produzidos por agências de pesquisas de opinião e analistas de mercado, supondo que seriam suficientes para o acompanhamento da realidade social e econômica. Aparentemente, taxa de desocupação e massa salarial, apurados mensalmente, com rigor estatístico, importam menos nas decisões públicas que informações, às vezes diárias, de natureza especulativa ou microeconômica, relacionada às decisões de investimentos de transnacionais e grupos financeiros.

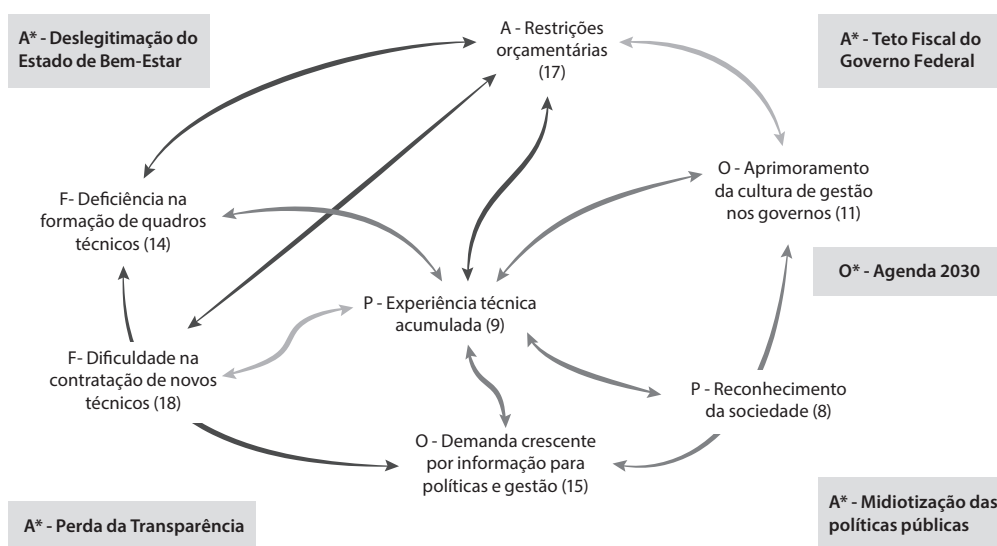
Há certamente que se aprofundar a análise da utilidade do “Big Data” em Sistemas Estatísticos, sem perder a referência dos atributos necessários ao dado oficial usado na formulação e avaliação de Políticas Públicas. “Evidências Estatísticas” podem ser complementadas por “Aparências Apontadas” nas mídias sociais, mas não substituídas. Mas, fato é que, mais recentemente, em vários âmbitos da Gestão Pública, as aparências parecem se sobrepor às evidências na orientação das Políticas Públicas. Em um contexto de regressão a um Estado Social Mínimo, pelas restrições apontadas por Rossi e Dweck (2016) aos gastos públicos nos próximos vinte anos- Teto Fiscal-, os fragmentos de informação providos por um Big Data anárquico talvez bastem. Nesse contexto, um Sistema Estatístico mais abrangente certamente geraria muito incômodo, ao desvelar as novas e velhas iniquidades na sociedade brasileira.

O aparente paradoxo entre a perda da institucionalidade das Estatísticas Públicas pela deslegitimação do Estado de Bem-Estar e a pressão por dinamização da produção frente à complexidade da realidade social, econômica e ambiental já havia sido apontado em um exercício de elaboração do Planejamento Estratégico das instituições subnacionais de Estatística em 2010 (CARVALHO; FIGUERÓA, 2010; JANNUZZI et al 2010). Na análise das respostas dos 21 participantes do Planejamento Estratégico produziu-se um diagnóstico interessante que veio a ser chamado de Paradoxo Existencial das Instituições de Estatística (Figura 2). A ambiguidade dos achados poderia ser resumida da seguinte forma: em que pesem, como Oportunidades, a valorização crescente da informação estatística pela sociedade, o aprimoramento da cultura de gestão por resultados na administração pública e o reconhecimento simbólico das instituições existentes

pelos sociedade civil (como Ponto Forte), elas continuam convivendo com um problema fundamental – a dotação orçamentária (Ameaça central) – que, em círculo vicioso, traz problemas para contratação e aprimoramento dos quadros técnicos, potencializando suas Fragilidades.

Se novo levantamento fosse realizado, a esse quadro deveriam ser acrescentada pelo menos quatro novas ameaças contextuais às instituições estatísticas (sinalizadas com A* na figura), que mitiguem as oportunidades então identificadas e acentuam as suas dificuldades de financiamento/instrumentalização. As Ameaças macro-contextuais como a Deslegitimação do Estado de Bem-Estar Social, as Consequências do Teto Fiscal, os riscos de Perda de Transparência e a “Midiotização” da Informação Pública têm efeitos muito mais deletérios para produção e disseminação das Estatísticas do que os impactos positivos e oportunos que a Agenda 2030 (Desenvolvimento Sustentável) traz para esse ambiente setorial.

Figura 2 - Paradoxo existencial das instituições estatísticas



Fonte: Reelaborado a partir de Jannuzzi *et al.* (2010).

Notas: 1. A = Ameaças; A* = Ameaças mais recentes; O = Oportunidades; O* = Oportunidade mais recente; F = Fragilidades; e P = Pontos fortes.

2. Entre parênteses está o número de entrevistados que citaram o fator, dentre os 21 participantes do Planejamento Estratégico.

Afinal, se as Políticas Públicas passam a ser questionadas em sua relevância, essência e desenho pela crítica liberal ao Estado Social, o que dizer dos instrumentos para sua formulação? Não há dúvidas de que o enfraquecimento do Estado afeta o imperativo da produção de informação abrangente e regular para as Políticas Públicas. Os cortes de recursos – e questionário- no Censo Agropecuário de 2016 e a discussão sobre a redução de 25% do Censo Demográfico 2020 são as primeiras evidências concretas desse fato no IBGE. De fato, não há dúvidas de que o Teto Fiscal em âmbito federal só vai acentuar a necessidade de mais recursos estaduais em várias políticas setoriais, pressionando os já combalidos orçamentos das instituições estaduais envolvidas com o Planejamento e as Estatísticas.

Não menos importante, existe a ameaça aos Sistemas Estatísticos vinda da perda da transparência dada pela restrição do acesso aos dados e informações produzidos. Desde a promulgação da Constituição de 1988, o Brasil tem trilhado um longo caminho na divulgação das informações públicas produzidas pelas diferentes esferas de governo e suas agências. Iniciando pela adoção da Lei nº 8.159/91, conhecida como “Lei dos Arquivos” (Polo, 2015) e passando pela

promulgação da Lei de Acesso à Informação⁷ (LAI), em novembro de 2011, onde se define em seu artigo 3º “...a publicidade como preceito geral e o sigilo como exceção e a divulgação da informações de interesse público, independente de solicitação...” (PEREIRA et al 2014) o país tem caminhado para um alinhamento, mesmo que de forma imperfeita ou incompleta, com os preceitos do movimento de Governo Aberto, onde a tomada de decisões e ações governamentais devem ser o mais transparentes e participativas para a sociedade civil como um todo (GONZALEZ-ZAPATA & HEEKS, 2014; MEIJER et al 2012).

Por fim o capital simbólico das instituições públicas produtoras de estatísticas e estudos têm sido minado pela profusão, nas mídias digital, televisiva e impressa, de factoides e notícias pouco embasadas e elaboradas acerca do desenho e resultados das Políticas Públicas. A regularidade, atratividade estética e aparente simplicidade dessa informação “midiotizada” e dos “fake-news” acerca da realidade social e das Políticas Públicas tende a ter grande apelo e receptividade de segmentos importantes da sociedade, menos afeitos ao padrão técnico dos relatórios e releases produzidos pelas instituições estatísticas. A midiotização parecer ser, entre as ameaças enumeradas, a que as instituições poderiam ter maior margem de reação ou contra-atauação. Mas isso implicaria em mudanças não só nas estratégias de disseminação e comunicação, como também de outras estratégias metodológicas de produção de informação e conhecimento.

Independentemente de onde venha o ataque, seja pelas restrições orçamentárias já apontadas (que limitam os recursos para a manutenção das estruturas tecnológicas necessárias ao processo de divulgação), seja pela mudança do paradigma legal (que imponha o sigilo como regra geral) a redução ou restrição da divulgação das informações produzidas pode representar uma “sombra” que ameaça, ainda mais, o processo de conhecimento das iniquidades (velhas ou novas) de nossa sociedade.

Há, pois, que se colocar em pauta de forma sistemática a pesquisa sobre o Sistema Estatístico, seu passado, seu presente, as pesquisas, os temas investigados, as lacunas da realidade não cobertas nos instrumentos, as publicações e estratégias de disseminação da informação, os avanços tecnológicos, os recursos humanos e financeiros envolvidos e o impacto do mesmo para a Sociedade. Esse texto é uma pequena contribuição para incentivar o engajamento de um número maior de pesquisadores nessa linha de pesquisa, crucial para o IBGE e indispensável para a construção de um Brasil mais justo e solidário.

⁷ Na prática, a LAI finalmente regula o acesso à informação que foi previsto no inciso XXXIII do artigo 5º, no inciso II do parágrafo 3º do artigo 37 e no parágrafo 2º do artigo 216 da Constituição Federal de 1988 (Cintrão & Bizelli, 2013).

Referências

- ATKINSON, T. *et al.* *Social indicators: the EU and Social Inclusion*. Oxford: Oxford Univ.Press, 2005.
- BELLO, L. A. *A mídia e os órgãos oficiais de estatísticas*. 2013. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2013.
- BLYTH, M. *Austeridade: a história de uma ideia perigosa*. São Paulo: Autonomia Literária, 2017.
- BRUNO, M.; CAFFE, R. Estado e financeirização no Brasil: interdependências macroeconômicas e limites estruturais ao desenvolvimento. *Economia e Sociedade*, Campinas, 26, N. especial, p.1025-1062, 2017.
- CARLEY, M. *Indicadores sociais: teoria e prática*. Rio de Janeiro: Zahar, 1985.
- CARVALHO, C. V.; FIGUERÔA, E. As instituições públicas estaduais brasileiras voltadas para a produção e disseminação de estatísticas públicas, estudos e pesquisas e planejamento e os desafios da contemporaneidade. *Boletim Estatísticas Públicas*. Salvador, p. 68-77, 2010.
- 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, p.1:66-95, 2011.
- CINTRÃO, L. P., & BIZELLI, J. L. Sistema de Monitoramento e Avaliação de Programas Sociais: revisitando mitos e recolocando premissas para sua maior efetividade na gestão. *Revista Brasileira de Monitoramento e Avaliação*, 5, p. 48–59, 2013.
- DARGENT, E. *et al.* *A quem importa saber?: a economia política da capacidade estatística na América Latina*. Washington: BID, 2018.
- ESPING-ANDERSEN, G. As três economias do Welfare State. *Lua Nova*. n. 24. 1991, p. 85-115.
- GONZALEZ-ZAPATA, F., & HEEKS, R. The multiple meanings of open government data: Understanding different stakeholders and their perspectives. *Government Information Quarterly*, 32(4), 441–452, 2014.
- IBGE. *Indicadores Sociais: 40 anos de história*. Rio de Janeiro, 2017.
- JANNUZZI, P. M. *Indicadores Sociais no Brasil: conceitos, fontes de dados e aplicações*. Campinas: Alínea, 6. ed. revista e ampliada, 2017.
- JANNUZZI, P. M.; SOUSA, M. F. Pobreza, desigualdade e mudança social no Brasil de 1992 a 2014: tendências empíricas para análise dos efeitos do Plano BSM e da Estratégia Brasileira de Desenvolvimento Inclusivo. *Cadernos de Estudos Secretaria de Avaliação e Gestão da Informação*, v. 25, p. 22-55, 2016.
- JANNUZZI, P. M.; VANETI, V. C.; LOUREIRO, S. Planejamento Estratégico da Anipes: registro do processo. *Boletim Estatísticas Públicas*. Salvador, p. 119-141, 2010.
- KERSTENETZKY, C. *O Estado do Bem-Estar Social na Idade da Razão*. Rio de Janeiro: Campus, 2012.
- KOEHLER, G. Looking back and looking forward: the case for a development welfare state. In: CIMADAMORE, A.; KOEHLER, G.; POGGE, T. (org.). *Poverty and the Millennium Development Goals*. 1. ed. Londres: Zed Books, v. 1, p. 229-257, 2016.
- MAURO, P. *et al.* A Modern History of Fiscal Prudence and Profligacy. *IMF Working Paper* 13/5, 2013.
- MEIJER, A. J.; CURTIN, D.; HILLEBRANDT, M. (2012). Open government: connecting vision and voice. *International Review of Administrative Sciences*, 78(1), 10–29. <https://doi.org/10.1177/0020852311429533>

MORENO, L. *et al.* *Los sistemas de bienestar europeos tras la crisis económica*. Madrid: EuroSocial/ IILA, 2014.

NEUFVILLE, J. *Social Indicators and Public Policy*. New York: Elsevier, 1975.

OLIVEIRA, L. A. P.; SIMÕES, C. C. S. O IBGE e as pesquisas populacionais. *Revista Brasileira de Estudos da População*, São Paulo, v. 22, n. 2, p. 291-302, jul./dez. 2005.

OLIVEIRA, J. S. “Brasil mostra a tua cara”: imagens da população brasileira nos censos demográficos de 1872 a 2000. *Textos para discussão - Escola Nacional de Ciências Estatísticas*. Rio de Janeiro, 2003.

PEREIRA, A. K.; PIRES, P. S.; PINTO, A. (2014). Pesquisas de Avaliação e Confidencialidade da Informação: Limites e Conflitos. *Revista Brasileira de Monitoramento e Avaliação*, 7, 82–99.

PNUD. *Relatório de Desenvolvimento Humano*. Nova York, 2014.

POLO, M. G. de O. (2015). Governo, sociedade civil e os desafios na publicação de dados abertos: o caso da base de dados do Programa Nacional de Apoio à Cultura no Brasil.

ROSSI, P.; DWECK, E. Impactos do Novo Regime Fiscal na saúde e educação. *Cad. Saúde Pública*, 32(12):1-5, 2016.

SANTOS, B. S. Os processos da globalização. In: _____ (org). *A globalização e as ciências sociais*. São Paulo: Cortez, 2002, p.25-102.

SAVIGNANO, V. N. M.; JANNUZZI, P. M. Disseminação de informação para a cidadania no Brasil: uma análise da cobertura das matérias sobre indicadores sociais na mídia impressa. *Transinformação*, v. 15, n. 3, 2003.

SENRA, N. *O saber e o poder das estatísticas*. Rio de Janeiro: IBGE, 2005.

SENRA, N. *et al.* *O Desafio de retratar o país: entrevistas com os presidentes do IBGE no período de 1985 a 2015*. Rio de Janeiro, IBGE, 2015.

SENRA, N.; *Tradição & renovação: uma síntese da história do IBGE*. Rio de Janeiro: IBGE, 2. ed., 2017.

SIMÕES, A.; ALKMIM, A. C.; SANTOS, C. Passado, presente e futuro da produção e análise dos Indicadores Sociais no IBGE. In: SIMÕES, A.; ALKMIM, A. C. *Indicadores sociais: passado, presente e futuro*. Rio de Janeiro: IBGE, 2017. p.17-27.

SOUZA, J. *A Elite do atraso: da escravidão à Lava-Jato*. São Paulo: Leya, 2017.

STIGLITZ, J. E.; SEN, A.; FITOUSSI, J. P. *Mis-measuring our lives: why GDP doesn't add up*. New York: New Press, 2014.

TANZI, V. The economic role of the state in the 21st Century. *Cato Journal*, v. 25, n. 3, p. 617-638, 2005.

TAVARES, M. C. O Planejamento em economias mistas. Cendec. *Seminário Internacional Estado e Planejamento: sonhos e realidade*. Brasília: Cendec, 1988, p. 25-45.

UNECE. *Value of official statistics: Recommendations on promoting, measuring and communicating the value of official statistics*. Genebra, 2017.

Big Data na produção de estatísticas públicas: conceitos, oportunidades e desafios*

Rafael Bassegio Caumo**

Resumo

A recente explosão no volume, na velocidade e na variedade de dados digitais produzidos, armazenados e disponibilizados, características do Big Data e determinantes da “revolução dos dados”, estão modificando a maneira como organizações produzem conhecimento. No que concerne especificamente aos produtores de estatísticas públicas e oficiais, os dados digitais se apresentam como uma oportunidade de inovação radical no âmbito dos processos tradicionais de produção, viabilizando possibilidades de contorno de uma série de dificuldades atualmente percebidas por esses produtores. Nesse contexto, o objetivo do artigo é apresentar o conteúdo do estado da arte da produção científica – por meio de uma revisão sistemática – sobre a utilização de métodos inovadores na produção de estatísticas públicas, em termos de identificação das oportunidades – com possíveis fontes de dados e aplicações – e dos desafios e problemáticas que se colocam. Espera-se que os resultados possam orientar a construção de agendas de trabalho dos produtores que vislumbrem alterar o paradigma de seus processos produtivos.

Palavras-chave: Estatísticas públicas e oficiais. Revolução dos dados. Big Data. Oportunidades. Desafios.

* Este capítulo é resultado de uma das revisões sistemáticas realizadas pelo autor durante a construção da sua tese de doutorado, em andamento, intitulada "Indicadores socioeconômicos produzidos a partir de Big Data: Um *framework* para avaliação da qualidade estatística aplicado ao turismo".

** Docente no Centro de Ciências da Administração e Socioeconômicas da Universidade do Estado de Santa Catarina e doutorando em Engenharia do Conhecimento na Universidade Federal de Santa Catarina. Concluiu o mestrado em Estudos Populacionais e Pesquisas Sociais em 2014.

Introdução

O fenômeno de explosão na quantidade de dados digitais sendo produzidos e armazenados, produto da popularização da internet, do avanço tecnológico e da consolidação da era digital (SCHMIDT; COHEN, 2013), tem modificado a maneira como muitas organizações norteiam seus processos de tomadas de decisão (MAYER-SCHONBERGER; CUKIER, 2013).

Atores que conduzem ações orientados por dados usufruíam, no passado, apenas dos ditos “dados analógicos”, produzidos predominantemente por levantamentos amostrais e censos baseados nos métodos científicos tradicionais (HACKL, 2016). Agora, dispõem de inúmeras novas fontes de dados – caracterizadas por bases de Big Data, *crowdsourcing data*, dados abertos, *scaled small data*, entre outras – em um contexto no qual o perfil majoritário dos dados disponíveis se modificou para o dos “dados digitais” – transformação que caracteriza o fenômeno da “revolução dos dados” (KITCHIN, 2015).

O valor econômico desses dados digitais tem sido percebido no ambiente empresarial, sendo recorrentemente tratados como uma nova classe de ativos, ou como “o novo petróleo”, tamanha sua interessante relação de custo benefício para processo de construção de conhecimento (WEF, 2011; BOSSOI, 2014). Como consequência, desempenham papel central na já mencionada quarta revolução industrial (SCHWAB, 2016), em um período de *data economy* (OPHER *et al.*, 2016).

Fora do ambiente dos negócios, as instituições públicas, acadêmicas e do terceiro setor já avaliam a relevância científica e de suporte social desenvolvimentista dessa nova classe de dados, percebendo um potencial de revolução nos modelos de todos os tipos de instituições. Os dados digitais têm aberto, portanto, uma oportunidade de inovação que pode alcançar os mais diversos setores da sociedade, gerando, por vezes, soluções radicais e até disruptivas para processos, produtos, serviços e modelos organizacionais tradicionais (MAYER-SCHONBERGER; CUKIER, 2013).

Os produtores de estatísticas públicas e/ou oficiais também estão atentos para esse fenômeno, avaliando a dimensão da oportunidade e do ônus e bônus da incorporação das novas fontes (ASQUER, 2013). E isso não se dá por acaso, mas sim porque muitos de seus métodos tradicionais de produção de estatísticas são fortemente concentrados em pesquisas amostrais ou censos via levantamentos em campo com coleta conduzida por intermediários humanos (BRAAKSMA *et al.*, 2014) que, apesar de apresentarem qualidade e validade científica reconhecida, estão sujeitos uma série de contratempos e dificuldades práticas – de ordem financeira, temporal e operacional (KITCHIN, 2015; HACKL, 2016; DAAS, 2014; CITRO, 2014).

Nesse contexto, considerando a complexidade dos modelos tradicionais de produção de estatísticas públicas e a oportunidade trazida pela “revolução dos dados”, surge o questionamento sobre a exequibilidade e a prosperidade de uma quebra de paradigma no status quo da produção de estatísticas. Será que a incorporação das novas fontes de dados nos modelos de produção dos agentes produtores pode representar uma inovação que traria economia de tempo e dinheiro, viabilizando produtos mais ágeis e responsivos ao contexto social evolutivo vigente?

Essa possível quebra de paradigma é um tema de crucial importância para produtores de estatísticas públicas, uma vez que pode estender a capacidade de produção de informação estatística para outras esferas, reduzindo a participação dos atuais principais agentes: os institutos oficiais de estatística (DEMUNTER, 2017).

Assim, objetivando explorar a relevância que o tema tem recebido pelo mundo, este artigo se propõe a apresentar o estado da arte da produção científica sobre a utilização de Big Data na produção de estatísticas públicas no que diz respeito à identificação de oportunidades – com possíveis fontes de dados e aplicações práticas já propostas – e de desafios – com problemáticas e riscos – já percebidos ou previstos. Para tanto, realizou-se uma busca sistemática de documentos pelas bases Scopus, Web of Science e Scielo. Espera-se que os resultados possam orientar a construção de agendas de trabalho dos produtores que vislumbrem considerar a utilização de Big Data.

Após essa introdução, o capítulo está organizado em quatro seções com os seguintes conteúdos: os principais referenciais teóricos que perpassam o estudo – estatísticas públicas e oficiais, a problemática da produção tradicional de estatísticas públicas, e a revolução dos dados;

a metodologia utilizada; os resultados da revisão sistemática, com identificação das oportunidades e desafios da utilização de Big Data na produção de estatísticas públicas; e, por último, algumas considerações finais e perspectivas de pesquisas futuras.

Referencial teórico

Estatísticas públicas e oficiais

Estatísticas são formas institucionalizadas de promover conhecimento para propiciar intervenções em uma realidade (FEIJÓ; VALENTE, 2005), correspondendo à complexas objetivações que revelam realidades previamente construídas e idealizadas, apoiadas por métodos de pesquisas avançados aceitos em várias ciências (SENRA, 2009).

Estatísticas oficiais – derivação da tradução do termo *official statistics* –, por sua vez, correspondem àquelas produzidas por agências governamentais (SCHWARTZMAN, 1997) ou organizações internacionais tidas como oficiais. No Brasil, o termo mais utilizado – muitas vezes empregado como sinônimo – é o de “estatísticas públicas”; mesmo que esse possa representar um universo mais amplo de estatísticas – incluindo aquelas não produzidas por órgãos governamentais e/ou não tidas como oficiais.

As estatísticas públicas, por vezes vistas como indicadores estatísticos em virtude da capacidade de substituir, quantificar ou operacionalizar um conceito social abstrato (JANUZZI, 2004) – ilustrando e comunicando um conjunto de fenômenos complexos de uma forma simples (OCDE, 2002; EEA, 2005; BOSSEL, 1999) –, são bens públicos essenciais para o desenvolvimento econômico, demográfico, social e ambiental de uma nação (UNECE, 1992).

Tratam-se de insumo fundamental para planejamento e formulação de políticas e estratégias no mundo contemporâneo (JANUZZI; GRACIOSO, 2002). No âmbito do poder público, por exemplo, podem subsidiar políticas e suas correspondentes ações, desde a formulação até a etapa de avaliação. Este processo, quando transparente, contribui para a consolidação da democracia e do exercício da cidadania à medida que permite que a sociedade civil em geral – incluindo a academia – atue com controladoria e desenvolvimento paralelo de conhecimento. No aspecto privado, não são menos importantes. O planejamento e as ações conduzidas por empresários são mais precisos e envolvem menores riscos quando realizados em posse de análises conjunturais bem fundamentadas. Ainda, estatísticas públicas podem agregar conhecimento relevante quando da avaliação das possibilidades de crescimento de mercados e identificação de oportunidades locais.

Por subsidiarem tantos processos de tomadas de decisão, se faz necessário que estatísticas públicas sejam confiáveis e de qualidade. Assim, é fundamental que apresentem caráter imparcial e que respeitem padrões que permitam comparabilidade ao longo do tempo e entre diferentes localidades. Portanto, espera-se que as instituições produtoras de estatísticas públicas atuem com rigor científico internacionalmente consolidado e autonomia dentro da estrutura governamental local.

Como fonte de conhecimento adicional, os artigos de Schwartzman (1997) e Feijó e Valente (2005) fazem uma importante reflexão sobre aspectos associados a legitimidade, a credibilidade e a confiabilidade das estatísticas públicas, contextualizando com o papel das instituições produtoras e, no caso do primeiro, apresentando um estudo de caso sobre o IBGE. Senra (1999), em tema correlato, discorre sobre a necessidade da existência de um órgão coordenador da produção nacional de estatísticas públicas.

Pesando nestes aspectos, a Comissão de Estatística das Nações Unidas (UNSC) adotou – em 1994 – balizadores para o processo de produção e disseminação de estatísticas públicas, mantidos até os dias atuais, conhecidos como Princípios Fundamentais das Estatísticas Oficiais (UNECE, 1992).

À luz dos Princípios Fundamentais das Estatísticas Oficiais, o Gabinete de Estatísticas da União Europeia (Eurostat) lançou, em 2005, posteriormente revisado em 2017 (EUROSTAT, 2017),

o *European Statistics Code of Practice* (CoP), que contempla um framework de orientação para as práticas estatísticas, o *Quality Assurance Framework* (QAF), documentado por Eurostat (2015).

Além do QAF, diversos agentes envolvidos e preocupados com produção de estatísticas públicas e públicas também já desenvolveram frameworks de qualidade. Dentre outros, pode-se citar: o *National Quality Assurance Framework* (NQAF), da Organização das Nações Unidas (UN, 2012), o *Data Quality Assessment Framework* (DQAF), do Fundo Monetário Internacional (FMI, 2003); o *Recommendation of the OECD Council on good statistical practice*, da Organização para a Cooperação e Desenvolvimento Econômico (OCDE, 2015); e o Código de Boas Práticas das Estatísticas do Instituto Brasileiro de Geografia e Estatística (IBGE, 2013).

A problemática da produção tradicional de estatísticas públicas

Produtores de estatísticas públicas existem para subsidiar processos de tomadas de decisão com informações (HARWOOD; MAYER, 2016). São encarregados de produzir e disponibilizar estatísticas sobre uma série de domínios e em diversos níveis de escala e granularidade (KITCHIN, 2015). Nesse contexto, devem servir a todos os setores da sociedade – públicos ou privados –, baseados nas melhores práticas e em critérios de qualidade, orientados por estatísticos com independência profissional e objetividade (EUROSTAT, 2014), conforme mencionado na subseção anterior.

Para tanto, tradicionalmente estruturam suas fontes de dados – que servirão de base para a confecção de seus produtos – a partir de: censos; pesquisas amostrais; e/ou registros administrativos (HACKL, 2016). O IBGE, maior produtor de estatísticas públicas do Brasil, responsável por grande parte das estatísticas oficiais do país, concentra grande parte da sua produção de indicadores sociais, econômicos, demográficos e ambientais em processos que envolvem censos ou pesquisas amostrais.

Tais tipos levantamentos possuem utilidade e qualidade reconhecida por conta do sustento metodológico estatístico cientificamente consolidado, rigorosamente desenvolvido e padronizado ao longo de séculos (KITCHIN, 2015). São métodos que permitem alto potencial analítico ao viabilizarem que covariáveis sejam investigadas em conjunto com as variáveis principais de interesse ao mesmo tempo que podem ser controlados pelo pesquisador (CITRO, 2014), fazendo com que boa parte das fontes de erro associadas aos produtos sejam conhecidas – já estudadas, conforme bem sintetizam Brackstone (1999) e Biemer *et al.* (2014) –, possibilitando mensuração em termos de qualidade, viabilizando processos inferenciais adequados para populações de interesses por meio de tratamentos corretivos e validações estatísticas.

Entretanto, estudos têm apontado para o crescimento de problemas de cunho operacional e metodológico até então não bem explorados e que podem impactar no conhecimento sobre a qualidade dos resultados gerados (MANSKI, 2014). Em geral, associados aos levantamentos de campo do tipo *survey*, especialmente quando domiciliares e de grande porte, que possuem operacionalização complicada – por vezes complexa – e sujeita a uma série de contratemos.

Em certa medida, a maioria dessas problemáticas operacionais acabam por impactar na taxa de aproveitamento – ou taxa de respostas – dos levantamentos (CITRO, 2014; DE LEEUW; DE HEER, 2002) – que corresponde à razão entre indivíduos efetivamente investigados e a quantidade inicialmente planejada. Aspecto problemático, pois, além de significar desperdício de recursos e fonte para uma série de diferentes erros, a tentativa de corrigi-lo pode ampliar o viés dos resultados (GROVES; PEYTCHEVA, 2008; FRICKER; TOURANGEAU, 2010). Em relação aos fatores que podem impactar na taxa de aproveitamento, pode-se citar:

- i. Aumento das recusas – totais ou em itens específicos – por descrença na pesquisa – no governo ou no entrevistador – ou simplesmente por fadiga – especialmente em investigações longas que acabam por desencadear o que é conhecido por *response burden* (HACKL, 2016; STRUJIS *et al.*, 2014);

- ii. As informações coletadas nem sempre são precisas pois suas metodologias e constructos teóricos por vezes se apresentam desatualizados frente a um contexto de constantes e cada vez mais frequentes transformações sociológicas ou também pelo fato de que os indivíduos não entenderam o questionamento, mentem ou simplesmente acreditam e comunicam ações que, na prática, não correspondem à realidade (CITRO, 2014; HAND, 2015);
- iii. Está cada vez mais difícil encontrar indivíduos nas residências, seja pelo fato de que as famílias estão menores ou porque passam mais tempo fora de casa;
- iv. O trabalho de entrevistador de campo exige esforços de treinamento, reciclagem e motivação constantes por parte do empregador e necessita de grande dedicação e múltiplas habilidades por parte do profissional. Em paralelo, geralmente é mal remunerado e acaba por apresentar alta rotatividade. Tais questões são problemáticas pois um satisfatório desempenho por parte do entrevistador é crucial para o alcance de maior aproveitamento e para a qualidade dos resultados gerados;
- v. A violência urbana pode intimidar e enclausurar moradores, fazendo com que não aceitem receber o entrevistador, da mesma forma que pode influenciar no acesso e no trânsito do entrevistador à determinadas áreas da cidade;
- vi. O crescimento na quantidade de condomínios fechados – verticais e horizontais – com regulamento de acesso rigidamente restrito tem dificultado cada vez mais o acesso dos entrevistadores à domicílios específicos.

No âmbito metodológico, parte dos problemas estão associados à dificuldade de se conseguir sistemas de referência completos da população alvo e à dificuldade de acesso a todas as parcelas populacionais de interesse – culminando, por vezes, em problemas de cobertura e representatividade (CITRO, 2014).

Muitas vezes, também, além das problemáticas metodológico-operacionais, os censos e as pesquisas amostrais são levantamentos de campo que demandam alto custo financeiro (HACKL, 2016) de planejamento e execução, da mesma forma que tomam tempo até a definitiva disponibilização dos resultados – especialmente se as pesquisas forem do tipo domiciliares. Nesse contexto, portanto, os produtos podem apresentar elevados custos e correm o risco de serem lançados com defasagem considerável em relação ao período que se referem.

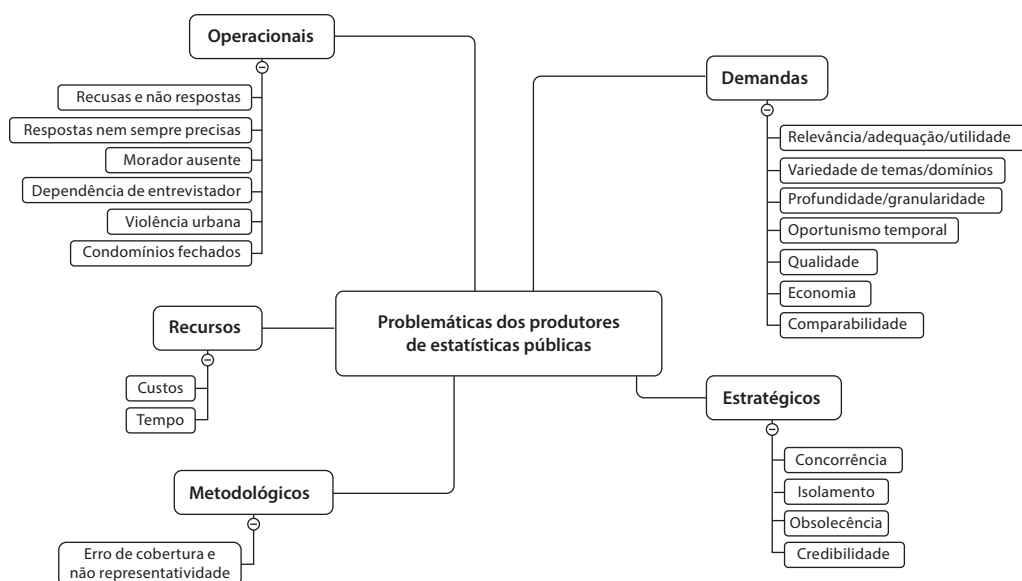
Em adição às questões metodológico-operacionais, de custo e de tempo, outros desafios têm se apresentado aos produtores pelo lado da demanda – e para que se mantenham relevantes, é fundamental que produtores continuem gerando produtos e serviços que atendam aos interesses de seus usuários (HARWOOD; MAYER, 2016).

Usuários, financiadores e provedores de estatísticas públicas cada vez mais desejam: estimativas que tratem sobre muitos aspectos e domínios; tenham profundidade para aberturas e unidades espaciais específicas; sejam oportunas temporalmente; de mais qualidade; mais baratas; de relevância ao retratar exatamente aquilo que se deseja; e menos onerosas (HOLT, 2007; CITRO, 2014) – características para as quais ainda se poderia adicionar a necessidade de comparabilidade entre países.

Tal qualificação da demanda pode ser vista como um reflexo da percepção dos indivíduos acerca do aumento do poder tecnológico disponível e de um conseqüente maior esclarecimento sobre a utilidade e o manuseio dos dados, criando a noção de indivíduos *data savvy* que cada vez mais demandam esclarecimentos mais sofisticados e exigem respostas mais rápidas (HARWOOD; MAYER, 2016).

Por fim, alguns desafios, aqui chamados de estratégicos, estão relacionados ao impacto da pressão das transformações externas e do surgimento de aspirantes a produtores de estatísticas públicas na sobrevivência de uma instituição produtora tradicional. Sobre esse aspecto, praticamente todas as problemáticas derivam da percepção de que novas tecnologias – e, conseqüentemente, processos mais rápidos e econômicos – são absorvidas de maneira mais ágil pelo setor privado, convertendo-se em ganho competitivo.

Figura 1 - Resumo de algumas dificuldades recentes dos produtores de estatísticas públicas com levantamentos do tipo de campo



Fonte: Elaborado pelo autor.

O contexto geral de dificuldades dos produtores de estatísticas públicas, resumido na Figura 1, é tema que já desperta atenção e está inserido na agenda de diversas instituições pelo mundo. No Brasil, por exemplo, o IBGE demonstrou estar atento a tais dificuldades ao tratar o tema como central na terceira edição da Conferência Nacional de Produtores e Usuários de Informações Estatísticas, Geográficas e Ambientais (INFOPLAN), realizada em 2016.

Para lidar com todo o contexto de problemáticas, alternativas trabalhadas pelos produtores tem caminhado no sentido: da correção estatística das problemáticas atuais (CITRO, 2014); da modernização dos processos tradicionais – especialmente com a informatização e a automatização de tarefas antes manuais; da substituição de algumas pesquisas e questões por uma maior utilização de dados disponíveis em registros administrativos; de tentativas de realização de pesquisas postais, via telefone e internet; de combinações de dados de distintas fontes; e até da comercialização de informações para captação de recursos a serem reinvestidos nos processos tradicionais.

Em paralelo, entretanto, uma nova oportunidade tem se oferecido como possibilidade de solução para algumas das dificuldades das práticas tradicionais: usufruir da “revolução dos dados” e do Big Data, conforme apresentado na próxima subseção.

Para Citro (2014), é obrigação mover-se do paradigma de produzir as melhores estimativas possíveis a partir de levantamentos tradicionais de campo – que cada vez mais se apresentam inadequadas para atender às expectativas do novo contexto de demandas (PINK *et al.*, 2009) e são estranguladas por questões de custos e diminuição nas taxas de respostas (BRAAKSMA *et al.*, 2014) – para o da produção que se aproveita de múltiplas fontes, beneficiando-se do contexto da “revolução dos dados”, para geração das melhores estimativas possíveis e que atendam às demandas dos usuários – resultados relevantes, oportunos, acurados e economicamente viáveis.

Não que o produto da revolução dos dados substituirá integralmente as fontes de dados atuais dos produtores de estatísticas públicas – até porque também está sujeito a problemáticas e complicadores –, mas deverá ter papel central como suplemento para os dados já existentes (HACKL, 2016) a partir de uma integração que poderá maximizar valor (HARWOOD; MAYER, 2016). Kitchin (2015), nesse contexto, é categórico ao reforçar a necessidade de uma significativa

inovação disruptiva na produção de estatísticas públicas a partir do surgimento de uma série de oportunidades. Demunter (2017) é ainda mais radical, afirmando que a utilização de Big Data é questão de sobrevivência para os agentes tradicionais, uma vez que o novo contexto pode estender a capacidade de produção de informação estatística em alguma medida para esferas que vão além dos produtores até então tradicionais.

Revolução dos dados e Big Data

Com a popularização da internet, o avanço das tecnologias e a consolidação da era digital, consagrando o fenômeno chamado de Nova Era Digital (SCHMIDT; COHEN, 2013), uma grande massa de indivíduos passou a ser não apenas consumidora de informações como também produtora, a medida que deixam “rastros” em dispositivos conectados em rede – uma das características, conforme Costa *et al.* (2009), do conceito de segunda geração da *World Wide Web* (Web 2.0), proposto por O’Reilly (2005).

A produção desse tipo de informação está associada ao registro de atividades e acontecimentos – na forma de dados digitais – em dispositivos inseridos no universo dos sensores, scanners e da internet das coisas, dos smartphones e aplicativos, das páginas da internet, dos mecanismos de buscas, das redes sociais virtuais, dos blogs e web fóruns, entre outros. Tais registros compreendem o produto tanto da interação consciente de indivíduos – ao deliberadamente executarem uma tarefa junto a esses dispositivos – quanto inconsciente – ao terem suas atividades capturadas de forma implícita ou despercebida.

No que diz respeito ao universo desses dados digitais – muitas vezes chamados simplesmente de Big Data, como faz UNECE (2013) ao apresentar e classificar¹ as possíveis fontes desse tipo de dado –, há algum tempo já se percebe uma explosão na dimensão da sua produção, ocasionando o que é chamado por Miller (2010) de “avalanche de dados”. Segundo Schmidt e Cohen (2013), se o ritmo atual da inovação tecnológica for mantido, a maioria da população da Terra, estimada em oito bilhões de pessoas, estará online em 2025. Este fenômeno vem contribuindo para que o “universo digital” crescesse em média 30% ao ano pelos últimos 30 anos – cerca de cinco vezes mais rápido que o crescimento da economia mundial (SAMMAN, 2015), em nível atual de crescimento do quantitativo de dados digitais no mundo na ordem de 100% ao ano – dobra a cada ano. Em 2015, por exemplo, conforme Helbing *et al.* (2016), o mundo produziu mais dados digitais do que havia produzido em toda história da humanidade até 2014. Segundo o mesmo autor, logo mais e mais coisas ao nosso redor estarão conectadas a internet, fazendo com que a quantidade de dados digitais no mundo dobre a cada 12 horas. (HELBING *et al.*, 2016).

Além da internet e dos dispositivos digitais, a explosão no volume de dados produzidos tem sido viabilizada por avanços tecnológicos que garantem maior capacidade de armazenamento e de processamento desses dados, fazendo com que informações possam ser trabalhadas com velocidade e em tempo real.

Esses dois “V’s”, volume e velocidade, quando adicionados ao “V” relacionado a característica de variedade dos dados digitais – no sentido de que derivam de diversas fontes e assumem distintos tipos e formas: quantitativos e qualitativos; estruturados, semi ou não estruturados – formam o núcleo base dos conceitos mais genéricos de Big Data – baseado nos 3 “V’s” (KITCHIN, 2013). Ou seja, dados que estão disponibilizados em tempo real, em larga escala, menos estruturados e em diferentes tipos de variáveis (EINAV; LEVIN, 2013). Outras definições focam no volume e colocam o Big Data como conjuntos de dados cujo tamanho ultrapassa a capacidade das ferramentas tradicionais de bancos de dados para capturar, armazenar, gerenciar e analisar (MANYIKA *et al.*, 2011).

¹ Em taxonomia proposta pela UNECE (2013), as fontes de Big Data podem ser classificadas em: redes sociais (informações de base humana); sistemas tradicionais de organizações (dados mediados por processos); e internet das coisas (dados gerados por máquinas).

Além do conceito de Big Data, Howe (2006) introduziu e Schenk e Guittard (2009) definiram o conceito de *crowdsourcing data* para se referir a dados de origem semelhantes aos do Big Data – extraídos da rede e baseados nas atividades dos indivíduos – mas sem impor uma restrição relacionada ao grande volume.

Em paralelo aos fenômenos do Big Data e do *crowdsourcing data*, os dados abertos – *open data* – tem recebido importante atenção na literatura relacionada. Caracterizados por dados que são tornados técnica e legalmente abertos, isto é, em formatos prontos para leitura computacional e com permissão de livre acesso para uso comercial e não comercial (WB, 2014), os dados abertos tem sido alvo de políticas de estímulo e incentivos a partir de esforços em regulação por parte de governos (JUG, 2014). Porém, essa já não tem sido uma diretriz restrita ao setor público, uma vez que a abertura de dados tem sido colocada também na agenda do setor privado. Wong (2012) traz alguns exemplos de abertura de dados no setor privado e alega que tal movimentação ocorre a partir da compreensão da importância dos negócios em redes, dos modelos de *open source* e da *open innovation*.

De maneira geral, tais os movimentos acontecem pelo entendimento de que os benefícios da disponibilização da maior quantidade possível de dados podem auxiliar tanto o setor público – na formulação e na avaliação de políticas públicas – quanto o setor privado – na gestão das atividades das empresas em busca de produtividade, competitividade e conhecimento de mercados – e a sociedade em geral (englobando o terceiro setor e a academia) – com aumento da produção de conhecimento, da transparência e das atividades de controladoria social, solidificando o processo democrático e o entendimento de informação pública enquanto bem-público.

Percebe-se, portanto, um processo de transformação nos dados, passando dos antigos *small data*, predominantemente analógicos, escassos e de acesso limitado, para o resultado de um dilúvio de produção, sintetizado por Kitchin (2015) como Big Data, com características de: vastos; velozes; variados; resolutos; relacionais; de baixo custo; e cada vez mais abertos e acessíveis. Este processo de transformação compõe o contexto do que está sendo chamado de “revolução dos dados” (KINTCHIN, 2014). Para Christensen (1997), o processo tem características de inovação de ruptura ao desafiar o *status quo* de como dados são produzidos, geridos, analisados, armazenados e utilizados. Dessa forma, modifica a maneira como o conhecimento é produzido, os negócios são conduzidos e a governança promulgada.

Portanto o Big Data, o *open data*, o *crowdsourcing data* e até mesmo o *scaled small data*², são todos elementos componentes da “revolução dos dados”, que se consolidam com auxílio dos avanços tecnológicos – especialmente no que se refere aos avanços em infraestrutura para dados e da nova onda dos TICs³.

Na sequência deste estudo, será adotada uma definição genérica de Big Data que engloba todos os tipos de dados digitais – inclusive registros administrativos –, considerando que se tratam de

vastas quantidades de dados digitais dinâmicos e variados, que são facilmente combinados, compartilhados e distribuídos por redes de tecnologia da informação e comunicação, analisados por uma nova geração de métodos e ferramentas de análise de dados desenvolvida para lidar com dados em abundância – em oposição ao métodos e ferramentas desenvolvidos antigamente desenvolvidos para lidar com dados em escassez (KITCHIN, 2014, p. XV).

Com o passar do tempo, a importância científica e social dessa explosão de dados vem sendo percebida, da mesma forma que seu valor econômico. Para Mayer-Schönberger e Cukier (2013), trata-se de um fenômeno que seguirá “chacoalhando” tudo, dos negócios e das ciências até os sistemas de saúde, os governos, a educação, a economia, as humanidades e todos os demais aspectos da sociedade.

No âmbito do setor privado, muitos negócios já têm percebido e utilizado o Big Data como uma ferramenta efetiva para superar a concorrência (UNGP, 2012). Entretanto, ao passo as

² Referente ao processo de digitalizar, combinar e ampliar bases de dados produzidas tradicionalmente (anteriores as de Big Data).

³ Caracterizada pela massificação: dos dispositivos digitais nas casas, ambientes de trabalho e espaços públicos; da computação distribuída e em nuvem; das redes sociais; e da internet das coisas (Kitchin, 2014). Em Jug (2014), o fenômeno associado é referido como “a terceira plataforma”.

empresas estão tendo sucesso demonstrando os benefícios que podem acumular a partir do Big Data, efeitos mais substantivos ou qualitativos – principalmente em termos de práticas sociais e de causas públicas – ainda estão apenas começando a surgir (ASQUER, 2013). O relatório de UN (2014), por exemplo, coloca a necessidade de os países construírem agendas que mobilizem esforços para a utilização do produto da revolução dos dados em práticas voltadas para o desenvolvimento sustentável.

Para o *World Economic Forum* (WEF), a revolução dos dados está colocando o Big Data na condição de “novo petróleo”, tornando-os um valioso recurso do século XXI, que surge como uma nova classe de ativos (WEF, 2011; BOSSOI, 2014). Como exemplos concretos, Manyika *et al.* (2011) e CEBR (2012) estimaram o valor adicional potencial do Big Data em diversos setores das economias dos Estados Unidos e do Reino Unido, respectivamente, encontrando cifras que estão na casa de trilhões de dólares ao ano. Como consequência, vem desempenhando papel central na dita quarta revolução industrial (SCHWAB, 2016), em um período de *data economy* (OPHER *et al.*, 2016).

Metodologia

Partindo do objetivo de apresentar o estado da arte da produção científica sobre a utilização de Big Data na produção de estatísticas públicas, com foco sobre oportunidades e desafios de aplicação, este estudo se propõe a ser uma pesquisa científica com fins de construção de conhecimento no âmbito das ciências empíricas sociais, conforme definições trazidas por Gil (2008). Em relação ao alcance do objetivo proposto, o presente estudo almeja alcançar o nível exploratório – objetivando proporcionar maior familiaridade com o problema, com vistas a torná-lo mais explícito (GIL, 2008). Em termos de sua natureza técnica e operacional, trata-se de uma pesquisa aplicada, que objetiva gerar conhecimentos para o aproveitamento prático e dirigidos à solução de problemas específicos, com delineamento do tipo pesquisa bibliográfica – elaborada a partir de materiais já publicados (GIL, 2008).

O levantamento bibliográfico é realizado de forma sistemática (FORBES, 1998), partindo de uma pergunta claramente formulada e utilizando métodos sistemáticos e explícitos para: identificar, selecionar e avaliar criticamente pesquisas relevantes; e coletar e analisar dados dos estudos (GREEN; HIGGINS, 2005), com critérios para a busca que seguem o método PRISMA (MOHER *et al.*, 2009).

Neste processo, foram percorridas – em 19 de setembro de 2017 – as bases Scopus (Elsevier), Web of Science (Clarivate Analytics) e Scielo.org. O critério de busca, por sua vez, considerou todos os documentos que contivessem no título, resumo ou palavras-chave, pelo menos uma dentre as expressões *official statistics* e *public statistics*, somada de pelo menos uma das expressões dentre: *big data*, *data revolution*, *data avalanche*, *digital data*, *digital era*, *internet of things*, *open data*, *crowdsourcing data* e *web 2.0*. A escolha dos termos foi baseada no conhecimento prévio deste autor – pelo entendimento de que tais termos estariam mais relacionados com o conteúdo a ser explorado.

Dos 178 documentos que retornaram às buscas, 37 eram repetidos e foram excluídos, fazendo com que a quantidade de documentos exclusivos considerados de tais bases fosse de 141. Destes, 85 foram excluídos em uma análise de conteúdo ou por terem sido julgados fora do escopo de interesse – não contribuindo para o atendimento do objetivo geral – ou por estarem escritos em língua diferente de inglês ou português. Assim, 56 documentos derivados das bases consultadas foram aqui trabalhados. Cabe destacar a ausência de documentos na base Scielo.org, após busca realizada com os termos também traduzidos para o português, sugerindo uma baixa ou inexistente produção sobre o tema em âmbito brasileiro.

Do quantitativo final selecionado, foram analisados dados referentes ao autor (nome, vínculo e nacionalidade) e ao texto (título, objetivo, metodologia, contexto, aplicação prática – fonte de dados, área de aplicação e finalidade –, resultados, e limitações, lacunas, desafios e problemá-

ticas). As análises serviram de base para a construção dos conteúdos que buscam responder ao objetivo central deste estudo – apresentados na seção de resultados.

Resultados

Esta seção está organizada em duas partes. Em um primeiro momento, são apresentadas oportunidades de inovação na produção de estatísticas públicas relativas à utilização de dados do tipo Big Data percebidas na literatura investigada, com destaque para possíveis fontes de dados e aplicações práticas. Posteriormente, são apresentadas problemáticas, desafios e riscos que permeiam tais oportunidades. Aspectos que demandam atenção e esforços futuros de pesquisa.

Oportunidades, fontes e aplicações

A utilização do produto da revolução dos dados – aqui sintetizado por Big Data – como fonte de dados tem sido percebida como uma oportunidade para os produtores de estatísticas públicas tentarem contornar diversas das problemáticas já aqui apresentadas (HACKL, 2016; CITRO, 2014; KITCHIN, 2015; BRAAKSMA *et al.*, 2014, HARWOOD; MAYER, 2016), resumidas na Figura 1.

Florescu *et al.* (2014) considera que as oportunidades podem ser analisadas em diferentes contextos que dependem do objetivo do processo de inovação: substituir total ou parcialmente as práticas tradicionais com interesse de geração de estatísticas equivalentes, com mesma perspectiva e utilidade (contexto de manutenção do propósito original); trazer uma nova perspectiva de análise para os domínios já trabalhados (contexto de adicionar ou complementar aos processos tradicionais); contribuir como informações auxiliares aos processos tradicionais (contexto de melhoramento das práticas existentes); ou alterar tanto a perspectiva de abordagem e análise quanto as fontes e os métodos de produção (contexto de propor uma alternativa ao *status quo* de análise de certo domínio).

Cabe destacar que o objetivo do processo de inovação não se define apenas por interesse do produtor, mas também por uma análise da qualidade dos insumos de Big Data – que podem ser problemáticos, conforme é apresentado na próxima subseção.

Em uma perspectiva genérica, partindo dos possíveis benefícios mencionados por Kitchin (2015) e considerando também as contribuições de Tam e Clarke (2014), Florescu *et al.* (2014), Eurostat (2014), UNGP (2012), Hand (2015), Braaksma e Zeelenberg (2015) e Hackl (2016), foi possível construir o resumo apresentado no Quadro 1. As oportunidades estão relacionadas com os tipos e problemáticas dos produtores – trazidos na Figura 1 – que podem auxiliar a resolver.

Nesses termos, a literatura investigada aponta, conforme apresentado no Quadro 1, que o Big Data pode contribuir, em alguma medida e nem sempre concomitantemente, para o tratamento de todas as problemáticas que produtores de estatísticas públicas enfrentam.

Tais oportunidades foram identificadas a partir das possibilidades de utilização de dados digitais advindos das fontes alternativas apresentadas no Quadro 2, construído a partir Kitchin (2015) e ampliada com contribuições dos autores relacionados no corpo do quadro. Pelas aplicações encontradas, percebe-se que as fontes alternativas de dados já se apresentam como extensa oportunidade de produção de conhecimento nas mais diversas áreas, com replicação imaginável para outras aqui não relacionadas.

Quadro 1 - Possíveis oportunidades trazidas pelos dados digitais para a produção de estatísticas públicas e sua relação com as problemáticas da Figura 1

Escopo	Problemática	Oportunidade
Operacional	Todos	Reduzir ou eliminar a necessidade de conduzir levantamentos de campo
Recursos	Custos	Reduzir custos dos processos de produção, hoje fortemente baseados em coletas em campo, contornando problemáticas de orçamento
	Tempo	Reduzir tempo dos processos de produção predominante atual
Metodológico	Erro de cobertura	Alcançar populações completas em algumas situações específicas
Demanda	Relevância	Captar informações mais aderentes ao contexto social evolutivo uma vez que são reflexos de transações, interações e comportamentos que os indivíduos de fato executam - e não simplesmente dizem, como em uma pesquisa tradicional
	Variedade	Gerar resultados para novas áreas/temas de investigação, atualmente não trabalhadas
	Profundidade	Gerar resultados mais detalhados e com aberturas para pequenas localidades e subgrupos populacionais específicos com o aumento da quantidade de informações disponíveis
	Oportunismo	Produzir estatísticas com maior frequência ou de forma contínua em tempo real, agregando valor pelo oportunismo e utilidade prática decisória (ações preventivas e de monitoramento rápido)
	Qualidade	Gerar estimativas mais precisas ao contornar a queda de aproveitamento das pesquisas tradicionais e alcançar uma quantidade maior de informações
	Economia	Otimizar processos de trabalho, gerando ganhos de eficiência percebidos pelo usuário
	Comparabilidade	Facilitar comparações entre regiões e países, permitindo que sejam construídas para quaisquer países, independente da sua estrutura
Estratégico	Concorrência	Desenvolver produtos de dados mais ricos a partir do link entre diversas bases de dados e/ou perspectivas longitudinais
		Liberar o quadro funcional para tarefas de maior valor agregado
		Promover maior visibilidade e utilização de estatísticas produzidas
		Explorar novos mercados como novas estatísticas e domínios
		Complementar, substituir, melhorar e adicionar às bases de dados já existentes
	Isolamento	Promover a aproximação, a colaboração e a co-criação com outros players da indústria de dados, produtores de dados digitais e especialistas das áreas das ciências da computação e de dados
	Obsolência	Explorar novas técnicas de análises de dados para obtenção de insights diferenciados
	Credibilidade	Servir de dados auxiliares para aumentar a qualidade de processos tradicionais de produção (para variáveis de estratificação, imputação de valores faltantes, tratamento de anomalias)
Reforçar a qualidade (validade e consistência) e a credibilidade de estatísticas tradicionais, funcionando como um contracheque		

Fonte: Elaborado pelo autor com base na literatura investigada.

Quadro 2 - Fontes de dados digitais por tipo e áreas de aplicação já trabalhadas na literatura investigada, com relação das correspondentes publicações

Tipo	Fonte de dados	Publicações relacionadas	Áreas de aplicação (1)
Comunicação móvel	Dados de telefones móveis (2)	Dattilo et al. (2016); deJonge et al. (2012); Heerschap et al. (2014); Tennekes; Offermans (2014); EUROSTAT (2014); Blumenstock et al. (2017); Ahas et al. (2010); Saluveer; Ahas (2010); Tiru (2014); Ahas (2008); Karlberg; Skaliotis (2013); SNZ (2012); UNECE (2014).	Turismo, população, migração, mobilidade, tráfego, redes de relacionamento, pobreza, saúde.
Internet e crowdsourcing	Redes sociais virtuais (3)	Chew (2010); Daas; Puts (2014); Daas et al. (2012); Java (2007); ACM (2007); UNECE (2014).	Comportamento, confiança do consumidor, consumo, economia, emprego, força de trabalho, inflação, informações sobre empresas, imóveis, migração, mobilidade, PIB, população, potencial de mercado, preços, qualidade dos serviços urbanos, renda, sociedade da informação, opinião, redes de relacionamento, saúde, sentimento, turismo, uso da terra.
	Comunidades de compartilhamentos de imagens e vídeos (4)	ESSC (2014).	
	Blogs e fóruns		
	Buscas na web	Ginsberg et al. (2008); Valdivia et al. (2010).	
	E-mails		
	Logs de utilização da web e de sistemas computacionais	EC (2012).	
	Sites de acomodação, hotelaria e passagens aéreas, e-commerce, empregos, imóveis, organizações públicas e privadas	Barcaroli et al. (2015); Heerschap et al. (2014); Giannini et al. (2015); Cavallo; Rigobon (2016); Cavallo (2013); Cavallo (2009); deHaan (2013); Hoekstra et al. (2012); Groves (2011); Jing (2013); Virgillito (2014); Summa (2014).	
	Informação Geográfica Voluntária (5)	ESSC (2014).	
Aplicativos			
Sensores e Internet das coisas	Metereologia e poluição (científicos)	UNECE (2014).	Agricultura, consumo, cultivo de drogas (crime), estilo de vida, emissão de poluentes, habitação, meteorologia, migração, mobilidade, pesca, silvicultura, tráfego, transportes, turismo, uso da terra.
	Automação domiciliar		
	Rastreamento veicular	UNECE (2014).	
	GPS		
	Segurança e vigilância		
	Trânsito	Daas et al. (2012); Ruth (2014).	
	Medidores inteligentes (6)	ESSC (2014).	
	Imagens de satélite	Brezzi; Sanchez-Serra (2014); Donkelaar et al. (2015); Schnorr-Bäcker (2016); Tam (2015); Figueroa; Mayorga (2015); UNECE (2014).	
	Identificação automática de embarcações	ESSC (2014).	
Movimentação de vôos	ESSC (2014).		
Registros e scanners	Transações comerciais (7)	UNECE (2014).	Comportamento, consumo, estilo de vida, economia, inflação, preços, vendas.
	Cartões de crédito	ESSC (2014).	
	Transações e estoques financeiros e bancários		
	Administrativos (8)		

Fonte: Elaborado pelo autor com base na literatura investigada.

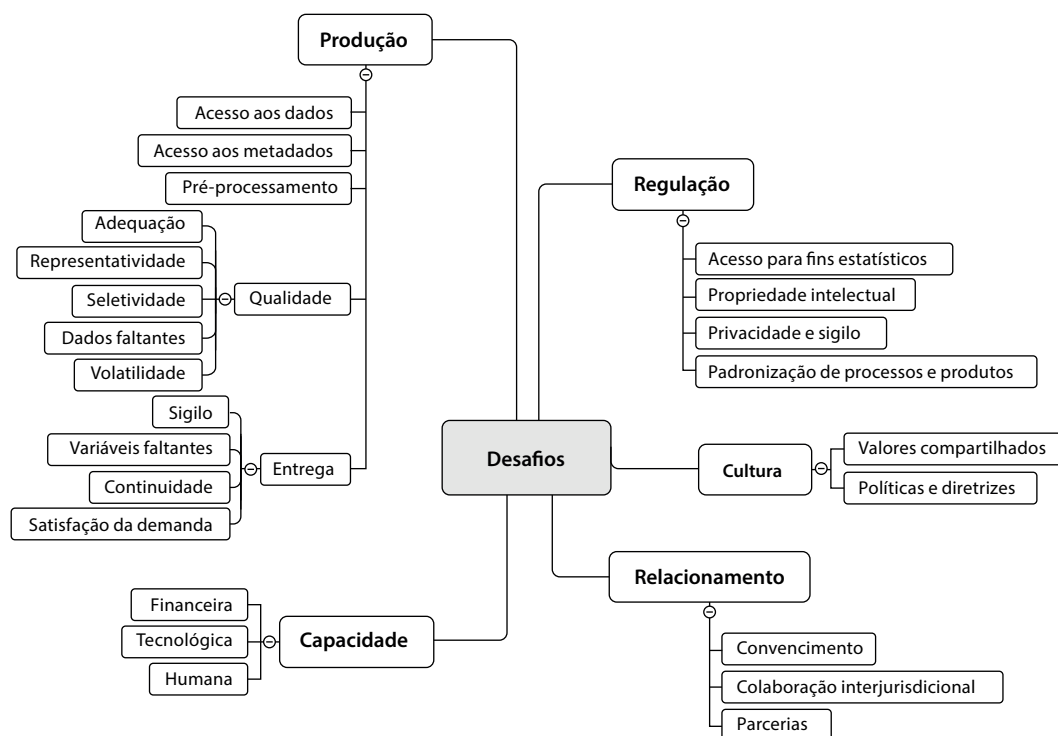
(1) Áreas listadas são aquelas com alguma aplicação já relatada na literatura investigada. (2) Registro de ligações, mensagens e posicionamento. (3) Facebook, Twitter, Instagram, LinkedIn, Tumblr etc. (4) YouTube, Flickr, Instagram, Picasa, Panoramio etc. (5) OpenStreetMap, Wikimapia, Geowiki etc. (6) Energia, entre outros. (7) Online, do varejo, de supermercados etc. (8) Médicos, de programas sociais etc.

Desafios, problemáticas e riscos

Conforme a literatura analisada, percebe-se, apesar do atraente contexto de oportunidades, que muitos desafios ainda estão colocados. Tais desafios derivam do fato de que dados digitais são predominantemente produzidos por instituições privadas para seus próprios interesses, sem a intenção de seguir padrões que permitam sua validade e utilização na forma de estatísticas públicas (KITCHIN, 2015). Assim, grande parte dos problemas passa pela questão de mensurar a qualidade dos dados digitais nos moldes tradicionalmente conhecidos (UNECE, 2014; REIMSBACH-KOUNATZE, 2015), uma vez que os dados digitais podem se apresentar cheios de ruídos, com dados estimulados, falsos e de cobertura parcial e seletiva da população (DASS *et al.*, 2013; KITCHIN, 2014).

Autores como Struijs *et al.* (2014), Tennekes & Offermans (2014) e Vaccari (2014) estabelecem uma classificação que distingue os desafios em seis grandes tipologias: legislação; privacidade; financeiro; gerencial; tecnológico; e metodológico. Conforme Vaccari (2014), em termos de legislação, a problemática se refere a possibilidade de acesso às bases que são produto da revolução dos dados – muitas vezes privadas. Sobre privacidade, as questões são relacionadas às garantias de que dados disponibilizados serão protegidos em termos de confidencialidade. Em termos financeiros, o custo pelo acesso aos dados – que por vezes serão buscados no setor privado – pode surpreender em certas ocasiões. Do ponto de vista da gestão, a utilização de fontes derivadas da revolução dos dados em instituições com ambientes de produção tradicional deverá demandar planejamento e gestão para novas práticas e modelos. Tecnicamente, os desafios estão na avaliação da capacidade computacional e da mudança necessária de paradigma requerido em termos de tecnologia da informação. E metodologicamente, por fim, as questões se referem ao impacto das novas fontes sobre os métodos tradicionais de coleta, processamento e disseminação dos dados, com uma provável e consequente adaptação e desenvolvimento de novos métodos de aferição da qualidade e de validação estatística.

Aqui, entretanto, em virtude da percepção de questões adicionais, optou-se por agrupar os desafios de outra maneira, conforme apresenta a Figura 2. Foram considerados como grandes grupos os aspectos: regulação e legislação; relacionamento externo; capacidades organizacionais; cultura organizacional; e o processo de produção de estatísticas.

Figura 2 - Resumo de desafios para produtores de estatísticas públicas no contexto da utilização de dados digitais

Fonte: Elaborado pelo autor.

Avanços no âmbito de regulação e legislação são talvez os pilares que podem consolidar e escalonar o processo de inovação na produção de estatísticas públicas. Nesse contexto, se faz necessária a formalização de normas que tratem sobre:

- i. Obrigatoriedade de certas organizações em disponibilizar bases de dados digitais para que produtores licenciados possam gerar produtos com fins de estatísticas públicas (VACCARI, 2014; HACKL, 2016);
- ii. Respeito à propriedade intelectual (KITCHIN, 2015);
- iii. Direito à privacidade, ao sigilo e à proteção contra intrusão e uso de dados autorizados e não autorizados pelo indivíduo fonte, também quando ele não está consciente dos dados que gera (VALE, 2015; STRUIJS; DAAS, 2013, DAAS *et al.*, 2015; DE JONGE *et al.*, 2012; HACKL, 2016; TAM; CLARKE, 2015);
- iv. Requisitos de padronização para os produtos gerados assim como de métodos e técnicas para os processos de produção, garantindo rigor e qualidade metodológica, viabilizando comparações temporais e entre diferentes países/regiões (KITCHIN, 2015).

Estruturado o aspecto normativo, produtores de estatísticas públicas vão precisar remodelar seus processos de gestão interna. Em termos de cultura organizacional, necessitarão direcionar esforços para:

- i. Trabalhar a mudança de visão estratégica, focando na disseminação contínua dos novos valores (KITCHIN, 2015);
- ii. Estabelecer políticas e diretrizes alinhadas com o novo contexto através de um processo de gestão que prepare a organização para assumir novas tarefas e responsabilidades (KITCHIN, 2015).

Ainda sobre a gestão institucional, porém agora considerando o relacionamento com o ambiente externo, surge a necessidade de se trabalhar aspectos no sentido de:

- i. Convencer detentores sobre a importância das estatísticas públicas (VACCARI, 2014) e consequente necessidade de liberação das bases de dados, nem sempre respeitada mesmo quando prevista legalmente (ESSNET, 2013), estabelecendo uma relação de confiança a partir da garantia de que os dados serão mantidos em sigilo e em segurança (TAM; CLARKE, 2015);
- ii. Estabelecer colaborações interjurisdicionais com outros produtores, no âmbito da *open innovation*, que permitam o enfrentamento aos desafios que se colocam, avanço conjunto e a homogeneização de processos de produção e produtos (VALE, 2015; KITCHIN, 2015; TAM; CLARKE, 2015);
- iii. Formalizar parcerias com detentores de Big Data, especialmente com o oferecimento de produtos de interesse mútuo (KRATKE; BYIERS, 2014; HACKL, 2016; VALE, 2015).

Para que a produção de estatísticas públicas alcance produtos que justifiquem a inovação, produtores vão precisar contornar problemáticas metodológico-operacionais que poderão surgir. Estas correspondem, talvez, ao mais preocupante entrave relacionado à utilização de Big Data, principalmente no quesito qualidade dos insumos – que podem apresentar anomalias inerentes e não passíveis de correção. Algumas problemáticas percebidas foram:

- i. Acesso às bases de dados, matéria prima para os produtos finais, de posse predominantemente do setor privado (TAM; CLARKE, 2015);
- ii. Acesso a aspectos metodológicos e metadados que auxiliam no entendimento da estrutura e do conteúdo das bases de Big Data (KITCHIN, 2015);
- iii. Mais tempo de pré-processamento das bases – compreensão e organização da estrutura, conteúdo, padrões e da qualidade dos dados (detecção de erros, anomalias e valores faltantes, etc.) – até que possam ser utilizados (HASSANI *et al.*, 2014; TENNEKES *et al.*, 2011; ZYKOPOULOS, 2012; FRY, 2008; DAAS, 2012; LIU *et al.*, 2013);
- iv. Requisitos associados à qualidade de seus produtos (VALE, 2015; HACKL, 2016). Avaliando a viabilidade metodológica do novo processo, certificando veracidade (acurácia, fidelidade), incerteza, erro (cobertura, seleção, mensuração e resposta), viés, confiança e calibração (KITCHIN, 2015; COUPER, 2013) uma vez que as bases de dados digitais podem conter ruídos, dados estimulados, falsos ou incompletos (DAAS; PUTS, 2014), aspectos na maioria das vezes desconhecidos (UNECE, 2014; REIMSBACH-KOUNATZE, 2015).
- v. Produtos que transmitam confiança (BRAAKSMA; ZEELENBERG, 2015), com relevância para o usuário através da garantia de que estejam mensurando e refletindo de maneira adequada os fenômenos aderentes às questões de pesquisa (KITCHIN, 2015; CITRO, 2014);
- vi. Representatividade dos resultados, uma vez que dados digitais não tiveram sua produção planejada e controlada como nos métodos tradicionais (HACKL, 2016; CITRO, 2014; TUFECKI, 2013; DAAS *et al.*, 2013; TAM; CLARKE, 2014);
- vii. Possível seletividade nas fontes de dados – que acaba por fazer com que representem apenas uma parcela da população alvo de interesse (BRAAKSMA; ZEELENBERG, 2015; DAAS *et al.*, 2015; ASA, 2014; BUELENS *et al.*, 2012; FLEKOVA; GUREVYCH, 2013; HASTIE *et al.* 2009);
- viii. Dados faltantes, que também podem afetar os dados digitais, especialmente em virtude de problemas com a rede ou com o servidor (DAAS, 2012);
- ix. Volatilidade devido à flutuação presente em alguns registros, especialmente de sensores, podendo fazer com que estimativas se apresentem distorcidas quando realizadas em alta resolução (BRAAKSMA; ZEELENBERG, 2015; DAAS *et al.*, 2015; DAAS; PUTS, 2014; O'CONNOR *et al.*, 2010);
- x. Segurança nos produtos e processos de modo que o sigilo das informações seja preservado (TAM; CLARKE, 2015);

- xi. Falta de variáveis chaves para integração com bases de dados tradicionais (BRAAKSMA; ZEELENBERG, 2015) ou covariáveis para análises completas;
- xii. Certificar de que as fontes de dados digitais não modificarão a estrutura dos dados ou descontinuarão o registro e a disponibilização dos dados (TAM; CLARKE, 2014; DAAS *et al.*, 2015; KITCHIN, 2015);
- xiii. Atendimento às expectativas do novo perfil de demandas, com coleta, processamento e disponibilização em tempo real de informações com valor agregado – riqueza e utilidade de conteúdo (PARISE *et al.*, 2012).

E, por fim, os objetivos serão alcançados apenas se os produtores estiverem capacidade em termos de:

- xiv. Recursos financeiros, principalmente para o caso de necessidade de aquisição de bases de dados digitais de empresas privadas (MILLS *et al.*, 2012; HACKL, 2016);
- xv. Recursos tecnológicos – infraestrutura de TI e ferramentas – com alta capacidade e eficiência para coletar e processar dados em tempo real – incluindo dados de streaming (NAS, 2013; SCHUTT; O'NEIL, 2013; PARISE *et al.*, 2012; SCANNAPIECO *et al.* 2013; STRUIJS *et al.* 2014; TAM; CLARKE, 2014; CERVERA *et al.*, 2014; HACKL, 2016; TAM; CLARKE, 2015; KITCHIN, 2015). Ao mesmo tempo, se faz necessário que os sistemas garantam a proteção e a segurança das informações armazenadas (DUNNE, 2013);
- xvi. Recursos humanos, com profissionais qualificados (DAVENPORT; PATIL, 2012; DAAS *et al.*, 2015; SCHUTT; O'NEIL, 2013; HACKL, 2016) que consigam avaliar e prover soluções – novos métodos e técnicas (LAVALLE, 2011) – tanto no âmbito tecnológico, para questões de TI (DUNNE, 2013) – com conhecimentos de engenharia de dados, computação de alta performance, data warehousing (HACKL, 2016), Hadoop e bases de dados SQL e noSQL (VALE, 2015) –, quanto metodológico e analítico, com habilidades estatísticas (VALE, 2015; HACKL, 2016) – para (1) soluções de desenvolvimento de métodos para *Big Data analytics* (KITCHIN, 2015; UNGP, 2012; EUROSTAT 2014; TAM; CLARKE, 2014), (2) avaliação da qualidade dos resultados (LAVALLE, 2011), (3) extração de conhecimento com visualização (KITCHIN, 2015) e mineração (DUNNE, 2013), (4) exploração inicial dos dados baseada em visualização (TENNEKES *et al.*, 2011; ZYKOPOULOS *et al.*, 2012; FRY, 2008; DAAS *et al.*, 2012; LIU *et al.*, 2013), (5) correção de dados faltantes (DAAS, 2012) e da seletividade (DAAS *et al.*, 2015) com métodos de predição (BUELENS *et al.*, 2012; FLEKOVA; GUREVYCH, 2013; HASTIE *et al.*, 2009) e mineração de dados (VALE, 2015), (6) tratamento da volatilidade com médias móveis e técnicas de filtragem avançada (DAAS *et al.*, 2015), e (7) garantia de sigilo das informações divulgadas (TAM; CLARKE, 2015). Além desses, é de fundamental importância a disposição de profissionais que entendam de ética e governança de dados e sejam criativos (VALE, 2015) para que possam propor soluções que ampliem a aplicabilidade, a utilidade e, conseqüentemente, o valor dos produtos (LAVALLE, 2011).

Encarar esses desafios compreende uma longa e incerta jornada para os produtores de estatísticas públicas que almejem aproveitar-se do Big Data para contornar, pelo menos em alguma medida, as problemáticas que se colocam no contexto dos processos e métodos tradicionais de produção.

Nesse caminho de mudança de paradigma, estão colocados inúmeros riscos a serem considerados, que, na visão de Kitchin (2015), contemplam: a possibilidade de desistência em meio à transição; a perda de reputação e de confiabilidade; a violação de privacidade e de segurança da informação; o acesso inconsistente às fontes alternativas e possíveis problemas de continuidade dos produtos gerados; uma possível resistência da população e de produtores das fontes de dados alternativos; a fragmentação das abordagens entre diferentes países e instituições; cortes e restrições de recursos disponíveis; e a competitividade com a possibilidade de transferência da produção para diferentes organizações do mercado.

Considerações finais

O presente artigo objetivou apresentar o estado da arte da produção científica que aborda a perspectiva de utilização de Big Data para a produção de estatísticas públicas e oficiais, especialmente no que concerne a oportunidades de contorno às problemáticas enfrentadas pelos produtores atuais – apresentando também possíveis fontes e domínios de aplicação – e aos desafios e problemáticas até então percebidos ou previstos e que podem se colocar àqueles que passarem a fazer uso de tais fontes alternativas.

O tema se apresenta de relevância extrema, pois perpassa uma possível quebra de paradigma que pode culminar no enfraquecimento e na perda de competitividade dos produtores atuais à medida que pode estender parte da capacidade de produção de informação estatística para outras esferas. Nesse sentido, considerar produzir estatísticas públicas a partir de processos inovadores e fontes alternativas – posicionando-se na vanguarda – se faz crucial e estratégico aos agentes atuais.

Nesse contexto, o presente artigo pode contribuir com produtores de estatísticas públicas que almejem construir uma agenda de trabalho em direção à mudança de paradigma em seus processos de produção. Ou seja, espera-se que o conteúdo aqui organizado e apresentado sirva para atentar e guiar produtores no que tange tanto a possibilidades de inovações e de contorno a problemáticas associadas aos métodos tradicionais de produção quanto aos desafios, riscos e incertezas que deverão ser considerados quando do planejamento de novas atividades, processos e produtos.

Como trabalhos futuros, coloca-se como elemento central a necessidade de pesquisa no âmbito dos novos desafios – agrupados em níveis de processo de produção (operacional-metodológico), regulação, cultura institucional, relacionamento externo e capacidade de produção – que aparecerão para aqueles que optarem por explorar processos inovadores de produção de estatísticas públicas.

Referências

- ASA. *Discovery With Data: Leveraging Statistics with Computer Science to Transform Science and Society*. July 2. 2014.
- ASQUER, A. *The Governance of Big Data: Perspectives and Issues*. Department of Financial and Management Studies SOAS, University of London, 2013.
- BIEMER, P.; TREWIN, D.; BERGDAHL, H.; LILLI, J. A system for managing the quality of official statistics, with discussion. *Journal of Official Statistics*, 30 (3, September), 381-442. 2014.
- BOSSEL, H. *Indicators for Sustainable Development: Theory, Method, Applications: A report to the Balaton Group*. Winnipeg: IISD, 1999.
- BOSSOI, R. A. C. *A proteção dos dados pessoais face às novas tecnologias: Direito e novas tecnologias*, Florianópolis: CONPEDI, 2014.
- BRAAKSMA, B.; DAAS, P.; OFFERMANS, M.; PUTS, M.; TENNEKES, M. *Big Data and official statistics: local experiences and international initiatives*. Paper for the 47th Scientific Meeting of the Italian Statistical Society, 11-13 June, Cagliari, Italy. 2014.
- BRAAKSMA, B.; ZEELENBERG, K. 'Re-make/Re-Model': Should Big Data Change the Modelling Paradigm in Official Statistics? *Statistical Journal of the IAOS*, 31 (2), pp. 193-202. 2015.
- BRACKSTONE, G. *Managing data quality in a statistical agency*. *Survey Methodology*, 25(2), 1-23. 1999.
- BUELENS, B.; BOONSTRA, H. J.; VAN DEN BRAKEL, J.; DAAS, P. *Shifting paradigms in official statistics: from design-based to model-based to algorithmic inference*. Discussion paper 201218, Statistics Netherlands, The Hague/Heerlen. 2012.
- CEBR. Data equity, unlocking the value of big data. *In: SAS Reports*, pp. 1-44, 2012.
- CERVERA, J. L.; VOTTA, P.; FAZIO, D.; SCANNAPIECO, M.; BRENNENRAEDTS, R.; VAN DER VORST, T. *Big data in official statistics*. ESS Big Data Event in Rome, 2014, Technical Workshop Report. 2014.
- CITRO, C. F. *From multiple modes for surveys to multiple data sources for estimates*. *Survey Methodology*, 40, pp. 137-161. 2014.
- COSTA, J.; FERREIRA, J. C.; DOMINGUES, L.; TAVARES, T.; DIEGUES, V.; COUTINHO, C. *Conhecer e utilizar a Web 2.0: Um estudo com professores do 2º, 3º Ciclos e Secundário*. X Congresso Internacional Galego-Português de Psicopedagogia (pp. 5614-5630). Braga: Universidade do Minho, 2009.
- COUPER, M. *Is the sky falling? New technology, changing media, and the future of surveys*. *Surv. Res. Meth.*, 7, 145-156. 2013.
- CHRISTENSEN, C. M. *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. Boston, MA: Harvard Business School Press, 1997.
- DAAS, P. *Big Data and official statistics*. Sharing Advisory Board, Software Sharing Newsletter 7, 2-3. 2012.
- DAAS, P.; TENNEKES, M.; DE JONGE, E.; PRIEM, A.; BUELENS, B.; VAN PELT, M.; VAN DEN HURK, P. *Data Science and the future of statistics*. Presentation for the first Data Science NL meetup, Utrecht, Nederland. 2012.
- DAAS, P. J. H.; PUTS, M. J.; BUELENS, B.; VAN DEN HURK, P. A. M. *Big Data and Official Statistics*. NTTs conference, Brussels, Belgium. 2013.

- DAAS, P. J. H.; PUTS, M. J. H. *Big data as a Source of Statistical Information*. The Survey Statistician 69, pp. 22-31. 2014.
- DAAS, P. J. H.; PUTS, M. J.; BUELENS, B.; VAN DEN HURK, P. A. M. Big Data as a Source for Official Statistics. *Journal of Official Statistics*, 31(2), 249–262. 2015.
- DAVENPORT, T.; PATIL, D. *Data scientist: the sexiest job of the 21st century* in Harvard business review 90.10: 70-77. 2012.
- DE JONGE, E.; VAN PELT, M.; ROOS, M. *Time patterns, geospatial clustering and mobility statistics based on mobile phone network data*. Discussion paper 201214, Statistics Netherlands. 2012.
- DE LEEUW, E. D.; DE HEER, W. *Trends in Household Survey Nonresponse: A Longitudinal and International Comparison*. R. M. Groves, D. A. Dillman, J. L. Eltinge and R. J. A. Little, eds. Survey Nonresponse, 41-54. New York: Wiley. 2002.
- DEMUNTER, C. *Tourism Statistics: Early Adopters Of Big Data?*, s.l.: UNWTO. Sixth UNWTO International Conference on Tourism Statistics. 2017.
- DUNNE, J. Big Data coming soon..... to an NSI near you. *Proceedings 59th ISI World Statistics Congress*, 25-30 August 2013, Hong Kong. 2013.
- EEA - EUROPEAN ENVIRONMENTAL AGENCY. *EEA core set of indicators: Guide*. Luxembourg: Office for Official Publications of the European Communities, 38p. 2005.
- EINAV, L.; LEVIN, J. *The data revolution and economic analysis*. NBER Conference, 2013.
- ESSnet. *Use of Administrative and Accounts Data in Business Statistics*. Deliverable 6.3: “Guidance on the accuracy of mixed-source statistics”. 2013.
- EUROSTAT. Big data: an opportunity or a threat to official statistics? Paper presented at the Conference of European Statisticians, 62nd plenary session, Paris. 2014.
- EUROSTAT. *The Quality Assurance Framework of the European Statistical System (ESS QAF)*. 2015.
- EUROSTAT. *European Statistics Code of Practices*. 2017.
- FEIJÓ, C.; VALENTE, E. As estatísticas oficiais e o interesse público. *Bahia Análise & Dados*, Salvador, v. 15, n. 1, p. 43-54, jun. 2005.
- FLEKOVA, L.; GUREVYCH, I. Can We Hide in the Web? Large Scale Simultaneous Age and Gender Author Profiling in Social Media. Paper for the evaluation lab on uncovering plagiarism, authorship, and social software misuse at Conference and Labs Evaluation Forum 2013, September 23–26, Valencia, Spain. 2013.
- FMI. *The Data Quality Assessment Framework (DQAF)*. 2003.
- FLORESCU, D.; KARLBERG, M.; REIS, F.; DEL CASTILLO, P. R.; SKALIOTIS, M.; WIRTHMANN, A. Will ‘big data’ transform official statistics? 2014.
- FMI. *The Data Quality Assessment Framework (DQAF)*. 2003.
- FORBES, D. A. Strategies for managing behavioural symptomatology associated with dementia of the Alzheimer type: A systematic overview. *Canadian Journal of Nursing Research*, 30(2), 67–86. 1998.
- FRICKER, S.; TOURANGEAU, R. Examining the relationship between nonresponse propensity and data quality in two national household surveys. *Public Opinion Quarterly*, 74(5), 935-955. 2010.
- FRY, B. *Visualizing Data: Exploring and Explaining Data with the Processing Environment*. Sebastopol, CA: OReilly Media Inc. 2008.
- GIL, A. C. *Métodos e Técnicas de Pesquisa Social*. 6. ed. São Paulo: Atlas, 2008.

- GREEN, S.; HIGGINS J. (editors). Glossary. *Cochrane Handbook for Systematic Reviews of Interventions* 4.2.5. 2005.
- GROVES, R. M.; PEYTCHEVA, E. The impact of nonresponse rates on nonresponse bias: A metaanalysis. *Public Opinion Quarterly*, 72(2), 167-189. 2008.
- IBGE. *Código de Boas Práticas das Estatísticas do IBGE*. Rio de Janeiro: IBGE. 2013.
- HACKL, P. Big Data: What Can Official Statistics expect? *Statistical Journal of the IAOS* 32. 43-52. 2016.
- HAND, D.J. Official Statistics in the New Data Ecosystem. Paper presented at the New Techniques and Technologies in Statistics conference, Brussels, March 10-12. 2015.
- HARWOOD, A.; MAYER, A. Big data and semantic technology: A future for data integration, exploration and visualisation. *Statistical Journal of the IAOS*, 32. 613-626. 2016.
- HASSANI, H.; SAPORTA, G.; SIRIMAL, S. E. Data Mining and Official Statistics: The Past, the Present and the Future. *Big Data* 2, 1-10. 2014.
- HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd ed. New York: Springer Science Business Media, LLC. 2009.
- HELBING, D.; FREY, B. S.; GIGERENZER, G.; HAFEN, E.; HAGNER, M.; HOFSTETTER, Y.; VAN DEN HOVEN, J.; ZICARI, R. V.; ZWITTER, A. *Behavioural Control or Digital Democracy?: A Digital Manifesto*. 2016.
- HOLT, D. T. The official statistics Olympics challenge: Wider, deeper, quicker, better, cheaper. *The American Statistician*, 61(1, February), 1-8. With commentary by G. Brackstone and J.L. Norwood. 2007.
- HOWE, J. The Rise of Crowdsourcing, *Wired*, 2006.
- JANNUZI, P. de M.; GRACIOSO, L. S. Produção e disseminação da informação estatística: agências estaduais no Brasil. *São Paulo em perspectiva*, São Paulo, 16 (3). pp. 92- 103. 2002.
- JANNUZZI, P. de M. *Indicadores sociais no Brasil: conceitos, medidas e aplicações*. 3. ed. Campinas: Alínea; Campinas: PUC, 2004.
- JUG, M. Information revolution from data to policy action in low-income countries: how can innovation help? PARIS21, Partnership in Statistics for Development in the 21st century. Discussion Paper No. 3. 2014.
- KITCHIN, R. Big data and human geography: opportunities, challenges and risks. *Dialogues in Human Geography* 3 (3): 262-267, 2013.
- KITCHIN, R. *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. Sage, London, 2014.
- KITCHIN, R. The opportunities, challenges and risks of big data for official statistics. *Statistical Journal of the IAOS*, 31 (3). pp. 471-481. 2015.
- KRATKE, F.; BYIERS, B. The political economy of official statistics: Implications for the data revolution in sub-Saharan Africa. PARIS21 Discussion Papers, No. 5, Partnership in Statistics for Development in the 21st Century, Paris. 2014.
- LAVALLE, S. Big data, analytics and the path from insights to value. *MIT Sloan Manage Rev.* 52. 2011.
- LIU Z.; JIANG B.; HEER J. imMens: real-time visual querying of big data. *Computer Graphics Forum* 32, 3, 421-430. 2013.

- MANSKI, C. F. Communicating uncertainty in official economic statistics. NBER Working Paper No. 20098. Cambridge, MA: National Bureau of Economic Research. 2014.
- MANYIKA, J.; CHUI, M.; BROWN, B.; BUGHIN, J.; DOBBS, R.; ROXBURGH, C.; BYERS, A. H. Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute, 2011.
- MAYER-SCHONBERGER, V.; CUKIER, K. *Big data: A Revolution that will transform how we live, work, and think*. New York: Houghton Mifflin Harcourt, 2013.
- MILLER, H. J. The data avalanche is here. Shouldn't we be digging? *Journal of Regional Science*, 50, 181–201, 2010.
- MILLS, S.; LUCAS, S.; IRAKLIOTIS, L.; RAPPA, M.; CARLSON, T.; PERLOWITZ, B. Demystifying Big Data: A Practical Guide to Transforming the Business of Government. Technical report. 2012.
- NAS. *Frontiers in Massive Data Analysis*. Washington, DC: The National Academies Press. 2013.
- MOHER, D.; LIBERATI, A.; TETZLAFF, J.; ALTMAN, D. Reprint – Preferred Reporting Items for Systematic Review and Meta-Analyses: The PRISMA Statement. *Physical Therapy*; 89(9): 873-80. 2009.
- OCDE. Rumo a um Ambiente Sustentável: Indicadores Ambientais. Série Cadernos de Referência Ambiental volume 9. Salvador: Organização para a Cooperação e Desenvolvimento Econômico. 2002.
- OCDE. Recommendation of the OECD Council on good statistical practice. 2015.
- O'CONNOR, B.; BALASUBRAMANYAN, R.; ROUTLEDGE, B. R.; SMITH, N. A. From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series. Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, May 23-26, Washington DC, USA. 2010.
- OPHER, A.; CHOU, A.; ONDA, A.; SOUNDERRAJAN, K. *The Rise of the Data Economy: Driving Value through Internet of Things Data Monetization*. USA: IBM Corporation. 2016.
- O'REILLY, T. *What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software*. O'Reilly Publishing, 2005.
- PARISE, S.; IYER, B.; VESSET, D. Four strategies to capture and create value from big data. *Ivey Business Journal*, 76(4), 1–5. 2012.
- PINK, B; BOROWIK, J.; LEE, G. The case for an international statistical innovation program – Transforming national and international statistics systems, *Statistical Journal of the IAOS*, 26(3), 125 –133. 2009.
- REIMSBACH-KOUNATZE, C. *The Proliferation of “Big Data” and Implications for Official Statistics and Statistical Agencies: A Preliminary Analysis*. OECD Digital Economy Papers, n. 245, OECD Publishing. 2015.
- SAMMAN, E. 10 things to know about development and the data revolution. ODI, 2015.
- SCANNAPIECO, M.; VIRGILLITO, A.; ZARDETTO, D. Placing Big Data in Official Statistics: A Big Challenge? Paper presented at New Techniques and Technologies in Statistics. 2013.
- SCHENK, E.; GUITTARD, C.: Crowdsourcing: What can be Outsourced to the Crowd, and Why? 2009.
- SCHMIDT, E.; COHEN, J. *A nova era digital: como será o futuro das pessoas, das nações e dos negócios*. Rio de Janeiro: Intrínseca, 2013.
- SCHNORR-BACKER, S. Statistical monitoring systems to inform policy decision-making, and new data sources. *Statistical Journal of the IAOS*, 33. 1-12. 2016.

- SCHUTT, R.; O'NEIL, C. *Doing Data Science: Straight Talk from the Frontline*. Sebastopol, CA: O'Reilly Media, 2013.
- SCHWAB, K. *A Quarta Revolução Industrial*. São Paulo: Edipro, 2016.
- SCHWARTZMAN, S. *Legitimidade, controvérsias e traduções em estatísticas públicas*. Teoria e Sociedade, Belo Horizonte, 1997.
- SENRA, N. Informação estatística: política, regulação, coordenação. IBICT, Ciência da Informação, v. 28, n. 2, 1999.
- SENRA, N. C. Estatísticas Formalizadas (C.1972–2002). *História das estatísticas brasileiras*. Rio de Janeiro: IBGE, vol. 4. 2009.
- STRUJIS, P.; DAAS, P. J. H. Big Data, Big Impact? Paper and presentation for the Seminar on Statistical Data Collection, Geneva, Switzerland. 2013.
- STRUJIS, P.; BRAAKSMA, B.; DAAS, P. Official statistics and Big Data. *Big Data and Society*, April–June, 1–6. 2014.
- TAM, S-M.; CLARKE, F. Big Data, Official Statistics and Some Initiatives by the Australian Bureau of Statistics. Paper presented at International Conference on Big Data for Official Statistics, Beijing, China. 2014.
- TAM, S-M.; CLARKE, F. Big data, official statistics and some initiatives by the Australian Bureau of Statistics. *International Statistical Review*, 83, 436–448. 2015.
- TENNEKES, M.; DE JONGE, E.; DAAS, P. J. H. *Visual Profiling of Large Statistical Datasets*. Statistics Netherlands, 2011.
- TENNEKES, M.; OFFERMANS, M. P. W. Daytime population estimations based on mobile phone metadata. Paper prepared for the Joint Statistical Meetings, Boston, 2014.
- TUFECKI, Z. Big Data: Pitfalls, Methods and Concepts for an Emergent Field. March 7. 2013.
- UN. National Quality Assurance Framework. 2012.
- UN Economic and Social Council. Big data and modernization of statistical systems, March 2014. E/CN.3/2014/11, 2014.
- UNECE. Fundamental Principles of Official Statistics in the UNECE region. Geneva, 1992.
- UNECE. Classification of Types of Big Data. 2013.
- UNECE. A Suggested Framework for National Statistical Offices for assessing the Quality of Big Data. 2014.
- UNGP. Big data for development: Challenges and opportunities. May 2012. Global Pulse, New York. 2012.
- VACCARI, C. Big Data in Official Statistics, Thesis University of Camerino, Italy. 2014.
- VALE, S. International collaboration to understand the relevance of Big Data for official statistics. *Statistical Journal of the IAOS*. 31. 159-163. 2015.
- WB. Open Data for Economic Growth. Transport & ICT Global Practice. 2014
- WEF. Report about Personal Data: The Emergence of a New Asset Class, World Economic Forum, 2011.
- WONG, D. Data is the next frontier, analytics the new tool: Five trends in big data and analytics, and their implications for innovation and organisations. Big Innovation Centre, London. 2012.

Estimação de fluxos brutos das situações ocupacionais com pesquisas amostrais repetidas: uma aplicação na Região Metropolitana de Salvador*

Priscila Pagung de Aquino Lapa**

Pedro Luis do Nascimento Silva***

Maria Luiza Barcellos Zacharias****

Resumo

A análise dos fluxos brutos consiste em compreender os movimentos dos indivíduos ao longo do tempo entre as situações ocupacionais de ocupado, desocupado e inativo, sendo uma poderosa ferramenta para identificar tendências e para entender a dinâmica do mercado de trabalho. Para realizar tal análise, são requeridas medidas repetidas em nível individual, configurando a necessidade de dados longitudinais. As pesquisas amostrais do tipo longitudinal ainda são pouco exploradas no Brasil com a finalidade de estimar os fluxos brutos de sua população em idade ativa, embora sejam comuns em diversos outros países. Este artigo tem como principal objetivo estimar e analisar os fluxos brutos das situações ocupacionais (ocupado, desocupado e inativo) da população em idade ativa, utilizando os dados da Pesquisa Mensal de Emprego (PME) realizada pelo Instituto Brasileiro de Geografia e Estatística no período de janeiro de 2008 a julho de 2013, considerando o caso específico da Região Metropolitana de Salvador. A estimação já foi feita usando modelos estatísticos que levam em consideração a ocorrência de não resposta e o emprego de amostragem complexa. Os resultados encontrados mostraram que o modelo que teve o melhor desempenho ao estimar os fluxos brutos pressupõe que a probabilidade de um indivíduo responder à pesquisa no próximo período de medição depende da situação ocupacional observada no período atual. Apesar disso, ao comparar os resultados obtidos do melhor modelo com uma análise ingênua em que a não resposta não é tratada, notou-se que as diferenças encontradas entre as estimativas foram pequenas. Também foi possível observar a trajetória da transição dos indivíduos entre as situações ocupacionais, permitindo analisar suas tendências no decorrer do tempo.

Palavras-chave: Fluxos brutos. Situação ocupacional. Modelos para fluxos brutos. Não resposta.

* Este capítulo é resultado da dissertação defendida na ENCE: "Estimando fluxos das situações ocupacionais com pesquisas amostrais repetidas."

** Docente da Faculdade do Centro Leste. Concluiu o mestrado em Estudos Populacionais e Pesquisas Sociais em 2015.

*** Pesquisador da ENCE e Docente do seu programa de pós-graduação desde sua criação em 1998.

**** Foi pesquisadora do IBGE até 2017, onde atuou na Coordenação de Métodos e Qualidade. Foi professora colaboradora do programa de pós-graduação da ENCE entre 2013 e 2017.

Introdução

As análises de fluxos brutos das situações ocupacionais são importantes para compreender o movimento dos indivíduos que estão em idade ativa entre dois períodos de tempo nas seguintes categorias: ocupado, desocupado e inativo (ILG, 2005). Constituem também uma poderosa ferramenta para análise da dinâmica do mercado de trabalho (BASSI; TORELLI; TRIVELLATO, 1998). Contudo, os fluxos brutos só podem ser estimados a partir de amostras nas quais se entrevistam os mesmos indivíduos em dois períodos de tempo (ROJAS, 2014).

A pesquisa do tipo longitudinal possibilita a análise de fluxos brutos, pois se caracteriza por coletar informações dos mesmos indivíduos em vários pontos no tempo (DUNCAN; KALTON, 1987). Pesquisas amostrais que usam um desenho do tipo painel rotativo, no qual os indivíduos que compõem a amostra são entrevistados várias vezes antes de saírem da pesquisa, têm se tornado cada vez mais comuns em diversos países. A *Canadian Labour Force Survey*, a *Current Population Survey*, a *UK Labour Force Survey* e a Pesquisa Mensal de Emprego, realizadas respectivamente no Canadá, Estados Unidos, Inglaterra e Brasil, são exemplos de pesquisas em que o foco consiste em analisar a conjuntura da força de trabalho de sua população, possibilitando captar amplamente suas características e suas formas de inserção no mercado produtivo.

Geralmente, os órgãos oficiais de estatística, como o Instituto Brasileiro de Geografia e Estatística, divulgam mensalmente estimativas das taxas de desocupação e de ocupação nas regiões de abrangência da pesquisa. No entanto, estas taxas não permitem a análise das mudanças em nível individual entre as situações ocupacionais. Por exemplo, não é possível saber o número de pessoas que mudam da situação de ocupado no mês atual para a de desocupado no mês seguinte.

A título de ilustração, suponha que no mês atual, os dados de estoque apresentaram um aumento líquido de indivíduos desocupados em relação ao mês anterior. As estimativas de fluxos brutos poderiam indicar que este aumento foi decorrente da transição de uma parcela de indivíduos que passaram a ser desocupados e de outra parcela que deixaram de ser desocupados. Observa-se que a transição de indivíduos que passaram a ser desocupados e que deixaram de ser desocupados podem apresentar diferentes configurações, embora a variação líquida seja igual (HUGHES E PEITZMEIER, 1989).

A pesquisa sobre a força de trabalho da Inglaterra (*UK Labour Force Survey*) realiza análises das situações ocupacionais por meio de fluxos brutos. Um resultado encontrado nesta pesquisa, comparando-se o trimestre de janeiro a março com o trimestre de abril a junho do ano de 2013, aponta que o fluxo bruto para a situação de ocupado aumentou ligeiramente no último trimestre. Isto ocorreu devido ao fato de que o número de desocupados que encontraram trabalho superou o dos ocupados que perderam trabalho. Em relação ao fluxo bruto para a inatividade, notou-se uma diminuição no fluxo dos ocupados que se tornaram inativos, parcialmente compensada pelo aumento no fluxo dos desocupados que passaram à inatividade. Informações como essas podem ser utilizadas para analisar as variações nas características do mercado de trabalho ao longo de dois períodos considerados (SIMPKINS, 2013).

A análise dos fluxos brutos entre as três categorias da situação ocupacional fornece, portanto, informações mais detalhadas sobre as transições dos indivíduos entre as categorias ao longo do tempo, possibilitando conhecer melhor a dinâmica da força de trabalho de um país ou de uma região de interesse.

O objetivo principal deste artigo consiste em estimar e analisar os fluxos brutos das situações ocupacionais (ocupado, desocupado e inativo) da população em idade ativa, utilizando os dados da Pesquisa Mensal de Emprego (PME) realizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE) no período de janeiro de 2008 a julho de 2013, considerando o caso específico da Região Metropolitana de Salvador.

Base de dados

O conjunto de dados utilizado é composto pelos microdados da PME para os meses de janeiro do ano de 2008 até julho do ano 2013, considerando a amostra referente à Região Metropolitana de Salvador. O recorte geográfico foi motivado pelas altas taxas de desocupação que recentemente vinham atingindo essa região, quando comparadas com as demais regiões metropolitanas contempladas pela pesquisa (PME, 2015). Assim, foram incluídas na base de dados em estudo as informações de todos os indivíduos residentes na Região Metropolitana de Salvador que foram entrevistados pela PME, com idade igual ou superior a 14 anos.

No esquema de rotação da amostra da PME, a cada mês dois painéis se renovam, sendo que um painel sempre se inicia na pesquisa pela primeira vez. Os painéis considerados no estudo foram os que iniciaram pela primeira vez na pesquisa em cada mês considerado. O motivo desse recorte consiste em selecionar os indivíduos que ainda não tiveram nenhuma participação anterior com a pesquisa buscando, assim, minimizar o viés de resposta potencial devido à permanência na pesquisa.

A variável de interesse para o estudo é a situação ocupacional do indivíduo que está categorizada em ocupado, desocupado e inativo, e corresponde à variável VD1 encontrada nos microdados da PME. Foram consideradas também as variáveis referentes à estrutura do plano amostral complexo usado para obtenção e ponderação da amostra.

Metodologia

Rojas (2014) propôs o uso de modelos de cadeia de Markov em duas etapas para descrever os fluxos brutos entre as diferentes categorias da situação ocupacional na população, considerando a possibilidade de não resposta. Propôs também estimadores consistentes para estimar os fluxos brutos sob tais modelos. Estes modelos e estimadores foram aqui aplicados aos dados da Região Metropolitana de Salvador.

Para observar os fluxos brutos entre dois períodos de tempo, obtidos de uma investigação exhaustiva (censo) sobre a população sob estudo, os mesmos indivíduos são entrevistados no tempo (mês) atual t e também no tempo (mês) posterior $t+1$. Suponha que o resultado de cada entrevista realizada possa levar à classificação dos indivíduos pertencentes àquela população em uma das G possíveis categorias mutuamente exclusivas de interesse em cada um dos períodos considerados. Neste trabalho, as categorias de interesse são Ocupado, Desocupado e Inativo, logo $G = 3$.

O modelo de Markov para a primeira etapa foi proposto para descrever os fluxos brutos populacionais na ausência de não resposta, situação em que seria possível obter uma classificação completa de todos os indivíduos pertencentes à população sob estudo, e obter os fluxos brutos representados na matriz $G \times G$, apresentada na Tabela 1.

Tabela 1 - Fluxos brutos populacionais em dois períodos de tempo consecutivos

Tempo t	Tempo $t + 1$						Total
	1	2	...	j	...	G	
1	X_{11}	X_{12}	...	X_{1j}	...	X_{1G}	X_{1+}
2	X_{21}	X_{22}	...	X_{2j}	...	X_{2G}	X_{2+}
⋮	⋮	⋮	...	⋮	...	⋮	⋮
i	X_{i1}	X_{i2}	...	X_{ij}	...	X_{iG}	X_{i+}
⋮	⋮	⋮	...	⋮	...	⋮	⋮
G	X_{G1}	X_{G2}	...	X_{Gj}	...	X_{GG}	X_{G+}
Total	X_{+1}	X_{+2}	...	X_{+j}	...	X_{+G}	N

Fonte: Adaptado de Rojas (2014, p.34).

Nota: 1. Processo não observável

2. Os símbolos "..." ou ":", nesta tabela, são apenas reticências, independente da orientação horizontal ou vertical.

Considerando esta representação, o tamanho total da população N satisfaz:

$$N = \sum_{i=1}^G \sum_{j=1}^G X_{ij} = \sum_{i=1}^G X_{i+} = \sum_{j=1}^G X_{+j}$$

onde X_{ij} representa o número de indivíduos na população com classificação i no tempo t e classificação j no tempo $t+1$; X_{i+} representa o número de indivíduos da população com classificação i no tempo t ; e X_{+j} representa o número de indivíduos da população com classificação j no tempo $t+1$.

O modelo de Markov para o processo cujos dados são descritos na Tabela 1 requer dois conjuntos de parâmetros:

- η_i ($i=1, \dots, G$) para representar a probabilidade de um indivíduo estar classificado na categoria i no tempo t ;
- p_{ij} ($i, j \in \{1, \dots, G\}$) para representar a probabilidade de transição do indivíduo entre a categoria i no tempo t para a categoria j no tempo $t+1$;

Sob um modelo de Markov, a probabilidade de um indivíduo ser classificado na categoria i no tempo t e na categoria j no tempo $t+1$ é dada por $\eta_i \cdot p_{ij}$.

A Tabela 2 apresenta as probabilidades de observação dos fluxos brutos para o modelo de Markov da primeira etapa.

Tabela 2 - Probabilidades dos fluxos brutos populacionais não observáveis para dois períodos de tempo consecutivos

Tempo t	Tempo $t+1$						Total
	1	2	...	j	...	G	
1	$\eta_1 p_{11}$	$\eta_1 p_{12}$...	$\eta_1 p_{1j}$...	$\eta_1 p_{1G}$	η_1
2	$\eta_2 p_{21}$	$\eta_2 p_{22}$...	$\eta_2 p_{2j}$...	$\eta_2 p_{2G}$	η_2
⋮	⋮	⋮	...	⋮	...	⋮	⋮
i	$\eta_i p_{i1}$	$\eta_i p_{i2}$...	$\eta_i p_{ij}$...	$\eta_i p_{iG}$	η_i
⋮	⋮	⋮	...	⋮	...	⋮	⋮
G	$\eta_G p_{G1}$	$\eta_G p_{G2}$...	$\eta_G p_{Gj}$...	$\eta_G p_{GG}$	η_G
Total	$\sum_{i=1}^G \eta_i p_{i1}$	$\sum_{i=1}^G \eta_i p_{i2}$...	$\sum_{i=1}^G \eta_i p_{ij}$...	$\sum_{i=1}^G \eta_i p_{iG}$	1

Fonte: Adaptado de Rojas (2014, p.31-34).

Nota: Os símbolos "..." ou ":", nesta tabela, são apenas reticências, independente da orientação horizontal ou vertical.

Os modelos propostos por Rojas (2014) para descrever os fluxos brutos entre as diferentes situações ocupacionais do indivíduo num determinado tempo t para o tempo posterior $t+1$ consideram também a possibilidade de ocorrência de não resposta. A não resposta pode ocorrer devido ao fato do indivíduo selecionado não ter sido localizado em algum dos períodos de interesse ou devido a dificuldades para o emparelhamento dos dados, causadas pelo uso de procedimentos não exatos de emparelhamento (a PME não disponibiliza chaves de emparelhamento longitudinal únicas para os indivíduos pesquisados).

A Tabela 3 ilustra os fluxos brutos em dois meses consecutivos para o processo populacional que seria observável sob um censo, mas sujeito a não resposta. Nessa tabela N_{ij} repre-

senta o número de indivíduos na população respondentes em ambos os tempos com classificação i no mês t e classificação j no mês $t+1$; R_i representa o número de indivíduos na população que responderam apenas no mês t com classificação i ; C_j representa o número de indivíduos na população que responderam apenas no mês $t+1$ com classificação j ; M corresponde ao número de indivíduos que não responderam à pesquisa em nenhum dos dois períodos.

Tabela 3 - Fluxos brutos em dois períodos de tempo consecutivos sob o Censo sujeito a não-resposta

Mês t	Mês $t + 1$				Complemento linha
	1	2	...	G	
1	N_{11}	N_{12}	...	N_{1G}	R_1
2	N_{21}	N_{22}	...	N_{2G}	R_2
⋮	⋮	⋮	...	⋮	⋮
G	N_{G1}	N_{G2}	...	N_{GG}	R_G
Complemento coluna	C_1	C_2	...	C_G	M

Fonte: Rojas (2014, p.20).

Nota: 1. Processo observável.

2. Os símbolos "..." ou ":", nesta tabela, são apenas reticências, independente da orientação horizontal ou vertical.

Considerando a notação introduzida na Tabela 3, o total de unidades da população sob estudo pode ser escrito como:

$$N = \sum_{i=1}^G \sum_{j=1}^G N_{ij} + \sum_{i=1}^G R_i + \sum_{j=1}^G C_j + M$$

Como é praticamente impossível fazer com que todos os indivíduos da população respondam à pesquisa, seja no tempo atual (t) ou no tempo posterior ($t+1$), a possibilidade de ocorrência de não resposta deve ser considerada.

Suponha que o resultado de cada tentativa de entrevista realizada possa levar a uma das seguintes situações: o indivíduo responde em ambos os tempos t e $t+1$ e tem classificação i e classificação j observadas; o indivíduo responde somente no tempo t e tem a classificação i observada; o indivíduo responde somente em $t+1$ e tem a classificação j observada; ou também, o indivíduo pode não responder nem no tempo t e nem no tempo $t+1$ e faltarem ambas as classificações.

O modelo de Markov proposto por Rojas (2014) para a segunda etapa busca justamente modelar o processo de ocorrência de resposta (não resposta) à pesquisa. Sob este modelo, denota-se por: $\psi(i,j)$ a probabilidade de observar no tempo t a classificação de um indivíduo, quando este indivíduo tem classificação i no tempo t e j no tempo $t+1$. Seja $\rho_{RR}(i,j)$ a probabilidade condicional de um indivíduo que teve sua situação ocupacional observada no tempo t também ter a situação ocupacional observada no tempo $t+1$, quando este indivíduo tem classificação i no tempo t e j no tempo $t+1$. Seja $\rho_{MM}(i,j)$ a probabilidade condicional de um indivíduo que não teve sua situação ocupacional observada no tempo t também não ter a situação ocupacional observada no tempo $t+1$, quando este indivíduo tem classificação i no tempo t e j no tempo $t+1$. O modelo de superpopulação proposto é o Markoviano e esses são os parâmetros do modelo de Markov para o processo de resposta (ou não resposta).

Sob o modelo Markoviano considerado para descrever o processo de resposta, as probabilidades da matriz de dados para o modelo em duas etapas estão dispostas na Tabela 4.

Tabela 4 - Probabilidades dos fluxos brutos populacionais observáveis para dois meses consecutivos

Mês t	Mês t + 1			
	1	2	...	Complemento linha
1				
2				
⋮				
i	$\rho_{RR}(i, j)\psi(i, j)p_{ij}\eta_i$			$\sum_{j=1}^G(1 - \rho_{RR}(i, j))\psi(i, j)p_{ij}\eta_i$
⋮				
G	$\sum_{j=1}^G(1 - \rho_{RR}(i, j))\psi(i, j)p_{ij}\eta_i$			$\sum_{i=1}^G \sum_{j=1}^G \rho_{MM}(i, j)(1 - \psi(i, j))p_{ij}\eta_i$
Complemento coluna				

Fonte: Rojas (2014, p.24).

Nota: Os símbolos "..." ou ":", nesta tabela, são apenas reticências, independente da orientação horizontal ou vertical.

Para detalhes da derivação das probabilidades contidas nessa tabela, consultar Rojas (2014). Assim especificado, o modelo não é identificável, pois há excesso de parâmetros. Foi necessário então especificar modelos simplificados denominados por Modelo A, Modelo B, Modelo C e Modelo D. Estes, por sua vez, caracterizam-se por diferentes suposições quanto ao mecanismo de resposta / observação. Estas simplificações são necessárias para tornar os modelos mais parcimoniosos, mediante a redução do número de parâmetros distintos que precisam ser estimados. Todos fazem hipóteses sobre os mecanismos de resposta ou de observação que descrevem as probabilidades de observar a situação ocupacional dos indivíduos nos períodos analisados.

O modelo A considera que a probabilidade de observar a classificação de um indivíduo no mês t não depende da situação ocupacional do indivíduo e nem de qualquer outro atributo do indivíduo, isto é, $\psi(i, j) = \psi \forall i, j$. Além disso, a probabilidade de um indivíduo que teve sua situação ocupacional observada no mês t também ter a situação ocupacional observada no mês t+1 é constante, e não depende de qualquer atributo do indivíduo ($\rho_{RR}(i, j) = \rho_{RR} \forall i, j$). A probabilidade de um indivíduo que não teve sua situação ocupacional observada no tempo t também não ter a situação ocupacional observada no tempo t+1 também é constante, e não depende de qualquer atributo do indivíduo ($\rho_{MM}(i, j) = \rho_{MM} \forall i, j$).

O modelo B considera que a probabilidade de observar a classificação de um indivíduo no mês t depende da situação ocupacional do indivíduo no mês t ($\psi(i, j) = \psi(i) \forall i, j$). A probabilidade de um indivíduo que teve sua situação ocupacional observada no mês t também ter a situação ocupacional observada no mês t+1 é constante, e não depende de qualquer atributo do indivíduo ($\rho_{RR}(i, j) = \rho_{RR} \forall i, j$). A probabilidade de um indivíduo que não teve sua situação ocupacional observada no mês t também não ter a situação ocupacional observada no mês t+1 é constante, e não depende de qualquer atributo do indivíduo ($\rho_{MM}(i, j) = \rho_{MM} \forall i, j$).

O modelo C considera que a probabilidade de observar a classificação de um indivíduo no mês t não depende da situação ocupacional do indivíduo e nem de qualquer outro atributo do indivíduo ($\psi(i, j) = \psi \forall i, j$). A probabilidade de um indivíduo que teve sua situação ocupacional observada no mês t também ter a situação ocupacional observada no mês t+1 depende

da situação ocupacional do indivíduo no mês t ($\rho_{RR}(i,j)=\rho_{RR}(i)\forall i,j$). A probabilidade do indivíduo que não teve sua situação ocupacional observada no mês t também não ter a situação ocupacional observada no mês $t+1$ depende da situação ocupacional do indivíduo no tempo t ($\rho_{MM}(i,j)=\rho_{MM}(i)\forall i,j$).

Por fim, o modelo D considera que a probabilidade de observar a classificação de um indivíduo no tempo t não depende da situação ocupacional do indivíduo e nem de qualquer outro atributo do indivíduo ($\psi(i,j)=\psi\forall i,j$). A probabilidade de um indivíduo que teve sua situação ocupacional observada no mês t também ter a situação ocupacional observada no mês $t+1$ depende da situação ocupacional do indivíduo no mês $t+1$ ($\rho_{RR}(i,j)=\rho_{RR}(j)\forall i,j$). A probabilidade de um indivíduo que não teve sua situação ocupacional observada no mês t também não ter a situação ocupacional observada no mês $t+1$, depende da situação ocupacional do indivíduo no tempo $t+1$ ($\rho_{MM}(i,j)=\rho_{MM}(j)\forall i,j$).

O Quadro 1 ilustra resumidamente as suposições simplificadoras adotadas para cada um dos quatro modelos reduzidos.

Quadro 1 - Probabilidades de resposta do modelo da segunda etapa referentes aos quatro modelos reduzidos considerados

Modelos	Probabilidade de observação no tempo t		Probabilidade de transição da situação de observação	
	Não depende da situação ocupacional em t	Depende da situação ocupacional em t	Não depende da situação ocupacional em t ou $t+1$	Depende da situação ocupacional em t ou $t+1$
Modelo A	$\psi(i, j) = \psi$	-	$\rho_{RR}(i, j) = \rho_{RR}$ $\rho_{MM}(i, j) = \rho_{MM}$	-
Modelo B	-	$\psi(i, j) = \psi(i)$	$\rho_{RR}(i, j) = \rho_{RR}$ $\rho_{MM}(i, j) = \rho_{MM}$	-
Modelo C	$\psi(i, j) = \psi$	-	-	$\rho_{RR}(i, j) = \rho_{RR}(i)$ $\rho_{MM}(i, j) = \rho_{MM}(i)$
Modelo D	$\psi(i, j) = \psi$	-	-	$\rho_{RR}(i, j) = \rho_{RR}(i)$ $\rho_{MM}(i, j) = \rho_{MM}(i)$

Fonte: Adaptado de Rojas (2014, p.19-30).

Nota: O símbolo "-", neste quadro, indica ausência de informação.

O procedimento que não considera a possibilidade de não resposta é chamado aqui de 'método ingênuo'. Portanto, as estimativas das probabilidades de transição que seriam obtidas sem considerar os modelos A, B, C e D, ignorando então as informações dos indivíduos selecionados que não responderam / não foram observados em ao menos uma das ocasiões, são denominadas por estimativas ingênuas das probabilidades de transição.

A qualidade do ajuste dos modelos foi analisada por meio do teste Qui-quadrado com correção de primeira ordem de Rao-Scott, que considera os efeitos do plano amostral complexo (ver Rao e Scott (1981)). O teste consiste em comparar as frequências (relativas) esperadas sob o modelo com as frequências (relativas) estimadas a partir dos dados.

Para seleção do melhor modelo, adotou-se como critério aquele cujo ajuste forneceu o menor valor da estatística do teste de Qui-quadrado com correção de Rao-Scott para a maioria dos meses analisados.

Procedimento de emparelhamento dos dados da PME

Autores como Ribas e Soares (2008, 2009), Rojas (2014) descreveram detalhadamente os procedimentos que desenvolveram/ adotaram para o emparelhamento dos indivíduos, uma vez que a PME acompanha o domicílio, não necessariamente as mesmas pessoas. O primeiro passo para reconstruir o painel é saber como identificar o mesmo domicílio em dois períodos de entrevista. As variáveis disponíveis nos microdados da PME que possibilitam identificar os domicílios são: a região metropolitana (V035); o número de controle (V040) que identifica um setor selecionado para a amostra; o número de série (V050) que identifica um domicílio selecionado num setor; o painel (V060) que é o conjunto de domicílios selecionados que toma valores de A a Z; e o grupo rotacional (V063) que é a divisão dos setores selecionados que toma valores de 1 a 8. Ao juntar as variáveis descritas se cria uma chave ou identificador do domicílio denominado de (VDom). Por exemplo, $V\text{Dom} = 262600133000001L7$, isto é, $(V035 = 26 + V040 = 26001330 + V050 = 00001 + V060 = L + V063 = 7)$.

A partir da identificação do domicílio, se cria um código único de identificação dos residentes do domicílio, já que a PME não reporta um código que identifique corretamente a mesma pessoa em dois períodos distintos. Para isso, os indivíduos precisam ter: igual identificador de domicílio (VDom), mesmo sexo (V203), mesmo dia de nascimento (V204), mesmo mês de nascimento (V214), mesmo ano de nascimento (V224) e mesmo número de ordem (V201). Este é o primeiro critério para o emparelhamento dos indivíduos. Ao juntar as variáveis descritas se cria uma chave ou identificador da pessoa denominado de (VPer). Por exemplo, $V\text{Per} = 262600133000001L721801194101$, isto é, $(V\text{Dom} = 262600133000001L7 + V203 = 2 + V204 = 18 + V214 = 01 + V224 = 1941 + V201 = 01)$. Esta pessoa identificada no domicílio 262600133000001L7 é uma mulher (V203), que nasceu no dia 18 (V204) do mês de janeiro (V214) do ano de 1941 (V224).

Os indivíduos que não foram emparelhados pelo primeiro critério, passam por um segundo critério e devem satisfazer as seguintes condições: mesma chave domiciliar; mesmo sexo; mesmo dia; mês e ano de nascimento.

O terceiro critério é aplicado caso o indivíduo não tenha sido emparelhado nos dois primeiros critérios, as condições correspondem: ao mesmo identificador domiciliar; mesmo número de ordem; mesmo dia e mês de nascimento.

Ribas e Soares (2009) constataram que pessoas mais velhas têm mais dificuldade em recordar corretamente a sua idade, portanto, os próximos critérios estipulam um limite máximo para a diferença da idade informada pelo indivíduo em dois períodos distintos. A função para impor esse limite é dada por $\exp(\text{idade}/30)$ para indivíduo de idade maior que vinte e cinco anos, essa função nos diz que, conforme o indivíduo seja mais velho, se aceita uma diferença maior da idade informada pelo indivíduo entre dois períodos.

O quarto critério é destinado somente a chefes, cônjuges e filhos com vinte cinco anos ou mais de idade, dado que os critérios anteriores não foram satisfeitos, assim, as condições correspondem: mesmo identificador domiciliar; mesmo sexo; até quatro dias de diferença no dia do nascimento; até dois meses de diferença no mês de nascimento; se a idade for menor de vinte cinco anos, a diferença na idade presumida até dois anos, caso contrário é de $\exp(\text{idade}/30)$.

O quinto critério também é designado somente a chefes, cônjuges e filhos com vinte e cinco anos ou mais de idade e se o dia de nascimento for não observado. As seguintes condições devem ser satisfeitas: mesmo identificador domiciliar, mesmo sexo, até dois meses de diferença no mês de nascimento; se a idade for menor de vinte e cinco anos, a diferença na idade presumida até dois anos, caso contrário é de $\exp(\text{idade}/30)$; até um ciclo de diferença no nível de escolaridade.

O sexto critério também é destinado somente a chefes, cônjuges e filhos com vinte e cinco anos ou mais de idade, sendo que o mês de nascimento do entrevistado não foi observado. Deve satisfazer as seguintes condições: mesmo identificador domiciliar; mesmo sexo; até quatro dias de diferença no dia de nascimento; se a idade for menor de vinte e cinco anos, a diferença na idade presumida até dois anos, caso contrário é de $\exp(\text{idade}/30)$; até um ciclo de diferença no nível de escolaridade.

O sétimo critério também é designado a chefes, cônjuges e filhos com vinte e cinco anos ou mais de idade, sendo que o dia e o mês de nascimento não foram observados. Assim, devem-se satisfazer as seguintes condições: mesmo identificador domiciliar; mesmo sexo; se a idade presumida for menor de vinte e cinco anos de idade, a diferença na idade presumida até dois anos, caso contrário é de $\exp(\text{idade}/30)$; até um ciclo de diferença no nível de escolaridade.

Os programas em linguagem R utilizados para a realização do procedimento de emparelhamento dos indivíduos foram desenvolvidos por Rojas (2014) e cedidos aos autores da presente aplicação.

Procedimento de ajuste dos modelos com os dados da PME

A primeira etapa consiste na leitura dos microdados da PME. Conforme já indicado na descrição da base de dados, foram considerados os microdados da PME para os meses de janeiro do ano de 2008 até julho do ano 2013, considerando a amostra referente à Região Metropolitana de Salvador. Assim, foram incluídas na base de dados em estudo as informações de todos os indivíduos residentes na Região Metropolitana de Salvador que foram entrevistados pela PME, com idade igual ou superior a 14 anos. Os painéis considerados no estudo foram os que iniciaram pela primeira vez na pesquisa em cada um dos meses do período considerado.

Após a leitura dos microdados de cada um dos meses, realizou-se o emparelhamento dos indivíduos de acordo com os critérios descritos na seção anterior. Assim, tem-se a base de dados com todas as informações dos entrevistados na pesquisa ao longo do tempo. Os indivíduos podem ter desde uma medição a no máximo oito. Os indivíduos com um único registro não poderiam ser emparelhados durante o processo, contudo as pessoas com mais de uma medição podem ser emparelhadas

Para realizar comparações entre pares de meses se tem os seguintes grupos que caracterizam a variável de interesse: respondentes do Tipo 0, Tipo 1, Tipo 2 e Tipo 3.

Os respondentes (Tipo 0): referem-se aos indivíduos que responderam à pesquisa em ambos os períodos; os respondentes apenas na primeira ocasião são denominados de (Tipo 1) e se referem aos indivíduos que responderam na primeira ocasião, mas não na segunda; os respondentes apenas na segunda ocasião (Tipo 2) se referem aos indivíduos que responderam na segunda ocasião, mas não na primeira; e os não respondentes em ambos os períodos (Tipo 3) se referem aos indivíduos que não responderam à pesquisa em nenhum dos períodos de interesse.

Os procedimentos de leitura, emparelhamento e classificação dos tipos de respostas foram utilizados para realizar os ajustes dos modelos de Cadeia de Markov (modelo A, B, C e D). Os códigos (em linguagem computacional – R) para ajustes dos modelos A, B, C e D, utilizados neste trabalho são de autoria de Rojas (2014).

Análise dos resultados

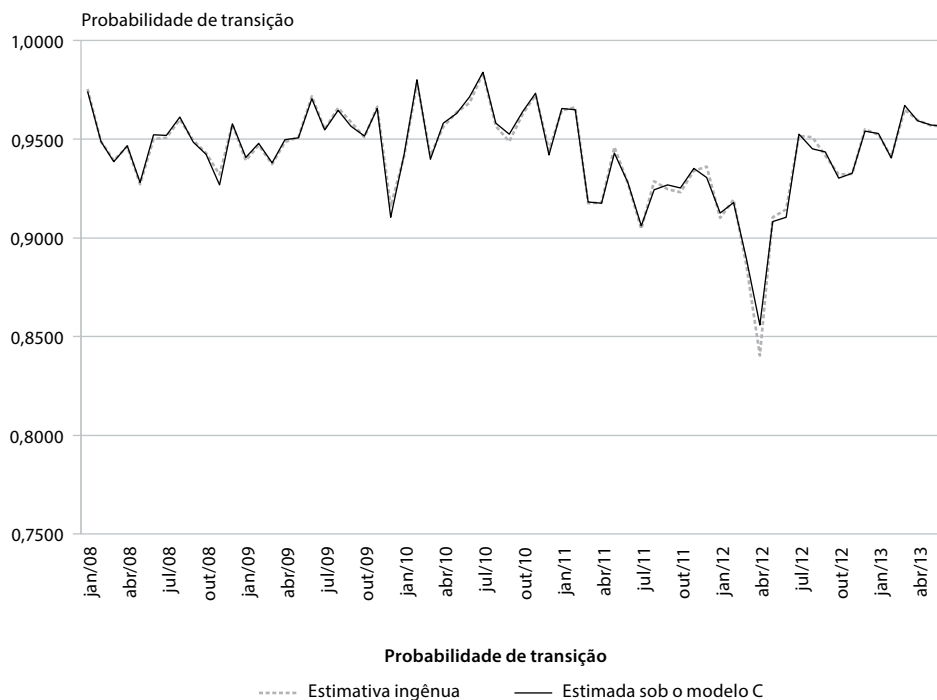
As análises dos fluxos brutos foram realizadas por meio da aplicação dos modelos de Cadeia de Markov conforme descritos na metodologia. Dos quatro modelos (A, B, C e D) ajustados, observou-se, de maneira global pelo teste Qui-quadrado com correção de Rao-Scott, que o modelo C foi o que melhor se ajustou aos dados.

Relembrando, o modelo C considera que a probabilidade inicial de que um indivíduo seja respondente no mês t é a mesma para todas as classificações contempladas na pesquisa, isto é, $\psi(i, j) = \psi$. Contudo, a probabilidade de um indivíduo que teve sua situação ocupacional observada no mês t também ter a situação ocupacional observada no mês $t+1$ depende da situação ocupacional do indivíduo no mês t , isto é, $\rho_{RR}(i, j) = \rho_{RR}(i)$. A probabilidade de um indivíduo que não teve sua situação ocupacional observada no mês $t+1$ também depende da situação ocupacional do indivíduo no mês t , isto é, $\rho_{MM}(i, j) = \rho_{MM}(i)$.

Os gráficos a seguir apresentam as séries das probabilidades de transição dos fluxos brutos entre as categorias da situação ocupacional estimadas pelo modelo C, sempre comparadas com as séries das probabilidades de transição estimadas considerando uma análise ingênua.

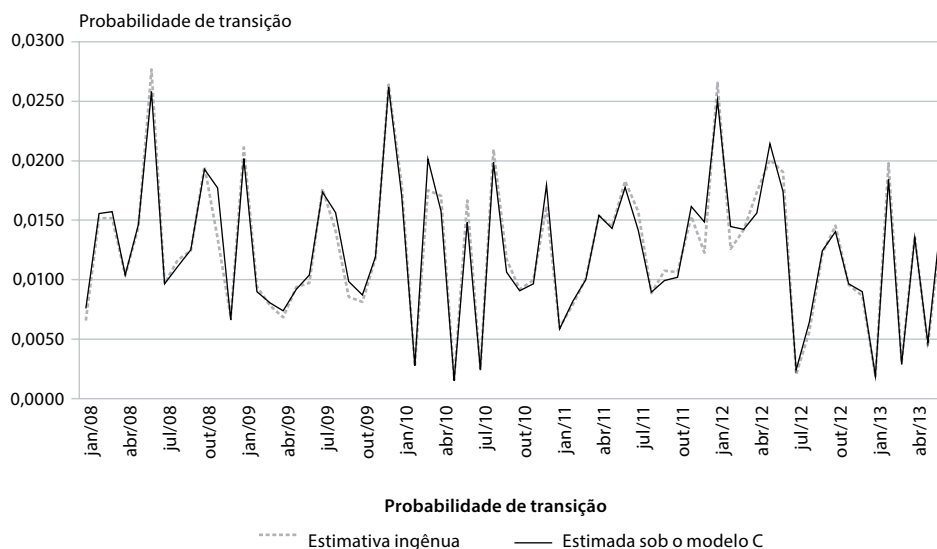
Os Gráficos 1, 2 e 3 revelam as trajetórias das probabilidades de transição de um indivíduo na situação de ocupado no mês atual (t) para ocupado, desocupado e inativo no mês seguinte (t+1), respectivamente.

Gráfico 1 - Probabilidade estimada do indivíduo ocupado no mês atual permanecer ocupado no mês seguinte (p_{11}) - Região Metropolitana de Salvador



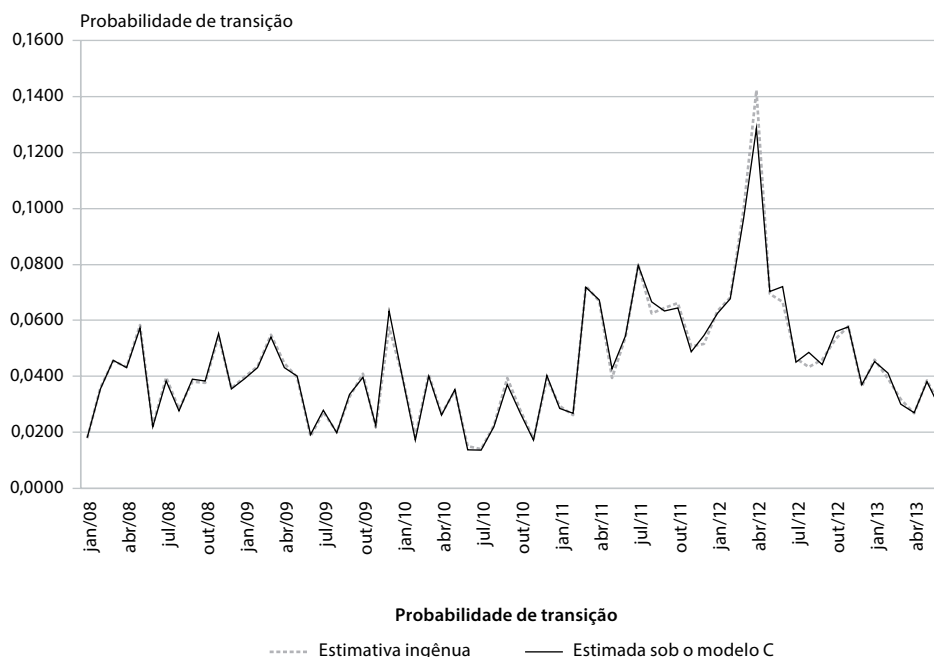
Fonte: IBGE, Microdados da Pesquisa Mensal de Emprego, 2008-2013.

Gráfico 2 - Probabilidade estimada de transição do indivíduo ocupado no mês atual para desocupado no mês seguinte (p_{12}) - Região Metropolitana de Salvador



Fonte: IBGE, Microdados da Pesquisa Mensal de Emprego, 2008-2013.

Gráfico 3 - Probabilidade estimada de transição do indivíduo ocupado no mês atual para inativo no mês seguinte (p_{13}) -Região Metropolitana de Salvador



Fonte: IBGE, Microdados da Pesquisa Mensal de Emprego, 2008-2013.

Pelo Gráfico 1, observa-se que de janeiro de 2008 a abril de 2010 as séries das probabilidades de permanecer ocupado estimadas com o modelo C e com a análise ingênua não apresentaram oscilações abruptas, exceto pelos meses de novembro de 2009 a fevereiro de 2010. Dos meses de abril de 2010 a abril de 2012, a probabilidade de um indivíduo se manter ocupado apontou uma tendência de decrescimento, atingindo o mínimo em abril de 2012 e, logo em seguida, retomando o patamar histórico a partir de julho de 2012. Destaca-se que nestas séries de probabilidade de transição não se notou volatilidade acentuada.

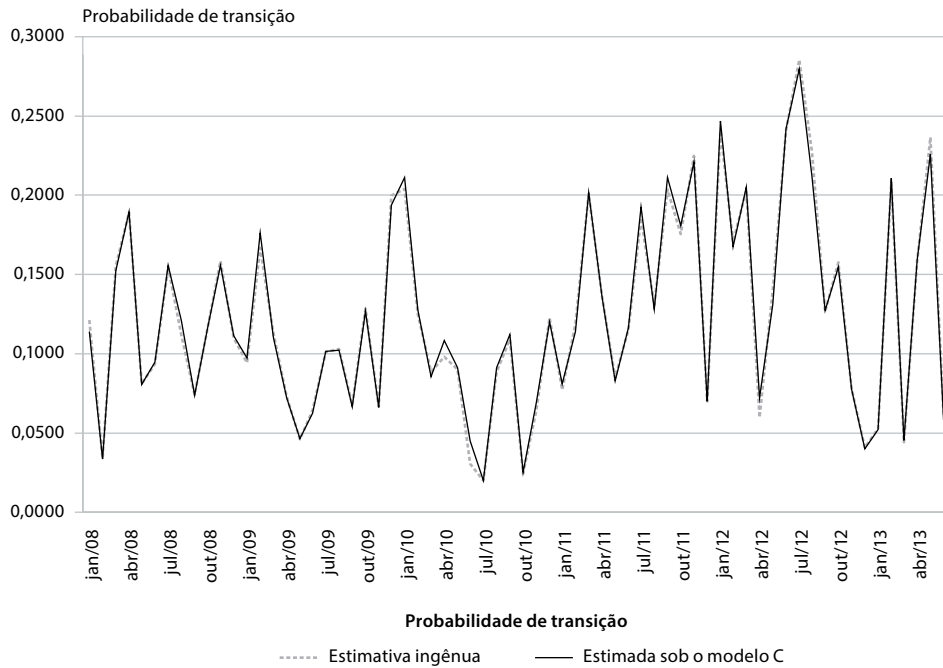
Pelo Gráfico 2, não se nota tendência nítida na trajetória das séries das probabilidades de transição de ocupado a desocupado estimadas pelo modelo C e com uma análise ingênua (sem modelo). Contudo, esta série apresenta uma volatilidade expressiva ao longo do tempo, indicando que a estimativa da probabilidade de um indivíduo ocupado no mês atual passar a desocupado no mês seguinte varia muito de um mês para o outro.

Pelo Gráfico 3, do mês de janeiro de 2008 a janeiro de 2011 as séries das probabilidades de transição de ocupado a inativo estimadas pelo modelo C e com uma análise ingênua apresentaram um comportamento basicamente estável. Do mês de janeiro de 2011 a abril de 2012 se nota uma tendência de crescimento, destacando-se o mês de abril de 2012, com a maior probabilidade de transição de ocupado a inativo observada na série. Em seguida, a probabilidade de transição apresenta uma tendência de decrescimento, retomando ao patamar histórico a partir de julho de 2012.

Um aspecto notável também deriva da comparação das estimativas de probabilidades de transição de ocupado a desocupado (Gráfico 2) com as probabilidades de transição de ocupado a inativo (Gráfico 3). Verifica-se que as últimas são maiores que as primeiras (isto ocorreu em 64 dos 66 meses do período estudado), sendo a mediana das razões destas probabilidades estimadas igual a 3,5, indicando que as transições ao sair da situação de ocupado para inatividade são cerca de 3,5 vezes mais frequentes que saídas para a desocupação. Esta constatação para o período recente na Região Metropolitana de Salvador merece destaque já que a atenção principal da literatura se volta para transições de ocupado a desocupado, sem dar tanta atenção às transições de ocupado a inativo.

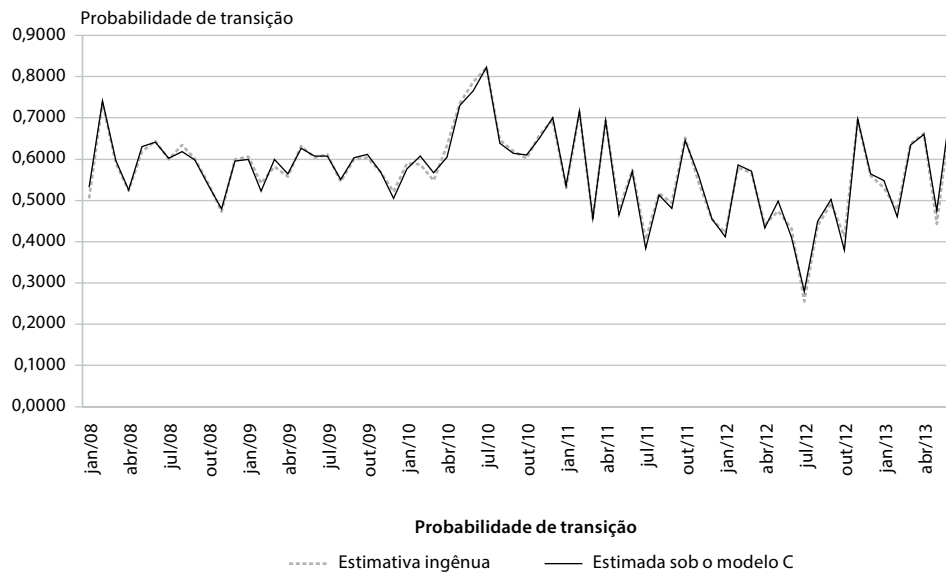
Os Gráficos 4, 5 e 6 revelam as trajetórias das probabilidades de transição de um indivíduo na situação de desocupado no mês atual (t) para ocupado, desocupado e inativo no mês seguinte (t+1), respectivamente.

Gráfico 4 - Probabilidade estimada de transição do indivíduo desocupado no mês atual para ocupado no mês seguinte (p_{21}) -Região Metropolitana de Salvador



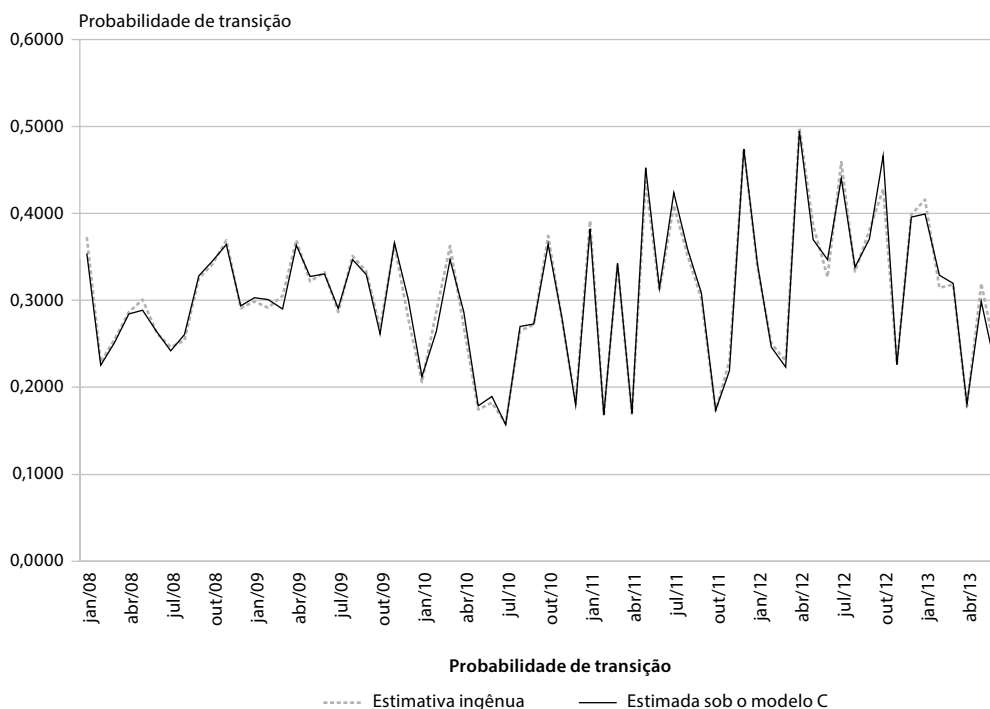
Fonte: IBGE, Microdados da Pesquisa Mensal de Emprego, 2008-2013.

Gráfico 5 - Probabilidade estimada do indivíduo desocupado no mês atual permanecer desocupado no mês seguinte (p_{22}) -Região Metropolitana de Salvador



Fonte: IBGE, Microdados da Pesquisa Mensal de Emprego, 2008-2013.

Gráfico 6 - Probabilidade estimada de transição do indivíduo desocupado no mês atual para inativo no mês seguinte (p_{23}) -Região Metropolitana de Salvador



Fonte: IBGE, Microdados da Pesquisa Mensal de Emprego, 2008-2013.

Pelo Gráfico 4, as séries das estimativas das probabilidades de transição de desocupado a ocupado estimadas pelo modelo C e através da análise ingênua (sem modelo) apresentam volatilidade expressiva, isto é, oscilam muito ao longo do tempo.

Do mês de janeiro de 2008 a outubro de 2010 as séries apresentaram um comportamento basicamente estável, apesar da volatilidade. Do mês de outubro de 2010 a julho de 2012 a série revela uma tendência de crescimento, embora apresente volatilidade. Além disso, de julho de 2012 a dezembro de 2012 a série apresenta uma tendência de decréscimo.

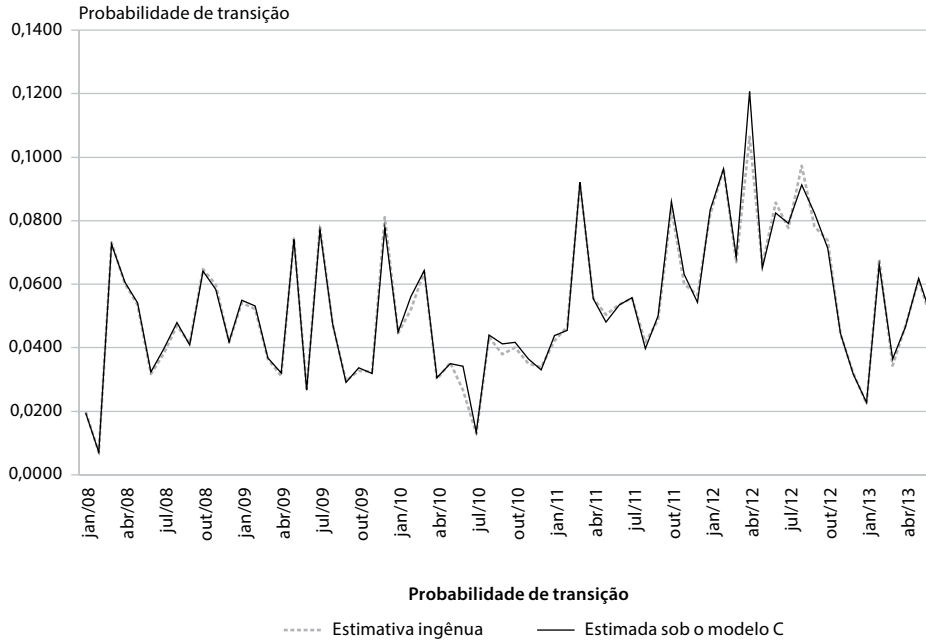
Pelo Gráfico 5 considerando as séries das estimativas das probabilidades de continuar desocupado estimadas pelo modelo C e através da análise ingênua (sem modelo), do mês de janeiro de 2008 a março de 2010 são basicamente estáveis. Contudo, do mês de março de 2010 a julho de 2010 houve um crescimento na probabilidade do indivíduo se manter desocupado. A partir de julho de 2010 a julho de 2012 nota-se uma tendência de decréscimo da probabilidade de manter-se desocupado, apesar de ter ocorrido um aumento na volatilidade. De julho de 2012 a julho de 2013 a série apresenta uma tendência de crescimento da probabilidade de um indivíduo se manter desocupado.

Pelo Gráfico 6, a série das estimativas das probabilidades de transição de desocupado a inativo estimadas sob o modelo C e a série ingênua (sem modelo), de janeiro de 2008 a julho de 2013 apresenta uma tendência estável, apesar da grande volatilidade.

Outro aspecto notável também deriva da comparação das estimativas de probabilidades de transição de desocupado a ocupado (Gráfico 4) com as probabilidades de transição de desocupado a inativo (Gráfico 6). Verifica-se que as últimas são maiores que as primeiras (isto ocorreu em 64 dos 66 meses do período estudado), sendo a mediana das razões destas probabilidades estimadas igual a 2,7, indicando que as transições ao sair da situação de desocupado para inatividade são cerca de 2,7 vezes mais frequentes que saídas para a ocupação. Esta constatação para o período recente na Região Metropolitana de Salvador também merece destaque, por revelar aspecto pouco discutido da dinâmica da situação ocupacional.

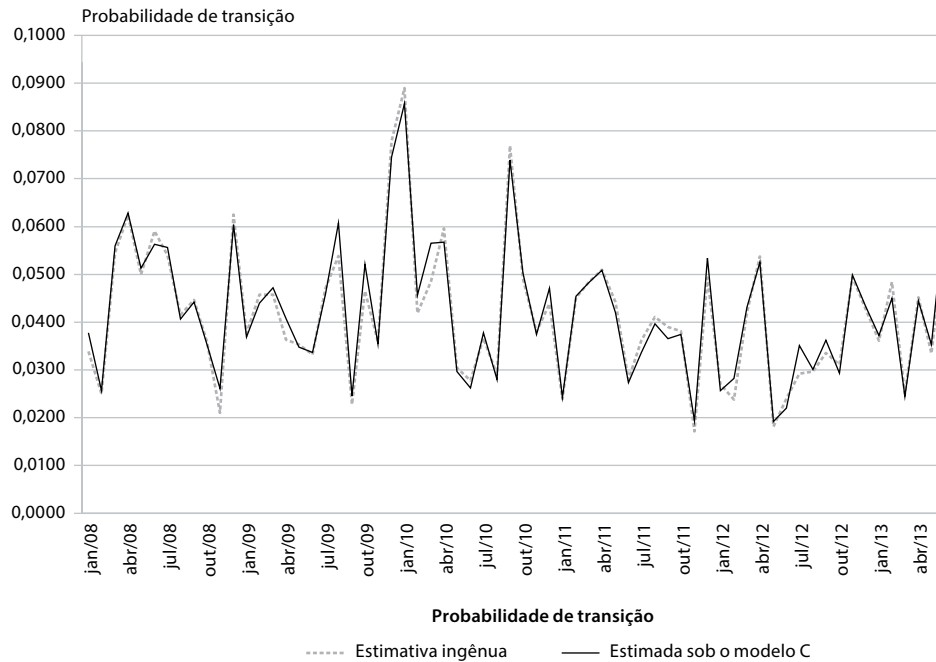
Os Gráficos 7, 8 e 9 revelam a trajetória das estimativas das probabilidades de transição de um indivíduo inativo no mês atual (t) para ocupado, desocupado e inativo no mês seguinte (t+1), respectivamente.

Gráfico 7 - Probabilidade estimada de transição do indivíduo inativo no mês atual para ocupado no mês seguinte (p_{31}) - Região Metropolitana de Salvador



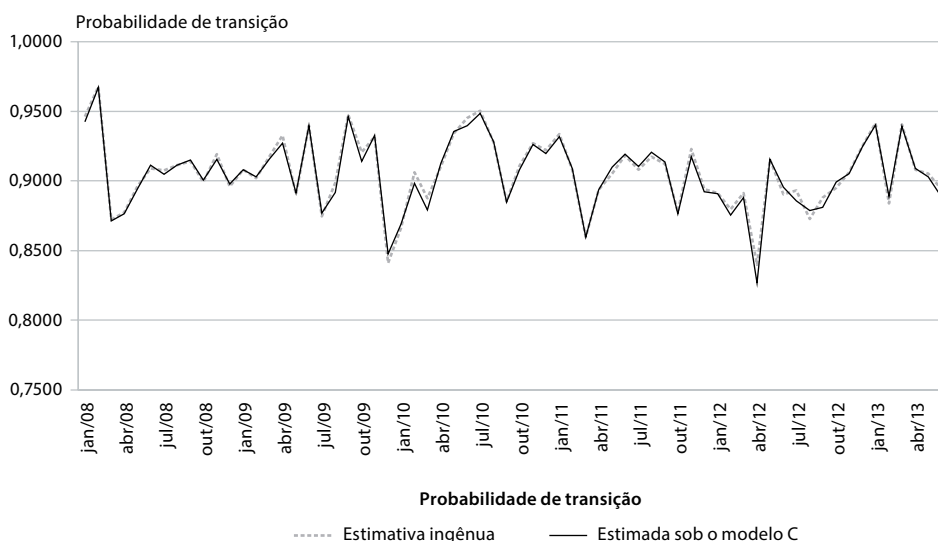
Fonte: IBGE, Microdados da Pesquisa Mensal de Emprego, 2008-2013.

Gráfico 8 - Probabilidades estimadas de transição do indivíduo na situação de inativo no mês atual para desocupado no mês seguinte (p_{32}) - Região Metropolitana de Salvador



Fonte: IBGE, Microdados da Pesquisa Mensal de Emprego, 2008-2013.

Gráfico 9 - Probabilidade estimada de transição do indivíduo inativo no mês atual permanecer inativo no mês seguinte (p_{33}) - Região Metropolitana de Salvador



Fonte: IBGE, Microdados da Pesquisa Mensal de Emprego, 2008-2013.

Pelo Gráfico 7, ao observar as séries das estimativas das probabilidades de transição de inativo a ocupado estimadas pelo modelo C e com uma análise ingênua (sem modelo), de maneira geral, as trajetórias das séries são bem parecidas. Do mês de fevereiro de 2008 a março de 2008 ocorreu um crescimento na probabilidade de transição de inativo para ocupado. De março de 2008 a julho de 2010 as séries são basicamente estáveis, apesar da volatilidade. De julho de 2010 a julho de 2012 as séries apresentaram uma tendência de crescimento e de julho de 2012 a janeiro de 2013 uma tendência de decréscimo e, logo em seguida, retomam o patamar histórico.

Pelo Gráfico 8, não se notam tendências nítidas na trajetória das séries das probabilidades de transição de inativo a desocupado estimada pelo modelo C e com uma análise ingênua (sem modelo). Contudo, estas séries apresentam volatilidade, significando que as estimativas da probabilidade de um indivíduo inativo no mês atual passar a estar desocupado no mês seguinte variam muito ao longo do tempo.

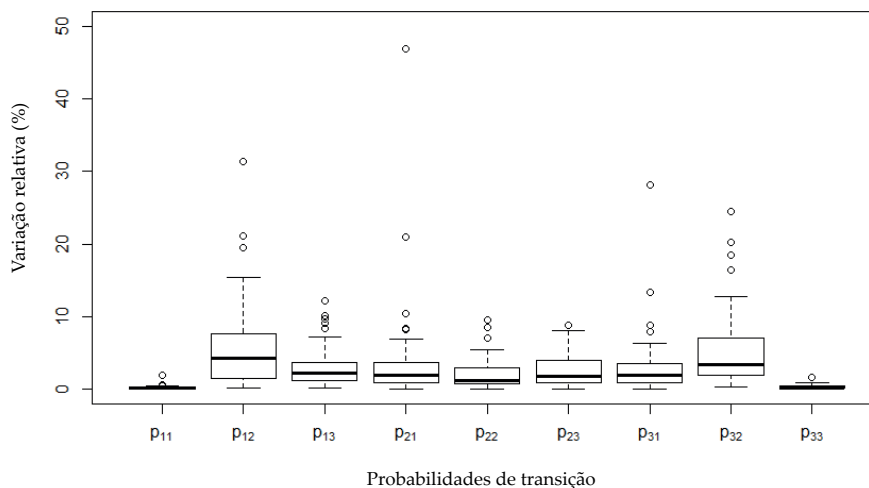
Pelo Gráfico 9, não se notam tendências nítidas na trajetória da série das estimativas das probabilidades de continuar inativo estimadas pelo modelo C e com uma análise ingênua (sem modelo), indicando que a probabilidade de se manter inativo é basicamente estável ao longo do período analisado. No entanto, essas séries das probabilidades de transição apresentam volatilidade, significando que as estimativas da probabilidade de um indivíduo inativo continuar inativo no mês posterior oscilam muito ao longo do período.

Comparação das estimativas das probabilidades de transição obtidas pelo modelo C e pelo método ingênuo

A observação das séries das estimativas das probabilidades de transição sugere que as diferenças nas estimativas obtidas com o modelo C e o método ingênuo são pequenas. O Gráfico 10 apresenta as distribuições das diferenças relativas em valor absoluto entre as estimativas das probabilidades obtidas sob os dois métodos.

O Gráfico 10 mostra que as diferenças entre as estimativas das probabilidades de transição pelos dois métodos são geralmente pequenas, inferiores a 10% (em valor absoluto).

Gráfico 10 - Boxplots das diferenças relativas das probabilidades de transição estimadas pelo modelo C e o método ingênuo - Região Metropolitana de Salvador



Fonte: IBGE, Microdados da Pesquisa Mensal de Emprego, 2008-2013.

As probabilidades p_{11} e p_{33} apresentam as menores variações, indicando que as estimativas do modelo C e as do método ingênuo são bem parecidas. Para as probabilidades p_{13} , p_{21} , p_{22} , p_{23} e p_{31} as variações relativas têm comportamento similar, com as estimativas de ambos os modelos próximas. No entanto, para as probabilidades p_{12} e p_{32} nota-se as maiores variações relativas, indicando que as estimativas são mais distintas.

Também se observa que para todas as probabilidades de transição há valores atípicos, ou seja, para alguns meses específicos os valores estimados pelo modelo C e pelo método ingênuo são mais distantes. Contudo, esta diferença nunca é superior a 50%. Em síntese, o Gráfico 2 fornece evidências de que as estimativas do modelo C e do método ingênuo apresentam um comportamento similar para a maioria das probabilidades de transição estimadas durante o período analisado, mas que não podem ser consideradas idênticas.

Considerações finais

Ainda são escassos no Brasil os estudos que analisam os fluxos brutos das situações ocupacionais, embora diversos países realizem tais análises de forma rotineira. Comumente, as informações disponíveis mais utilizadas são as taxas de ocupação e desocupação. Estas, por sua vez, não informam com detalhes o contingente populacional que transitou entre as diferentes situações ocupacionais ou que permaneceram nas mesmas situações ocupacionais em períodos consecutivos.

As análises dos fluxos brutos objetivam, por exemplo, estimar o número de pessoas que no período atual estão trabalhando e que ainda permanecem trabalhando no período seguinte. Outro objetivo possível é estimar o número de pessoas que estão desocupadas atualmente e que se mantém desocupadas no período seguinte. Outra situação de interesse consiste em estimar o número de pessoas que atualmente estão ocupadas e que no período seguinte passaram a estar desocupadas, e assim por diante. Com estas análises dos fluxos brutos é possível observar a trajetória dos indivíduos que transitaram entre as situações ocupacionais, possibilitando averiguar suas tendências no decorrer do tempo. Além disso, a importância de realizar análises dos fluxos brutos entre as situações ocupacionais é que tais análises proporcionam melhor compreensão da dinâmica do mercado de trabalho ao longo do tempo.

Neste sentido, o objetivo principal deste trabalho foi estimar e analisar os fluxos brutos das situações ocupacionais (ocupado, desocupado e inativo) da população em idade ativa da Região

Metropolitana de Salvador, considerando a modelagem da não resposta e, também, o delineamento complexo usado para coletar as informações contidas na amostra no processo de inferência.

O estudo aqui desenvolvido modelou os fluxos brutos utilizando os dados da amostra da PME na Região Metropolitana de Salvador, a partir dos modelos de cadeia de Markov em duas etapas desenvolvidos por Rojas (2014) e Gutiérrez, Trujillo e Silva (2014). Cada um dos modelos reduzidos (A, B, C e D) tem diferentes interpretações para a resposta (ou não resposta). O ajuste dos modelos foi avaliado por meio do teste Qui-quadrado com correção de Rao-Scott. Além disso, para fins de comparação, os fluxos brutos também foram estimados sem considerar os modelos para o mecanismo de resposta.

Os resultados encontrados mostram que o modelo C obteve o melhor ajuste, pelo teste Qui-quadrado com correção de Rao-Scott ao estimar os fluxos brutos das situações ocupacionais para o caso específico da Região Metropolitana de Salvador, embora não se deva descartar os demais modelos pelo teste estatístico. O modelo C é caracterizado por considerar que a probabilidade de um indivíduo ter sua situação ocupacional observada ou não no período de medição subsequente, dado que foi observado no tempo t depende da situação ocupacional observada no tempo t .

Ao realizar uma análise comparativa entre as trajetórias das séries das probabilidades de transição obtidas por meio do modelo C e do método ingênuo, de maneira geral, se observa que as estimativas são próximas, evidenciando que as diferenças entre as estimativas do modelo C que trata a não resposta e do método ingênuo são pequenas para a Região Metropolitana de Salvador. Apesar disso, como o modelo C se ajustou bem aos dados, deve ser dada preferência às estimativas obtidas sob este modelo, já que incorpora o tratamento dos efeitos da não resposta diferencial detectada na pesquisa.

Ao observar a variação relativa das probabilidades de transição entre as situações ocupacionais, nota-se que as diferenças nas estimativas do modelo C com o método ingênuo são maiores quando os indivíduos são ocupados no tempo t e passam para desocupados no tempo $t+1$ ou de inativos no tempo t e passam para a situação de desocupado no tempo $t+1$. As diferenças encontradas nas estimativas do modelo C com o método ingênuo são menores quando o indivíduo se mantém na mesma situação ocupacional entre dois períodos de medição.

Pelos resultados observados, nota-se claramente que a modelagem aplicada, ou seja, o modelo Markoviano em duas etapas apresenta grande contribuição na modelagem dos fluxos brutos, pois leva em consideração o tratamento da resposta (ou não resposta) e, também, o plano amostral complexo em suas estimativas dos fluxos brutos nos períodos considerados. No caso da aplicação do método ingênuo esses pressupostos são desconsiderados na estimativa dos fluxos brutos.

Por último, a estimação das probabilidades de transição entre as categorias consideradas de situação ocupacional revelou que para a Região Metropolitana de Salvador e no período observado, as transições de ocupado e desocupado para a situação inativo são mais importantes do que as transições ocupado-desocupado e desocupado-ocupado. Tal constatação sugere que seria importante estudar com mais atenção tais transições e talvez conhecer melhor os determinantes das transições, para análises mais aprofundadas da dinâmica do mercado de trabalho ao longo do tempo.

Referências

- BASSI, F., TORELLI, N.; TRIVELLATO, U. Data and Modelling Strategies in Estimating Labour Force Gross Flows Affected by Classification Errors. *Survey Methodology*. Canada, v. 24, n. 2, p. 109-122, 1998.
- DUNCAN, G. J.; KALTON, G. Issues of Design and Analysis of Surveys Across Time. *International Statistical Review*, v. 55, n. 1, p. 97-117, 1987.
- GUTIÉRREZ, A.; TRUJILLO, L.; SILVA, P. L. N. The estimation of Gross flows in complex surveys with random nonresponse. *Survey Methodology*. Canada, v. 40, n. 2, p. 285-321, 2014.
- HUGHES, A. L.; PEITZMEIER, F. K. *Weighting and Imputation Methods for Nonresponse in CPS Gross Flows Estimation*. Bureau of Labor Statistics, Washington, p. 279-285, 1989. Disponível em: http://www.amstat.org/sections/SRMS/Proceedings/papers/1989_047.pdf. Acesso em: jun. 2014.
- ILG, R. Analyzing CPS data using Gross flows. *Monthly Labor Review*, 2005. Disponível em: <http://www.bls.gov/opub/mlr/2005/09/art2full.pdf>. Acesso em: jun. 2014.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE. *Pesquisa Mensal de Emprego. Série Relatórios Metodológicos*, v. 23, Rio de Janeiro, 2ª Ed, 2007. Disponível em: ftp://ftp.ibge.gov.br/Trabalho_e_Rendimento/Pesquisa_Mensal_de_Emprego/Metodologia_da_Pesquisa/srmpme_2ed.pdf. Acesso em: jan. 2014.
- LAPA, P. P. A. *Estimando fluxos das situações ocupacionais com pesquisas amostrais repetidas*. Dissertação (Mestrado). Escola Nacional de Ciências Estatísticas, 2015.
- RAO, J. N. K.; SCOTT, A. J. The analysis of categorical data from complex sample surveys: chi-squared tests for goodness-of-fit and independence in two way tables. *Journal of the American Statistical Association*, v. 76, p. 221-230, 1981.
- RIBAS, R. P.; SOARES, S. S. D. Sobre o Painel da Pesquisa Mensal de Emprego PME do IBGE: problemas e soluções para o emparelhamento usando microdados. *Revista Brasileira de Estatística*, Rio de Janeiro, v. 70, n. 233, p. 75-108, 2009.
- RIBAS, R. P.; SOARES, S. S. D. Sobre o Painel da Pesquisa Mensal de Emprego (PME) do IBGE. *Texto para discussão n° 1348*. IPEA, 2008. Rio de Janeiro, agosto 2008. Disponível em: http://www.ipea.gov.br/agencia/images/stories/PDFs/TDs/td_1348.pdf. Acesso em: jan. 2014.
- ROJAS, H. A. G. *Modelos para estimar cambios brutos en encuestas rotativas con ausencia de respuesta en diseños de muestreo complejos*. Tese (Doutorado). Universidad Nacional de Colombia, 2014. Disponível em: <http://www.bdigital.unal.edu.co/39564/1/hugoandresgutierrezrojas.pdf>. Acesso em: fev. 2015.
- SIMPKINS, S. *Labour Market Flows*. Office for National Statistics, agosto 2013. Disponível em: http://www.ons.gov.uk/ons/dcp171766_322921.pdf. Acesso em: jun. 2014.

Proposta de compatibilização de dados censitários para estudos longitudinais: o caso das áreas de ponderação do município do Rio de Janeiro nos Censos 2000 e 2010

Vivian Alves da Costa Rangel Gomes*

Resumo

É de grande relevância entre os temas da investigação sociodemográfica a abordagem de questões referentes às condições de vida da população, sua distribuição, seus deslocamentos, sua evolução, entre diversos outros temas cuja magnitude pode ser apurada de forma a compreender todos esses fenômenos segundo suas dinâmicas e diferenciações socioespaciais. Este trabalho trata de uma proposta que permite a compatibilização de dados espaciais para análises longitudinais. É indicado para a elaboração de uma análise sociodemográfica e espacial a partir dos geodados (dados censitários segundo áreas pré-estabelecidas) disponibilizados pelo IBGE (Instituto Brasileiro de Geografia e Estatística), portanto indispensável no âmbito da produção e utilização das Estatísticas Públicas. Essa proposta metodológica foi testada para o município do Rio de Janeiro segundo áreas de ponderação dos Censos 2000 e 2010.

Palavras-chave: Dados espaciais. Análise Longitudinal. Dados Censitários. Geoestatística.

* Doutoranda em Demografia pelo Instituto de Filosofia e Ciências Humanas da Universidade Estadual de Campinas (IFCH/UNICAMP) e Mestre em Estudos Populacionais e Pesquisa Social pela Escola Nacional de Ciências Estatística (ENCE) e pelo Instituto Brasileiro de Geografia e Estatística (IBGE).

Introdução

É de grande relevância entre os temas da investigação sociodemográfica a abordagem de questões referentes às condições de vida da população, sua distribuição, seus deslocamentos, sua evolução, entre diversos outros temas cuja magnitude pode ser apurada de forma a compreender todos esses fenômenos segundo suas dinâmicas e diferenciações socioespaciais. Constitui-se um grande desafio relacionar muitas variáveis na busca de explicar causas e efeitos, sobretudo quando essas apresentam influência direta na configuração espacial recente.

Este trabalho trata de uma proposta que permite a compatibilização de dados espaciais para análises longitudinais. É indicado para a elaboração de uma análise sociodemográfica e espacial a partir dos geodados (dados censitários segundo áreas pré-estabelecidas) disponibilizados pelo IBGE (Instituto Brasileiro de Geografia e Estatística), portanto indispensável no âmbito da produção e utilização das Estatísticas Públicas.

Isto é importante porque estudos dessa natureza nos permitem explorar como vantagens ou desvantagens socioeconômicas se acumulam no espaço, por exemplo. A análise espacial já se confirma como ferramenta metodológica de grande amplitude no que tange sua contribuição potencial para evidenciar processos socioterritoriais.

Esta proposta de compatibilização permite que novos enfoques sejam incorporados à análise no amplo campo das Ciências Sociais, pois ao investigar diversos fenômenos é necessário conhecer, pelo menos, “onde” e “como” se encontram distribuídos. Nessa perspectiva, a Geografia vem compartilhar sua expertise sobre o “pensar espacial”.

No caso dos estudos demográficos, esse diálogo é propício e adequado, os enfoques espaciais se direcionam para indicar o desenvolvimento da população no espaço e no tempo, o que enriquece os estudos que privilegiam essa abordagem. Melhores explicações dos fenômenos são proporcionadas e maiores são as contribuições para a elaboração de políticas públicas em função da associação de um contexto espacial aos dados empregados nas pesquisas.

Este texto conta com outras cinco seções principais e as considerações finais. Na primeira abordamos a informação espacial e as variáveis regionalizadas; na segunda, as áreas de ponderação como unidades mínimas de análise; na terceira, uma proposta metodológica para a compatibilização das áreas de ponderação dos Censos Demográficos 2000 e 2010; na quarta, a composição da malha vetorial para áreas de ponderação do Censo 2000; e, na quinta, reconhecemos o motivo fundamental da compatibilização.

Informação espacial e variáveis regionalizadas

Para possibilitar a elaboração de uma análise espacial com a espacialização da ocorrência dos eventos registrados por área, no intuito de identificar e compreender possíveis padrões de distribuição de ocorrências e a sua relação com as diferentes variáveis selecionadas é necessário a utilização de uma ferramenta que permita trabalhar os dados censitários e os dados espaciais.

Para isso, contamos com a utilização do Sistema de Informações Geográficas (SIG), por se tratar de uma ferramenta que possibilita a execução da investigação e, neste trabalho, utilizou-se o software livre Quantum GIS (QGIS) versão 2.18. Las Palmas e o GEoDA¹. O contexto se remete à aplicação do SIG no estudo sociodemográfico intraurbano com informações censitárias como um caminho técnico e metodológico viável para ser empregado em diversas pesquisas. O SIG na Demografia auxilia, e muito, o ganho de informação para melhor compreensão da localização populacional em diversas áreas, sobretudo para estudos da distribuição espacial da população em escala intraurbana (UMBELINO; BARBIERI, 2008).

Neste trabalho tratamos a localização residencial equiparada a uma variável regionalizada, pois “qualquer variável dependente do espaço que apresente, além do caráter aleatório, um caráter estrutural, pode ser tratada como variável regionalizada e sofrer uma análise segundo o

¹ GeoDa é um software GIS gratuito de análise espacial desenvolvido pelo Professor Doutor Luc Anselin da Universidade de Chicago, Estados Unidos.

formalismo desenvolvido pela Geoestatística”, (YAMAMOTO; LANDIM; 2013, p. 10). Esse detalhe implica a adoção de técnicas adequadas à exploração de variáveis que apresentam resultados agregados por área. Trata-se de uma peculiaridade da informação espacial que aplicamos em conjunto aos dados socioeconômicos e demográficos.

A área de estudo é o município do Rio de Janeiro (contorno político administrativo) subdividido em unidades territoriais mínimas, as áreas de ponderação criadas pelo IBGE na data de referência de cada Censo Demográfico considerado.

Áreas de ponderação como unidades mínimas de análise

Define-se área de ponderação como sendo uma unidade geográfica, formada por um agrupamento de setores censitários, para a aplicação dos procedimentos de calibração das estimativas com as informações conhecidas para a população como um todo (IBGE, 2010, p.62).

A área de ponderação como unidade mínima de análise permite um nível de desagregação maior do que o dado municipal e possibilita uma investigação mais detalhada, mantendo-se o nível de confiança estatística. Apesar de menos desagregada que os setores censitários, as áreas de ponderação garantem abrangência, confiabilidade e periodicidade como todos os dados demográficos do IBGE.

Esse recorte espacial e operacional possibilita o alcance específico do território de estudo, mas vale ressaltar que essa divisão territorial foi criada pelo IBGE para que os dados da amostra dos Censos Demográficos pudessem ser divulgados obedecendo a critérios estatísticos. Na verdade, a área de ponderação é um coletivo de setores censitários, exclusivos, sem a possibilidade de dupla contagem, criada para calibrar as estimativas populacionais. Estas áreas serão representadas por polígonos na análise espacial.

Outros países também já adotam a criação de áreas operacionais para a agregação de dados como os censos recentes do Reino Unido² com a criação das *Output areas* para dar suporte à coleta de dados. A relevância de explorar esta unidade territorial é poder identificar as características da população residente. No entanto, essas unidades podem ser alteradas entre um Censo e outro, devido a critérios técnicos.

No caso do Brasil, o IBGE disponibiliza a descrição das áreas de ponderação criadas em cada Censo contemplando informações como um código específico de cada uma dessas áreas, o tipo de área (todas as áreas de ponderação do município do Rio de Janeiro são do tipo 70 ou “definida pelo usuário”), número de setores reunidos em cada área e um rótulo que se refere ao bairro de cobertura dessa área de ponderação.

Dependendo do número de setores censitários definidos para cada bairro, esse reunirá um número maior ou menor de áreas de ponderação para cobrir a extensão de seu território, pois segundo critérios técnicos, uma área de ponderação não deveria reunir mais que 50 setores censitários, mas encontramos exceções relacionadas a bairros considerados mais densamente povoados ou com área extensa. Sendo assim encontramos bairros nos Censos analisados que reúnem mais de 7 e até 11 áreas de ponderação, como é o caso de Campo Grande em 2010 e, por outro lado, encontramos também áreas de ponderação que reúnem mais de um bairro, como é o caso de Cosme Velho e Santa Teresa no Censo 2000.

Apesar do território administrativo do município não ter passado por qualquer alteração de seus limites, o IBGE optou, em função da operacionalização técnica, pela redefinição de alguns setores censitários que resultou na criação de mais 2.088 setores implicando a agregação de mais 30 áreas de ponderação no município para o Censo Demográfico 2010. Assim, os dados do ano 2000 para o Rio de Janeiro apresentam 8.145 setores censitários e 170 áreas de ponderação e os dados para o ano de 2010 apresentam 10.233 setores censitários e 200 áreas de ponderação.

² Disponível em <https://www.gov.uk/government/statistics/announcements/2011-area-classification-for-super-output-areas>

Reconhecer as alterações ocorridas no critério técnico referente à amostragem e à expansão da amostra entre os Censos é relevante para que a confiabilidade estatística desses dados seja mantida. Por isso, apresentamos aqui uma breve descrição de algumas alterações identificadas.

Desde o Censo 2000 o IBGE disponibilizou um arquivo chamado “Divisão Territorial Brasileira” onde é possível encontrar as informações como o código das Unidades da Federação (UF), das mesorregiões, da microrregião, do município, do distrito e do subdistrito. No arquivo “Lista de Áreas de Ponderação – Brasil” encontra-se a descrição das áreas de ponderação do Censo segundo o código atribuído a cada área, o tipo de área, o número de setores que constituem cada área, o número de domicílios particulares ocupados na amostra e o número de pessoas no universo.

No arquivo referente à “Documentação”, o instituto disponibiliza as “Notas Metodológicas” e diversas outras informações quanto à data de referência, a base territorial, o período de coleta, aspectos da amostragem, entre outras. Nesse arquivo as informações sobre a expansão da amostra e a ponderação de unidades da amostra revelam que a metodologia utilizada é aplicada, independentemente, para cada uma das áreas de ponderação.

No Censo 2010, as áreas de ponderação foram criadas segundo os mesmos critérios aplicados ao Censo 2000 abaixo listados:

- Conjugação de tamanho, contiguidade e homogeneidade do conjunto de características populacionais e de infraestrutura.
- O maior nível geográfico é o Município.
- O menor tamanho de uma área de ponderação é de 400 domicílios particulares ocupados na amostra.

O conceito de área de ponderação foi mantido, no entanto, no Censo 2010, 14 municípios tiveram suas áreas redefinidas, uma espécie de ajuste de configuração da representação espacial, e o município do Rio de Janeiro se encontra entre as unidades alteradas.

É importante destacar que para desenvolver um estudo longitudinal que compare os dados agregados por área é necessário que essas áreas sejam compatíveis, ou seja, seu conteúdo, em última análise deve ter referência espacial similar no que tange ao espaço físico de sua cobertura, pois os dados uma vez agregados não podem mais passar por processo de desagregação³ e para serem comparados devem refletir a mesma “realidade espacial” do campo de coleta.

Sabe-se que são os setores censitários as referências mínimas para o desenvolvimento de estudos com os dados censitários, no entanto, além da necessidade de inovar e explorar um pouco mais as áreas de ponderação em estudos acadêmicos, levamos em consideração críticas relevantes quanto a possíveis inconsistências na base de dados dos setores censitários (SOUZA, 2012).

O Instituto Pereira Passos⁴ detectou inconsistências na base dos setores censitários do IBGE, tais como deslocamentos e a existência de áreas socialmente e economicamente heterogêneas em um mesmo setor, que prejudicavam a qualidade da informação gerada (Souza, 2012, p. 2).

O problema geral da base de dados de setores censitários, identificado pelos pesquisadores do Instituto Pereira Passos (IPP), envolvia “inconsistências geográficas” e “incompatibilidade dos desenhos dos setores em relação às descrições de seus perímetros”, além de setores que incluíam em um mesmo perímetro áreas com características socioeconômicas muito distintas criando “imprecisões” nas informações geradas.

Por isso, o IPP propôs uma espécie de revisão da base de dados para refinar sua qualidade para que o território da cidade se tornasse melhor compreendido. No ano de 2008 foi iniciada uma parceria entre o IPP e o IBGE com o objetivo de construção da malha digital do Censo 2010. Os critérios operacionais do IBGE foram mantidos e empregou-se a inovação tecnológica que possibilitou basear o Censo em uma malha digital georreferenciada de setores censitários,

³ Esta impossibilidade de desagregação dos dados foi contestada por uma metodologia de compatibilização de setores censitários baseada em imagens Landsat que afirma ser possível desagregar dados de setores por meio do auxílio dessas imagens de satélite (FEITOSA; MONTEIRO; CÂMARA, 2005).

⁴ O Instituto Pereira Passos (IPP), autarquia da Prefeitura da Cidade do Rio de Janeiro, é o organismo municipal responsável pela produção e consolidação de informações estatísticas, geográficas e cartográficas da cidade, que subsidiam políticas setoriais pelos mais diversos entes da administração pública municipal.

gerando a vantagem de associação direta dos dados coletados aos setores georreferenciados com ganho na escala de detalhes e precisão cartográfica superior aos anos anteriores.

De forma geral, os ajustes consistiram na edição da base original do Censo 2000, para a melhoria da identificação do perímetro de cada setor. Em um segundo estágio dessa revisão, o IPP propôs subdivisões, a criação de novos setores para a coleta do Censo 2010. O pressuposto principal das subdivisões foi “a noção de que os setores deveriam conter em seu recorte características socioeconômicas semelhantes”, e assim se justifica a recriação de setores censitários para o Censo 2010 (SOUZA, 2012, p.6).

Os limites dos bairros também foram corrigidos de forma que coincidisse com os limites adotados pela Prefeitura e aqueles considerados pelo IBGE. Foi muito comum adequações em que “alguns setores que equivocadamente abrangiam áreas de mais de um bairro tiveram que ser subdivididos”. Também foram criados “setores vazios” que se referem a grandes áreas não ocupadas dentro de alguns setores, “no caso das áreas vazias, e isso nos ajuda no estudo mais preciso da ocupação do solo, separando quando possível as áreas ocupadas de áreas não ocupadas”, Souza (2012, p. 7).

Como as áreas de ponderação são agregados de setores censitários, a malha vetorial de 2010 é a mais adequada para ser a nossa base para a elaboração de uma malha para o ano 2000.

Uma proposta metodológica para compatibilização das áreas de ponderação dos Censos Demográficos 2000 e 2010

Para desenvolver um estudo demográfico em área intraurbana a partir de informações censitárias com diferentes datas de referência foi necessário que um exercício de compatibilização fosse desenvolvido para tornar possível a sua realização. As unidades mínimas de análise passaram por uma adequação conforme apresentaremos a seguir.

Estudos anteriores mostraram em sua maioria a compatibilização dos setores censitários dos anos de 1991, 2000 e 2010 como contribuição para análise temporal, orientando a realização da compatibilização por meio da utilização de softwares SIG diversos e até mesmos por meio de imagens de satélite Landsat (FEITOSA;MONTEIRO;CÂMARA,2005). Esta metodologia foi realizada quase que por completo de forma manual a partir do estudo das informações contidas na documentação dos microdados do Censo 2000 e nas notas metodológicas do Censo 2010. Basicamente, relacionamos os principais itens necessários para alcançar áreas compatíveis. Reparamos que um dos primeiros indícios da necessidade de compatibilização era a inexistência da malha de áreas de ponderação para o ano 2000.

A malha vetorial disponível para o município do Rio de Janeiro é a malha de setores censitários. Diante desse fato, uma solução plausível seria então usar o SIG para criar uma malha vetorial de áreas de ponderação do Rio de Janeiro para o ano 2000 baseada na malha de 2010.

A partir do arquivo “composição das áreas de ponderação”, contido na documentação do Censo 2000 foi possível relacionar os 8.145 setores censitários representados na malha vetorial a 170 áreas de ponderação criadas para este Censo e então elaborar a geometria poligonal mais similar possível à malha de áreas de ponderação do Censo 2010.

Empregamos uma metodologia baseada no reagrupamento das áreas de ponderação existentes, visto a verificação de que não houve mudanças na subordinação do território do município, ou seja, os limites políticos administrativos foram mantidos, reforçamos a hipótese de que as áreas alteradas se adequaram a um caráter operacional da calibração amostral.

Sendo assim, as áreas em ambas as datas deveriam então coincidir, mantendo seus limites e sua subordinação compatíveis possibilitando a comparação dos dados, ou seja, as áreas devem apresentar contornos semelhantes e estarem subordinadas ao mesmo bairro em ambas as datas ou reunir o mesmo conjunto de bairros em seu interior.

Para certificar de que esse requisito estava sendo atendido, partimos para o arquivo de descrição dos setores e das áreas de ponderação. Comparamos o agrupamento de setores da

malha do Censo 2000 com a descrição das áreas de ponderação da malha do censo 2010. Os principais elementos identificados como necessários na compatibilização foram:

- Reagregar ou reagrupar áreas contíguas;
- Seguir a composição das áreas de ponderação em ambas as datas;
- Dar mais enfoque ao conteúdo coletado na área do que a sua geometria.

Consideramos que toda a mudança entre setores e áreas foram mudanças operacionais, isso significa que é possível agregar os setores de forma que as áreas de ponderação formadas sejam semelhantes, tudo isso do ponto de vista do conteúdo, pois o desenho ou redesenho da geometria deve representar a área de cobertura, mas devemos lembrar que se trata apenas de uma representação. Sendo assim, assumimos critérios de compatibilização de conteúdo dessas áreas de ponderação.

Podemos citar, como exemplo, as informações dos arquivos de descrição dos setores censitários e a composição das áreas de ponderação. Apresentamos a compatibilização do conteúdo da área que convencionamos chamar de “Centro Antigo”. Na data de referência do ano 2000, tratava-se da área de ponderação do “Caju” com código 3304557999024 e da área de ponderação que reunia os bairros da “Gamboa, Santo Cristo e Saúde” com código 3304557999141. No ano de 2010, essas duas áreas se tornaram apenas uma com o nome “Caju, Gamboa, Santo Cristo e Saúde” sob o código 3304557005001.

Observamos que segundo a descrição dos setores censitários, a área de cobertura no Censo 2010 era exatamente a mesma para as duas áreas do Censo 2000. Por isso, decidimos agregá-las em 2000 tornando-as uma única área compatível com a área do ano de 2010 e definimos o mesmo rótulo para identificá-las. A diferença em relação aos setores censitários é que no ano 2000, a área do “Caju” era constituída por 21 setores e a área da “Gamboa, Santo Cristo e Saúde” era composta por 42 setores, totalizando em 2000 para a área “Centro Antigo” uma composição de 63 setores censitários. Em 2010 essa mesma área foi composta por 73 setores o que indica que 10 novos setores censitários foram criados, mas foram criados na mesma área física dos bairros, na verdade sua criação ocorre a partir da subdivisão de setores censitários já existentes apenas como forma de facilitar a distribuição de setores entre os agentes recenseadores no momento da operação de coleta de dados, fazendo com que a subordinação dos bairros que estão contidos na área de ponderação seja mantida, possibilitando a compatibilização.

Com essa solução conseguimos adequar o conteúdo das áreas de ponderação para a elaboração de um estudo longitudinal. As alterações necessárias para essa compatibilização de conteúdo estão listadas no quadro a seguir.

Nesse quadro podemos visualizar que diversas áreas de ponderação do ano 2000 foram agregadas formando apenas uma no ano 2010 apesar do número de setores censitários ter aumentado. Também existe o caso, por exemplo, de áreas do Censo 2010 estarem subdivididas, como “Cosme Velho” sendo a área de número 3304557005002 e “Santa Teresa” com o número de identificação 3304557005144, que foram unidas em uma área chamada de “Cosme Velho” para se tornar compatível com o conteúdo de dados que temos para a área do ano 2000 em que “Cosme Velho e Santa Teresa” estão agregadas na área de ponderação de número 3304557999139. As áreas que foram mantidas em dimensão e conteúdo entre os censos receberam uma nomenclatura compatível em ambas às datas e não foram alteradas.

Quadro 1 - Descrição da compatibilização das áreas de ponderação - Município do Rio de Janeiro - 2000/2010

Áreas de ponderação compatibilizadas – áreas únicas	Áreas correspondentes no Censo 2000	Áreas correspondentes no Censo 2010
Centro Antigo	Caju/Gamboa, Santo Cristo e Saúde (2 áreas)	Caju, Gamboa, Santo Cristo e Saúde (1 área)
Cosme Velho	Cosme Velho e Santa Teresa (1 área)	Cosme Velho/Santa Teresa (2 áreas)
Glória	Glória/Catete (2 áreas)	Glória, Catete (1 área)
Guaratiba	Guaratiba, Barra de Guaratiba e Pedra de Guaratiba (1 área)	Guaratiba 1, 2, 3/Pedra de Guaratiba/Barra de Guaratiba (3 áreas)
Gávea	Gávea/Jardim Botânico (2 áreas)	Jardim Botânico e Gávea (1 área)
Estácio	Estácio/Catumbi/Rio Comprido (3 áreas)	Rio Comprido/Estácio, Catumbi e Cidade Nova (2 áreas)
São Francisco Xavier	Jacaré, Rocha, Sampaio/Riachuelo/São Francisco Xavier/ Mangueira (4 áreas)	Jacaré, Rocha, Sampaio, Riachuelo e São Francisco Xavier (1 área)
Encantado	Abolição/ Encantado e Água Santa (2 áreas)	Encantado, Abolição e Água Santa (1 área)
Vicente de Carvalho	Vila Kosmos/Vicente de Carvalho (2 áreas)	Vicente de Carvalho e Vila Kosmos (1 área)
Irajá	Irajá 1, 2 e 3 (4 áreas)	Irajá e Vista alegre/Irajá 1, 2 e 3 (5 áreas)
Vaz Lobo	Turiaçu	Turiaçu, Vaz Lobo (1 área)
Cacuaia	Cacuaia e Ribeira/Zumbi, Pitangueira e Praia da Bandeira (2 áreas)	Pitangueiras, Cacuaia e Praia da Bandeira (1 área)
Bancários	Freguesia/Bancários (2 áreas)	Freguesia e Bancários (1 área)
Costa Barros	Barros Filho/Costa Barros (2 áreas)	Costa Barros e Barros Filho (1 área)
Jacarepaguá	Jacarepaguá	Jacarepaguá 1, 2 e 3/ Rio das Pedras 1 e 2 (5 áreas)
Anil	Gardênia Azul/Anil (2 áreas)	Anil e Gardênia Azul (1 área)
Senador Camará	Senador Camará 1, 2 e 3 (3 áreas)	Fazenda Coqueiro, Coreia e Senador Camará 1 e 2 (4 áreas)
Bangu	Bangu 1, 2, 3, 4, 5 e 6 (6 áreas)	Vila Kennedy/ Bangu 1, 2, 3, 4, 5 e 6 (7 áreas)
Del Castilho	Maria da Graça e Del Castilho/Higienópolis (3 áreas)	Del Castilho, Higienópolis e Maria da Graça (1 área)
Botafogo	Botafogo 1, 2, 3 e 4 (4 áreas)	Botafogo e Urca/Botafogo 1, 2, 3 (4 áreas)
Penha	Penha 1 e 2/ Penha Circular 1 e 2 (4 áreas)	Complexo da Penha e Caricó/Penha/ Penha circular (3 áreas)

Fonte: Elaboração própria a partir dos dados dos Censos 2000 e 2010.

Nota: O símbolo “/” indica áreas diferentes, ou seja, no caso das áreas correspondentes no Censo 2000, por exemplo, o “Centro Antigo” no ano 2000 foi composto por duas áreas de ponderação, conforme já foi explicado, a área do “Caju” foi somada à área da “Gamboa, Santo Cristo e Saúde”. Já no ano de 2010 a área correspondente já incluía “Caju, Gamboa, Santo Cristo e Saúde” em um único polígono.

Este quadro apresenta as principais alterações executadas no processo de compatibilização. As áreas de ponderação que foram identificadas como “compatíveis” nos dois Censos não sofreram alterações.

Um exemplo de como elaboramos as correspondências para unir as áreas de ponderação considerando a composição e a descrição dos setores censitários disponibilizados pelo IBGE é a possibilidade de aproveitar a descrição de cada setor e, a partir de seu ponto inicial e da informação sobre o perímetro que o recenseador deve percorrer para a coleta de dados, identificar os setores que não foram alterados entre os censos.

Após um trabalho extenso com as correspondências tabulares para a compatibilização dos conteúdos das áreas de ponderação, partimos para a constituição de uma malha vetorial para as áreas de ponderação do ano 2000.

Composição da malha vetorial para as áreas de ponderação do Censo 2000

Conforme já mencionada, não há malha vetorial para as áreas de ponderação do Censo 2000 disponibilizada pelo IBGE. No entanto, diante da abrangência e riqueza de detalhes contidos nos documentos sobre os microdados é possível para o pesquisador a constituição de uma malha vetorial a partir dessas informações.

Devemos apreender que a malha vetorial trata-se de uma representação geográfica para os dados censitários que são apresentados de forma agregada para que a privacidade dos respondentes seja respeitada e mantida. O setor censitário enquanto unidade cadastral do Censo e a área de ponderação enquanto elemento de calibração amostral são unidades que podem ser representadas por polígonos fechados. A rigor, como se trata da representação espacial de unidades operacionais, estas não foram elaboradas com o fim de apresentar a distribuição espacial da população, no entanto, sua sistematização permite seu uso para esse fim desde que alguns procedimentos sejam respeitados.

Referimos-nos ao problema da unidade de área modificável. É necessário que o pesquisador seja cauteloso diante dos resultados gerados para uma representação espacial que sofre influência da fronteira das unidades territoriais representadas. Por isso, o pesquisador deve ser manter atento quanto a esse problema (*Modifiable Areal Unit Problem – MAUP*) que está relacionado à necessidade de tomar algumas decisões para a compreensão dos resultados obtidos a partir da regionalização escolhida pelo (PERES DE ÁVILA; MONASTERIO, 2008).

A principal questão que enfatizamos aqui é a interferência desse “problema” na leitura dos resultados em análise. Optamos por empregar o MAUP para considerar o risco existente dos resultados sofrerem influência do grau de agregação escolhido. Como o objeto de análise que propomos são as unidades agregadas, buscamos evitar inferir equivocadamente um comportamento individual em função de uma base de dados agregados. O MAUP ainda pode se desdobrar em dois tipos, sendo um referente aos diferentes níveis de agregação espacial (escala) e, o outro, que decorre do recorte das unidades (partição) (HAINING, 2003, p. 150).

É válida a reflexão sobre a influência desse “problema” na pesquisa demográfica, assim como para a Economia e para Geografia e as Ciências Sociais cujas análises devem considerar essencialmente o MAUP. Ao inserir a concepção do MAUP na Demografia permitimos o estudo das variáveis em uma escala menor, por meio da otimização da função do dado agregado, para evitar que a análise caia no que os pesquisadores consideram como “falácia ecológica”. A falácia ecológica é um erro de análise. Isto ocorre quando análises são elaboradas a partir de dados agregados por área e o pesquisador a considera no nível individual. O problema é que este tipo de inferência pode gerar análises incorretas sobre um fenômeno.

É importante lembrar que a análise espacial será uma das ferramentas muito útil nos estudos recentes que visam conhecer a distribuição da população e as características atribuídas a ocupação do território, sendo assim, as fronteiras das áreas são importantes, sua definição espacial é fundamental e a estatística gera resultados cuja validade depende da unidade da área em análise, pois a agregação de indivíduos tende a reduzir suas flutuações. O MAUP é muito conhecido pelos geógrafos mais ainda pouco estudado pelos Demógrafos. No desenvolvimento de estudos com abordagem empírica é muito relevante fazer essa consideração, pois se busca alcançar objetivos como a identificação de padrões e tendências espaciais nos estudos populacionais.

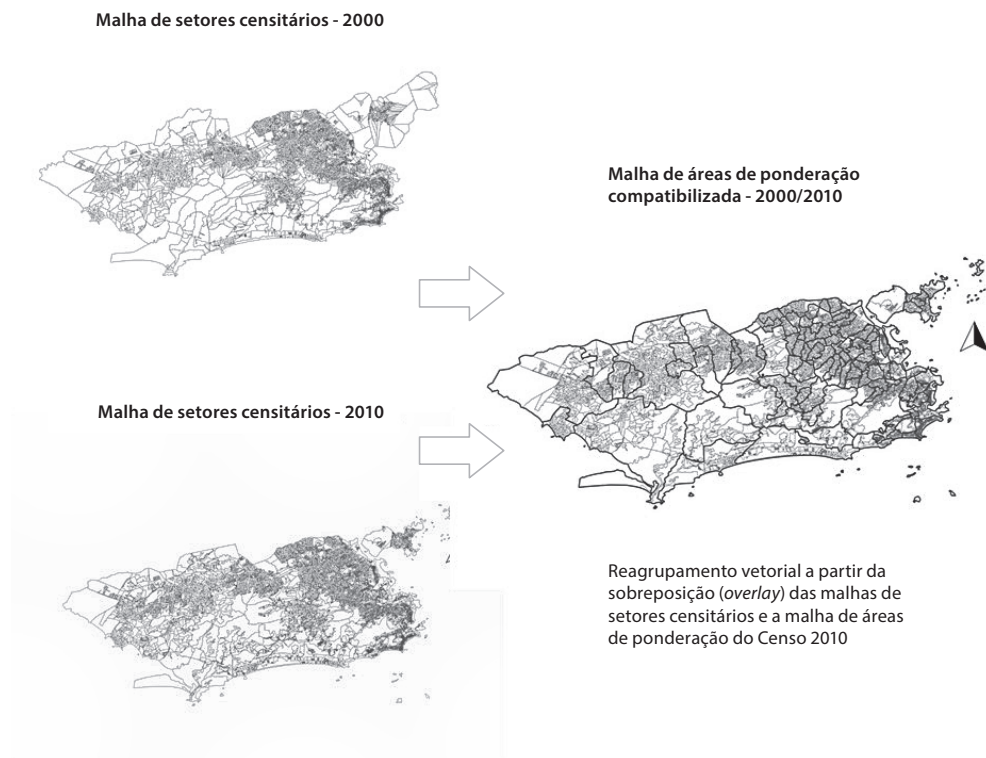
O motivo fundamental da compatibilização

A insistência na elaboração dessa metodologia de compatibilização foi reforçada diante da necessidade de evitar que a análise dos dados fosse acometida pela incerteza em relação aos seus resultados. Sem que as áreas estejam compatibilizadas se torna inviável entender e confiar nos resultados gerados que podem refletir apenas distorções relativas às diferenças das fronteiras das unidades territoriais consideradas (FEITOSA; MONTEIRO; CÂMARA, 2005).

O procedimento metodológico que sugerimos tem como primeiro passo a identificação da composição das áreas de ponderação do Censo 2000, ou seja, verificar quais eram os setores censitários incluídos em cada área. Após esse processo, identificamos quais áreas tinham sido modificadas entre um Censo e outro. Nos casos em que as áreas contemplavam e respeitavam os limites administrativos dos bairros, por exemplo, reagregamos de forma expressiva criando um novo bloco, uma área de ponderação mais extensa, mas com cobertura dos dados em ambos os censos. Por exemplo, podemos citar as áreas de ponderação do bairro de “Campo Grande”. No Censo 2000, foram criadas 9 áreas para cobrir os dados coletados nesse bairro. No Censo 2010, foram criadas 12 áreas. Neste caso, assim como as áreas de Bangu, Realengo, Copacabana, Botafogo, em função do limite administrativo ter sido respeitado, foi possível redimensionar a geometria reagrupando-as, criando as áreas correspondentes entre os Censos.

Outro passo essencial foi criar a correspondência entre os setores censitários do censo 2000 com suas áreas de ponderação e, no software Qgis, trabalhar na tabela de atributos da malha vetorial de setores censitários do ano 2000 relacionando-os a essas áreas.

Figura 1 - Sobreposição da malha vetorial de setores censitários na malha de áreas de ponderação compatibilizadas - 2000/2010



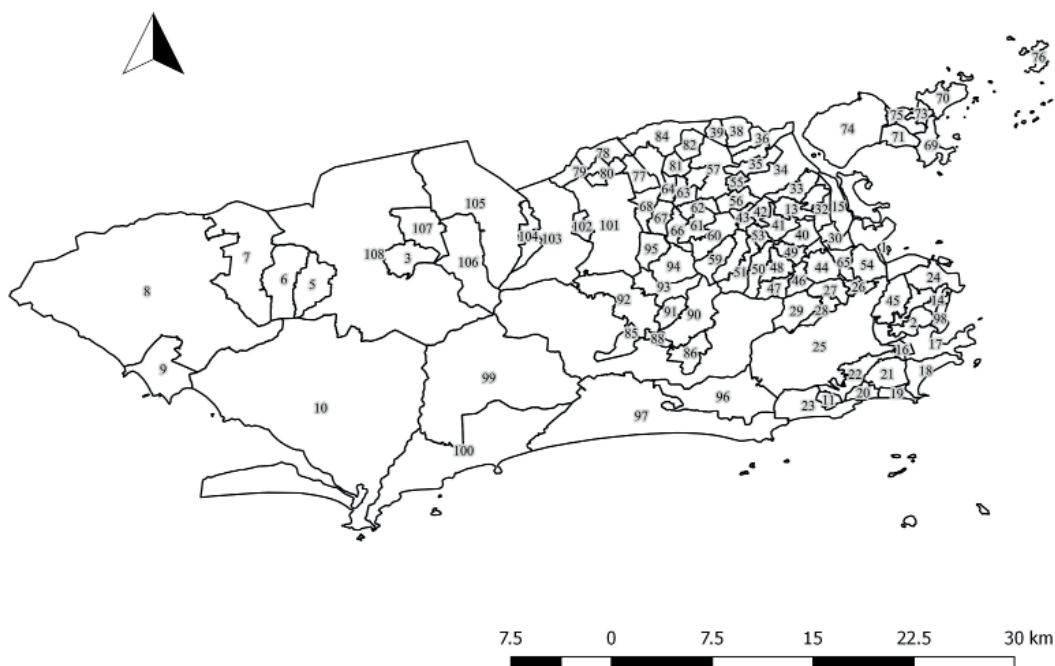
Fonte: Elaboração própria.

Feito isto, tomamos a malha vetorial das áreas de ponderação do Censo 2010 como base para a criação de uma malha para 2000. Uma das maiores preocupações nesse procedimento era manter as áreas sem distorções e que as fronteiras refletissem seu conteúdo. Por isso, levamos em conta, na tabela de atributos, a extensão das áreas e verificamos também a representação da geometria dos polígonos. Identificamos problemas na geometria dos polígonos que representavam setores censitários dos bairros da Ilha do Governador. A representação espacial era muito pouco precisa, enquanto para o Censo de 2010 esses problemas já haviam sido resolvidos. Para corrigir, agregamos as áreas referentes a esse bairro e, na malha de áreas de ponderação 2010, calculamos a área para os dados de 2000.

Descobrimos que apesar da geometria distorcida em 2000 o conteúdo das áreas contemplavam extensão similar e assim a compatibilização foi um exercício de reagrupamento. Esse reagrupamento pode gerar perda de detalhes se comparados aos dados originais, mas como é necessário compatibilizar, optamos por trabalhar de forma que as áreas compatíveis fossem suscetíveis à comparação. Alcançamos as áreas compatíveis para os dois Censos conforme mapa a seguir.

Concluimos essa importante etapa técnica com a apresentação de uma malha vetorial compatível para as áreas de ponderação dos Censos 2000 e 2010, representando o município do Rio de Janeiro. A compatibilização resultou num total de 109 áreas de ponderação comuns.

Mapa 1 - Áreas de ponderação compatibilizadas - Município do Rio de Janeiro - Censos 2000 e 2010



Fonte: Elaboração própria.

**Quadro 2 - Rótulos das áreas de ponderação
compatibilizadas – município do Rio de Janeiro**

Código	Área	Código	Área	Código	Área
1	Centro Antigo	38	Jardim América	75	Paquetá
2	Laranjeiras	39	Del Castilho	76	Guadalupe
3	Senador Vasconcelos	40	Inhaúma	77	Anchieta
4	Inhoaíba	41	Engenho da Rainha	78	Parque Anchieta
5	Cosmos	42	Tomás Coelho	79	Ricardo de Albuquerque
6	Paciência	43	São Francisco Xavier	80	Coelho Neto
7	Santa Cruz	44	Estácio	81	Acari
8	Sepetiba	45	Engenho Novo	82	Costa Barros
9	Guaratiba	46	Lins	83	Pavuna
10	Rocinha	47	Méier	84	Jacarepaguá
11	Jacarezinho	48	Cachambi	85	Anil
12	Alemão	49	Engenho de Dentro	86	Cosme Velho
13	Glória	50	Encantado	87	Cidade de Deus
14	Maré	51	Piedade	88	Curicica
15	Humaitá	52	Pilares	89	Freguesia
16	Botafogo	53	São Cristóvão	90	Pechincha
17	Copacabana	54	Vila da Penha	91	Taquara
18	Ipanema	55	Vicente de Carvalho	92	Tanque
19	Leblon	56	Irajá	93	Praça Seca
20	Lagoa	57	Colégio	94	Vila Valqueire
21	Gávea	58	Quintino	95	Itanhangá
22	Centro	59	Cascadura	96	Barra
23	Vidigal	60	Madureira	97	Flamengo
24	Tijuca	61	Vaz Lobo	98	Vargens
25	Maracanã	62	Rocha Miranda	99	Recreio
26	Vila Isabel	63	Honório Gurgel	100	Deodoro
27	Andaraí	64	Benfica	101	Magalhães Bastos
28	Grajaú	65	Oswaldo Cruz	102	Realengo
29	Manguinhos	66	Bento Ribeiro	103	Padre Miguel
30	Ramos	67	Marechal Hermes	104	Bangu
31	Bonsucesso	68	Cacuia	105	Jacarepaguá
32	Olaria	69	Bancários	106	Santíssimo
33	Penha	70	Jardim Guanabara	107	Campo Grande
34	Brás de Pina	71	Jardim Carioca	108	Leme
35	Parada de Lucas	72	Tauá	109	Higienópolis
36	Cordovil	73	Galeão		
37	Vigário Geral	74	Portuguesa		

Fonte: Elaboração própria.

Considerações finais

Não podemos deixar de considerar a heterogeneidade interna das áreas de ponderação nas próximas etapas da pesquisa. É necessário cautela no trabalho de representação dos fenômenos nos mapas coropléticos, pois estes tem propensão à generalização dos valores extremos reduzindo a heterogeneidade espacial dos dados (CRUZ, 2015).

A área de ponderação, unidade de calibração amostral, tem em sua delimitação a suposição de homogeneidade interna, ou seja, é um agrupamento em que os elementos componentes tendem a ser semelhantes internamente. No entanto, essa premissa não se confirma como verdadeira, pois como unidade para fins operacionais, não há garantia de uma distribuição homogênea do evento no interior dessas áreas (FRANÇA et al, 2008).

Assim, a ampliação da cobertura geográfica proporcionada pelas áreas de ponderação em relação aos setores censitários também aumenta sua heterogeneidade interna. Como exemplo podemos citar a área de número 24 definida como “Tijuca”, área que reúne bairros como “Tijuca”, “Praça da Bandeira” e “Alto da Boa Vista”. Nessa área devem ser consideradas as heterogeneidades referentes ao bairro de alta renda (Alto da Boa Vista) e o bairro da Tijuca, um bairro característico de classe de renda média, cercado de comunidades de baixa renda.

A análise dos dados de áreas com heterogeneidades como essa deve ser cuidadosa e devemos considerar relativa perda de informação, fruto da agregação da unidade territorial. Caso ocorra a necessidade de uma investigação mais precisa das áreas com tal característica, deve-se recorrer à análise de dados segundo setores censitários (AZEVEDO, 2002).

Podemos adicionar a existência de uma heterogeneidade ainda mais peculiar do que essa apresentada anteriormente, trata-se de uma heterogeneidade em nível de comunidades que existem como nichos diferenciados em determinadas áreas da cidade. Destacaremos o que acontece na área de ponderação 17 referente à “Copacabana”, bairro conhecido internacionalmente como grande representação do Rio de Janeiro, bairro que hospeda a população de alta renda e de idade mais avançada. No entanto, a área de Copacabana também reúne comunidades carentes como “Morro da Babilônia” e “Chapéu Mangueira”, além de parte da comunidade do morro “Pavão pavãozinho” que se localiza entre Copacabana e Ipanema. Com isso mostramos que é necessário que o pesquisador tenha o conhecimento dessa disposição espacial do interior do município para captar de forma mais adequada as informações geradas pelos dados recenseados.

Por fim, a técnica de compatibilização de áreas operacionais entre as pesquisas censitárias é de vital potencialidade e pode ser replicada para todas as áreas de ponderação com cobertura de todo o território nacional. É inquestionável a possibilidade que se abre para a utilização dos dados censitários em pesquisas longitudinais.

Referências

- AZEVEDO, S. J. S. Território e Análise Sociodemográfica: contribuições para a definição de demandas sociais e o exemplo das telecomunicações e da saúde pública em Campinas. In: CUNHA, J. M. P. (org.) *Novas Metrópoles Paulistas: população, vulnerabilidades e segregação*, 2002.
- CRUZ, A. I. Compatibilização de limites através de Sistemas de Informações Geográficas – Censo Demográfico 2010: o caso do Distrito de Tatuapé. São Paulo, USP, 2015.
- CUNHA, J. (org.). *Novas metrópoles paulistas: população, vulnerabilidade e segregação*. Campinas: NEPO/UNICAMP, 2006. 611 p.
- CUNHA, J. *et al.* Expansão metropolitana, mobilidade espacial e segregação nos anos 90: o caso da Região Metropolitana de Campinas. In: ENCONTRO NACIONAL DA ANPUR. 11, 2005, Salvador. *Planejamento, soberania e solidariedade: perspectivas para o território e a cidade*. Salvador, Bahia: ANPUR, 2005. (Disponível em CD-ROM).
- FEITOSA, F. da F.; MONTEIRO, A. M. V.; CÂMARA, G. Compatibilização dos dados censitários para análises temporais com auxílio de imagens Landsat. XII Simpósio Brasileiro de Sensoriamento Remoto, Goiânia, INPE, 2005.
- FRANÇA, C. J.; ALVES, C. E.; FREITAS, M. W. D; BERGAMASCHI, R. B. Compatibilização dos bairros e unidades de planejamento com os setores censitários no Estado do Espírito Santo. *Geografares*, n. 6, 2008.
- HAINING, R. *Spatial Data analysis: theory and practice*. London: Cambridge University, 2003.
- IBGE. Documentação dos Microdados da Amostra. *Censo Demográfico 2000*. Rio de Janeiro: 2002.
- IBGE. Notas Metodológicas. *Censo Demográfico 2010*. Rio de Janeiro: 2010.
- PERES DE ÁVILA, R.; MONASTERIO, L. M. O MAUP e a análise espacial: um estudo de caso para o Rio Grande do Sul (1991-2000). *Revista Análise Econômica*. UFRGS/FCE, v. 26, n. 49, 2008. Disponível em: <http://www.fee.rs.gov.br/3eeg/Artigos/m23t01.pdf>
- REIS, E. J.; PIMENTEL, M.; ALVARENGA, A. I.; SANTOS, M. do C. H.. Áreas Mínimas Comparáveis para os períodos intercensitários de 1872 a 2000. I Simpósio Brasileiro de Cartografia Histórica. Paraty, Rio de Janeiro, 2011.
- SOUZA, L. G. Preparação da Base Espacial do Censo Demográfico IBGE 2010 para o município do Rio de Janeiro. IV Simpósio Brasileiro de Ciências Geodésias e Tecnologias da Geoinformação. Recife-PE, 06-09 de maio de 2012 – p.001-009.
- UMBELINO, G.; BARBIERI, A. Metodologia para a compatibilização de setores censitários e perímetros urbanos entre os censos de 1991, 2000 e 2010. XVI Encontro Nacional de Estudos Populacionais, Caxambu – MG – 2008.
- YAMAMOTO, J. K.; LANDIM, P. M. B. *Geoestatística: conceitos + aplicações*. São Paulo: Oficina de Textos, 2013.

Indicadores de representatividade na pesquisa sobre o uso das Tecnologias de Informação e Comunicação nos domicílios brasileiros*

Mayra Pizzott Rodrigues dos Santos**

Marcelo Trindade Pitta***

Denise Britz do Nascimento Silva****

Resumo

A ocorrência de não resposta é um fenômeno inerente a pesquisas e precisa ser considerado na avaliação de qualidade. No caso de pesquisas domiciliares, isso acontece quando os indivíduos ou domicílios selecionados para a amostra não fornecem as informações requeridas. O indicador R – de representatividade da amostra – foi desenvolvido para quantificar o impacto da não resposta nos resultados de uma pesquisa. Neste estudo, utiliza-se o indicador R para avaliar e analisar a representatividade da amostra da Pesquisa sobre o uso das Tecnologias de Informação e Comunicação nos Domicílios Brasileiros (TIC Domicílios), realizada pelo NIC.br, para produção de estatísticas por Unidade da Federação (UF). A amostra da pesquisa sobre TIC Domicílios é originalmente planejada para fornecer resultados por Grandes Regiões, não sendo possível assegurar a adequação da amostra para a estimação de estatísticas por UF. Para o cumprimento do objetivo, utilizou-se o Indicador R, cuja finalidade usual é quantificar a distância entre a amostra de respondentes e a amostra planejada/selecionada, para avaliar se, mesmo na fase de planejamento – antes da operação de coleta, a amostra desenhada era suficientemente informativa e de qualidade para produção de estatísticas por UF. No estudo são apresentadas análises comparativas do indicador R para diferentes níveis territoriais e entre diferentes bases de dados. Os resultados dos indicadores estimados evidenciaram que a amostra de respondentes e a amostra planejada da TIC Domicílios podem ser consideradas representativas a nível Brasil e Grandes Regiões. Algumas UFs, no entanto, apresentam um distanciamento entre as amostras planejadas e/ou de respondentes e a população-alvo. Os resultados deste trabalho serão incorporados num estudo de estimação em pequenas áreas para a pesquisa TIC Domicílios com foco na produção de estimativas para as Unidades da Federação.

Palavras-chave: Não resposta. Indicador R. Representatividade da amostra. TIC Domicílios.

* Este capítulo é resultado de projeto de pesquisa realizado no âmbito do Acordo de Cooperação Técnica entre a ENCE e o NIC.br com vigência de maio de 2014 a maio de 2020.

** Estatística Jr. da área de Métodos Quantitativos do Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação - Cetic.br/NIC.br. Concluiu o mestrado em População, Território e Estatísticas Públicas na ENCE em 2017.

*** Coordenador da área de Métodos Quantitativos do Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação - Cetic.br/NIC.br. Concluiu o mestrado em Estudos Populacionais e Pesquisas Sociais na ENCE em 2003.

**** Pesquisadora da ENCE e Docente do seu programa de pós-graduação desde sua criação em 1998.

Introdução

A demanda por estatísticas públicas para o planejamento, monitoramento e avaliação de políticas públicas é cada vez maior. Essa demanda caracteriza-se pela necessidade de estatísticas mais frequentes, para territórios “menores”, ou ambas: dados mais frequentes e mais desagregados (ou passíveis de maior desagregação). Esse movimento mundial em busca de mais dados é fruto de um entendimento de que políticas e ações públicas devem ser baseadas em informações quantitativas. Reconhecimento refletido, por exemplo, na nova agenda dos Objetivos de Desenvolvimento Sustentável (ODS) e seus respectivos indicadores de monitoramento (UN, 2017; ONUBR, 2018). Apesar de uma demanda maior e diferenciada, tais estatísticas devem ter a mesma qualidade das estatísticas que são hoje disponibilizadas. A compatibilização das necessidades com a manutenção da qualidade é um grande desafio. Nesse artigo introduzimos o cálculo do indicador R para estatísticas da pesquisa TIC Domicílios realizada anualmente pelo CETIC.br, departamento do NIC.br, com o objetivo de avaliar a qualidade das estatísticas para as divisões territoriais consideradas no planejamento da pesquisa e outras desagregações não contempladas na fase de planejamento amostral.

Os indicadores R, em que R significa “representatividade”, foram desenvolvidos a partir da necessidade de concepção de medidas de qualidade para pesquisas amostrais que pudessem informar sobre o possível risco de produção de estimativas viesadas. Esses indicadores medem quão representativa é a amostra de respondentes da pesquisa, ou melhor, como a composição da amostra de respondentes difere da amostra planejada.

Os indicadores R podem ser usados de várias maneiras. Uma delas é na análise dos dados da pesquisa após a conclusão do trabalho de campo. Outra maneira é para o monitoramento durante a coleta de dados, por exemplo, no período do trabalho de campo, no qual os esforços de coleta podem ser direcionados para obter uma amostra de respondentes cuja composição não seja tão diferente da amostra planejada (BETHLEHEM *et al.*, 2008).

Em pesquisas amostrais, o fenômeno da não resposta tem duas principais consequências para pesquisas amostrais: a redução do tamanho da amostra e a alteração da composição da amostra originalmente planejada e selecionada (impedindo assim a efetiva realização do seu desenho amostral). Enquanto o tamanho da amostra está diretamente relacionado à precisão das estimativas, uma modificação na composição da amostra pode causar vício. Nesse caso, as probabilidades de observação dos dados dos indivíduos selecionados são afetadas e a possibilidade de correção do viés de não resposta depende da disponibilidade de informação auxiliar sobre as unidades amostrais. Caso contrário, há pouco o que se possa fazer (SCHOUTEN; COBBEN, 2007).

Segundo Schouten *et al.* (2010; 2012), os indicadores R podem ser usados para comparar a resposta entre diferentes pesquisas que compartilham a mesma população-alvo, comparar a resposta em ondas de uma pesquisa repetida ou longitudinal, monitorar a resposta a uma pesquisa durante a coleta de dados, e para adaptar o procedimento de coleta de dados com base em dados históricos, dados de cadastro disponíveis e parados¹.

Propõe-se o emprego do indicador R para avaliar a adequação de uma amostra planejada para produzir estimativas para áreas ou domínios que não foram considerados no seu desenho amostral.

A pesquisa cuja base de dados é utilizada nesse estudo é a Pesquisa sobre Uso e Acesso a Tecnologias de Comunicação e Informação nos Domicílios brasileiros – TIC Domicílios. Essa pesquisa é conduzida pelo CGI – Comitê Gestor da Internet do Brasil, por meio do NIC.br (Núcleo de Informação e Coordenação do Ponto.br) e de seu departamento de pesquisas e produção de estatísticas sobre a Internet no Brasil: CETIC.br (Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação). Os dados do Cetic.br relativos às pesquisas sobre Internet, comunicação e tecnologia, para diversas populações – domicílios, pessoas, estabelecimentos de saúde e educação, empresas, etc. – são utilizados desde 2005 pelo governo, iniciativa privada

¹ Informações referentes ao processo da coleta de dados da pesquisa.

e sociedade civil no planejamento e monitoramento de políticas e programas que visam à evolução com qualidade da Internet e das tecnologias de informação e comunicação no país².

O objetivo deste trabalho é avaliar a amostra da pesquisa TIC Domicílios como fonte de informação para obtenção de estimativas com boa precisão por Unidade da Federação (UF). Destaca-se que a amostra da referida pesquisa, cujo desenho é estratificado e conglomerado em diversos estágios, é originalmente planejada para fornecer resultados com precisão controlada por Grandes Regiões. No seu plano amostral, as UFs foram utilizadas como estratos de seleção, mas a alocação da amostra em cada um desses estratos é efetuada de forma a minimizar custos e não para controlar a precisão das estimativas por UF. Isso acarreta uma redução do tamanho da amostra planejada em algumas UFs (usualmente nas regiões Norte e Nordeste), garantindo resultados de qualidade apenas no nível da Grande Região.

Considerando o objetivo de produzir estimativas por UF a partir da amostra de respondentes, e que a amostra planejada foi elaborada para produzir estimativas por Grande Região, as composições da amostra planejada e da amostra de respondentes são comparadas com a população-alvo da qual se selecionou a amostra, ou seja, a população investigada no Censo 2010. Esse conjunto de comparações permitirá identificar se a falta de representatividade, no que se refere à produção de estimativas por UF, já estava presente na amostra planejada, antes mesmo da coleta dos dados, ou se é resultante de não resposta diferenciada entre as UFs.

Considerando a comparação com a população-alvo (Censo 2010 nesse caso), o indicador R é calculado no nível de setores censitários (unidade mínima para qual há dados de toda a população) com as seguintes configurações/comparações: população versus amostra planejada, população versus amostra de respondentes e amostra planejada versus amostra de respondentes.

Portanto, o indicador é obtido utilizando as informações da população (a partir dos dados do Censo 2010), da amostra planejada e de respondentes da pesquisa TIC Domicílios nos anos 2015 a 2017. Nesse caso, deseja-se avaliar o quanto a amostra de respondentes e a amostra planejada são representativas da população da qual a amostra foi selecionada considerando o objetivo de produzir estimativas dos indicadores de interesse por UF.

Os resultados deste trabalho servirão como informação para um estudo de viabilidade de estimação em pequenas áreas³ (Unidades da Federação) na pesquisa TIC Domicílios que, como já destacado, atualmente produz resultados apenas para as Grandes Regiões.

Metodologia

Esta seção apresenta um resumo da metodologia descrevendo o indicador R, a fonte dos dados e os procedimentos utilizados para obter os resultados apresentados no artigo.

Fonte dos dados

A fonte de dados deste estudo é a Pesquisa sobre Uso e Acesso a Tecnologias de Informação e Comunicação nos Domicílios brasileiros – TIC Domicílios conduzida pelo CGI, por meio do NIC.br e seu departamento de pesquisas e produção de estatísticas sobre a Internet no Brasil, CETIC.br. O trabalho foi desenvolvido com os dados referente aos anos 2015 a 2017. A pesquisa tem como objetivo principal medir a posse e o uso das TIC entre a população residente no Brasil com idade de 10 anos ou mais (NIC.br, 2017).

Na definição do desenho amostral da pesquisa TIC Domicílios é utilizada a base de setores censitários do Censo Demográfico 2010 do IBGE. O plano amostral pode ser descrito como amostragem estratificada de conglomerados em múltiplos estágios. O número de estágios de conglomeração depende da seleção de municípios, pois alguns municípios

² Mais detalhes em: https://www.cetic.br/media/docs/publicacoes/2/tic_dom_2017_livro_eletronico.pdf.

³ Uma área é considerada pequena se o tamanho da amostra na área não é suficientemente grande para produzir estimativas com a precisão desejada.

são incluídos com certeza na amostra, enquanto outros são selecionados por sorteio, com probabilidades especificadas.

No caso dos municípios incluídos na amostra com certeza (isto é, com probabilidade de seleção igual a um), a amostragem é feita em três estágios: os municípios funcionam como estratos para seleção de setores censitários nos quais, por sua vez, são selecionados domicílios e, no último estágio, moradores. Os demais municípios não incluídos com certeza na amostra funcionam como unidades primárias de amostragem (UPAs). Nesses casos, a amostragem probabilística apresenta quatro etapas: seleção de municípios, de setores censitários, de domicílios e, por último, seleção de moradores⁴.

Indicador R

Os indicadores R fornecem uma medida de quanto a composição da amostra de respondentes de uma pesquisa se desvia da amostra planejada. Se as probabilidades de resposta de todas as unidades amostrais na amostra coletada forem iguais, não haverá diferenças sistemáticas entre a composição da amostra de respondentes e a da amostra planejada. Por outro lado, se as probabilidades de resposta não forem iguais, é importante estabelecer até que ponto a composição da amostra de respondentes é afetada. Isso pode ser obtido definindo-se uma função de distância que mede o quanto as probabilidades de resposta das unidades diferem da probabilidade média de resposta (BETHLEHEM et al., 2008). Esses indicadores foram propostos por Schouten e Cobben (2007) e são objeto de um projeto denominado Representative Indicators for Survey Quality (RISQ⁵) realizado por instituições de alguns países da União Europeia.

Suponha que uma amostra probabilística s de tamanho n é selecionada sem reposição de uma população finita. A amostra pode ser representada por um vetor de indicadores $s = (s_1, s_2, \dots, s_i, \dots, s_N)$, no qual s_i assume valor 1 quando o elemento i é selecionado na amostra e 0, caso contrário (COBBEN; SCHOUTEN, 2008). A resposta à pesquisa pode ser representada pelo vetor de indicadores $r = (r_1, r_2, \dots, r_N)$ no qual r_i assume valor 1 se o elemento i foi selecionado na amostra e respondeu e 0, caso contrário. Cada elemento i na população possui probabilidade desconhecida ρ_i de responder quando selecionado para a amostra, ou seja, $\rho_i = P(r_i = 1 | s_i = 1)$.

Como ρ_i é desconhecida, é preciso estimar seu valor. Fazemos isso usando um vetor de variáveis auxiliares que está disponível para todos os elementos da amostra. Segundo Bethlehem et al. (2008), podem ser utilizadas várias técnicas para estimação da propensão de resposta, por exemplo, modelos logísticos ou *probit* (AGRESTI, 2002), ou árvores de classificação CHAID (KASS, 1980).

Seja então $\hat{\rho}_i$ um estimador de ρ_i que utiliza um conjunto de informações auxiliares e seja $\hat{\rho}$ a média ponderada das probabilidades de resposta estimadas, tal que

$$\hat{\rho} = \frac{1}{N} \sum_{i=1}^N \hat{\rho}_i \frac{s_i}{\pi_i}, \quad (1)$$

sendo π_i a probabilidade de inclusão de primeira ordem da unidade amostral i .

A propensão de resposta ρ_i é definida como a esperança condicional de r_i dados os valores do vetor \mathbf{X} com m variáveis auxiliares.

$$\rho_i = E(r_i = 1 | \mathbf{X} = \mathbf{x}_i \text{ e } s_i = 1) = P(r_i = 1 | \mathbf{X} = \mathbf{x}_i \text{ e } s_i = 1).$$

⁴ Maiores informações podem ser encontradas em: <https://www.cetic.br/>.

⁵ <https://www.cmist.manchester.ac.uk/research/projects/representative-indicators-for-survey-quality/>.

Uma classe geral de modelos que representam a relação de ρ_i e \mathbf{x}_i pode ser expressa na forma:

$$g(\rho_i) = \mathbf{x}'_i \boldsymbol{\beta}$$

na qual $g(\cdot)$ é uma determinada função de ligação, $\boldsymbol{\beta}$ é um vetor de parâmetros desconhecidos a serem estimados, e \mathbf{x}_i é o vetor com valores das variáveis auxiliares para a unidade i . De acordo com Skinner et al. (2009), uma escolha padrão da função de ligação é a função *logit*. O modelo de regressão logística para a propensão de resposta pode então ser expresso da seguinte forma:

$$\text{logit}(\rho_i) = \log\left(\frac{\rho_i}{1 - \rho_i}\right) = \mathbf{x}'_i \boldsymbol{\beta} \quad (2)$$

Sendo assim, a probabilidade do elemento i responder (ρ_i), reescrevendo (2), é dada por:

$$\rho_i = \frac{\exp(\mathbf{x}'_i \boldsymbol{\beta})}{1 + \exp(\mathbf{x}'_i \boldsymbol{\beta})} \quad (3)$$

Schouten e Cobben (2007) definiram três indicadores R : um com base no desvio padrão das probabilidades de resposta (R_1), um com base na variância das probabilidades de resposta (R_2) e outro com foco na redução do erro de predição da propensão de resposta individual (R_3).

$$\hat{R}_1(\rho) = 1 - 2 \hat{S}(\hat{\rho}) = 1 - 2 \sqrt{\frac{1}{N-1} \sum_{i=1}^N \frac{S_i}{\pi_i} (\hat{\rho}_i - \hat{\rho})^2} \quad (4)$$

$$\hat{R}_2(\rho) = 1 - 4 \hat{S}^2(\hat{\rho}) = 1 - 4 \frac{1}{N-1} \sum_{i=1}^N \frac{S_i}{\pi_i} (\hat{\rho}_i - \hat{\rho})^2 \quad (5)$$

$$\hat{R}_3(\rho) = 1 - \mathfrak{R}^2 \quad (6)$$

A estatística \mathfrak{R}^2 na equação (6) é a proporção da variância explicada pelo modelo usualmente utilizada em regressão linear e, de forma análoga, a estatística pseudo \mathfrak{R}^2 foi desenvolvida para modelos com variáveis dependentes categóricas.

O indicador obtido do desvio padrão das probabilidades de resposta - R_1 - admite valores no intervalo [0, 1], no qual o valor 1 indica forte representatividade e valores mais baixos correspondem a uma resposta menos representativa. Neste trabalho, decidiu-se usar o Indicador R proposto por Schouten e Cobben (2007) com base no desvio padrão das probabilidades de resposta estimadas, que é definido em (4), por ser o indicador mais utilizado na literatura para avaliação de representatividade de amostras.

Na obtenção do indicador $\hat{R}_1(\rho)$ há dois processos aleatórios envolvidos: o primeiro processo é a amostragem da população, e o segundo é o mecanismo de resposta das unidades da amostra (COBBEN; SCHOUTEN, 2008).

É preciso portanto estimar o erro padrão (σ_R) do indicador $\hat{R}_1(\rho)$ pois o cálculo não é direto devido à necessidade de obter estimativas das probabilidades de resposta. De acordo com Heij et al. (2015), são calculadas aproximações analíticas do erro padrão para o indicador $\hat{R}_1(\rho)$. Detalhes podem ser encontrados em Shlomo et al. (2012).

Indicador R para o caso da TIC Domicílios

A TIC Domicílios é uma pesquisa cuja amostra é planejada para fornecer resultados por Grandes Regiões, porém o objetivo deste estudo é avaliar se essa amostra, e a amostra de res-

pondentes, são representativas por UF. Nesta seção é abordada a metodologia do indicador R aplicado à Pesquisa TIC Domicílios.

Na análise de resultados, são apresentados apenas os indicadores estimados a partir das seguintes comparações: População versus Amostra de respondentes e População versus Amostra planejada. Os níveis territoriais avaliados são: Brasil, Grandes Regiões e Unidades da Federação.

Para o estudo em questão, o modelo⁶ relaciona a propensão de resposta no setor censitário com as seguintes variáveis⁷ relacionadas aos domicílios e indivíduos:

X_1 : Média do número de moradores em domicílios particulares permanentes (DPP⁸) no setor,

X_2 : Valor do rendimento nominal médio mensal das pessoas de 10 anos ou mais de idade (com e sem rendimento) no setor.

As variáveis são numéricas e foram categorizadas da seguinte forma:

- X_1 : 1 a 3 moradores, mais de 3 a 5 moradores, mais de 5 a 7 moradores e mais de 7 moradores.
- X_2 : R\$ 1,00 a R\$ 650,00, R\$ 650,01 a R\$ 950,00, R\$ 950,01 a R\$ 1500,00 e mais de R\$ 1500,00.

A escolha das variáveis auxiliares foi feita a partir das variáveis disponíveis para a população. As variáveis devem ser as mesmas para todos os modelos nas duas comparações (População versus Amostra planejada e População versus Amostra de respondentes), nos três anos de estudo (2015, 2016 e 2017) e para os três níveis geográficos (Brasil, Grandes Regiões e UFs), dessa forma podem ser feitas comparações entre níveis geográficos, uma vez que os fatores considerados para a avaliação de não resposta serão os mesmos. Um exemplo dos resultados dos modelos é apresentado no Apêndice A.

A variável de interesse (R_i) para o modelo logístico de propensão de resposta é definida como um indicador de resposta:

$$R_i = \begin{cases} 1, & \text{se há resposta para o } i - \text{ésimo setor censitário} \\ 0, & \text{caso contrário} \end{cases}$$

Por exemplo, na comparação População versus Amostra planejada, a variável resposta assume valor 1 se o setor censitário está presente no Censo e na Amostra planejada e 0, se está presente apenas no Censo.

No modelo de regressão logística, a probabilidade de sucesso ρ e o conjunto de variáveis explicativas se relacionam através da função de ligação logit da seguinte forma:

$$\text{logit}(\rho_{jki}) = \log\left(\frac{\rho_{jki}}{1 - \rho_{jki}}\right) = \mu + \tau_j + \beta_k \quad \begin{cases} j = 1, \dots, 4. \\ k = 1, \dots, 4. \\ i = 1, \dots, n. \end{cases} \quad (7)$$

na qual

μ representa a probabilidade de sucesso nas categorias de referência,

τ_j representa o efeito do nível j da variável média do número de moradores em DPP,

β_k representa o efeito do nível k da variável valor do rendimento nominal médio mensal das pessoas de 10 anos ou mais de idade (com e sem rendimento),

ρ_{jki} é a probabilidade de resposta do setor censitário i nos níveis j e k .

⁶ Outros modelos com mais variáveis foram ajustados, mas as demais variáveis não foram consideradas significantes, ou ocorreram problemas de convergência na estimação dos parâmetros.

⁷ Variáveis explicativas para setores censitários do Censo Demográfico de 2010.

⁸ DPP corresponde aos domicílios particulares permanentes, ou seja, domicílios que foram construídos a fim de servir exclusivamente para habitação e, na data de referência do Censo Demográfico 2010, tinham a finalidade de servir de moradia a uma ou mais pessoas (IBGE, 2016).

Assim, obtém-se a probabilidade de resposta (ρ_{jki}) do setor censitário i nos níveis j e k reescrevendo (7) como:

$$\rho_{jki} = \frac{\exp(\mu + \tau_j + \beta_k)}{1 + \exp(\mu + \tau_j + \beta_k)} \quad (8)$$

Como o indicador R não está sendo calculado de forma usual, o significado de ρ_{jki} varia de acordo com a comparação realizada. No caso da comparação População versus Amostra de respondentes, ρ_{jki} é a probabilidade de ocorrência de resposta no setor censitário i nos níveis j e k . Por outro lado, no caso da utilização do indicador R para comparação da População com a Amostra planejada, ρ_{jki} é a probabilidade do setor censitário i nos níveis j e k estar presente (ter sido selecionado) na amostra planejada.

Santos (2017) incorporou o desenho complexo da amostra da pesquisa na estimação dos modelos de propensão. Portanto, neste estudo decidiu-se também incorporar o efeito do plano amostral utilizando a função *svyglm* do pacote *survey* do software R para obter as estimativas das probabilidades de resposta.

No caso do cálculo do indicador R para comparações com a população-alvo da pesquisa, as unidades contidas apenas no censo têm peso 1 enquanto as unidades na amostra recebem o respectivo peso definido pelo desenho amostral.

Resultados

A pesquisa TIC Domicílios é uma pesquisa de abrangência nacional, com uma amostra estratificada e conglomerada em vários estágios. A amostra planejada da pesquisa tem tamanho total de 33.210 domicílios em 2.114 setores censitários e 350 municípios. O espalhamento da amostra no território brasileiro atende a restrições de custo e de representatividade populacional. A Tabela 1 apresenta a alocação dos setores na amostra segundo estratos geográficos.

De acordo com o plano amostral da pesquisa, são selecionados 15 domicílios por setor censitário. Verifica-se, na Tabela 1, que as regiões Norte e Centro-Oeste têm uma amostra de municípios e setores muito menor que as demais regiões. Na região Norte essa amostra menor está espalhada em mais estados do que na região Centro-Oeste. Na região Norte, a média de setores censitários por UF na amostra planejada é de 25, enquanto na região Sudeste a média é superior a 125. Essa distribuição assimétrica reflete a distribuição de população no território e o alto custo da realização de entrevistas presenciais nas regiões mais remotas do país.

Os indicadores R foram calculados para comparar três bases de dados (universo, amostra planejada e amostra de respondentes) nos três anos de estudo da pesquisa (2015, 2016 e 2017) e para os três níveis geográficos (Brasil, Grandes Regiões e UFs). A Tabela 2 apresenta o indicador R e o respectivo intervalo de confiança calculados no nível Brasil e Grandes Regiões para cada ano de estudo da pesquisa TIC Domicílios. Na Tabela 2, o indicador é obtido a partir da comparação do universo de interesse com a amostra de respondentes e do universo de interesse com a amostra planejada.

Tabela 1 - Alocação da amostra da pesquisa TIC Domicílios

Estrato	Amostra		
	Setores	Municípios	Entrevistas planejadas
Norte			
Rondônia	18	4	270
Roraima	15	4	225
Acre	15	4	225
Amapá	15	6	225
Tocantins	15	4	225
Amazonas	38	8	570
Pará - RM Belém	27	4	405
Pará - Interior	57	9	855
Nordeste			
Maranhão	71	12	1 065
Piauí	36	7	540
Ceará - RM Fortaleza	42	6	630
Ceará - Interior	55	8	825
Pernambuco - RM Recife	41	6	615
Pernambuco - Interior	57	10	855
Rio Grande do Norte	39	7	585
Paraíba	45	11	675
Alagoas	35	7	525
Sergipe	28	6	420
Bahia - RM Salvador	44	6	660
Bahia - Interior	122	19	1 830
Sudeste			
Minas Gerais - RM BH	63	8	945
Minas Gerais - Interior	146	27	2 190
Espírito Santo	47	8	705
Rio de Janeiro - RM RJ	136	13	2 040
Rio de Janeiro - Interior	53	7	795
São Paulo - RM São Paulo	206	18	3 090
São Paulo - Interior	226	42	3 390
Sul			
Paraná - RM Curitiba	42	6	630
Paraná - Interior	88	15	1 320
Santa Catarina	82	13	1 230
Rio Grande do Sul - RM Porto Alegre	50	7	750
Rio Grande do Sul - Interior	84	14	1 260
Centro-Oeste			
Mato Grosso do Sul	32	5	480
Mato Grosso	41	7	615
Goiás	70	11	1 050
Distrito Federal	33	1	495

Fonte: CGI.br/NIC.br, Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação (Cetic.br), Pesquisa sobre o Uso das Tecnologias de Informação e Comunicação nos domicílios brasileiros - TIC Domicílios 2016.

Nota: RM = Região Metropolitana.

Considerando os resultados para o Brasil apresentados na Tabela 2, observa-se que os indicadores da comparação População versus Amostra planejada são maiores do que na comparação População versus Amostra de respondentes, porém não há diferenças signifi-

cativas entre os indicadores considerando-se os intervalos de confiança das medidas. Os indicadores *R* estimados podem ser considerados altos, apontando que, tanto a amostra planejada, como a de respondentes podem ser consideradas representativas quando comparadas à população. Nota-se também um acréscimo no valor do indicador de 2015 para 2016 e um decréscimo de 2016 para 2017, indicando que a amostra de 2016 é um pouco mais representativa, porém essa diferença não é estatisticamente significativa de acordo com os intervalos de confiança das medidas.

Tabela 2 - Indicador *R* e respectivo IC_{95%}, segundo Grandes Regiões - 2015-2017

Grandes Regiões	Indicador <i>R</i> e respectivo IC _{95%}	
	População versus amostra planejada	População versus amostra de respondentes
2015		
Brasil	0,918 (0,880 - 0,955)	0,910 (0,872 - 0,947)
Norte	0,839 (0,694 - 0,983)	0,832 (0,693 - 0,972)
Nordeste	0,895 (0,868 - 0,923)	0,893 (0,858 - 0,928)
Sudeste	0,919 (0,865 - 0,972)	0,902 (0,852 - 0,953)
Sul	0,904 (0,824 - 0,984)	0,902 (0,821 - 0,983)
Centro-Oeste	0,863 (0,775 - 0,952)	0,861 (0,770 - 0,953)
2016		
Brasil	0,961 (0,912 - 1,009)	0,958 (0,914 - 1,003)
Norte	0,918 (0,762 - 1,073)	0,919 (0,765 - 1,073)
Nordeste	0,911 (0,877 - 0,945)	0,908 (0,871 - 0,946)
Sudeste	0,938 (0,887 - 0,988)	0,931 (0,881 - 0,981)
Sul	0,924 (0,824 - 1,024)	0,925 (0,824 - 1,026)
Centro-Oeste	0,901 (0,726 - 1,076)	0,847 (0,680 - 1,014)
2017		
Brasil	0,929 (0,891 - 0,966)	0,924 (0,887 - 0,961)
Norte	0,808 (0,690 - 0,926)	0,806 (0,687 - 0,925)
Nordeste	0,939 (0,875 - 1,003)	0,924 (0,862 - 0,985)
Sudeste	0,930 (0,868 - 0,992)	0,921 (0,867 - 0,974)
Sul	0,917 (0,799 - 1,034)	0,924 (0,801 - 1,047)
Centro-Oeste	0,862 (0,773 - 0,951)	0,806 (0,650 - 0,962)

Fonte: CGL.br/NIC.br, Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação (Cetic.br), Pesquisa sobre o Uso das Tecnologias de Informação e Comunicação nos domicílios brasileiros – Elaboração dos autores.

Nota: IC_{95%} = Intervalo de confiança de 95%.

Ao analisarmos as amostras de setores no que se refere à representatividade para as Grandes Regiões brasileiras, observa-se que todas as regiões possuem valores de indicador *R* mais altos que 0,80 nos três anos da pesquisa, indicando que a composição da amostra planejada e da amostra de respondentes não diferem muito da população. A comparação População versus Amostra planejada, em geral, possui estimativas pontuais dos indicadores *R* maiores do que as obtidas para a comparação População versus Amostra de respondentes, porém as diferenças não são significativas.

Tabela 3 - Indicador R e respectivo IC_{95%}, segundo Unidades da Federação - 2015

Unidades da Federação	Indicador R e respectivo IC _{95%}	
	População versus amostra planejada	População versus amostra de respondentes
Amapá	0,534 (0,405 - 0,662)	0,534 (0,405 - 0,662)
Roraima	0,549 (0,379 - 0,720)	0,549 (0,379 - 0,720)
Espírito Santo	0,599 (0,428 - 0,770)	0,599 (0,428 - 0,770)
Acre	0,668 (0,517 - 0,819)	0,668 (0,517 - 0,819)
Mato Grosso do Sul	0,718 (0,588 - 0,848)	0,706 (0,569 - 0,842)
Tocantins	0,731 (0,566 - 0,895)	0,731 (0,566 - 0,895)
Rondônia	0,745 (0,436 - 1,054)	0,745 (0,436 - 1,054)
Mato Grosso	0,758 (0,517 - 0,999)	0,758 (0,517 - 0,999)
Amazonas	0,758 (0,583 - 0,932)	0,742 (0,596 - 0,888)
Maranhão	0,779 (0,668 - 0,890)	0,777 (0,660 - 0,894)
Paraíba	0,799 (0,651 - 0,946)	0,799 (0,651 - 0,946)
Goiás	0,801 (0,651 - 0,950)	0,804 (0,647 - 0,961)
Alagoas	0,818 (0,721 - 0,915)	0,818 (0,721 - 0,915)
Santa Catarina	0,821 (0,685 - 0,956)	0,816 (0,681 - 0,951)
Pará	0,823 (0,616 - 1,031)	0,823 (0,616 - 1,031)
Rio Grande do Norte	0,843 (0,656 - 1,030)	0,843 (0,656 - 1,030)
Pernambuco	0,846 (0,728 - 0,964)	0,852 (0,726 - 0,978)
Distrito Federal	0,869 (0,614 - 1,123)	0,869 (0,614 - 1,123)
São Paulo	0,873 (0,815 - 0,931)	0,801 (0,741 - 0,860)
Ceará	0,874 (0,787 - 0,962)	0,793 (0,676 - 0,909)
Sergipe	0,880 (0,715 - 1,044)	0,880 (0,715 - 1,044)
Rio Grande do Sul	0,895 (0,692 - 1,097)	0,895 (0,692 - 1,097)
Paraná	0,895 (0,764 - 1,027)	0,891 (0,761 - 1,020)
Minas Gerais	0,908 (0,800 - 1,017)	0,912 (0,800 - 1,024)
Bahia	0,922 (0,813 - 1,031)	0,922 (0,813 - 1,031)
Rio de Janeiro	0,930 (0,810 - 1,051)	0,903 (0,777 - 1,029)
Piauí	0,940 (0,743 - 1,138)	0,940 (0,743 - 1,138)

Fonte: CGI.br/NIC.br, Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação (Cetic.br), Pesquisa sobre o Uso das Tecnologias de Informação e Comunicação nos domicílios brasileiros – Elaboração dos autores

Nota: IC_{95%} = Intervalo de confiança de 95%.

Tabela 4 - Indicador R e respectivo IC_{95%}, segundo Unidades da Federação - 2016

Unidades da Federação	Indicador R e respectivo IC _{95%}	
	População versus amostra planejada	População versus amostra de respondentes
Amapá	0,535 (0,321 - 0,748)	0,535 (0,321 - 0,748)
Roraima	0,597 (0,427 - 0,767)	0,597 (0,427 - 0,767)
Tocantins	0,657 (0,432 - 0,883)	0,657 (0,432 - 0,883)
Mato Grosso do Sul	0,703 (0,592 - 0,814)	0,632 (0,480 - 0,784)
Acre	0,705 (0,443 - 0,968)	0,705 (0,443 - 0,968)
Rondônia	0,716 (0,432 - 0,999)	0,716 (0,432 - 0,999)
Alagoas	0,726 (0,551 - 0,902)	0,726 (0,551 - 0,902)
Mato Grosso	0,728 (0,529 - 0,927)	0,728 (0,529 - 0,927)
Piauí	0,728 (0,562 - 0,893)	0,730 (0,558 - 0,903)
Amazonas	0,759 (0,561 - 0,956)	0,759 (0,561 - 0,956)
Espírito Santo	0,777 (0,430 - 1,123)	0,750 (0,403 - 1,096)
Maranhão	0,807 (0,691 - 0,924)	0,799 (0,682 - 0,917)
Rio Grande do Sul	0,816 (0,696 - 0,936)	0,825 (0,699 - 0,952)
Pará	0,825 (0,693 - 0,957)	0,831 (0,708 - 0,954)
Paraíba	0,831 (0,684 - 0,978)	0,831 (0,684 - 0,978)
Goiás	0,837 (0,567 - 1,108)	0,812 (0,558 - 1,065)
Pernambuco	0,838 (0,752 - 0,924)	0,838 (0,752 - 0,924)
Sergipe	0,844 (0,675 - 1,012)	0,844 (0,675 - 1,012)
Paraná	0,846 (0,726 - 0,965)	0,843 (0,722 - 0,964)
Distrito Federal	0,861 (0,609 - 1,112)	0,867 (0,620 - 1,115)
Rio Grande do Norte	0,878 (0,696 - 1,060)	0,878 (0,696 - 1,060)
Bahia	0,897 (0,792 - 1,002)	0,897 (0,792 - 1,002)
São Paulo	0,899 (0,827 - 0,971)	0,901 (0,834 - 0,968)
Minas Gerais	0,911 (0,801 - 1,022)	0,911 (0,801 - 1,022)
Santa Catarina	0,914 (0,714 - 1,115)	0,943 (0,690 - 1,196)
Ceará	0,917 (0,812 - 1,021)	0,839 (0,757 - 0,921)
Rio de Janeiro	0,938 (0,811 - 1,065)	0,947 (0,812 - 1,083)

Fonte: CGL.br/NIC.br, Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação (Cetic.br), Pesquisa sobre o Uso das Tecnologias de Informação e Comunicação nos domicílios brasileiros – Elaboração dos autores

Nota: IC_{95%} = Intervalo de confiança de 95%.

Tabela 5 - Indicador R e respectivo IC_{95%}, segundo Unidades da Federação - 2017

Unidades da Federação	Indicador R e respectivo IC _{95%}	
	População versus amostra planejada	População versus amostra de respondentes
Roraima	0,539 (0,412 - 0,666)	0,539 (0,412 - 0,666)
Mato Grosso do Sul	0,588 (0,422 - 0,754)	0,588 (0,422 - 0,754)
Rondônia	0,599 (0,368 - 0,830)	0,599 (0,368 - 0,830)
Bahia	0,692 (0,476 - 0,907)	0,689 (0,503 - 0,875)
Tocantins	0,709 (0,587 - 0,831)	0,709 (0,587 - 0,831)
Santa Catarina	0,714 (0,507 - 0,920)	0,602 (0,298 - 0,906)
Amapá	0,717 (0,530 - 0,905)	0,717 (0,530 - 0,905)
Acre	0,730 (0,570 - 0,891)	0,730 (0,557 - 0,903)
Rio de Janeiro	0,745 (0,533 - 0,957)	0,788 (0,560 - 1,017)
Mato Grosso	0,746 (0,635 - 0,858)	0,747 (0,633 - 0,860)
Alagoas	0,759 (0,526 - 0,993)	0,759 (0,526 - 0,993)
Pará	0,764 (0,560 - 0,969)	0,724 (0,521 - 0,926)
Maranhão	0,773 (0,673 - 0,873)	0,758 (0,642 - 0,875)
Ceará	0,790 (0,587 - 0,992)	0,810 (0,571 - 1,048)
Piauí	0,811 (0,692 - 0,930)	-
Minas Gerais	0,844 (0,670 - 1,017)	0,792 (0,656 - 0,929)
Paraíba	0,852 (0,666 - 1,037)	0,852 (0,666 - 1,037)
Sergipe	0,856 (0,662 - 1,051)	0,788 (0,638 - 0,938)
São Paulo	0,864 (0,761 - 0,968)	0,864 (0,761 - 0,968)
Paraná	0,868 (0,795 - 0,940)	0,861 (0,762 - 0,961)
Rio Grande do Sul	0,868 (0,757 - 0,978)	0,868 (0,757 - 0,978)
Amazonas	0,878 (0,743 - 1,014)	0,771 (0,652 - 0,889)
Goiás	0,882 (0,781 - 0,983)	0,882 (0,781 - 0,983)
Pernambuco	0,897 (0,795 - 0,998)	0,895 (0,794 - 0,996)
Espírito Santo	0,928 (0,705 - 1,150)	0,928 (0,705 - 1,150)
Distrito Federal	0,928 (0,858 - 0,998)	0,935 (0,858 - 1,013)
Rio Grande do Norte	0,933 (0,745 - 1,120)	0,933 (0,745 - 1,120)

Fonte: CGI.br/NIC.br, Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação (Cetic.br), Pesquisa sobre o Uso das Tecnologias de Informação e Comunicação nos domicílios brasileiros – Elaboração dos autores

Nota: IC_{95%} = Intervalo de confiança de 95%.

O indicador *R* calculado para as Unidades da Federação é apresentado nas Tabelas 3, 4 e 5. Os valores dos indicadores das UFs estão separados de acordo com intervalos de valores dos indicadores *R*: abaixo de 0,60, de 0,60 a 0,70, mais de 0,70 e 0,85 e acima de 0,85.

Pode-se observar que diversos estados da região Norte, ao longo dos três anos, possuem indicadores *R* fora do intervalo de melhor representatividade (acima de 0,85). Em 2015 e 2016, nenhum estado dessa região se encontra nesse intervalo.

Além da região Norte, a região Centro-Oeste e Nordeste também apresentam alguns estados nos intervalos de representatividade intermediários, entre 0,60 e 0,85. O estado de Roraima possui menor representatividade em relação às outras UFs do Brasil nos três anos.

Considerações finais

A TIC Domicílios é uma pesquisa anual criada em 2005 e de grande importância para o país. É a principal fonte de informação sobre Tecnologia da Informação e Comunicação nos domicílios brasileiros, além de ter cobertura nacional. A definição de medidas de qualidade, como o indicador *R*, proporciona avanços na relevância da informação produzida pela pesquisa. Neste estudo buscou-se avaliar, com base no indicador *R*, a amostra planejada e a amostra de respondentes da TIC Domicílios.

Os resultados dos indicadores estimados fornecem evidências de que a amostra de respondentes e a amostra planejada da TIC Domicílios podem ser consideradas representativas no nível Brasil e Grandes Regiões por apresentarem valores altos do indicador *R*, como esperado de acordo com o desenho amostral da pesquisa. Ao avaliar a possibilidade de utilizar os dados da pesquisa para obter estimativas diretas (a partir do desenho amostral) para indicadores por UF, verifica-se que algumas UFs apresentam problemas de representatividade. Esses problemas são causados por não resposta ou pela ocorrência de uma amostra planejada de tamanho muito reduzido. Tais resultados apontam para a necessidade de ajuste no desenho amostral da pesquisa para produzir estimativas por UF ou para a utilização de métodos de estimação em pequenas áreas que possibilitem contornar o problema de representatividade nesse nível de desagregação geográfica.

Um exemplo da utilização futura dos resultados deste trabalho é a incorporação dessa informação (estimativas do indicador *R*) para compor uma metodologia de estimação em pequenas áreas para produção de estimativas com precisão controlada para todas as Unidades da Federação. O valor do indicador *R* pode ser utilizado para determinação de pesos em estimadores compostos que, por sua vez, são obtidos por uma combinação linear de estimativas diretas e de estimativas obtidas com estimadores sintéticos (RAO, 2003, p. 57). O valor do indicador *R* poderia então ser incorporado no peso associado à estimativa direta da UF (proveniente do desenho amostral) na definição de um estimador composto. Quanto menor a representatividade da amostra por UF – (menor o valor do indicador *R*) – menor deve ser o peso do estimador direto da UF e maior a importância dada ao estimador sintético que pode ser obtido em níveis geográficos mais agregados para os quais a amostra foi efetivamente desenhada (região, por exemplo). Sendo assim, as evidências obtidas neste estudo formam a base do desenvolvimento de métodos de estimação para ampliação do uso dos resultados produzidos pela pesquisa TIC Domicílios.

Referências

- AGRESTI, A. *Categorical Data Analysis*. John Wiley & Sons, Second Edition, 2002.
- BETHLEHEM, J., COBBEN, F.; SCHOUTEN, B. *Indicators for the representativeness of survey response*. Statistics Canada Symposium, Gatineau, Canada, 2008.
- COBBEN, F.; SCHOUTEN, B. *An empirical validation of R-indicators*. Discussion Paper 08006, Statistics Netherlands, Voorburg, The Netherlands, 2008.
- HEIJ, V. de.; SCHOUTEN, B.; SHLOMO, N. *RISQ manual 2.1: Tools in SAS and R for the computation of R-indicators, partial R-indicators and partial coefficients of variation*. Representativity Indicators for Survey Quality, 2015.
- IBGE. *Metodologia do Censo Demográfico 2010*. Série Relatórios Metodológicos, 2ª ed., v. 41, 2016.
- KASS, G. V. *An Exploratory Technique for Investigating Large Quantities of Categorical Data*. Applied Statistics, No. 2., p. 119-127, Vol. 29, 1980.
- NIC.br. *Pesquisa sobre o uso das Tecnologias de Informação e Comunicação nos domicílios brasileiros - TIC Domicílios 2016*. Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação (Cetic.br), 2017.
- ONUBR. *Nações Unidas no Brasil*. Disponível em: <https://nacoesunidas.org/pos2015/agenda2030>. Acesso em: out. 2018.
- RAO, J. N. K. *Small Area Estimation*. Wiley Series in Survey Methodology, John Wiley & Sons, Inc., 2003.
- SANTOS, M. P. R. *Indicadores para monitoramento de representatividade no caso da Pesquisa Nacional por Amostra de Domicílios Contínua*. Dissertação (Mestrado em População, Território e Estatísticas Públicas) – Escola Nacional de Ciências Estatísticas (ENCE), Rio de Janeiro, 2017.
- SCHOUTEN, B.; COBBEN, F. *R-indexes for the comparison of different fieldword strategies and data collection modes*. Discussion Paper 07002, Statistics Netherlands, Voorburg, The Netherlands, 2007.
- SCHOUTEN, B.; SHLOMO, N.; SKINNER, C. *Indicators for representative response*. Paper presented at Q2010, Helinski, Finland, 2010.
- SCHOUTEN, B.; BETHLEHEM, J.; BEULLENS, K.; KLEVEN, O.; LOOSVELDT, G.; LUITEN, A.; RUTAR, K.; SHLOMO, N.; SKINNER, C. *Evaluating, Comparing, Monitoring, and Improving Representativeness of Survey Response Through R-Indicators and Partial R-Indicators*. International Statistical Review, p. 382–399, v. 80, 2012.
- SHLOMO, N.; SKINNER, C.; SCHOUTEN, B. *Estimation of an indicator of the representativeness of survey response*. Journal of Statistical Planning and Inference 142, p. 201-211, 2012.
- SKINNER, C.; SHLOMO, N.; SCHOUTEN, B.; ZHANG, L.; BETHLEHEM, J. *Measuring survey quality through representativeness indicators using sample and population based information*. NTTS Conference, Brussels, Belgium, 2009.
- UN. Statistics Division. UN Department of Economic and Social Affairs. *Data Disaggregation and the SDGs: An Overview*. 4th Expert Group Meeting of IAEG-SDG's Working Group on Geospatial Information. New York, USA, 2017.

Apêndice A – Resultados dos modelos

Tabela 6 - Estimativas dos coeficientes, respectivos erros padrão e estatísticas de teste do modelo ajustado para a comparação população versus amostra de respondentes, segundo Grandes Regiões - 2017

(continua)

Grandes Regiões	Estimativa	Erro padrão	Estatística t	P-valor
Brasil				
Intercepto	-0,225	0,162	-1,386	0,166
x ₂ : 650,01 a 950 reais	0,042	0,093	0,458	0,647
x ₂ : 950,01 a 1 500 reais	0,192	0,119	1,613	0,107
x ₂ : Mais de 1 500,00 reais	-0,133	0,150	-0,884	0,377
x ₂ : Sem rendimento	-12,438	0,157	-79,086	<2e-16
x ₁ : 3,01 a 5 moradores	0,193	0,147	1,316	0,188
x ₁ : 5,01 a 7 moradores	-0,432	0,435	-0,994	0,320
x ₁ : Mais de 7,01 moradores	-12,875	0,199	-64,851	<2e-16
Norte				
Intercepto	-0,287	0,512	-0,561	0,575
x ₂ : 650,01 a 950 reais	0,359	0,327	1,097	0,273
x ₂ : 950,01 a 1 500 reais	0,268	0,346	0,775	0,438
x ₂ : Mais de 1 500,00 reais	-0,455	0,466	-0,976	0,329
x ₂ : Sem rendimento	-14,369	0,296	-48,543	<2e-16
x ₁ : 3,01 a 5 moradores	0,369	0,492	0,749	0,454
x ₁ : 5,01 a 7 moradores	-0,338	0,723	-0,467	0,640
x ₁ : Mais de 7,01 moradores	-14,844	0,531	-27,952	<2e-16
Nordeste				
Intercepto	-0,274	0,262	-1,048	0,295
x ₂ : 650,01 a 950 reais	-0,178	0,244	-0,731	0,465
x ₂ : 950,01 a 1 500 reais	-0,112	0,288	-0,388	0,698
x ₂ : Mais de 1 500,00 reais	-0,390	0,271	-1,437	0,151
x ₂ : Sem rendimento	-12,139	0,235	-51,745	<2e-16
x ₁ : 3,01 a 5 moradores	0,267	0,272	0,981	0,326
x ₁ : 5,01 a 7 moradores	-0,595	0,838	-0,710	0,478
x ₁ : Mais de 7,01 moradores	-12,075	0,337	-35,791	<2e-16
Sudeste				
Intercepto	-0,352	0,166	-2,115	0,034
x ₂ : 650,01 a 950 reais	0,114	0,109	1,044	0,297
x ₂ : 950,01 a 1 500 reais	0,267	0,151	1,762	0,078
x ₂ : Mais de 1 500,00 reais	-0,074	0,198	-0,375	0,708
x ₂ : Sem rendimento	-11,022	0,247	-44,579	<2e-16
x ₁ : 3,01 a 5 moradores	0,274	0,138	1,983	0,047
x ₁ : 5,01 a 7 moradores	0,411	0,990	0,415	0,678
x ₁ : Mais de 7,01 moradores	-11,021	0,275	-40,127	<2e-16

Tabela 6 - Estimativas dos coeficientes, repectivos erros padrão e estatísticas de teste do modelo ajustado para a comparação população versus amostra de respondentes, segundo Grandes Regiões - 2017

Grandes Regiões	Estimativa	Erro padrão	Estatística t	P-valor	(conclusão)
Sul					
Intercepto	0,043	0,431	0,100	0,920	
x ₂ : 650,01 a 950 reais	0,016	0,226	0,070	0,944	
x ₂ : 950,01 a 1 500 reais	0,317	0,287	1,105	0,269	
x ₂ : Mais de 1 500,00 reais	0,091	0,338	0,270	0,787	
x ₂ : Sem rendimento	-12,396	0,592	-20,949	<2e-16	
x ₁ : 3,01 a 5 moradores	-0,057	0,322	-0,178	0,859	
x ₁ : 5,01 a 7 moradores	-12,399	0,436	-28,434	<2e-16	
x ₁ : Mais de 7,01 moradores	-12,396	0,655	-18,918	<2e-16	
Centro-Oeste					
Intercepto	-0,828	0,254	-3,266	0,001	
x ₂ : 650,01 a 950 reais	-0,232	0,224	-1,033	0,302	
x ₂ : 950,01 a 1 500 reais	-0,421	0,409	-1,028	0,304	
x ₂ : Mais de 1 500,00 reais	-0,152	0,398	-0,382	0,702	
x ₂ : Sem rendimento	-11,711	0,580	-20,188	<2e-16	
x ₁ : 3,01 a 5 moradores	0,757	0,320	2,365	0,018	
x ₁ : 5,01 a 7 moradores	-13,489	0,304	-44,352	<2e-16	
x ₁ : Mais de 7,01 moradores	-13,327	0,400	-33,358	<2e-16	

Fonte: CGL.br/NIC.br, Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação (Cetic.br), Pesquisa sobre o Uso das Tecnologias de Informação e Comunicação nos domicílios brasileiros - Elaboração dos autores.

Seção 2

População e território

População, gênero e cor/raça na produção cinematográfica brasileira contemporânea*

Paula Alves de Almeida**

José Eustáquio Diniz Alves***

Denise Britz do Nascimento Silva****

Resumo

O cinema influenciou o entretenimento no mundo e fundamentou padrões visuais que passaram a ditar referências de moda, de comportamentos e de estilos de vida, e a forma de construção das representações sociais. O cinema é considerado há muito como objeto de estudo, e também como método de pesquisa, pelas ciências sociais, pela história, psicanálise, filosofia e por outros campos do conhecimento. No entanto, existem pouquíssimos trabalhos sobre as possibilidades de relação entre o cinema e a demografia, cuja maior parte se dedica ao estudo do público de cinema, mas raríssimos se debruçam sobre as equipes ou as personagens – a população que filma e a população filmada – e os processos e meios de produção audiovisual. Este artigo apresenta a composição por sexo e cor/raça de equipes e protagonistas da produção brasileira de longas-metragens das últimas décadas e reflexões sobre a relação entre as representações cinematográficas e os contextos sociais e culturais em que estão inseridas.

Palavras-chave: Demografia. Cinema. Mercado de trabalho. Relações de gênero. Relações raciais.

* Este capítulo é resultado da tese de doutorado concluída em 2019 por Paula Alves de Almeida, intitulada "Cinedemografia, população que filma e população filmada: relações e hierarquias de gênero e raciais na produção cinematográfica brasileira contemporânea", sob orientação do Prof. Dr. José Eustáquio Diniz Alves e da Prof^a. Dr^a. Denise Britz do Nascimento e Silva.

** Pesquisadora no GRAPPA - Grupo de Análises de Políticas e Poéticas Audiovisuais e no GRUA - Grupo de Reconhecimento de Universos Artísticos/Audiovisuais. Diretora no Instituto de Cultura e Cidadania Femina. Concluiu o doutorado em População, Território e Estatísticas Públicas pela ENCE em 2019.

*** Doutor em Demografia e pesquisador independente.

**** Pesquisadora da ENCE e docente do programa de pós-graduação desde sua criação em 1998.

Introdução

O cinema é uma atividade que movimentava altas quantias de dinheiro em cada uma de suas etapas – produção, distribuição, difusão, exibição –, e gera muitos empregos diretos e indiretos. Uma indústria, um meio de comunicação de massa, uma arte, que reproduz as relações sociais e hierárquicas dos contextos em que está inserido, representando “realidades” ou transformando a realidade, mas, com certeza, transmitindo e reforçando valores, preconceitos, perspectivas, categorizações, estereótipos e relações de poder. As imagens em movimento filmadas de forma não cronológica, e posteriormente editadas, provocaram uma desconfiança que influenciou a credibilidade do cinema como registro da realidade e documento histórico. Apesar disto, os antropólogos perceberam o filme como um potente instrumento de etnografia e passaram a utilizar a câmera em seu trabalho de campo, como meio de documentar, guardar e exibir outras culturas. O cinema há muito vem sendo considerado como objeto de estudo pelas ciências sociais, pela história, psicanálise, filosofia e outros campos do conhecimento, no entanto, existem pouquíssimos trabalhos sobre a possibilidade de relação do cinema com a demografia.

O cinema é a arte por excelência da modernidade e veio saciar o desejo da sociedade por uma representação que combinasse com o novo tempo: em movimento acelerado e ávido por novas tecnologias. Também foi no bojo da modernidade que aconteceu a transição demográfica e uma série de mudanças na dinâmica social, espacial e na estrutura etária da população. O cinema e a demografia são campos de ação e disciplinas que caminham paralelamente na contemporaneidade, possuem pontos de interseção, embora suas interligações não sejam plenamente reconhecidas e adequadamente descritas.

Este artigo apresenta uma análise da produção cinematográfica brasileira das últimas décadas, sob a perspectiva demográfica, de gênero e de cor/raça. Através de uma base de dados, elaborada a partir de diferentes fontes, o objetivo é traçar um panorama das equipes e personagens dos filmes – a população que filma e a população filmada.

As “populações” do cinema

Um texto que praticamente inaugura as reflexões sobre a relação do cinema com a demografia é *Pour une ciné-démographie*, de Serge Daney (1997, publicado originalmente em 1988). Para o autor, quando o cinema surgiu, assistir a um filme devia causar uma sensação de pertencimento ao mundo. Os filmes continham estrelas e astros, atores coadjuvantes e figurantes, contavam histórias de multidões, lutas de classes, povos em guerra. Havia, portanto, muitas pessoas nas salas de cinema assistindo a filmes com muitas pessoas nas telas. Para Daney (1997), não foram somente as salas de cinema que se esvaziaram, mas os filmes ficaram cada vez menos povoados, numa crítica ao cinema intimista. Ele sugere então um estudo das populações dos filmes, fazendo-se necessária uma demografia dos seres filmados.

Inspirados na “cinedemografia” de Daney (1997), sugerimos chamar as personagens dos filmes de *população filmada*, e o conjunto de indivíduos que compõem as equipes dos filmes de *população que filma*. Descrever a composição das populações filmada e que filma, assim como as relações entre ambas e com a população real, nos permite refletir sobre as posições e ausências de determinados grupos sociais nas representações cinematográficas. Sobre a possível associação entre a população que filma e a população filmada, podemos supor diferentes relações, como:

- 1) a população que filma se parece com a população filmada, quando acontece o que Daney (1997) sugere, o cineasta se volta para os grupos sociais aos quais pertence;
- 2) a população que filma reinterpreta a população filmada, no caso do cineasta não filmar seus semelhantes, mas “o outro”; e
- 3) as ausências, quando determinadas populações ficam de fora tanto da população que filma, quanto são excluídas das narrativas cinematográficas.

No primeiro caso, quando os cineastas passam a representar a si próprios ou a seus semelhantes, pode-se imaginar que a população filmada se assemelhe à população que filma. A questão é se essas duas populações incorporam a diversidade da população real, e se esta é representada em sua pluralidade pelos filmes. No segundo caso, quando os cineastas se voltam para outros grupos sociais aos quais não pertencem, os filmes podem privilegiar as perspectivas dos diretores e retratar temáticas e contextos sociais a partir de seus valores e preconceitos. E as ausências acontecem quando grupos sociais estão sub-representados nas equipes dos filmes, ou seja, no mercado de trabalho do cinema, e também nas telas. Novamente, vale questionar se partes da população também não estariam de fora das representações cinematográficas, e se não seriam as mesmas que são excluídas dos cargos de comando no mundo social e político.

O cinema foi um dos primeiros e, ainda é, um dos maiores produtores de padrões visuais de massa. Sempre desempenhou importante papel na construção e difusão de imagens de gênero, de pertencimento racial ou étnico, de classe e outras distinções sociais, especialmente considerando-se o poder de alcance do cinema clássico narrativo¹ em quase todo o mundo. O cinema tem papel fundamental na reprodução de valores, costumes e divisão de papéis tradicionais entre os atores sociais e, por outro lado, na transformação desses valores e sugestão de novas divisões (ALVES, 2011).

Num mundo cada vez mais transpassado por informação audiovisual, o domínio da imagem e do som por determinados grupos sociais permite sua manutenção no centro de decisão e poder na sociedade contemporânea. Por isso, o fato de uma parcela da população permanecer sem representatividade e representação, seja nas equipes ou nas telas do cinema brasileiro, indica ser fundamental e urgente a discussão sobre as desigualdades de gênero e cor/raça, utilizando o cinema para uma análise da sociedade que o produz.

Relações e hierarquias de gênero e raciais no cinema

Assim como nos espaços de poder e em outras áreas do mercado de trabalho – nas quais a presença feminina em cargos de direção e gerência ainda é limitada, – também no cinema a participação da mulher desempenhando funções-chave (como direção, roteiro, produção, entre outras) ainda está expressivamente abaixo da presença masculina (ALVES, 2011). Nos campos social e político, quanto mais avançadas são as ocupações em termos de prestígio e recursos financeiros, maiores os limites impostos a mulheres e negros à sua plena participação.

O cinema foi criado e estruturado no contexto de uma sociedade desigual e androcêntrica. Consequentemente, a representação da mulher e a construção das personagens femininas foram elaboradas, essencialmente, a partir da perspectiva masculina. A representação do corpo feminino pelo cinema como objeto de consumo, e a relação não incomum da imagem feminina aos papéis de menor prestígio social, reafirmam a distinção de representações de homens e mulheres, não só refletindo a sociedade como influenciando-a, num círculo vicioso (ALVES; ALVES; SILVA, 2011).

Segundo Buet (1999), nas primeiras décadas do cinema, as mulheres estavam presentes principalmente como atrizes e assistentes, mas eram raras² na direção cinematográfica. De acordo com a autora, o número de diretoras começa a crescer consideravelmente na Europa nos anos 1960 e 1970.

Stacy Smith (SIGNIFICANCE, 2018) analisou a diferença de gênero entre os 100 filmes norte-americanos de maior bilheteria em 2017. Segundo sua pesquisa, 33% desses filmes foram protagonizados ou coprotagonizados por mulheres, 7,3% dirigidos por mulheres, 10,1% tiveram os roteiros assinados por mulheres, e 18,2% foram produzidos por mulheres. Segundo a pesquisadora, os filmes com diretoras ou roteiristas mulheres tendiam a ter mais personagens femininas.

¹ O cinema clássico narrativo se refere normalmente ao cinema hollywoodiano que segue determinadas normas e uma estrutura narrativa linear. Como o cinema estadunidense influenciou os cinemas de todo o mundo, a expressão pode se referir não somente aos filmes norte-americanos, mas a todos que sigam o mesmo padrão. O cinema clássico narrativo é apontado como o responsável pelo sucesso e hegemonia do cinema hollywoodiano no mundo e pela criação da indústria do cinema como a conhecemos hoje.

² Comparativamente com os homens e com a presença de mulheres em outras funções, as diretoras de cinema eram raras. Mas, evidentemente, havia exceções como as pioneiras Lois Weber, Germaine Dulac e Alice Guy Blaché. Esta última é apontada como a **primeira pessoa que realizou um filme de ficção narrativo** (em 1896), apesar de por mais de um século a história do cinema ter creditado esse feito a homens.

Martha Lauzen (2019) realizou pesquisa semelhante com os 250 filmes de maior bilheteria nos EUA. Segundo sua pesquisa, em 2018, as mulheres representaram 8% de diretores, 26% entre os produtores, 21% dos editores, 16% de roteiristas, e 4% entre os cinegrafistas.

As relações sociais e de poder no mundo diegético³ reproduzem as relações encontradas na sociedade, tanto quanto contribuem para uma reorganização da estrutura social quando elaboram discursos que reconfiguram as referências e os códigos simbólicos (ALVES; COELHO, 2015). O padrão de beleza instituído pelo cinema hollywoodiano, através do *star system*, disseminou imagens de heróis e mocinhas que passaram a ser ideais desejados, e inalcançáveis, para as mulheres e os homens na sociedade. Esses padrões limitam não só as carreias de atores e atrizes, mas influenciam autoestimas, a vivência da sexualidade, o mercado de trabalho, as relações afetivas e sociais. Mulheres brancas e negras, por exemplo, no cinema brasileiro ocupam papéis marcados por diferentes estereótipos. Ceiza Ferreira (2016) destaca que a mulher negra está representada nos filmes nacionais notadamente por personagens sexualizadas ou trabalhadoras domésticas, quando não simplesmente ausente.

Conforme demonstrado nos estudos de Luis Felipe Hirano (2013), há uma enorme desigualdade racial na distribuição de cargos e funções no cinema brasileiro, que segue as desigualdades existentes no acesso à educação, saúde, justiça, trabalho, etc. Para este autor, o cinema brasileiro produz e reproduz as desigualdades e hierarquias entre negros e brancos existentes em nossa sociedade, reforçando as manifestações do racismo como ele atua no país, através da gradação da cor de pele. João Carlos Rodrigues (2011) afirma que, de modo geral, o cinema brasileiro, em seus diversos estilos e épocas, não tem refletido e captado a verdadeira realidade da população negra, que comumente é apresentada como personagens arquetípicos. O autor afirma também que a presença de negros na direção, roteiro e protagonismo no cinema brasileiro é ainda muito reduzida.

A presença de indígenas na produção cinematográfica brasileira é *ainda mais rara e* também marcada por estereótipos. Caio Lazaneo (2015) discorre sobre o que poderia ser considerado cinema indígena. Segundo ele, não basta um filme abordar temáticas indígenas, mas ser de autoria de um realizador indígena ou de autoria de não indígena de algum modo negociada. De toda forma, a presença indígena no cinema brasileiro é muito pequena mesmo considerando-se o total dos filmes sobre indígenas realizados por não indígenas e os cinemas indígenas segundo esta definição.

Para Petrucelli (2013a), o pensamento racial brasileiro reflete a maneira como a ideologia racial determinou desigualdades nas relações sociais a serviço da colonização e da escravidão, definindo as posições e espaços hierarquizados de grupos na sociedade segundo a cor/raça/etnia a eles atribuídos por quem os observa. Segundo esta definição, a atribuição de cor/raça estaria baseada no olhar de terceiros e está relacionada inicialmente à necessidade de legitimação da exploração de povos não brancos.

Partindo do postulado formulado por Bourdieu (1980 apud PETRUCCELLI, 2013a), de que a realidade social e cultural é relacional, Petrucelli (2013a, p. 20) conclui que é preciso passar a conceber as práticas discriminatórias baseadas em propriedades biológicas ou culturais, supostamente inerentes aos indivíduos ou grupos, como “produtos de conflitos simbólicos” entre indivíduos e grupos em posições desiguais. A atribuição de cor ou raça a alguém, portanto, é um processo relacional (e arbitrário) entre indivíduos em disputa de posições.

Da mesma forma, para Edward Telles (2003), as atribuições raciais feitas por terceiros determinam discriminações, oportunidades na vida das pessoas e as formas como se darão as interações sociais. Para o autor, as desigualdades de poder e privilégios nas interações sociais baseadas em raça ou cor de pele não derivam apenas de um passado colonial e escravocrata, ou de grandes desigualdades de classe, mas resultam também de uma contínua prática social através da reprodução de estereótipos pela mídia, redes sociais, pelo sistema educacional, por práticas de consumo, pelas políticas públicas, etc. Para Telles, a discriminação racial

³ Diegese é o lugar ou a dimensão (o universo espaço-temporal) onde se passa a narrativa do filme, onde vivem as personagens e se desenvolvem as ações, é o “mundo do filme”.

no Brasil funciona como um poderoso mecanismo que mantém uma enorme desigualdade social e permite que um grupo tenha privilégios perante os demais.

De acordo com Petrucci (2013b), os estudos sobre categorias de classificação de cor ou raça se intensificaram especialmente após a realização da III Conferência Mundial de Combate ao Racismo, Discriminação Racial, Xenofobia e Intolerância Correlata, realizada em Durban, África do Sul, em 2001. O IBGE realizou, em 2008, a Pesquisa das Características Étnico-raciais da População (PCERP)⁴ que, além da autoclassificação do entrevistado, incluiu um quesito no qual o entrevistador realizava a classificação de cor/raça da pessoa entrevistada. Autoclassificação é a atribuição de uma categoria de cor ou raça escolhida pela própria pessoa. *Alter* ou heteroclassificação é a atribuição por outra pessoa de uma categoria de cor ou raça a alguém. Petrucci (2013b) registra que alguns estudos apontam para discrepâncias entre a auto e a heteroclassificação, enquanto outros apontam bons níveis de concordância entre ambas. No caso da PCERP 2008, a comparação entre a distribuição das respostas de auto e de heteroclassificação para o conjunto das Unidades da Federação (UFs) investigadas mostrou um alto grau de consistência.

Saperstein (2008 apud Muniz, 2010) comparou a cor/raça autodeclarada com a atribuída pelo entrevistador em uma pesquisa nos Estados Unidos e encontrou uma consistência entre as duas formas de atribuição de aproximadamente 96%. A autora defende uma abordagem multidimensional para a mensuração da cor/raça, não apenas a utilização de uma única medida. Ela realizou uma análise de classes latentes para demonstrar que a raça não é somente como nos percebemos, mas também como os outros nos percebem.

Segundo Osorio (2003 apud PETRUCCELLI, 2013b), a diferença entre a auto ou a heteroclassificação de cor ou raça seria a diferença entre a subjetividade da própria pessoa classificada e a subjetividade da pessoa que a classifica. Portanto, as duas formas de classificação seriam subjetivas. Para Campos e Longo (2006), mais de 70% das declarações de cor e/ou raça em pesquisas domiciliares no Brasil tidas como autodeclarações, na realidade são *alter* declarações, já que são feitas por terceiros – pessoas da família ou não, presentes no domicílio e respondentes das pesquisas.

Diante da crescente discussão sobre a representação de não brancos pelo cinema no mundo, consideramos fundamental a inclusão desta dimensão em nossa análise. Todavia, as informações sobre cor/raça de profissionais da produção cinematográfica brasileira não estão disponíveis em nenhuma fonte de dados. Conforme exposto, a heteroatribuição é mais comum do que se imagina, reflete as raízes do surgimento da ideia de raça e o desenvolvimento do racismo – que está relacionada a como terceiros veem os indivíduos e não como os próprios indivíduos se veem – e se faz, portanto, uma alternativa nem inferior, nem superior, mas diferenciada da autodeclaração. Sendo assim, realizamos a classificação da cor/raça dos indivíduos que desempenharam funções-chave nos filmes de nossa base de dados por heteroatribuição.

Métodos e estratégia de pesquisa

Para avaliar empiricamente as desigualdades comentadas anteriormente, foi elaborada uma base de dados com cerca de 4.950 filmes brasileiros de longa-metragem (com 60 minutos ou mais), lançados entre 1960 e 2017, contendo variáveis como: sexo e cor/raça de diretores, protagonistas e roteiristas, sexo de produtores, editores, fotógrafos, diretores de arte e outras funções-chave, principais temáticas abordadas pelos filmes, informação se o filme aborda temáticas demográficas (e quais), ano de produção e lançamento, UF sede da empresa produtora majoritária, gênero cinematográfico, valores captados por mecanismo de incentivo, público e renda de bilheteria em salas comerciais, duração em minutos, sinopse e outras características dos filmes.

As principais fontes de dados consultadas foram: a Agência Nacional do Cinema - ANCI-NE, o portal sobre o mercado exibidor brasileiro Filme B, o *Dicionário de Filmes Brasileiros - Longa*

⁴ Para saber mais sobre a Pesquisa das Características Étnico-raciais da População (PCERP): <<https://www.ibge.gov.br/estatisticas-novportal/sociais/populacao/9372-caracteristicas-etnico-raciais-da-populacao.html?=&t=o-que-e>>.

Metragem, de Antônio da Silva Neto (2009); o *Dicionário de Cinema Brasileiro*, de Mauro Baladi (2013), a lista de filmes lançados em festivais e mostras de cinema do pesquisador Eduardo Valente (não publicada), além de catálogos e sites de festivais de cinema, de empresas produtoras e distribuidoras, portais especializados em cinema, materiais de divulgação, e os próprios filmes.

Variáveis como o sexo dos protagonistas, temáticas abordadas pelo filme, gênero cinematográfico, entre outras foram classificadas pela primeira autora deste artigo com base nas sinopses e visionamento dos filmes⁵. O sexo dos indivíduos que desempenharam as funções selecionadas foi atribuído a partir do nome e foto dos mesmos. O sexo é classificado segundo três categorias de resposta: “homens” e “mulheres” quando a função foi desempenhada exclusivamente por homens ou por mulheres, e “ambos” para funções desempenhadas por homens e mulheres conjuntamente.

Desta forma, foi possível classificar as funções na direção, roteiro, produção ou fotografia em três categorias de resposta: “homens”, “mulheres” e “ambos”; ou em duas categorias, neste caso de duas maneiras diferentes: “homens” ou “mulheres”, não considerando a categoria “ambos”; ou “homens” ou “mulheres + ambos” (que também podemos chamar de “pelo menos uma mulher”).

Preparamos uma massa de teste para avaliar e realizar ajustes no processo de heteroclassificação de cor/raça dos profissionais da produção cinematográfica brasileira. A partir de uma população de controle de 463 indivíduos da qual tivemos acesso, num primeiro momento, somente a fotos digitalizadas numeradas sequencialmente, fez-se a heteroclassificação. Posteriormente tivemos acesso às autodeclarações de pertencimento racial desses mesmos indivíduos. Comparou-se então a heteroclassificação realizada com a autodeclaração para mensurar e analisar as concordâncias e discrepâncias. A concordância observada entre a hetero e a autoclassificação foi de 95,6%. Para avaliar estatisticamente o grau de concordância entre as duas medidas, utilizamos o coeficiente de concordância *Kappa*. A estatística *K* fornece evidências para indicar se a concordância existe (é legítima), ou se pode ser considerada aleatória (ao acaso), a partir do cálculo da concordância observada e da concordância esperada sob a hipótese de independência entre duas medidas de uma variável nominal (PAES; SILVA, 2012). A estatística *K* varia de valores menores que 0 a 1, sendo que valores abaixo de 0,20 indicam não haver concordância além do acaso, e valores próximos de 1 representam a concordância perfeita⁶. Vários autores classificam valores acima de 0,80 como concordância quase perfeita ou excelente. Registra-se que o resultado encontrado para o índice de *Kappa* no experimento realizado foi de 0,89, concluindo-se em favor da existência de concordância entre a auto e a heteroclassificação de cor/raça no caso em questão.

Após os testes com a população controle, elaboramos uma base de fotos dos diretores dos filmes estudados, a partir das quais realizou-se a heteroclassificação de cor/raça. Outros pesquisadores foram convidados a fazer também a atribuição de cor/raça dos diretores a partir das mesmas fotos. O intuito era não apenas comparar a classificação realizada pela autora com a dos demais pesquisadores, mas permitir encontrar uma heteroatribuição modal a fim de reduzir vieses de classificação. O grupo de 13 colaboradores foi formado por estudantes de graduação ou pós-graduação, mestres e doutores, e contou com homens e mulheres, que se declararam brancos, pardos, pretos, amarelos e indígenas (residentes em diferentes UFs nas regiões Nordeste, Centro-Oeste e Sudeste). Para a classificação, foram consideradas as categorias de cor/raça adotadas pelo IBGE: “cor amarela”, “cor branca”, “etnia indígena”, “cor parda” e “cor preta”.

Para obtenção de alguns resultados e análises, trabalhamos com uma seleção de filmes cujos critérios foram a renda de bilheteria e o público em salas comerciais. Foram selecionados os 30 filmes que obtiveram maior renda de bilheteria por ano em salas comerciais no país entre 1995 e 2016, os 30 filmes que atingiram maior público por ano no circuito

⁵ O trabalho de classificação de variáveis e visionamento dos filmes foi realizado apenas pela autora Paula Alves de Almeida.

⁶ Um valor de *Kappa* menor que zero sugere que a concordância encontrada foi menor do que a esperada pelo acaso, no entanto, um valor negativo não tem interpretação em termos de intensidade de discordância.

comercial, os 600 filmes com maior renda de bilheteria⁷ no total do período, e os 600 filmes com maior público em exibições comerciais no total do período. O conjunto selecionado é composto por 688 filmes que foram dirigidos por 419 diretores, considerando que alguns filmes foram codirigidos por mais de um diretor, e um mesmo diretor pode ter dirigido mais de um filme no período estudado. Neste artigo são apresentados os resultados para este conjunto de filmes referentes às variáveis: sexo/gênero de diretores, roteiristas e protagonistas, cor/raça dos diretores classificada pela autora e pelo grupo de colaboradores, cor/raça dos protagonistas e roteiristas (classificados apenas pela autora).

Resultados

Apresentamos na Tabela 1 os resultados referentes aos 419 diretores dos 688 filmes de maior público e renda de bilheteria em salas comerciais, por sexo e cor/raça atribuída pela hetero-classificação realizada pela autora e pelo grupo de colaboradores. A partir das classificações atribuídas pelos pesquisadores convidados, verificamos a atribuição modal, ou seja, a categoria de cor/raça que ocorreu com maior frequência para cada indivíduo classificado. Não houve divergências entre as atribuições realizadas pela autora e a modal nas categorias amarela e indígena. Nas categorias branca e parda a divergência foi de apenas 1 diretor/a em cada, e na categoria preta de 2 diretores. A concordância observada foi de 99,5% e o valor da estatística K foi 0,96, indicando uma concordância excelente.

Tabela 1 - Distribuição percentual dos diretores dos filmes brasileiros de maior público e bilheteria, por sexo e heteroatribuição, segundo cor/raça - Brasil - 1995-2016

Cor/raça	Distribuição percentual dos diretores dos filmes brasileiros de maior público e bilheteria (%)					
	Heteroatribuição					
	Autora			Modal		
	Total	Homem	Mulher	Total	Homem	Mulher
Total	100,0	85,3	14,7	100,0	85,3	14,7
Amarela	0,4	0,2	0,2	0,4	0,2	0,2
Branca	93,1	79,5	13,6	93,4	79,8	13,6
Indígena	0,0	0,0	0,0	0,0	0,0	0,0
Parda	4,8	4,1	0,7	5,0	4,3	0,7
Preta	1,0	1,0	0,0	0,5	0,5	0,0
Sem informação	0,7	0,5	0,2	0,7	0,5	0,2
Diretores	419	357	62	419	357	62

Fontes: ADOROCINEMA; ACADEMIA BRASILEIRA DE CINEMA; ANCINE, 2018; BALADI, 2013; CINEMATECA BRASILEIRA; FEMINA; FILME B, 2018; FILMOW; MOSTRA INTERNACIONAL DE CINEMA; SILVA NETO, 2009.

Ambas as classificações (realizadas pelos colaboradores e pela autora) apontam uma enorme desigualdade tanto de gênero como de cor/raça entre os diretores brasileiros que realizaram longas-metragens nas últimas décadas, sendo quase 80% deles homens brancos.

Também analisamos as informações sobre o sexo/gênero e a cor/raça dos roteiristas deste mesmo conjunto de 688 filmes de longa-metragem com maior bilheteria e público no circuito comercial no país entre 1995 e 2016. Neste caso, contamos apenas com a heteroclassificação da auto-

⁷ Os valores das rendas de bilheteria foram corrigidos segundo o Índice Nacional de Preços ao Consumidor Amplo – IPCA para o ano 2018.

ra, cujos resultados são apresentados na Tabela 2. Os homens representaram 76% dos roteiristas desse conjunto de filmes, sendo que quase 70% deles classificados como brancos. As posições de roteiro e direção são muito próximas no cinema independente (fora dos grandes estúdios estadunidenses), algumas vezes sendo desempenhadas pelos mesmos profissionais. Sendo assim, é de se esperar uma distribuição de sexo e cor/raça semelhante dos indivíduos nessas funções. No período estudado, o total de pessoas classificadas como brancas na direção foi de cerca de 93% (Tabela1), enquanto que no roteiro obteve-se valor similar (92%). Identificamos uma pessoa transgênero assinando um dos roteiros dessa seleção de filmes (uma mulher trans).

Tabela 2 - Distribuição percentual dos roteiristas dos filmes de maior público e bilheteria, por sexo/gênero, segundo cor/raça - Brasil - 1995-2016.

Cor/raça	Distribuição percentual dos roteiristas dos filmes de maior público e bilheteria (%)			
	Total	Sexo/Gênero		
		Homem	Mulher	Trans
Total	100,0	76,0	23,9	0,1
Amarela	0,4	0,3	0,1	0,0
Branca	91,9	69,7	22,1	0,1
Indígena	0,0	0,0	0,0	0,0
Parda	3,8	2,9	0,9	0,0
Preta	1,2	1,2	0,0	0,0
Sem informação	2,7	1,9	0,8	0,0
Roteiristas	752	571	180	1

Fontes: ADOROCINEMA; ACADEMIA BRASILEIRA DE CINEMA; ANCINE, 2018; BALADI, 2013; CINEMATECA BRASILEIRA; FEMINA; FILME B, 2018; FILMOW; MOSTRA INTERNACIONAL DE CINEMA; SILVA NETO, 2009.

Nota: Cor/raça dos roteiristas atribuídos pela autora

As Tabelas 1 e 2 acima apresentam os dados de direção e roteiro tendo como unidade de análise os indivíduos nessas funções. Na Tabela 3 a seguir apresentamos o número de filmes por sexo/gênero e cor/raça dos protagonistas. Além das categorias “homem” e “mulher”, observamos a categoria “ambos” para quando a função foi desempenhada por homens e mulheres conjuntamente. A categoria que denominamos de “trans” agrega as personagens transgênero, transexuais, travestis e não gênero. A atribuição de cor/raça das personagens considerou as seguintes categorias: amarela, branca, indígena e negra (pardos + pretos). É importante ressaltar que nem sempre a cor/raça das personagens é a mesma das atrizes e atores que as interpretam. Dos filmes com protagonistas homens indígenas: um filme tem um ator não indígena representando a personagem indígena, um tem atores indígenas e não indígenas representando indígenas, e um filme é de animação (a imagem do protagonista é um desenho, e a voz é de um ator não indígena). Dos filmes com protagonistas mulheres indígenas: três filmes apresentam indígenas de fato e um filme tem uma atriz não indígena representando a personagem indígena. Foram encontrados dois filmes com personagens importantes indígenas, mas não protagonistas, sendo ambos representados por atores não indígenas, e quatro filmes com temática indígena, mas não protagonizados por indígenas. Da mesma forma, há quatro filmes com personagens importantes amarelos, mas que não são protagonistas.

Encontramos uma participação maior de negros entre os protagonistas do que entre os diretores e roteiristas da mesma seleção de filmes. Considerando as categorias parda e preta das Tabelas 1 e 2, observa-se 5,8% (classificação da autora) dos diretores e 5% dos roteiristas classificados como negros. De toda forma, tanto a população que filma (representada aqui pelos diretores e roteiristas) quanto a população filmada (representada pelos protagonistas), nessa seleção de filmes, são majoritariamente brancas. Lembrando que estamos considerando os filmes

de maior público e bilheteria das últimas décadas, ou seja, os filmes mais vistos (ao menos pelos espectadores do circuito comercial). Se a grande maioria dos protagonistas no cinema brasileiro é branca, podemos afirmar que uma expressiva parcela da população brasileira (e dos espectadores) não se reconhece nos filmes. O protagonista é o elemento central de um filme, com o qual as demais personagens se relacionam, a partir de quem acompanhamos a história contada, a personagem que tem mais tempo de tela, e conseqüentemente, mais tempo de fala – uma das questões mais importantes na utilização da análise de protagonistas como *proxy* da representação de grupos sociais pelo cinema. Normalmente, os filmes contam a história de um ou mais protagonistas. Portanto, é essencial refletir sobre o sexo/gênero e a cor/raça dos protagonistas no tocante à identificação dos espectadores com o filme, e como aspecto da sensação de pertencimento ao mundo.

Tabela 3 - Distribuição percentual de filmes de longa-metragem brasileiros, por gênero, segundo cor/raça dos protagonistas - Brasil - 1995-2016

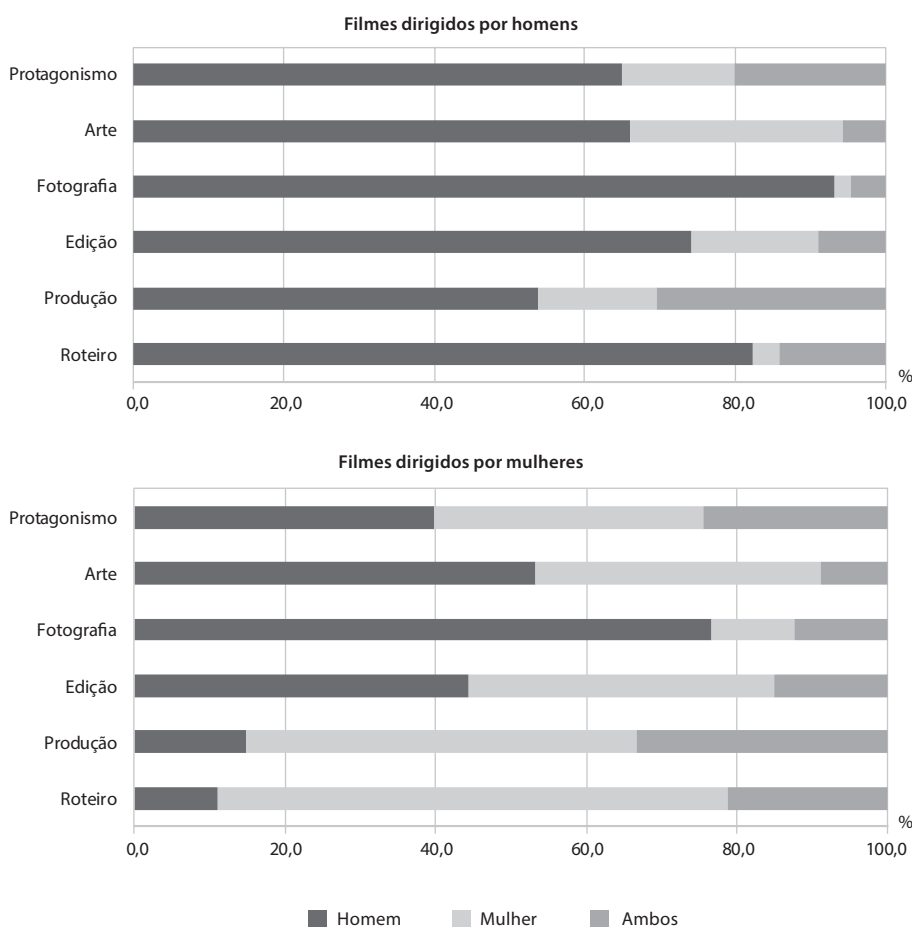
Cor/raça dos protagonistas	Distribuição percentual de filmes de longa-metragem brasileiros (%)				
	Total	Gênero			
		Homem	Mulher	Ambos	Trans
Total	100,0	59,7	23,3	16,8	0,2
Amarela	0,3	0,1	0,1	0,1	0,0
Branca	81,6	48,7	20,1	12,7	0,1
Indígena	1,0	0,4	0,6	0,0	0,0
Negra	17,1	10,5	2,5	4,0	0,1
Protagonistas	688	411	160	115	2

Fontes: ADOROCINEMA; ACADEMIA BRASILEIRA DE CINEMA; ANCINE, 2018; BALADI, 2013; CINEMATECA BRASILEIRA; FEMINA; FILME B, 2018; FILMOW; MOSTRA INTERNACIONAL DE CINEMA; SILVA NETO, 2009.

Nota: Cor/raça dos protagonistas atribuída pela autora.

Apresentamos no Gráfico 1 a distribuição por sexo nas funções selecionadas nos filmes dirigidos por homens e mulheres. A expressão arte está associada à função de diretor de arte, fotografia à direção de fotografia e operadores de câmera, já a denominação produção indica produtores, produtores executivos e diretores de produção. Os resultados se referem ao número total de filmes lançados entre 1991 e 2016 (2.619 filmes). O gráfico aponta associação entre o sexo dos indivíduos que desempenham posições destacadas na produção do conjunto de filmes analisados com o sexo dos diretores desses filmes. As mulheres estão mais presentes como roteiristas, produtoras, editoras, fotógrafas, diretoras de arte e protagonistas nos filmes dirigidos por mulheres. Por outro lado, a participação dos homens é maior nessas funções nos filmes dirigidos por homens.

Com exceção da produção e do roteiro, os homens são maioria nas demais funções também nos filmes dirigidos por mulheres. Além disso, a participação de homens em todas as funções nos filmes dirigidos por mulheres é maior do que a participação das mulheres em todas as funções nos filmes dirigidos por homens. Da mesma forma, a porcentagem de filmes com as funções assinadas por ambos é maior nos filmes dirigidos por mulheres do que nos filmes dirigidos por homens. Ou seja, os filmes dirigidos por homens priorizam mais homens em cada uma das demais funções e nas equipes. Verifica-se que 72,8% do total das funções-chave analisadas nos filmes dirigidos por homens foram desempenhadas exclusivamente por homens, 15,1% por homens e mulheres, e 12,1% somente por mulheres, enquanto que nos filmes dirigidos por mulheres a distribuição das funções por sexo é menos desigual: 41,6% do total das funções-chave analisadas neste caso foram desempenhadas somente por mulheres, 37,6% por homens, e 20,8% por homens e mulheres conjuntamente.

Gráfico 1 - Distribuição percentual dos filmes de longa-metragem, segundo o sexo nas funções-chave e o sexo dos diretores - Brasil - período 1991-2016

Fontes: ADOROCINEMA; ACADEMIA BRASILEIRA DE CINEMA; ANCINE, 2018; BALADI, 2013; CINEMATECA BRASILEIRA; FEMINA; FILME B, 2018; FILMOW; MOSTRA INTERNACIONAL DE CINEMA; SILVA NETO, 2009.

Adicionalmente, nossos resultados apontam relação entre a cor/raça dos diretores dos filmes da amostra com a cor/raça dos roteiristas e protagonistas. Dos filmes dirigidos por pessoas classificadas como amarelas, 16,7% tinham funções-chave desempenhadas por pessoas classificadas como amarelas, 75% por brancas, 8,3% por brancas e negras conjuntamente. Dos filmes dirigidos por pessoas classificadas como brancas, 87,8% das funções-chave foram desempenhadas por pessoas brancas, 8,7% por negras (soma de pardas e pretas), 2,7% por brancas e negras conjuntamente, 0,4% por indígenas, 0,2% por amarelos, e 0,2% por brancas e amarelas conjuntamente. E dos filmes dirigidos por pessoas classificadas como pardas e pretas que agregamos na categoria negra, temos 50% de funções-chave desempenhadas por pessoas brancas, 21,5% por pessoas negras, 28,5% por brancas e negras.

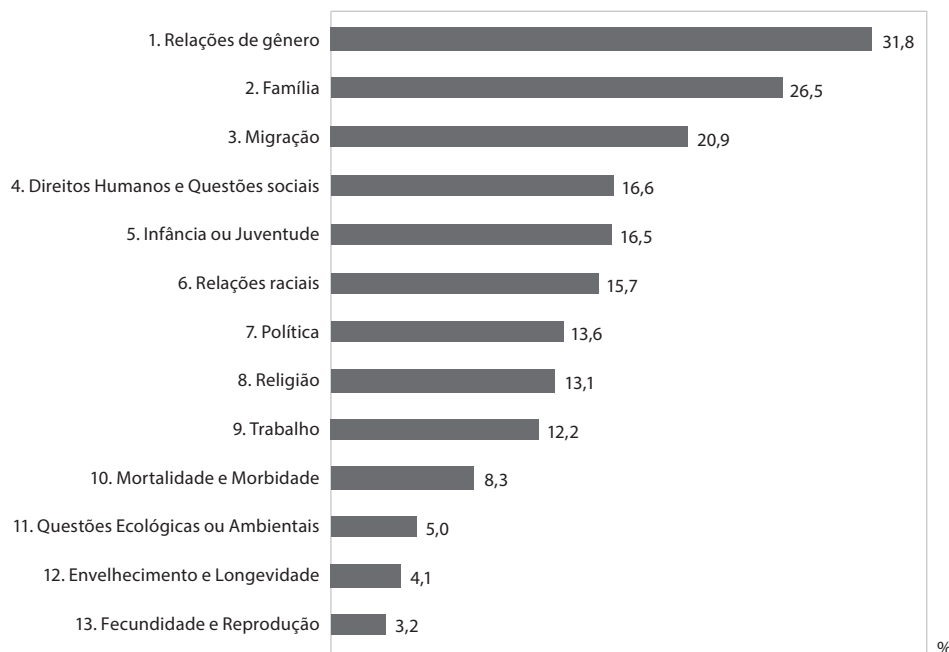
Apenas para fins de reflexão, comparamos o percentual de participação dos diretores por heteroclassificação de cor/raça nos filmes brasileiros de maior público e renda em salas comerciais, lançados entre 1995 e 2016, com a distribuição da população brasileira quanto à autoclassificação de cor/raça segundo os Censos Demográficos 1991, 2000 e 2010 e a estimada na Pesquisa Nacional por Amostra de Domicílios 2015 (IBGE). Enquanto a população que se autoclassificou como parda variou de 43% a 45% entre 1991 e 2015, apenas 4,8% dos diretores que realizaram filmes de longa-metragem no Brasil foram classificados como pardos (Tabela 1 - classificação da autora). A parcela da população que se autodeclarava como preta variou de 5% a 9% no período indicado, enquanto 1% dos diretores de cinema foram assim classificados (Tabela 1 - classificação

da autora). Consequentemente, a mesma disparidade é evidenciada no que se refere a pessoas de cor/raça branca que constituem cerca de 93% dos diretores de longas-metragens enquanto a participação da população autodeclarada branca variou de 52% a 45% de 1991 a 2015.

Foram ajustados modelos estatísticos para dados categóricos para verificar a associação entre o sexo dos indivíduos que exercem a direção e o sexo das personagens protagonistas dos filmes brasileiros. Incluímos também no modelo o gênero cinematográfico. Nos filmes de documentário, as chances da protagonista ser mulher é 370% maior quando a diretora é mulher em relação aos filmes dirigidos por homens. Já nos filmes de ficção, quando a diretora é mulher as chances da protagonista ser feminina aumentam em 60% em relação aos filmes de diretores homens. Este resultado comprova o que encontramos nas análises descritivas que mostraram que as mulheres estão mais presentes como protagonistas nos filmes de ficção do que nos de documentário. O que o modelo nos mostra é que, nas ficções, o sexo dos diretores influencia menos o sexo dos protagonistas do que nos documentários. Ou seja, exatamente nos filmes onde as mulheres não têm o papel principal, ter uma diretora pode aumentar o protagonismo feminino.

Também procuramos avaliar se, e o quanto, o cinema brasileiro aborda temáticas demográficas. Para tanto criamos variáveis na base de dados que correspondem a questões estudadas pela demografia, com as opções de resposta: “principal” – indicando que o filme aborda aquela temática de interesse como uma de suas temáticas principais, “secundária” – indicando que o filme aborda aquela temática, mas não como uma de suas principais, e “não” para o caso do filme não abordar. A ordenação das 13 temáticas pesquisadas quanto à abordagem nos longas-metragens de maior público e renda de bilheteria no circuito comercial, entre 1995 e 2016, é apresentada no Gráfico 2.

Gráfico 2 - Percentual de filmes brasileiros de maior público e bilheteria, que abordam cada uma das 13 temáticas pesquisadas - Brasil - período 1995-2016



Fontes: ADOROCINEMA; ACADEMIA BRASILEIRA DE CINEMA; ANCINE, 2018; BALADI, 2013; CINEMATECA BRASILEIRA; FEMINA; FILME B, 2018; FILMOW; MOSTRA INTERNACIONAL DE CINEMA; SILVA NETO, 2009.

No que se refere às temáticas que mais aparecem nos filmes, destacamos que o tema relações de gênero é abordado como temática principal em 15,5% dos filmes, e como temática secundária em 16,3%. O tema família aparece como temática principal em 8,1% dos filmes e como secundária em 18,4%, e o tema migração se destaca em 6,5% dos filmes como temática principal, enquanto que em 14,4% como temática secundária.

Para definir como e quais questões de cada assunto são abordadas nos filmes, analisamos com mais detalhes cada uma das temáticas. No caso dos filmes que abordam a temática relações de gênero, 61,7% têm uma abordagem feminista ou se referem especificamente à mulher, 16,7% têm temática ou protagonista gay ou lésbica, 3,1% têm temática ou protagonista transgênero, transexual, travesti ou não gênero, 15,4% se referem à sexualidade no sentido de prazer e/ou prática sexual, 15,4% abordam a violência contra a mulher, 1,2% abordam homofobia, 9,3% abordam violência sexual/estupro, e 1,9% abordam masculinidades ou violência contra homens (esses percentuais somam mais que 100% porque um filme pode abordar mais de um aspecto).

Examinando os filmes que abordam o tema migração, concluímos que 35,8% abordam movimentos migratórios nacionais, 48,6% abordam movimentos internacionais, 6,4% movimentos relacionados ao processo de colonização, 33,0% movimentos motivados por questões econômicas e/ou trabalho, 14,7% migração de retorno, 9,2% questões relacionadas a exílio ou refúgio. Daqueles que abordam movimentos migratórios nacionais, 30,8% se referem a movimentos da região Nordeste para o Sudeste, 23,1% movimentos rural-urbano, 43,6% movimentos de cidades pequenas para cidades grandes, indicando como os cineastas estão atentos às questões que permeiam a sociedade brasileira e representam fatos históricos, demográficos e sociais.

Consideramos também a abordagem das temáticas selecionadas pelos filmes de maior público e renda de bilheteria, lançados de 1995 a 2016, por quinquênio, para observar se houve alguma mudança ao longo desses 22 anos. Podemos dizer que a abordagem de temáticas como relações de gênero, fecundidade e reprodução, envelhecimento e longevidade, infância e juventude permaneceu praticamente estável no período analisado. Temáticas como migração, relações étnico-raciais e religião passaram a ser menos abordadas nos filmes lançados no último quinquênio do que no início do período analisado. Por outro lado, temas como família, mortalidade e violência se tornaram mais presentes nos filmes lançados no final do mesmo período. Observamos também se houve alteração na abordagem da temática religião ao longo das últimas décadas. Verificamos que os filmes passaram a abordar mais as religiões evangélicas e espíritas, as religiões de matriz africana tiveram uma leve queda, enquanto que as religiões católicas perderam protagonismo. É importante citar a transição religiosa no Brasil descrita por Alves et al (2014), cuja principal característica é o aumento do número de filiações às religiões evangélicas e a queda das pessoas que se declaram católicas. Mais uma vez, percebemos uma sincronia do cinema com a realidade social.

Segundo Francisco de la Peña (2014), os filmes expressam os medos, os anseios, preconceitos e crenças, costumes e mentalidades das sociedades que os produzem que, por sua vez, estão baseados no passado histórico e nas tradições culturais dos povos. Ainda de acordo com Peña (2014), os gêneros cinematográficos, as temáticas e as formas de representação variam de país para país, pois correspondem a tradições, conflitos, dinâmicas e estruturas de cada sociedade. Não à toa os filmes brasileiros que abordam a temática migração, por exemplo, se voltam para temas que remetem a fatos que marcam a sociedade brasileira, como o êxodo rural, a migração de cidades pequenas para cidades grandes e o movimento do Nordeste para o Sudeste. Da mesma forma, os filmes que abordam temas religiosos foram sensíveis às transformações em curso neste campo.

O Brasil tem passado por grandes transformações sociais e demográficas, com a transição urbana e epidemiológica, a transição demográfica, nos tipos de família e nas relações intrafamiliares, a transição religiosa, e outras. Para entender as novas configurações estruturais da sociedade brasileira não bastam retratos transversais de momentos específicos, mas sim imagens dinâmicas das transições que ocorrem de maneira concomitante no espaço e no tempo. Nesta perspectiva, o cinema é uma das formas de representar a realidade mais próxima das movimentadas alterações das relações sociais. Deste modo, entender a evolução e o desenvolvimento da relação entre o cinema e a demografia pode contribuir para a compreensão da realidade cultural em constante alteração no país.

Considerações finais

Os resultados aqui apresentados demonstram que há, inegavelmente, uma distribuição desigual na produção cinematográfica brasileira contemporânea em termos de gênero e cor/raça dos postos de direção, roteiro e protagonismo, e ainda de gênero em outras funções nas equipes dos filmes. Os resultados obtidos indicam também possibilidades de utilização do cinema para uma análise dos contextos em que é produzido através das relações entre a população que filma e a população filmada, e da observação das temáticas abordadas nos filmes.

A produção audiovisual refletiu mudanças comportamentais das sociedades produtoras de imagens, como o individualismo e outras características da modernidade, bem como de aspectos relacionados a uma suposta pós-modernidade, indicando que o cinema seja influenciado por mudanças sociais. Dessa forma, se os dados apresentados indicam uma distribuição de papéis desigual para negros, indígenas, amarelos e mulheres, quando comparados com o espaço ocupado por homens brancos no cinema brasileiro, supomos que este fato seja uma reprodução das desigualdades que ocorrem na sociedade, bem como da priorização das perspectivas brancas e masculinas em detrimento das dos demais grupos sociais.

Apesar das mudanças culturais que aconteceram nas últimas décadas no país, ainda permanecem desigualdades de gênero e raciais no mercado de trabalho. A divisão social no trabalho apresenta dois níveis de segmentação: um relacionado ao setor de atividade – que determina por tradições culturais que mulheres e negros atuem em determinadas atividades culturalmente atribuídas à identidade feminina e negra – e outro relacionado à hierarquia dos cargos – mulheres e negros ainda possuem baixa representação nas ocupações de chefia e gerenciais. No cinema, a diferença nas funções de tomada de decisão é explícita. Embora um filme precise de uma equipe de dezenas de profissionais envolvidos em sua realização, a concepção artística, a representação dos grupos sociais, a escolha das temáticas e da forma como serão abordadas nos filmes são decisões tomadas principalmente pelos diretores, função em que verificamos a predominância de homens e brancos. Se a direção cinematográfica é predominantemente ocupada por homens brancos, a gerência do imaginário, da representação de gêneros, raças e sexualidades, a disseminação de valores, modismos e ideais políticos estão sob domínio de uma perspectiva majoritariamente masculina e branca.

O cinema está em sintonia com as mudanças sociais, e é rápido em sua representação dessas mudanças, de modo que a relação entre o cinema e a demografia pode ser útil para um exame da sociedade, mais do que analisando o filme sob uma perspectiva demográfica, mas utilizando-se da narrativa audiovisual como método de observação das dinâmicas e transformações sociais e populacionais. Para a demografia, a análise fílmica se mostra oportuna ao apropriar-se de narrativas sobre temas que lhe são caros, além de trabalhar com a composição demográfica das personagens dos discursos cinematográficos (a população filmada). Por outro lado, a perspectiva demográfica se faz muito útil para os estudos de cinema, que não se limitam ao conteúdo dos filmes, mas incluem os bastidores sociais, políticos e econômicos da produção audiovisual – que envolvem desde a criação e regulamentação de políticas audiovisuais, a captação de recursos, a distribuição e a exibição, os contextos em que ocorre a produção do filme. Uma das contribuições da análise demográfica sobre o cinema ocorre justamente na abordagem extratelas: na análise de equipes e outros aspectos socioeconômicos da produção audiovisual, orçamentos, recursos e políticas.

Ao buscar capturar a dinâmica social e demográfica ou as subjetividades historicamente construídas, a própria produção cinematográfica se vê no meio de um vórtice de transformações culturais que afeta todo o processo de filmagem, tanto em relação à composição sociodemográfica da população por detrás das câmeras, quanto às mudanças de olhar sobre a população que está na frente das câmeras. O cinema brasileiro se desenvolveu em uma sociedade marcada por profundas desigualdades de classe, gênero e raça que afetam a população que filma e a representação da população filmada. Recentemente o cinema vem sendo cobrado a fazer parte das reivindicações por igualdades e direitos democráticos da sociedade brasileira, de forma a dar espaço à sociodiversidade do país em todas as suas etapas de produção e distribuição.

Referências

- ACADEMIA BRASILEIRA DE CINEMA. Disponível em: <http://academiabrasileiradecinema.com.br>.
- ADOROCINEMA. Disponível em: <http://www.adorocinema.com>.
- AGÊNCIA NACIONAL DO CINEMA (ANCINE). *Listagem completa dos filmes com os mecanismos de incentivo (2018)*. Disponível em: http://oca.ancine.gov.br/producao_.htm.
- ALVES, J. E.; BARROS, L.; CAVENAGHI, S. A transição religiosa brasileira e o processo de difusão das filiações evangélicas no Rio de Janeiro. *Revista Horizonte*, v. 12, n. 36, p.1055-1085, out./dez. 2014. Disponível em: <http://dx.doi.org/10.5752/P.2175-5841.2014v12n36p1055>.
- ALVES, P. *O cinema brasileiro de 1961 a 2010 pela perspectiva de gênero*. 2011. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) - Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2011. Disponível em: http://www.ence.ibge.gov.br/images/ence/doc/mestrado/dissertacoes/2011/Dissertacao_2011_Paula_Alves.pdf.
- ALVES, P.; ALVES, J. E.; SILVA, D. Mulheres no cinema brasileiro. *Revista Caderno Espaço Feminino*, v. 24, n. 2, p. 365-394, 2011. Disponível em: <http://www.seer.ufu.br/index.php/neguem/article/view/13661>.
- ALVES, P.; COELHO, P. Discursos, performatividades e padrões visuais no cinema: reflexões sobre as representações de gênero, o mercado cinematográfico e o cinema de mulheres. *Aceno - Revista de Antropologia do Centro-Oeste*, v. 2, n. 3, p.159-176, 2015. Disponível em: <http://periodicoscientificos.ufmt.br/ojs/index.php/aceno/article/view/2529>.
- BALADI, M. *Dicionário de cinema brasileiro: filmes de longa-metragem produzidos entre 1909 e 2012*. São Paulo: Martins Fontes, 2013.
- BUET, J. *Films de Femmes: six générations de réalisatrices*. Paris: Editions Alternatives, 1999.
- CAMPOS, M.; LONGO, L. Auto ou alter-declaração? Uma análise da informação de raça/cor nas pesquisas domiciliares. In: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, 15, 2006, Caxambu. *Anais...* Caxambu, 2006.
- CINEMATECA BRASILEIRA. Bases: filmografia brasileira. Disponível em: <http://bases.cinemateca.gov.br>.
- DANEY, S. Pour une ciné-démographie. In: DANEY, Serge. *Devant la recrudescence des vols de sacs à main*. Lyon: Aléas éditeur, 1997.
- FESTIVAL INTERNACIONAL DE CINEMA FEMININO - FEMINA. Disponível em: <http://www.feminafest.com.br>.
- FILME B. Database Brasil. *Ranking filmes nacionais 1995-2017 por renda*. Disponível em: <http://www.filmeb.com.br/conteudo-exclusivo?r=node/385735>.
- FILMOW. Disponível em: <https://filmow.com/>.
- HIRANO, L. F. *Uma interpretação do cinema brasileiro através de Grande Otelo: raça, corpo e gênero em sua performance cinematográfica (1917-1993)*. 2013. Tese (Doutorado em Antropologia Social) - Universidade de São Paulo, São Paulo, 2013.
- FERREIRA, C. *Mulheres negras e (in)visibilidade: imaginários sobre a intersecção de raça e gênero no cinema brasileiro (1999-2009)*. 2016. Tese (Doutorado em Comunicação) - Universidade de Brasília, Brasília, 2016.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICAS (IBGE). Sistema IBGE de Recuperação Automática - SIDRA. Disponível em: <https://sidra.ibge.gov.br/home/ipp/brasil>.

LAUZEN, M. *The Celluloid Ceiling: Behind-the-Scenes Employment of Women on the Top 100, 250, and 500 Films of 2018*. Disponível em: https://womenintvfilm.sdsu.edu/wp-content/uploads/2019/01/2018_Celluloid_Ceiling_Report.pdf.

LAZANEO, C. Cinemas indígenas e trânsitos da representação. *Revista Agreste*, n. 1, 2015. Disponível em: http://agre.st/edicao/1/Cinemas_Indigenas.

MOSTRA INTERNACIONAL DE CINEMA. Disponível em: <http://www.mostra.org>.

MUNIZ, J. Sobre o uso da variável raça-cor em estudos quantitativos. *Revista de Sociologia e Política*, v. 18, n. 36, p. 277-291. Curitiba, jun/2010.

PAES, A.; SILVA, R. Teste de concordância Kappa. *Revista Educação Continuada em Saúde Einstein*, v. 10, n. 4, p. 165-166, Instituto Israelita de Ensino e Pesquisa Albert Einstein, São Paulo, 2012.

PEÑA, F. de la. *Por un análisis antropológico del cine*. Imaginarios fílmicos, cultura y subjetividad. México: Ediciones Navarra, 2014.

PETRUCELLI, J. L. Raça, identidade, identificação: abordagem histórica conceitual. In: PETRUCELLI, J. L.; SABOIA, A. L. (org). *Características Étnico-raciais da População Classificações e identidades*. Estudos e Análises - Informação Demográfica e Socioeconômica, número 2. Rio de Janeiro: IBGE, 2013a.

_____. Autoidentificação, identidade étnico-racial e heteroclassificação. In: PETRUCELLI, J. L.; SABOIA, A. L. (Org). *Características Étnico-raciais da População Classificações e identidades*. Estudos e Análises - Informação Demográfica e Socioeconômica, número 2. Rio de Janeiro: IBGE, 2013b.

RODRIGUES, J. C. *O negro brasileiro e o cinema*. 4. ed. Rio de Janeiro: Pallas, 2011.

SILVA NETO, A. *Dicionário de Filmes Brasileiros: longa-metragem – 2. edição revista e atualizada*. São Bernardo do Campo: Ed. do Autor, 2009.

SIGNIFICANCE. On the web: Hollywood's gender inequality in numbers. *Significance Magazine*, v. 15, n. 5, p. 48, out 2018. Disponível em: <https://rss.onlinelibrary.wiley.com/doi/epdf/10.1111/j.1740-9713.2018.01198.x>.

TELLES, E. Repensando as relações de raça no Brasil. *Teoria & Pesquisa: Revista de Ciência Política*, v.1, n. 42/43, p. 131-159, Universidade Federal de São Carlos, Programa de Pós-Graduação em Ciência Política, São Carlos, jan/jul 2003. Disponível em: <http://www.teoriaepesquisa.ufscar.br/index.php/tp/article/view/59/49>.

Breves considerações sobre a dinâmica dos espaços urbanos no mundo contemporâneo*

Letícia de Carvalho Giannella**
Ana Carolina Chaves Católico***
Annelize de Souza Pereira****
Bruna de Castro Dias Bicalho*****
Cauan Braga da Silva Cardoso*****
Davi Bovolenta*****
Eloá Nascimento dos Santos*****
Raphael Henriques da Rocha*****
Ulisses Carlos Silva Ferreira*****

Resumo

Este capítulo investiga se a compreensão teórica sobre a produção do espaço urbano e o momento contemporâneo da metropolização pode se configurar como uma chave metodológica fundamental para interpretar o período atual de acirramento das contradições do capitalismo e apontar para possibilidades de sua superação. Discute-se como intervenções urbanas em áreas centrais se relacionam com a produção da metrópole, e demonstra-se que há uma lógica no modo como as políticas econômicas e financeiras se articulam com as políticas urbanas e sociais. Por fim, conclui-se que os modelos neoliberais de planejamento urbano são concebidos visando a acumulação de capital e resultam na configuração de metrópoles que se pautam pela segregação socioespacial.

Palavras-chave: Metrópole. *City marketing*. Segregação socioespacial. Políticas urbanas. Rio de Janeiro.

* Este capítulo é resultado de um trabalho coletivo realizado por docente e discentes ao final da disciplina "Dinâmica dos Espaços Urbanos e Regionais", ministrada no terceiro trimestre de 2016 no Programa de Pós-Graduação em População, Território e Estatísticas Públicas da ENCE/IBGE.

** Doutora em Geografia e docente do População, Território e Estatísticas Públicas desde 2015.

*** Doutora em População, Território e Estatísticas Públicas (2020). Pesquisadora no Grupo de Estudos do Setor Elétrico (GESEL/UFRJ).

**** Mestre em Ensino de Ciências (IFRJ) e Especialista em Análise Ambiental e Gestão do Território pela ENCE (2017).

***** Doutora em População, Território e Estatísticas Públicas (2020).

***** Mestre em População, Território e Estatísticas Públicas (2018) e atualmente doutorando no programa.

***** Mestre em População, Território e Estatísticas Públicas (2018).

***** Mestre em População, Território e Estatísticas Públicas (2018). Assistente de Pesquisa do Instituto de Pesquisa Econômica Aplicada - IPEA.

***** Jornalista e Especialista em Análise Ambiental e Gestão do Território pela ENCE (2017).

***** Mestre em População, Território e Estatísticas Públicas (2018) e atualmente doutorando no programa.

Introdução

O presente capítulo tem como objetivo central apresentar e debater alguns conceitos fundamentais para a compreensão da dinâmica dos espaços urbanos no mundo contemporâneo. A motivação para elaborá-lo esteve nas discussões processadas no escopo da disciplina “Dinâmica dos espaços urbanos e regionais” ministrada anualmente, desde 2016, no Programa de Pós-Graduação em População, Território e Estatísticas da ENCE.

O texto que se apresenta, escrito a deztoito mãos, é a culminância de um trabalho coletivo que teve como ponto de partida ricos debates realizados em sala de aula baseados em leituras sobre a temática. A proposta que nos norteou, assim, foi a sistematização do conteúdo debatido, que se insere no campo dos estudos urbanos críticos de filiação teórico-filosófica marxista.

Em tempos de grave crise político-econômica, torna-se vital compreender a dinâmica espacial brasileira, interpretando as transformações espaciais recentes e traçando caminhos futuros. A fim de alcançar, por sua vez, tal dinâmica, é preciso se municiar de conceitos fundamentais que nos auxiliam a revelar a dialética espaço-sociedade, especialmente no contexto mais recente. Este texto, assim, ambiciona trazer elementos que possam contribuir para esta compreensão.

Enxergamos a crise enquanto oportunidade que nos coloca a necessidade de superação capitalista, e partimos da ideia de que, como afirma Henri Lefèbvre, para mudar, é necessário mudar o espaço.

Frente a um cenário de acirramento das contradições do capitalismo com o fortalecimento do neoliberalismo combinado com o crescimento de pautas ultraconservadoras, é preciso promover esforços rumo ao entendimento do papel da produção do espaço nesse cenário para construir alternativas.

Assim, o trabalho parte da seguinte hipótese: a compreensão sobre a produção do espaço urbano e o momento contemporâneo da metropolização pode se configurar como uma chave fundamental para interpretar o momento atual de acirramento das contradições do capitalismo e apontar para possibilidades de sua superação.

Sendo assim, a questão central que orientou a escrita deste texto foi: em que medida o processo de (re)produção do espaço urbano revela e é relevado pelas contradições do capitalismo?

A primeira seção, neste sentido, discorre sobre o conceito de espaço, trazendo a perspectiva relacional para o centro da análise. Entendemos que só se pode compreender a produção do espaço urbano dentro de um quadro conceitual relacional, esforço que questiona análises que privilegiam uma leitura do espaço a partir de uma perspectiva absoluta, quando muito relativa.

Em seguida, partimos da discussão a respeito do espaço para um enfoque na produção do espaço urbano, compreendendo o espaço a partir da dialética valor de uso/valor de troca, para então desembocar em um breve debate sobre renda fundiária e raridade do espaço. É impossível entender a produção do espaço urbano no mundo contemporâneo sem compreender a centralidade da propriedade privada na produção capitalista do espaço e o papel que a renda desempenha neste processo.

Logo após, enfatizaremos o momento contemporâneo e suas especificidades. Discutiremos, assim, a financeirização e suas implicações na produção do espaço e, especificamente, na metropolização brasileira. Esta discussão está relacionada ao debate sobre os instrumentos da cidade neoliberal e o papel que eles desempenham na formação das metrópoles.

Por fim, discutiremos o modo com o qual a metropolização se relaciona com a segregação espacial, evidenciando o fenômeno como a essência da produção do espaço no modo de produção capitalista atual.

A perspectiva relacional do espaço

O espaço absoluto é o espaço como “coisa em si mesma”, com existência independente da matéria, possuindo uma estrutura passível de uso para classificação ou distinção de fenômenos (HARVEY, 1980, p. 10). É aquele que registra eventos ou ações dentro de determinada forma que o constitui e que pode ser marcado, fechado, delimitado. Geometricamente, é o espaço euclídi-

no. Socialmente, é o espaço da propriedade privada e de outras entidades territoriais delimitadas (Estados, unidades administrativas, planos urbanos, grades urbanas) (HARVEY, 2012, p. 10).

O espaço relativo é o espaço compreendido como uma relação entre objetos que existe pelo próprio fato dos objetos existirem e se relacionarem (HARVEY, 1980, p. 13). O espaço é relativo em dois sentidos: de que há múltiplas geometrias que podemos escolher (entre as quais, geometrias não-euclidianas) e de que o quadro espacial depende estritamente daquilo que está sendo relativizado e por quem (HARVEY, 2012, p. 10-12). Dessa forma, é impossível compreender o espaço independentemente do tempo, o que implica no surgimento do conceito uno de espaço e tempo, o espaço-tempo.

É daí que surgem as ideias a respeito das localizações relativas e fricções diferenciais da distância, que nos auxiliam a compreender, por exemplo, as questões de mobilidade nas metrópoles vinculadas às remoções de moradores de favelas e ocupações situadas em áreas centrais para empreendimentos habitacionais construídos em bairros geograficamente periféricos. O espaço-tempo dos deslocamentos pendulares dessas populações removidas sofre variações com fortes implicações em suas vidas cotidianas. Ainda, é a noção de espaço relativo que nos auxilia a compreender a diferença entre molduras espaço-temporais de alguém que vive em bairros periféricos, mas se desloca para o trabalho nas áreas centrais, por exemplo, de carro, e alguém que vence a mesma distância utilizando um ônibus, seguido de trem e metrô. Ou ainda, entre aquele que se desloca de carro, porém em horários alternativos, e aquele que precisa se deslocar na chamada “hora do *rush*”.

O espaço relacional também implica na impossibilidade de separar os conceitos de tempo e espaço, entretanto, está contido nos objetos e processos. Segundo Harvey (2012, p. 12), “não há coisas como espaço ou tempo fora dos processos que os definem”. Os processos, assim, definem seu próprio quadro espacial e não reduzem um evento a aquilo que acontece em determinado ponto, isto é, “uma grande variedade de influências diferentes que turbilham sobre o espaço no passado, no presente e no futuro concentram e congelam em um certo ponto para definir a natureza daquele ponto” (HARVEY, 2012, p. 12-13). Como veremos mais adiante, é impraticável compreender os fundamentos da produção capitalista do espaço e do processo de metropolização sem que consideremos sua relacionalidade, uma vez que é o espaço relacional que nos auxilia a esclarecer as diversas formas de extração da renda da terra e a especulação imobiliária vinculada a este processo. Ademais, é somente sob a perspectiva relacional do espaço-tempo que podemos compreender o papel das memórias coletivas, subjetividades e ideologias na produção do espaço urbano.

O espaço, nesta perspectiva, desnaturaliza-se. A partir do momento em que se deixa de considerar o espaço como um ente neutro que contém simplesmente as relações sociais, inicia-se a sua compreensão enquanto algo intrínseco à própria produção das relações de produção, algo “imaneente à produção da sociedade no movimento (histórico) de sua reprodução” (CARLOS, 2011, p. 9), o que implica dizer que, para transformar o modo de produção capitalista, é preciso transformar o espaço.

É aqui que podemos passar deste nível mais geral e abstrato de conceitualização do espaço — segundo Haesbaert (2010, p. 163) — para um nível mais concreto da análise e produção do espaço urbano. A mediação entre estes níveis pode se dar com o recurso à conhecida tríade lefebvriana da produção do espaço, composta pelas dimensões das práticas espaciais, representações do espaço e espaços de representação.

A primeira dimensão refere-se à materialidade do espaço, aquela vinculada ao espaço percebido pelas pessoas em seu cotidiano. De acordo com Harvey (2012, p. 19), trata-se do “espaço da experiência e da percepção aberto ao toque físico e à sensação” decorrente das práticas espaciais do dia a dia, associando realidade cotidiana (emprego do tempo) e realidade urbana (percursos e redes religando os lugares do trabalho, da vida privada, do lazer, etc).

As representações do espaço, por sua vez, são aquelas que procuram descrever os espaços através das palavras, mapas, imagens e gráficos a partir da apropriação da realidade material percebida (HARVEY, 2012, p. 22). Tal perspectiva se conecta à noção de espaço dominante em uma

sociedade, aquele vinculado às relações de produção, possuindo caráter técnico e ideológico. É o espaço, portanto, da dominação. Trata-se do espaço concebido, ligado às relações de produção e à ordem que elas impõem, aos signos, códigos etc., que elas implicam. É o espaço dos cientistas, dos planejadores, dos urbanistas, dos tecnocratas que retalham o espaço, espaço este que tende, portanto, para um sistema de signos elaborado intelectualmente (LEFÈBVRE, 2013, p. 214).

Já o espaço vivido refere-se ao modo de viver e sentir a realidade; são aqueles espaços de representação segundo as emoções, medos e sensações dos indivíduos e grupos sociais. Está vinculado aos significados subjetivos dos indivíduos em relação às suas práticas sociais no cotidiano, que transcendem a forma urbana. Os espaços de representação referem-se à apropriação do espaço, aos sonhos, devaneios e à imaginação, jamais se sujeitando à coerência do espaço concebido. Estão ligados ao lado mais clandestino e subterrâneo da vida social, e também à arte. É o espaço dos usuários e daqueles que o descrevem, como os escritores e filósofos. Trata-se do espaço dominado, portanto, submetido, que a imaginação tenta modificar e apropriar, e tende para sistemas mais ou menos coerentes de signos que não são intelectuais.

Lefèbvre considera que o espaço produzido a partir desta trílice dialética cuja predominância está nas representações do espaço e na dominação engendra um espaço abstrato que comporta em si, entretanto, contradições. Assim, ocorreria um duplo movimento com a dissolução de relações e nascimento de novas relações, o que faz com que seja engendrado um novo espaço que Lefèbvre chamará de espaço diferencial. Enquanto o espaço abstrato é repetitivo e homogêneo, o espaço diferencial acentua as diferenças e, para o autor, somente a luta de classes impede que o espaço abstrato se estenda a todo o planeta (LEFÈBVRE, 2013).

Tal proposição nos parece essencial para que possamos alcançar o objetivo proposto por este trabalho. Para compreendê-la com maior clareza em seu caráter de mediação entre a concepção relacional do espaço e a análise da produção do espaço urbano, entretanto, é importante que examinemos com um pouco de atenção aquilo que é, de certa forma, a base do pensamento de Lefèbvre e também de Harvey: a incorporação da teoria marxiana do valor na consideração sobre a produção do espaço urbano.

Valor de uso e valor de troca na produção do espaço urbano

Uma vez que partimos de uma concepção de espaço que pressupõe seu caráter relacional, entendemos que o espaço é produzido e reproduzido pelas relações sociais de produção capitalistas e, ao mesmo tempo, as (re)produz.

Como se dá este processo? Ou seja, como se estabelece esta relação dialética entre relações sociais de produção e produção do espaço?

Considerado o precursor da formulação da teoria do valor, Adam Smith, na obra *Riqueza das Nações*, afirma que

[...] o valor tem dois significados diferentes: algumas vezes expressa a utilidade de algum objeto particular, e algumas vezes o poder de compra de outros bens que a posse daquele objeto transmite. O primeiro pode ser chamado 'valor de uso', e o segundo, 'valor de troca' (SMITH, 1996, p. 85).

Marx tenta resolver as ambiguidades presentes nas abordagens clássicas. Ele considera que o duplo caráter da mercadoria é derivado do trabalho, que, por sua vez, apresenta outro duplo caráter: concreto e abstrato. Sendo o valor de uso advindo do uso concreto do produto pelos indivíduos, o valor de troca advém da interação entre objetos produzidos pelo caráter abstrato do trabalho.

Todo trabalho é, por um lado, dispêndio da força de trabalho do homem no sentido fisiológico, e nessa qualidade de trabalho humano igual ou trabalho humano abstrato gera o valor da mercadoria. Todo trabalho é, por outro lado, dispêndio de força de trabalho do homem sob a forma especificamente adequada a um fim, e nessa qualidade de trabalho concreto útil produz valores de uso. (MARX, 1985, p. 53)

A partir daí, Marx critica as abordagens anteriores que consideram o valor de uso e de troca enquanto “escalas fixas e separadas”, e concebe os conceitos de forma dialética e relacional.

Compreende-se a mercadoria enquanto duplo aspecto da expressão da sociedade capitalista burguesa: valor de uso e valor de troca. Segundo Marx, portanto, os valores de uso referem-se ao produto na medida em que forem necessários como meio de existência. Já o valor de troca, à primeira vista, aparece como uma relação “quantitativa” resultante da proporção entre valores de uso dos produtos trocados entre si.

A abordagem marxista compreende que a economia capitalista contemporânea transforma tudo em mercadoria. Não seria diferente quando se trata do solo e das benfeitorias nele realizadas. No entanto, aplicando-se os conceitos de valor de uso e de troca da mercadoria, seis aspectos particulares se fazem presentes ao solo e às suas benfeitorias: não podem circular livremente como as outras mercadorias; são mercadorias indispensáveis a todos os indivíduos, uma vez que todos necessitam morar, trabalhar etc., em algum espaço localizado; têm pouca liquidez; o solo é permanente e as benfeitorias têm, de maneira geral, longa vida útil; a troca do solo se faz num curto período de tempo, mas o uso se realiza em longos prazos; ambos (solo e benfeitorias) têm numerosos e diferentes usos para os diversos tipos de ocupantes – sendo esta uma peculiaridade do valor de uso do solo (HARVEY, 1980, p. 135-136).

A maioria dos estudos sobre uso do solo urbano costuma desconsiderar essas especificidades e a dialética valor de uso/valor de troca, concentrando-se ora em um, ora em outro (GOTTDIENER, 1993). Tais abordagens desconsideram, por exemplo, a diversidade de atores presentes no mercado imobiliário e suas diferentes relações com os valores de uso e de troca, como: os usuários (que obtêm valor de uso); os corretores de imóveis, proprietários rentistas, incorporadoras e o setor financeiro (que obtêm valor de troca); as indústrias e os governos (que obtêm ambos os valores); e, principalmente, desconsideram os resultados das interações entre os diversos atores.

É justamente a contradição entre valor de uso e de troca, em sua interpenetração dialética, que constitui o cerne da produção e reprodução do urbano. Para produzir valor de troca, é necessário produzir valor de uso para outrem, e para adquirir valor de uso, por sua vez, é necessário negociar valores de troca, ou melhor, “o que é valor de uso para um é valor de troca para outro, e cada um concebe o valor de uso diferencialmente” (HARVEY, 1980, p. 142).

Vimos na seção anterior que o espaço é produzido a partir da relação dialética entre as práticas espaciais, representações do espaço e espaços de representação. A luta pelo espaço, segundo esta abordagem, se dá na medida em que os espaços vividos são sufocados pelos espaços concebidos, dinâmica que leva à configuração dos espaços diferenciais (ou contraespaços). Uma vez que partimos de uma concepção de espaço que enfatiza sua relacionalidade, entendemos que, em uma sociedade capitalista, os espaços são concebidos mental e materialmente tendo como objetivo central possibilitar a reprodução do capital. Deste modo, podemos vincular a dimensão (ou momento) da produção do espaço que Lefèbvre caracteriza como espaço concebido à possibilidade de realização do valor de troca. Por outro lado, os espaços vividos podem ser vinculados ao valor de uso. A luta pelo espaço, assim, se configura como a luta do uso contra a troca, da apropriação contra a dominação.

Em tempos de reestruturação produtiva do capital, a dialética valor de uso/valor de troca ganha novos contornos que só podem ser compreendidos a partir do entendimento sobre a renda fundiária urbana, objetivo da próxima seção.

Renda fundiária e raridade do espaço

O capitalismo, ao possibilitar a compra e venda da terra e o usufruto dos seus bens, faz com que esta adquira valores de uso e troca, passando a ser monopolizável e alienável e ganhando o caráter de mercadoria (HARVEY, 2013). Vale ressaltar que se trata de uma mercadoria especial, uma vez que apenas as melhorias nela implementadas são produtos do trabalho humano. Além disso, a terra não pode ser reproduzida, ou seja, ela é limitada espacialmente (HARVEY, 2013, p. 432).

Ainda de acordo com o autor, os valores de uso que a terra possui podem ser retirados, usados como sustentação para a reprodução constante e podem, ainda, impulsionar a produção

como potência da natureza. Assim, ao fixar um capital específico como aplicação em determinada propriedade para aumentar a sua produtividade, o aumento obtido referente a essa aplicação será permanente. Mediante isso, tem-se que as benfeitorias implementadas no solo são produtos da atividade humana que intervêm diretamente na produção de valores de uso e de troca.

Para entender a questão da renda, é preciso compreender que há uma divisão do valor obtido pela produção: uma parcela do valor é o salário, que equivale à medida do valor pago pela força de trabalho; outra parcela é a do trabalho excedente, denominada mais valia, onde o lucro se apresenta como o valor obtido além do salário. Assim, a renda fundiária está numa categoria diferenciada desses valores alcançados pela produção, pois é inerente à propriedade privada da terra, sendo um elemento relativamente extrínseco ao processo de produção (HARVEY, 2013). Porém, a renda só pode ser paga ao proprietário por meio do lucro obtido na produção. O desafio, deste modo, é enxergar a relação de produção que existe por trás da renda, uma vez que não é a terra em si, mas sim as relações de produção, que são o fator ativo no processo de obtenção da renda.

A terra e o trabalhador são os mananciais de toda a riqueza, ou seja, a renda se fundamenta na exploração da classe trabalhadora e não da classe capitalista pelos proprietários, como costuma-se compreender. Com o capitalismo, a terra passa a ser vista como mercadoria, e a renda começa a aparecer como juros sobre o investimento na sua aquisição ou utilização. Negociasse, desta forma, o direito sobre lucros futuros do uso da terra, o que não deixa de ser o direito sobre o trabalho futuro. A terra torna-se, assim, capital fictício, tal como outros instrumentos financeiros cada vez mais presentes na economia global (ativos diversos), porém com o agravante de ser um instrumento monopolizável.

Mediante essas considerações, chegamos às formas de renda. A renda diferencial resulta da concorrência entre os capitalistas e não existe, portanto, em si mesma, mas sim somente em relação com outras parcelas do espaço. Sua obtenção se dá através da diferença entre o preço individual de produção e o preço de produção no pior solo, por isso só pode ser considerada dentro de uma concepção de espaço relacional. Dentre as possibilidades para a obtenção da renda diferencial, tem-se que ela pode independer do capital investido, sendo somente resultado da fertilidade natural do solo e sua localização (renda diferencial I); ou pode ser o efeito de investimentos de capital que modificam suas condições de concorrência frente a outras parcelas de terra (renda diferencial II).

A renda absoluta resulta da condição monopolizável da terra. O proprietário pode colocar a terra para produzir somente mediante tributos, sendo esta a razão da existência da renda nos solos de pior condição. Ela é obtida mediante a elevação artificial dos preços acima do preço de produção geral (pior solo). O lucro obtido, assim, não é fração do trabalho excedente dos trabalhadores daquela terra, mas sim da massa de mais-valia global. A renda de monopólio, por sua vez, realiza-se ao ter o seu preço determinado pelo desejo e capacidade de pagamento dos consumidores (HARVEY, 2013, PAULANI, 2016).

Para Harvey (2013, p. 461), “a apropriação da renda e a existência da propriedade privada na terra são condições socialmente necessárias para a perpetuação do capitalismo”. O autor afirma, entretanto, que a existência da renda se torna uma fenda de contradição, confusão e irracionalidade justamente por estar fundada na existência da propriedade privada.

Dois implicações estão intimamente conectadas a este processo de apropriação da renda na produção do espaço urbano cujo domínio está na renda absoluta em torno do valor de troca que surge quando o espaço se configura como uma condição para a produção capitalista. A primeira delas se relaciona com o crescimento urbano com a incorporação do capital como capital fixo, o que contribui para a compreensão dos processos de metropolização. A segunda se refere à escassez do espaço, socialmente determinada, que Carlos vai analisar em termos da raridade do espaço:

[...] diante das necessidades impostas pela reprodução do capital, o espaço produzido socialmente - e tornado mercadoria, no processo histórico - é apropriado privatamente, criando limites à sua própria reprodução. Nesse momento, o espaço, produto da reprodução da sociedade, entra em contradição com as necessidades do desenvolvimento do próprio capital. O que significa dizer que a “raridade” é produto do próprio processo de produção do espaço ao mesmo tempo que sua limitação - o que se configura como uma contradição do espaço. (CARLOS, 2001, não paginado).

Esta raridade ocorre em determinados pontos da cidade que estão associados geralmente à centralidade, no sentido locacional e no sentido da infraestrutura de transportes e comunicações existentes, e este processo se relaciona intimamente com a questão da renda fundiária. As pessoas, as corporações e o Estado que são proprietários dos terrenos situados nas áreas centrais das grandes cidades, por exemplo, são conscientes de que as porções do espaço que lhes pertencem são espaços de raridade cujas formas não cumprem mais as funções a que foram planejadas, se constituindo contemporaneamente mais como rugosidades no processo da produção do espaço urbano. Para manter a inserção dessas porções do espaço no circuito da troca e manter o processo de obtenção da renda torna-se necessário ressignificar essas formas e dotá-las de novas funções, sendo este o caso de projetos de revitalização de áreas centrais e portuárias, por exemplo. No contexto neoliberal, como veremos, em que se estabelece o planejamento estratégico das cidades, pontos do território são selecionados para sofrerem tais intervenções.

Deste modo, a extração da renda fundiária – associada à produção da escassez nos espaços urbanos – está vinculada à intervenção do Estado, que se vale de estratégias sob o discurso da modernização. Tal processo implica em várias consequências, tais como: remoções de favelas, expulsão de populações de uma determinada área, destruição de bairros inteiros e transformações nas leis de zoneamento da cidade (CARLOS, 2001).

Lefèbvre (1999), já nos anos 70, teria nomeado este processo como a implosão-explosão da metrópole. Ao mesmo tempo em que o capitalismo se reproduz por meio da expansão geográfica absoluta da urbanização, configurando as metrópoles como a forma urbana contemporânea por excelência, é preciso que outras estratégias sejam criadas no que tange às contratendências à tendência de queda da taxa de lucro do capital. Entre essas estratégias está a renovação de áreas centrais e portuárias, que por sua vez resulta da formação de *rent gaps* (SMITH, 2007). Ambos os processos são duas faces da mesma dinâmica de reprodução do capital e estão, deste modo, intimamente relacionados.

Como abordado anteriormente, a dinâmica valor de uso/valor de troca na produção do espaço urbano, as formas de extração da renda da terra e a formação de espaços de raridade ganham novos contornos no cenário de neoliberalização e financeirização da economia.

Financeirização e produção do espaço urbano

De acordo com Calvete (2009), a financeirização é o fenômeno caracterizado pelo aumento da influência das finanças na economia, avançando também sobre as relações políticas e sociais. Historicamente, esse fenômeno se inicia nos anos 70 do século XX, com a transição dos chamados anos dourados do fordismo para um capitalismo financeirizado baseado na especulação, nos riscos e na instabilidade.

O reconhecimento do processo de financeirização – não compreendido de forma limitada como o domínio do crédito sobre o setor produtivo, mas sim de modo abrangente, ou seja, como um fenômeno que atinge as relações sociais como um todo – é crucial, uma vez que a urbanização vem se dando cada vez mais sob forte influência da especulação e “do capital portador de juros” (SANFELICI, 2013, p. 28).

Como aponta Harvey (2011), a produção do espaço urbano é essencial para o desenvolvimento do capitalismo financeiro, pois grandes quantidades de capitais são movidas por empréstimos de longo prazo. Ainda, tais formas de investimentos, baseadas em endividamentos de diversos matizes, jogam um papel central na formação das crises do capitalismo.

Segundo Sanfelici (2013), as reformulações do sistema financeiro, principalmente após os anos 80 do século passado, são relevantes para explicar a expansão do crédito pessoal, principalmente para o financiamento habitacional.

No Brasil, alguns autores vêm se debruçando sobre aspectos da financeirização e suas implicações na urbanização, especialmente no estado de São Paulo.¹ Suas conclusões apontam

¹ Tais como Mariana Fix (2001, 2007), Flávia Martins (2011) e o já citado Daniel Sanfelici (2013).

para alguns processos que parecem indicar certo padrão, quais sejam: a intensa participação do Estado na criação de condições institucionais para o fortalecimento do sistema de finanças na produção do espaço urbano; a atuação de fundos de investimentos imobiliários compostos por agentes financeiros globais capazes de extrair rendas da terra na escala local, promovendo o que Sanfelici denomina como “reordenamento escalar do processo de urbanização” (2013, p. 43); e

um recrudescimento do processo de fragmentação socioespacial, sobretudo em razão da ênfase dos grandes grupos do setor imobiliário na oferta de mega-projetos que apontam para graus mais ou menos elevados de ruptura com o tecido urbano como decorrência de sua natureza ‘introvertida’ – separação por muros e cercas, integração de funções diferenciadas (moradia, trabalho etc.), segurança e controle de acesso etc. (SANFELICI, 2013, p. 43)

O mesmo autor destaca “o papel das grandes construtoras/incorporadoras como elos mediadores entre os fluxos financeiros globais e a captura de rendas do solo na escala local” (SANFELICI, 2013, p. 44).

Assim, entende-se que o processo de urbanização leva à transferência e concentração da renda fundiária para poucos agentes que atuam em diversas escalas por meio da dominação do espaço por meio do espaço concebido – associado aos planos, projetos e processos de revitalização – que tende a sufocar os espaços vividos, ou da sobreposição da troca sobre o uso. Interessa também articular este processo com a metropolização, mostrando que existe um movimento que associa as intervenções urbanas em áreas centrais com a produção da metrópole.

Intervenções em áreas centrais e portuárias são exemplos emblemáticos do que se denomina empresariamento urbano (HARVEY, 1996, 2005), que está vinculado, por sua vez, ao planejamento estratégico das cidades (VAINER, 2009).

Barbosa (2006) mostra que, neste contexto, as empresas passam a indicar lugares otimizados para a reprodução do capital, exigindo as condições fiscais necessárias para os empreendimentos e a flexibilização das leis trabalhistas, o que tem como consequência o estabelecimento de uma

[...] competição na escala dos lugares, que por sua vez, aprofunda a competição no interior das redes urbanas e, ao mesmo tempo, torna-se mais predatória, pois se realiza de modo cada vez mais coercitivo, fragmentário e pouco previsível (BARBOSA, 2006, p.139).

Outro ponto que chama a atenção é a participação de órgãos internacionais e suas diretrizes para a liberação de financiamentos, onde as relações com o empresariamento urbano ganham maior dimensão. O Banco Interamericano de Desenvolvimento (BID), por exemplo, exige uma relação de redução do Estado e maior participação das empresas privadas na produção da infraestrutura urbana, principalmente nos serviços essenciais das cidades como habitação, saneamento e transporte, para a liberação do crédito (BARBOSA, 2006, p. 138).

A competição entre cidades leva ao recrudescimento de desigualdades e fragmentações socioespaciais na escala da metrópole, uma vez que, neste jogo, as cidades que costumam ter condições de enfrentar a concorrência no intuito de se inserir no rol das chamadas cidades globais são sempre os nós centrais das regiões metropolitanas. Mesmo na escala da cidade, como dissemos, parcelas específicas do espaço são recortadas e eleitas para receberem investimentos a fim de que a cidade possa participar deste jogo que se dá em escala global.

Neste cenário de empresariamento urbano, surge uma emblemática concepção de cidade-mercadoria, na qual muitas cidades estão à venda e, inclusive, podem ser objetos de luxo. Nelas, os atributos comercializados são os mais variados, dependendo dos interesses e demandas de cada grupo social. Esse processo contribui para explicar o chamado marketing urbano ou *city marketing*, que será descrito na próxima seção.

O papel do *city marketing* na construção da metrópole financeirizada

A criação de uma identidade é complexa e, não por menos, também é complexo o mecanismo aplicado à construção de uma identidade artificial sobre um território destinado à venda. A arquiteta Fernanda Sanchez (2010), ao estudar os mecanismos de inserção competitiva das cidades no mercado global dentro do contexto do empresariamento urbano, revela – a partir da análise de Barcelona

e Curitiba – os modos pelos quais uma cidade passa a ser concebida como produto, alicerçada pelos rotineiros discursos midiáticos de reestruturação econômica ou regeneração urbana.

Com o que vimos até o momento, entende-se que caso a produção do espaço urbano se mostrasse em sua essência, ou seja, como um processo conflituoso entre valor de uso e de troca, entre apropriação e dominação, que tende à vitória opressiva do primeiro sobre o segundo por meio de concepções de espaço que tentam sufocar os espaços de representação, e se o espaço urbano fosse compreendido pelo senso comum dentro de uma perspectiva relacional, não aceitaríamos, por exemplo, que a renda fundiária obtida por meio de investimento público em determinadas porções do espaço fosse transferida em sua totalidade para detentores do capital financeiro (e setores do capital produtivo) que atuam em diversas escalas.

Não aceitaríamos, também, que pessoas que historicamente produziram espaços que são alvos de revitalização, mantendo-os como espaços de raridade ao longo de décadas, não consigam se manter ali em meio a um processo de especulação decorrente do caráter relacional do espaço. Não aceitaríamos, ainda, que parcelas restritas do espaço sejam selecionadas para receber investimentos, em meio à guerra dos lugares, enquanto cidades inteiras se veem à margem do processo, configurando um processo de segregação socioespacial cada vez mais violento.

O que faz então com que processos como estes sejam “naturalmente” aceitos perante a sociedade? Para além da questão da ideologia – discussão complexa que não poderemos aprofundar aqui mas que pode superficialmente ser compreendida como, de certo modo, a universalização de interesses particulares –, ou melhor, associado a ela e como um dos seus instrumentos, assistimos à entrada em cena do marketing urbano, ou *city marketing*, que objetiva não somente converter a cidade como um todo em mercadoria – ainda que através da venda de projetos pontuais, à luz do planejamento estratégico – mas também construir o consenso e legitimidade em relação ao que está sendo proposto/implementado.

Transformar uma cidade em mercadoria e passível de venda num mercado global envolve uma sofisticada articulação de ideias e práticas que estejam o mais próximo possível de um sentimento coletivo que, de algum modo, reflita uma identidade local e, por consequência, traduza uma sensação de pertencimento. Conceber a cidade enquanto mercadoria é arrancá-la dos laços sociais que produzem diariamente o espaço e dotá-la de fetiches, ou algo desprovido de relações histórico-culturais (quicá a história torne-se mais um elemento mercadológico).

Sanchez revela-se intrigada e suscita a seguinte provocação: “o que está à venda na venda da cidade?” (2010, p. 58). A resposta à pergunta vem com a prévia explicação de que, se antigamente as cidades eram planejadas para as populações residentes, nos dias atuais elas estão cada vez mais pensadas quanto às influências que exercerão sobre o capital financeiro. Esta peculiaridade está intrinsecamente associada à criação e ao surgimento das cidades globais, cujos pilares estão, atualmente, fixados na recepção de megaeventos internacionais e na competitividade.

De modo paradoxal, todavia, a despeito de ser produzido e vendido como uma mercadoria, sempre haverá mecanismos de resistência nas escalas locais.

Um dos elementos que podemos mencionar a respeito das resistências, dentro deste processo que pode ser entendido como uma forma de luta pelo espaço, é em relação à categoria gentrificação. Muito se tem discutido a respeito da validade ou não do termo para a realidade brasileira, por exemplo. Neil Smith (2006), uma das referências centrais nos estudos sobre o tema, descreveu que os processos de renovação urbana estão profundamente associados à gentrificação. Ainda que os estudos de Smith tenham sido realizados em Nova Iorque (EUA), o fenômeno pôde ser observado em diversas outras cidades e países. Regra geral, os bairros periféricos serviram de refúgio para as camadas sensíveis da sociedade, ao passo que os bairros centrais tornaram-se alvo da especulação imobiliária com o advento das classes mais abastadas, elevando os preços dos imóveis e transformando o perfil socioeconômico. Ainda que este não figure como o padrão encontrado nas cidades brasileiras, e por este motivo necessitamos cautela no uso do termo, nos parece que a gentrificação vem sendo acionada,

nos últimos anos, como uma categoria em disputa na resistência contra processos de remoção de moradores de áreas centrais.²

De qualquer forma, Smith (2006) traz à tona uma reflexão que está fortemente associada à sociedade do consumo:

Do mesmo modo que a televisão popular fez as massas acreditarem que a vida das pessoas ricas e célebres representavam a norma social à qual todo mundo pode aspirar, a gentrificação produz agora uma paisagem urbana que veicula as mesmas aspirações (acreditar no consumo) (p.73).

Sanchez reflete que a produção da imagem das novas cidades, bem como os discursos que são veiculados nos meios de comunicação, provêm de vultosos orçamentos destinados aos meios de comunicação e equipes de marketing para a criação de uma “imagem da marca das cidades inventadas” (SANCHEZ, 2010, p. 83), a que a autora denomina de representações do espaço e transformações urbanas. Para ela, a mídia atua na produção de discursos que, por sua vez, servirão de base para a formulação de consensos. Os discursos e conteúdos veiculados na mídia agem na tentativa de aniquilar e/ou enfraquecer as visões e forças antagônicas aos discursos hegemônicos. Nesse ponto, a autora reconhece a importância das ações de comunicação como estratégia de amálgama da opinião pública e reflete que o próprio espaço urbano também é construído para se comunicar simbolicamente com os cidadãos. A autora recorre aos conceitos da sociedade do espetáculo de Guy Debord (1997) para compreender a utilização de discursos e imagens na formulação de ideais coletivos: o acesso à modernidade, os padrões de primeiro mundo, a sensação de segurança etc.

Por fim, Sanchez (2010) menciona que à construção dos discursos cabe ainda uma importante etapa: levá-lo a público por quem seja visto como autoridade e cuja mensagem tenha uma força (quase) inquestionável: técnicos de marketing, grandes empresários, famosos urbanistas, personalidades públicas carismáticas, engenheiros etc. Assim, sugere a autora, a sentença passa a ser aceita mais facilmente, uma vez que, supostamente, essas autoridades têm mais propriedade em suas falas previamente elaboradas.

Como culminância de todo o processo descrito até o momento, tem-se a configuração de cidades e regiões metropolitanas cada vez mais segregadas/segregadoras.

Metropolização e segregação socioespacial

A segregação emerge como condição, meio e produto de um modelo de produção e reprodução das relações sociais centrado na expansão e consolidação dos valores de troca (CARLOS, 2013).

As contradições inerentes ao próprio modo de produção do espaço urbano aparecem como elementos determinantes da segregação social. A noção de propriedade, apresentada no modelo capitalista como direito natural do ser humano, situa-se no cerne de todas as formas de desigualdades e enfrentamentos. Segundo Alvarez (2013, p. 113), apesar da produção do espaço ser concebida socialmente, é a propriedade privada que mediará as relações de apropriação e sua subordinação ao valor de troca. Neste sentido, os conflitos existentes entre a necessidade de apropriação do espaço como lugar de vida e como instrumento de manutenção da acumulação de riquezas permeiam as diversas relações sociais existentes.

O modelo empreendedor de planejamento estratégico e gestão da cidade fornece os subsídios básicos para a privatização de todos os conteúdos sociais da vida humana. Desta forma, a metrópole passa a ser vivida através de fragmentos, de experiências atomizadas e deslocadas de sua totalidade.

A segregação domina todas as esferas da vida social, desde a paisagem e a morfologia das cidades até as mais abstratas das relações. Existe, porém, um elemento comum capaz de mediar e orientar todas estas esferas e fragmentos: o Estado. Por meio da legitimação de normas, regulamentos e políticas públicas, o Estado direciona as formas e funções de uso do espaço, garantindo toda a infraestrutura e superestrutura necessárias à reprodução ampliada do capital (CARLOS, 2013). Tendo o valor de troca como fio condutor do processo de “produção/apropriação/reprodu-

² Ver, por exemplo, Giannella e De la Garza (2018).

ção” (CARLOS, 2013, p. 97) do espaço urbano, selecionam-se os sujeitos possibilitados de acessá-lo. Os consecutivos fenômenos de desvalorização, requalificação e expropriação associados à revitalização de áreas centrais e portuárias constituem exemplos das estratégias de viabilização da expansão do capital, como também evidenciam o papel e as alianças dos diversos atores sociais envolvidos (ALVAREZ, 2013).

Nas cidades brasileiras, o processo de rápida urbanização restringiu grupos sociais de menor renda ao acesso de “condições básicas de urbanidade” e impossibilitou-os, assim, de uma verdadeira inserção na cidade (ROLNIK, 2006, p. 199). Desta forma, a segregação passa a atuar em diferentes dimensões sociais sobrepostas, efetivando a completa desmobilização destes grupos. A falta de moradias legais desdobra-se na formação de assentamentos precários, ilegais e irregulares, localidades tidas como invisíveis para a gestão e administração estatais.

A cidade segregada é um cenário tão intrínseco ao cotidiano da sociedade moderna, que se converteu em uma realidade naturalizada. Os territórios periféricos e as favelas são facilmente associados à desordem e ao caos, ou seja, a um elemento social errático do espaço produzido. Uma percepção que desassocia completamente estes espaços de sua materialidade histórica. No entanto, ignorar o conteúdo social destes espaços é desprover o modo de produção dos espaços urbanos de sua essência, o que configura, na verdade, uma condição de existência.

Mas, se por um lado as cidades se configuram como espaços estratégicos para o empresariamento urbano, são nelas também que emergem os movimentos de resistência. Como destaca Lefèbvre (1991), é justamente no cotidiano que se encontra o limite do viver e a capacidade de lutar. O sociólogo enxerga, a partir da tendência das cidades à urbanização completa, um meio de superação através da sobreposição do valor de uso ao valor de troca, da imaginação à razão, da criação ao repetitivo e, principalmente, através de projetos coletivos revolucionários.

É preciso, portanto, direcionar nossos esforços para apreender aquilo que existe enquanto potência transformadora no processo de luta pelo espaço, e esmiuçar cada lampejo que aponte para a superação do modo capitalista de produção na vida cotidiana que, ao mesmo tempo em que se vê dominada, resiste e cria, produzindo o espaço urbano enquanto obra.

Considerações finais

Tendo percorrido brevemente sobre os temas que nos orientam no sentido da compreensão sobre a (re)produção do espaço urbano no contexto neoliberal, esperamos que tenha ficado claro para quem nos lê que, de fato, podemos considerar a dinâmica dos espaços urbanos como um elemento chave para a compreensão da sociedade capitalista. As transformações no sistema produtivo podem, deste modo, ser mais bem visualizadas e interpretadas à luz da relação dialética entre espaço e sociedade. Em outras palavras, o estudo do espaço urbano se configura como um instrumento para ultrapassar o véu da aparência dos processos econômico-sociais e buscar a sua essência, numa perspectiva que preza pela totalidade.

A lógica neoliberal pauta suas práticas “(...) pela tomada de terra urbana e da moradia pelas finanças globais” (ROLNIK, 2015, p. 368). O desencadeamento dos tópicos ao longo deste artigo buscou mostrar a articulação atual entre economia e políticas urbanas.

O domínio do valor de troca sobre o valor de uso com vistas à aceleração do fluxo global de capital faz emergir novos modos de relacionamento com o espaço.

A questão do monopólio da propriedade privada, cada vez mais, pressupõe a existência de políticas urbanas associadas à problemática da raridade do espaço, cujo efeito de ponta é a escalada dos preços dos imóveis e o surgimento do fenômeno da gentrificação.

As cidades-mercadorias se consolidam no apogeu do uso das técnicas de segurança, comunicação, urbanismo, informação e transporte logístico em favor do fluxo de capitais (e de alguns cidadãos selecionados). Como consequência, vivencia-se o acirramento das desigualdades sociais e o aumento de marginalizados.

Uma das principais conclusões a que se chega este texto é o fato de que as políticas urbanas associadas ao modo de produção capitalista em sua etapa neoliberal acabam por resultar em

idades cada vez mais segregadas e injustas, favorecendo poucos grupos sociais e relegando a maior parte da população a condições de vida insustentáveis e inviáveis em termos de reprodução da própria sociedade.

Todavia, os espaços vividos, embora tendam a ser sufocados, resistem e pautam práticas espaciais alternativas que necessitam ser melhor observadas, compreendidas e multiplicadas, se queremos transformar o mundo. É preciso, assim, olhar para o que está sendo produzido em luta dentro do eixo dominação-apropriação. Esta se desenha, portanto, como uma agenda de pesquisa fundamental dentro do campo dos estudos urbanos. Se para mudar o capitalismo é preciso mudar o espaço, sigamos por aí.

Referências

- ALVAREZ, I. P. A segregação como conteúdo da produção do espaço urbano. In: CORREA, R. L.; VASCONCELOS, P. A.; PINTAUDI, S. M. (orgs.). *A cidade contemporânea: segregação espacial*. São Paulo: Contexto, 2013.
- BARBOSA, J. L. O ordenamento territorial urbano na era da acumulação globalizada. In: OLIVEIRA, M. P.; HAESBAERT, Rogério; MOREIRA, Ruy (orgs.). *Território, territórios: ensaios sobre o ordenamento territorial*. Rio de Janeiro: DP&A, 2006 [2002].
- CALVETE, C. da S. Impactos da financeirização no mundo do trabalho. In: MACAMBIRA, J.; CARLEIAL, L. M. da F. (orgs.). *Emprego, trabalho e políticas públicas*. Fortaleza: Instituto de Desenvolvimento do Trabalho, 2009.
- CARLOS, A. F. A. A prática espacial urbana como segregação e o 'direito à cidade' como horizonte utópico. In: CORREA, R. L.; VASCONCELOS, P. A.; PINTAUDI, S. M. (orgs.). *A cidade contemporânea: segregação espacial*. São Paulo: Contexto, 2013.
- _____. *A condição espacial*. São Paulo: Contexto, 2011.
- _____. São Paulo hoje: as contradições no processo de reprodução do espaço. *Scripta Nova*, 2001, n. 88, sem paginação.
- FIX, M. *Parceiros da exclusão: duas histórias da construção de uma "nova cidade" em São Paulo: Faria Lima e Água Espraiada*. São Paulo: Boitempo, 2001.
- _____. *São Paulo Cidade Global: fundamentos financeiros de uma miragem*. São Paulo: Boitempo, 2007.
- GIANNELLA, L. de C.; DE LA GARZA, M. M. Conexões Rio-Barcelona: da produção da cidade neoliberal à articulação das resistências. In: OLIVEIRA, M. P.; MARTINS, F. E. S.; GIANNELLA, L. C. (orgs.). *Dominação e apropriação na luta por espaço urbano*. Rio de Janeiro: Consequência, 2018, pp. 359-392.
- GOTTDIENER, M. *A produção social do espaço urbano*. São Paulo: EdUSP, 1993.
- HAESBAERT, R. *Regional-global: dilemas da região e da regionalização na geografia contemporânea*. Rio de Janeiro: Bertrand Brasil, 2010.
- HARVEY, D. *O enigma do capital e as crises do capitalismo*. São Paulo: Boitempo, 2011.
- _____. O espaço como palavra-chave. *GEOgraphia*, v. 14, n. 28, 2012 [2006], p. 8-39.
- _____. *A produção capitalista do espaço*. São Paulo: Annablume, 2005 [2001].
- _____. Do gerenciamento ao empresariamento: a transformação da administração urbana no capitalismo tardio. *Espaço & Debates*, n. 39, 1996, p. 48-64.
- _____. *Os Limites do capital*. São Paulo: Boitempo, 2013 [1982].
- _____. *A justiça social e a cidade*. São Paulo: Hucitec, 1980 [1973].
- LEFÈBVRE, H. *La producción del espacio*. Madrid: Capitáin Swing, 2013, [1974].
- _____. *A revolução urbana*. Belo Horizonte: Ed. UFMG, 1999 [1970].
- _____. *A vida cotidiana no mundo moderno*. São Paulo: Ática, 1991 [1968].
- MARTINS, F. E. da S. *A (re)produção social da escala metropolitana: um estudo sobre a abertura de capitais nas incorporadoras e sobre o endividamento imobiliário urbano em São Paulo*. 2011. Tese de doutorado. São Paulo, Universidade de São Paulo.
- MARX, K. *O Capital: crítica da economia política*. São Paulo: Nova Cultural, 1985 [1867].

PAULANI, L. M. Acumulação e rentismo: resgatando a teoria da renda de Marx para pensar o capitalismo contemporâneo. *Revista de Economia Política*, v. 36, n. 3, 2016, pp. 514-535.

ROLNIK, R. A construção de uma política fundiária e de planejamento urbano para o país: avanços e desafios. *Boletim de políticas sociais*, IPEA, n. 12, 2006, p. 199-210.

_____. *Guerra dos lugares: a colonização da terra e da moradia na era das finanças*. São Paulo: Boitempo, 2015.

SANCHEZ, F. *A reinvenção das cidades para um mercado mundial*. Chapecó: Argos, 2010.

SANFELICI, D. Financeirização e a produção do espaço urbano no Brasil: uma contribuição ao debate. *EURE*. Santiago, v. 39, 2013, p. 27-46.

SMITH, N. A gentrificação generalizada: de uma anomalia local à 'regeneração' urbana como estratégia urbana global. In: BIDOU-ZACHARIASEN, C. (org.). *De volta à cidade: dos processos de gentrificação às políticas de "revitalização" dos centros urbanos*. São Paulo: Annablume, 2006 [2003].

_____. Gentrificação, a fronteira e a reestruturação do espaço urbano. In: *Geosp*, Espaço e tempo, n. 21, 2007 [1996], p. 15-31.

SMITH, A. *A Riqueza das Nações: investigação sobre sua natureza e suas causas*, v. 1. São Paulo: Nova Cultural, 1996 [1776].

VAINER, C. Pátria, empresa e mercadoria: notas sobre a estratégia discursiva do Planejamento Estratégico Urbano. In: ARANTES, O. B. F.; VAINER, C.; MARICATO, E. (orgs.). *A cidade do pensamento único: desmanchando consensos*. Petrópolis: Vozes, 2009 [2000].

A contribuição das componentes demográficas nas projeções populacionais: um exercício para o Brasil incorporando hipóteses de imigração internacional*

Helena Nobre de Oliveira**

César Augusto Marques da Silva***

Antonio Tadeu Ribeiro de Oliveira****

Resumo

As mudanças demográficas previstas para os próximos anos indicam um processo em direção ao envelhecimento populacional. Nesse contexto, alterações consideráveis nas componentes demográficas, principalmente a fecundidade e a migração, poderiam amenizar o envelhecimento. Dessas, a migração seria a componente com implicações mais imediatas, e, nesse sentido, a bibliografia discute os seus limites e potencialidades, considerando os efeitos sobre o tamanho populacional e as potenciais consequências sociais e econômicas. Diante disso, o presente trabalho, com ênfase metodológica, traz diferentes cenários de projeções demográficas para o Brasil a partir de mudanças na dinâmica migratória, analisando quais seriam os efeitos demográficos dessa dinâmica. Com isso, calculam-se, separadamente, quais seriam os efeitos da fecundidade, mortalidade, migração e da estrutura etária para o crescimento demográfico. Em todos os cenários mais factíveis observou-se que a principal contribuição ao crescimento populacional será da estrutura etária (momentum), e a principal contribuição ao decréscimo populacional será da queda da fecundidade.

Palavras-chave: Componentes demográficas. Imigração internacional. Transição demográfica.

* Este capítulo é um dos resultados da dissertação de mestrado concluída em 2017 por Helena Nobre de Oliveira, intitulada "Imigração internacional: uma alternativa para os impactos das mudanças demográficas no Brasil?", sob orientação de César Marques e Antônio Tadeu Oliveira.

** Cientista de Dados no Grupo Abril. Concluiu o mestrado em População, Território e Estatísticas Públicas em 2017.

*** Doutor em Demografia e docente no programa de pós-graduação da Escola Nacional de Ciências Estatísticas desde 2014. Coordenou o programa entre 2016 e 2020.

**** Doutor em Demografia e pós-doutor em Ecologia Humana e População pela Universidade Complutense de Madri (2014-2015). É pesquisador do IBGE e pesquisador associado do Observatório das Migrações Internacionais (OBMigra/UnB).

Introdução

Diversos países passaram, durante os últimos séculos e décadas, pelas diversas fases da transição demográfica, marcada pela queda da mortalidade e da fecundidade, passando de patamares elevados para níveis relativamente estáveis bem abaixo do ponto inicial. Enquanto isso, a migração, que é uma componente mais sensível e com amplitude de variações, sofreu oscilações relativas aos cenários econômicos e sociais das diversas regiões, países e seus respectivos contextos históricos.

No Brasil, esse processo teve início em 1940 e gradualmente foi modificando sua estrutura etária. Primeiramente houve o alargamento da base da pirâmide etária para que, posteriormente, ocorresse seu processo de estreitamento. Para os próximos anos, esse quadro terá continuidade com a pirâmide cada vez mais próxima de um formato retangular. Esse cenário levará o país para um contexto inédito de envelhecimento populacional, onde as idades mais avançadas terão pesos crescentes (WONG; CARVALHO, 2006).

Ao longo dos anos, os principais efeitos foram oriundos de ganhos em mortalidade e alterações na fecundidade associados às modificações na estrutura etária. A migração pouco afetou, já que a taxa líquida de migração (TLM) internacional no país teve valores próximos a zero a partir da segunda metade do século XX. Entretanto, para os próximos anos, diante do cenário descrito, essa componente poderia influir de modo maior. Conforme indicam Bacci (2012) e Oliveira (2014), o aumento do fluxo migratório poderia ser visto como uma das alternativas mais rápidas em busca da suavização do envelhecimento populacional.

Dado que as variações e interações das componentes demográficas, considerando as taxas vitais e os saldos migratórios, são os determinantes do ritmo de crescimento populacional, Bongaarts e Bulatao (1999) propuseram um conjunto de projeções que possibilitam o cálculo da contribuição de tais componentes demográficas no crescimento da população. Os autores ressaltam que esse tipo de cálculo é relevante para o direcionamento de políticas públicas ligadas ao bem-estar da população diante das mudanças demográficas em curso.

Dessa forma, o objetivo do presente artigo é analisar a contribuição das componentes demográficas para até o ano de 2050 no Brasil, incorporando diferentes hipóteses de imigração internacional em projeções demográficas. Para tal utiliza-se a proposta de Bongaarts e Bulatao (1999). A escolha de mudanças relacionadas a essa componente advém do fato de que a migração é, por vezes, subdimensionada em sua potencialidade. Conforme explicitado, a entrada de imigrantes pode ser vista como uma dinâmica importante na equalização de mudanças decorrentes do envelhecimento populacional. Contudo, ressalta-se que a discussão mais ampla sobre a factibilidade e os limites da inserção da migração internacional e sua interação com o processo de envelhecimento é discutida em Nobre (2017), sendo que esse trabalho prioriza a análise da contribuição das componentes demográficas em cada cenário das projeções.

Os cenários que incluem modificações na componente migratória são três: RDI 12, Atenuação da RDI e TLM da Alemanha. No primeiro a Razão de Dependência de Idosos (RDI) é mantida em 12,0 ao longo de todos os anos, mantendo o que Pereima e Porse (2013) consideram ideal em relação ao desenvolvimento econômico. O segundo buscou diminuir o ritmo de crescimento da RDI no país e o terceiro considerou que o Brasil adotaria as TLM (Taxas Líquidas de Migração) positivas a partir de uma experiência concreta. Nesse caso adotou-se a mesma TLM estimada para a Alemanha entre 2010 e 2015.

A interação das componentes demográficas no contexto geral e no Brasil

O termo transição demográfica, amplamente debatido na literatura demográfica, traz à tona a ideia de que as taxas vitais sofrem, ao longo dos anos, um complexo processo de transformação. Através de modificações de cunho social e econômico, níveis elevados de fecundidade e mortalidade gradualmente tendem à redução, alcançando um quadro com patamares mais baixos para esses indicadores.

Em termos gerais, o histórico da dinâmica demográfica brasileira não é diferente. Na década de 1940, a melhoria nas condições de higiene, do padrão de vida e os avanços médicos fizeram com que a taxa de mortalidade declinasse (ALVES, 2014; CAMARANO, KANSO e FERNANDES, 2014) e, assim, a população brasileira entrou em expansão, com considerável participação das idades mais jovens. Se em 1940 o Brasil tinha uma população total de 40 milhões de habitantes, em apenas vinte anos, essa passou a ser de cerca de 70 milhões.

A partir de 1960 notaram-se os primeiros sinais de tendência à redução na fecundidade brasileira, atribuída principalmente à urbanização e a mudanças comportamentais (FARIA e POTTER, 2002; SIMÕES, 2006). Assim, viu-se lentamente a base da pirâmide brasileira estreitar, pelo menos de maneira relativa. Em 1991, o total populacional brasileiro era de 147 milhões, chegando a 169 milhões em 2000 e 190 milhões em 2010. Como um todo, o peso da migração para esse crescimento foi praticamente nulo.

Considerando as próximas décadas, as projeções populacionais elaboradas pelo IBGE (2018) apontam que o número de indivíduos aumentará gradativamente até o ano 2047, quando alcançará o valor de 233 milhões. A partir de então a população declinará, chegando a 228 milhões em 2060. Ainda para os próximos anos as estimativas relacionadas à migração são discretas, trazendo pouca influência ao longo desse processo.

Tal cenário será inédito no país, tanto em relação ao envelhecimento populacional como com o posterior declínio pós transição demográfica. Assim, a análise do comportamento dessas componentes permite compreender a contribuição em cada uma delas, em especial a migração, muitas vezes pouco considerada. A compreensão desse comportamento torna possível a preparação para as mudanças demográficas em curso e ainda direcionar as componentes de acordo com cenários possíveis (BONGAARTS; BULATAO, 1999), já que a formulação de políticas públicas pode e deve se basear na dinâmica dessas variáveis.

A migração e suas potencialidades

Diante do quadro de envelhecimento populacional mais recentemente descrito pelo IBGE (2018), surgem discussões sobre as possibilidades de que esse seja minimizado. Bacci (2012) aponta que isso seria possível através de um aumento nos níveis da fecundidade ou do fluxo migratório. Apesar de ambas serem viáveis, o crescimento no número de nascimentos levaria anos para ter efeito sobre a estrutura etária, além de depender de políticas relacionadas a outras áreas, como criação de creches e a inserção da mulher no mercado de trabalho. Já a migração possui efeito mais imediato e impactaria diretamente o grupo em idade ativa, onde, tendencialmente, se concentram os imigrantes (ROSA, 2012).

Apesar da grande potencialidade dessa componente em relação às mudanças em curso, estimá-la e avaliar seu impacto é uma tarefa que envolve grande complexidade, já que esse é um processo bastante sensível a mudanças sociais e econômicas. Ainda assim, é possível, com base em hipóteses, elaborar distintos cenários que possibilitem verificar seu efeito.

Uma das principais questões levantadas em projeções que consideram crescimento do saldo migratório é relativa ao tamanho da população no futuro. Estimando a reversão do envelhecimento em uma série de países, a ONU (2000) chega a resultados com saldos migratórios inimagináveis e

absolutamente insustentáveis, como argumenta Coleman (2001). Ainda que numericamente a população na análise proposta pela ONU (2000) seja exorbitante, nada se analisou acerca da contribuição de cada uma das componentes para o crescimento populacional em termos percentuais.

Mesmo Bongaarts e Bulatao (1999) e a ONU (2013), em suas análises acerca da contribuição de cada uma das componentes demográficas em projeções populacionais, foram pouco audaciosos em suas estimativas para a migração. Em ambos os casos, as tendências migratórias seguiram aquelas observadas em anos anteriores, sem supor uma alteração econômica ou social considerável que pudesse impactar esse cenário.

Metodologia

A análise da contribuição de cada componente demográfica baseou-se no cálculo das estimativas para o futuro da população do Brasil. Para tal, foram elaborados seis cenários com base no Método das Componentes Demográficas. Esse método baseia-se em pressupostos acerca do comportamento das componentes demográficas, considerando um nível para cada um dos grupos de idade. Assim, considera-se que a entrada na população ao longo do tempo ocorre através de nascimentos ou imigração e que a saída se dá pelos óbitos ou emigração.

Para a fecundidade, a hipótese assumida pelo IBGE (2013) para as próximas décadas mostrou-se bastante modesta no que diz respeito ao nível dessa componente. Segundo o estudo, em meados da década de 2030, a TFT brasileira alcançaria o valor de 1,50, o mantendo até 2050¹. Já para a ONU (2015), no quinquênio 2045-2050, esse indicador seria de 1,67. Conforme observado em anos recentes (IBGE, 2018), a queda na fecundidade parece ter sido mais branda do que aquela prevista pelo IBGE (2013) e, portanto, esses dados demandam ajuste em relação ao nível para que a queda prevista seja suavizada. Com essa finalidade, fez-se uso do nível previsto pelo ONU (2015) para o Brasil, que é menos audaciosa na previsão de queda, e manteve-se a distribuição etária estimada pelo IBGE (2013).

A projeção pelo Método das Componentes Demográficas depende, além dos pressupostos acerca do nível da mortalidade, fecundidade e migração, do perfil etário dessas componentes. Para mortalidade e fecundidade, esse padrão seguiu as hipóteses do IBGE (2013). Já para a migração tomou-se como base a estrutura etária do saldo migratório previsto para o Brasil, mas foram realizados ajustes posteriores.

Considerando que os migrantes “laborais” possuem duas características: “*concentram-se nas idades ativas, também elas as mais férteis, e possuem, com frequência, níveis de fecundidade superiores*” (ROSA, 2012 p.183) e que um aumento relevante no número de imigrantes seria relacionado a incentivos à migração por meio de políticas públicas, assumiu-se que a quase a totalidade dos migrantes pertenceria aos grupos etário em idade ativa, restando aos demais grupos somente uma parcela residual. Nesse sentido a presente análise considerou que todos os imigrantes teriam idade entre 15 e 64 anos. Vale ressaltar que a projeção teve um intuito metodológico, e embora a concentração da migração nesses grupos seja tendencial, ela não seria assim concentrada em sua totalidade.

Cenários projetados

Toda a análise realizada no presente estudo faz uso de hipóteses acerca das tendências para os próximos anos para a fecundidade, mortalidade e migração. Para tal, baseia-se no uso de cenários que somente se concretizarão nas próximas décadas caso as hipóteses feitas se confirmem. Portanto, trata-se de um método determinístico, o que torna inviável saber o quão próximo da realidade encontra-se, sendo apenas um “*retrato detalhado do que o futuro poderá ser*” (BORGES *et al.*, 2006, p.8)

¹ Na última revisão das projeções, realizadas pelo IBGE, em 2018, o nível médio projetado para fecundidade, no período 2045-2050, foi de 1,68 filho por mulher, aproximando-se da estimativa da ONU, que foi incorporada neste artigo. As estimativas desse estudo foram feitas antes da publicação de tais projeções populacionais.

Além dos cenários necessários para aplicação da metodologia – Fechada, Reposição e Mortalidade constante – foram elaborados outros três cenários que alteram a componente migratória para que seja possível avaliar seu impacto.

Todos os cenários partiram da população de 2010 das projeções do IBGE (2013). Os pressupostos foram apenas iniciados a partir do ano de 2017. Dessa forma, entre 2010 e 2016, todos os cenários são idênticos.

Cenário População Fechada

O Cenário População Fechada considera que o saldo migratório será nulo ao longo de todo o período projetado. Além disso, a fecundidade desse cenário foi dada pela estimativa suavizada baseada nos dados da ONU (2015) e a mortalidade foi a estimada pelo IBGE (2013).

Cenário Reposição

O Cenário Reposição considera uma população fechada ao longo dos anos, com fecundidade ao nível de reposição desde o início do cálculo e a mortalidade estimada em IBGE (2013).

Cenário Mortalidade Constante

O Cenário Mortalidade Constante considera que o saldo migratório é nulo e a taxa de fecundidade permanece constante no nível de reposição ao longo de todo o período. Em relação à mortalidade, essa permanece constante nos níveis estimados pelo IBGE (2013) no ano de 2016.

Cenário RDI 12

O Cenário RDI 12 baseia-se na análise feita por Pereima e Porse (2013), que buscaram relacionar as variáveis crescimento do PIB, taxa de investimento e taxa de poupança à Razão de Dependência de Idosos (RDI). Nessa análise, através de um método não paramétrico, os autores verificaram uma relação positiva entre as três variáveis e o indicador para valores baixos da RDI e negativa para valores altos. Os autores concluíram que a inversão nessa relação ocorreu quando a RDI ultrapassa o valor de 8 para o crescimento do PIB e de 12 para as demais variáveis. Como já a priori a manutenção da RDI em 8 seria totalmente improvável (essa foi a RDI estimada para o Brasil em 1991), adotou-se a hipótese de RDI em 12 ao longo de todo o período. Para isso a migração foi fixada em valores que permitissem a manutenção de tal estrutura etária. A fecundidade para esse cenário será aquela resultante da suavização anteriormente explicitada e a mortalidade é a mesma estimada pelo IBGE (2013).

Cenário Atenuação da RDI

O Cenário Atenuação da RDI busca desacelerar o envelhecimento populacional no Brasil. Para tal, baseia-se na taxa de crescimento geométrico anual da RDI do Chile. O Chile foi escolhido por tratar-se de um país com algumas características semelhantes às do Brasil, tanto economicamente quanto socialmente, porém, com uma população que envelhece em ritmo menos acelerado.

Para que o indicador permanecesse em um nível tal que seu ritmo de crescimento fosse o mesmo do chileno, foi variada a componente migratória. Portanto, para a fecundidade foi utilizada a suavização, para mortalidade as estimativas do IBGE e, para a migração, foi estimado o nível de migração necessária para o controle do ritmo de crescimento da RDI.

Cenário TLM da Alemanha

O Cenário TLM da Alemanha buscou retratar um panorama com um acentuado crescimento no número de imigrantes, já que o país adotou, recentemente, um posicionamento político favorável à entrada de imigrantes. Sendo assim, o presente cenário teve como pressuposto para a

componente migratória a TLM observada na Alemanha no quinquênio 2010-2015, que foi de 3,1. A fecundidade é a suavizada e a mortalidade do IBGE.

Contribuição das componentes

A análise da contribuição de cada componente demográfica para o crescimento populacional foi proposta por Bongaarts e Bulatao (1999). Esse cálculo baseia-se na elaboração de três projeções distintas e simples e, através de cálculos algébricos, retira a influência de cada variável, possibilitando obter os resultados de interesse. O cenário mais básico é aquele que possui a população fechada (P_{fechada}). Em seguida, é necessário um cenário que, além da migração nula, possua também uma taxa de fecundidade no nível de reposição ao longo de todo o período ($P_{\text{reposição}}$). O terceiro cenário necessita, além das hipóteses para o cenário anterior, de uma mortalidade constante ao longo de todo o período projetado ($P_{\text{mortalidade}}$). Existe ainda a possibilidade de um quarto cenário, para estimar o impacto da migração ($P_{\text{migração}}$).

Após a elaboração dos cenários, para o cálculo da contribuição da componente migratória ($C_{\text{migração}}$), calcula-se a diferença entre o total de indivíduos obtido com base em uma população fechada e aquele que considerou a população aberta (1). O impacto da fecundidade ($C_{\text{fecundidade}}$) se dá pela diferença entre o cenário com população fechada e aquele que considera uma TFT no nível de reposição (2). Já a contribuição da mortalidade ($C_{\text{mortalidade}}$) se dá pela subtração entre a população fechada e a população com mortalidade constante (3). Por fim, também é estimado o efeito da estrutura etária sobre o crescimento populacional (C_{momentum}), calculado a partir da fixação de uma variação nula das demais componentes. O *momentum* é, na verdade, o efeito que o tamanho da população de mulheres em idade fértil tem na população, ainda que haja variação na taxa de fecundidade. Como exemplo, ainda que exista queda na TFT, uma população com relativa concentração de mulheres em idade ativa continuaria resultando em aumento da população por, ao menos, determinado período de tempo. Assim, o *momentum* resume-se como sendo o impacto da estrutura etária da população. Para esse cálculo considera-se a diferença entre a população com mortalidade constante e a população no ano inicial (P_0) para esse resultado (4). As fórmulas que resumem esses cálculos estão a seguir.

$$C_{\text{migração}} = P_{\text{fechada}} - P_{\text{migração}} \quad (1)$$

$$C_{\text{fecundidade}} = P_{\text{fechada}} - P_{\text{fecundidade}} \quad (2)$$

$$C_{\text{mortalidade}} = P_{\text{fechada}} - P_{\text{mortalidade}} \quad (3)$$

$$C_{\text{momentum}} = P_{\text{mortalidade}} - P_0 \quad (4)$$

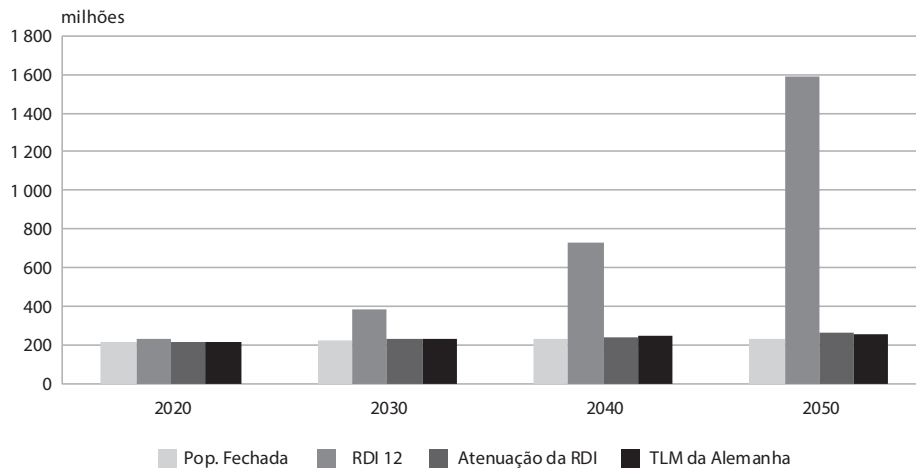
Dessa forma, torna-se possível estimar a contribuição de cada componente para o crescimento populacional do Brasil em relação à população inicial do ano de 2016 nos próximos anos.

Análise dos resultados

A interação entre a fecundidade, mortalidade e migração, assim como a estrutura etária da população, trazem mudanças sobre o crescimento demográfico de acordo com os pressupostos assumidos. Análises que permitam a distinção entre o impacto de cada um desses são relevantes para que ocorra direcionamento das políticas públicas, considerando de modo mais profundo as mudanças demográficas das próximas décadas.

Com base nos 6 cenários projetados e descritos anteriormente (População Fechada, Reposição, Mortalidade constante, RDI 12, Atenuação da RDI e TLM Alemanha) foram obtidos os resultados apresentados no Gráfico 1, que traz os respectivos totais populacionais.

Gráfico 1 - População total, por cenário projetado - Brasil - 2020/2050



Fonte: Elaboração própria.

Nota: População inicial da projeção em 2010 (IBGE, 2013).

Os resultados obtidos ao final do horizonte de projeção mostram que as hipóteses assumidas no Cenário RDI 12 levam a um volume populacional praticamente impossível para o país. Essa população ultrapassaria o valor de 1,5 bilhão em 2050 nesse cenário, que aproxima-se daquele estimado pela ONU (2000) e criticado por Coleman (2001). Nos demais cenários a população em 2050 variou entre 234 milhões de habitantes (na ausência de migração) e 260 milhões, que é o Cenário Atenuação da RDI.

Ainda analisando os resultados associados às populações encontradas através das projeções realizadas, elaborou-se o Gráfico 2, com as estruturas etárias dos diferentes cenários de migração.

Gráfico 2 - Composição da população segundo o sexo e a faixa etária, por cenários selecionados - Brasil - 2050



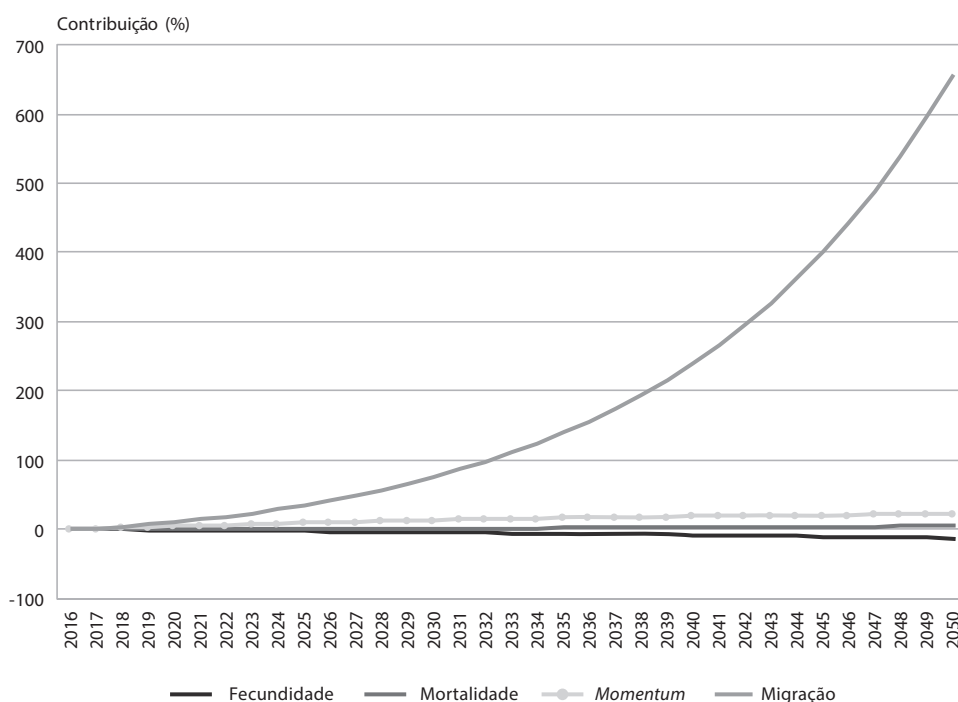
Fonte: Elaboração própria.

Nota: População inicial da projeção em 2010 (IBGE, 2013).

Dentre as estruturas etárias dos diferentes cenários, a referente ao cenário RDI 12 se destaca das demais devido à considerável concentração de indivíduos nas idades consideradas “ativas”, que resulta da estrutura etária dos imigrantes, que se concentram nas mesmas idades. Assim, para impedir o envelhecimento populacional, seria necessário cada vez mais indivíduos entrando no país, o que leva a uma distribuição pelos grupos de idade de tal forma. Além disso, é possível perceber o início de um processo de alargamento da base da pirâmide, resultado do número de nascimentos oriundo de uma população tão concentrada nas idades férteis. Os Cenários Atenuação da RDI e TLM da Alemanha, apesar de receberem uma parcela consideráveis de imigrantes, levaram a resultados mais discretos. Ainda que as idades jovens aumentem proporcionalmente, os efeitos desses pressupostos são mais amenos e, potencialmente, poderiam ser considerados mais factíveis.

Buscando verificar o impacto de cada componente sobre a população ao longo dos anos, foram elaborados os Gráficos 3, 4 e 5, que apresentam a contribuição de cada componente em relação à população de 2016 para os três cenários analisados.

Gráfico 3 - Mudança nas componentes demográficas ao longo do tempo no Cenário “RDI 12”, como proporção da população total em 2016 - Brasil - 2016-2050



Fonte: Elaboração própria.

Nota: População inicial da projeção em 2010 (IBGE, 2013).

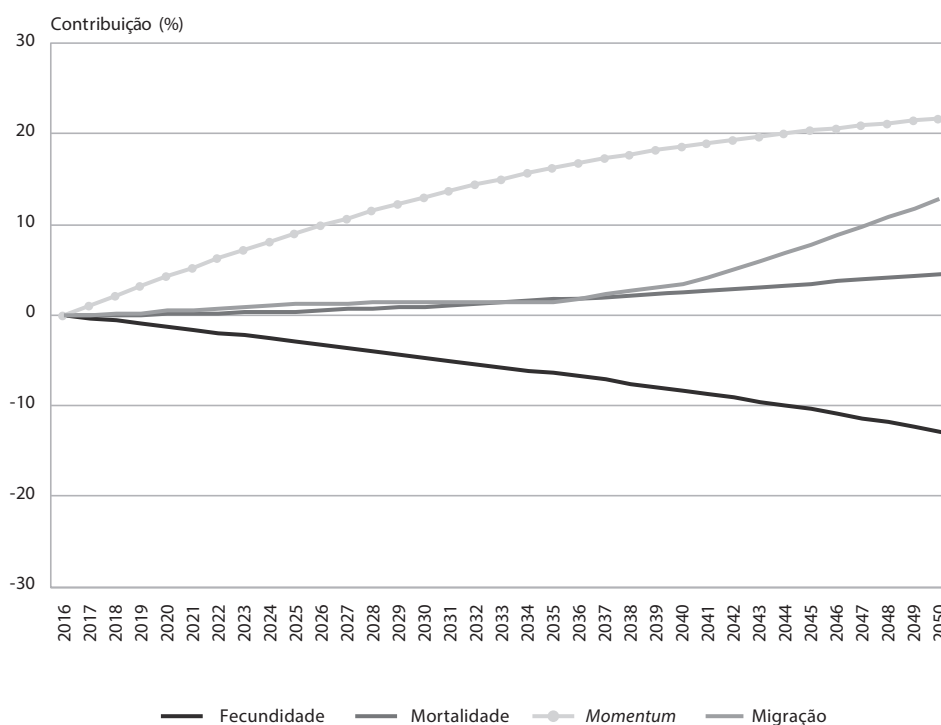
No Cenário RDI 12, a componente migratória é a principal responsável pelo crescimento exponencial da população futura: em 2050, mais de 600% do crescimento populacional seria dado pela migração. Como um todo, a combinação da contribuição das componentes levaria a um aumento de 770% da população em 2050, quando comparada ao ano de 2016. Esse resultado assemelha-se ao descrito pela ONU (2000) e tidos como utópico por Coleman (2001). De fato, interromper o envelhecimento populacional é uma tarefa bastante árdua e com pouca ou nenhuma chance de sucesso.

Já no Cenário Atenuação da RDI, com resultados dispostos no Gráfico 4, a migração já não foi a principal força no crescimento populacional. Nesse cenário, a estrutura etária da po-

pulação assume esse papel, mostrando que, mesmo com um número de entrada de imigrantes considerável, as demais variáveis demográficas tem um efeito considerável. Com essa dinâmica, 21% do crescimento populacional do período 2016-2050 resulta da distribuição etária observada no ano inicial da projeção. A migração contribui ao longo dos anos de maneira modesta, apenas aumentando sua contribuição já na metade da década de 2030, quando representa 11% do crescimento em relação à 2016. Entre 2016 e 2050, a população brasileira teria um aumento de 26% se observadas as hipóteses realizadas.

Tal aceleração do crescimento da contribuição da componente migratória pode ser explicada justamente pelo fato de que nesse período a RDI avança com mais intensidade, “forçando” a entrada de um número maior de imigrantes para que o pressuposto do cenário ocorra.

Gráfico 4 - Mudança nas componentes demográficas ao longo do tempo no Cenário “Atenuação da RDI”, como proporção da população total em 2016 - Brasil - 2016-2050

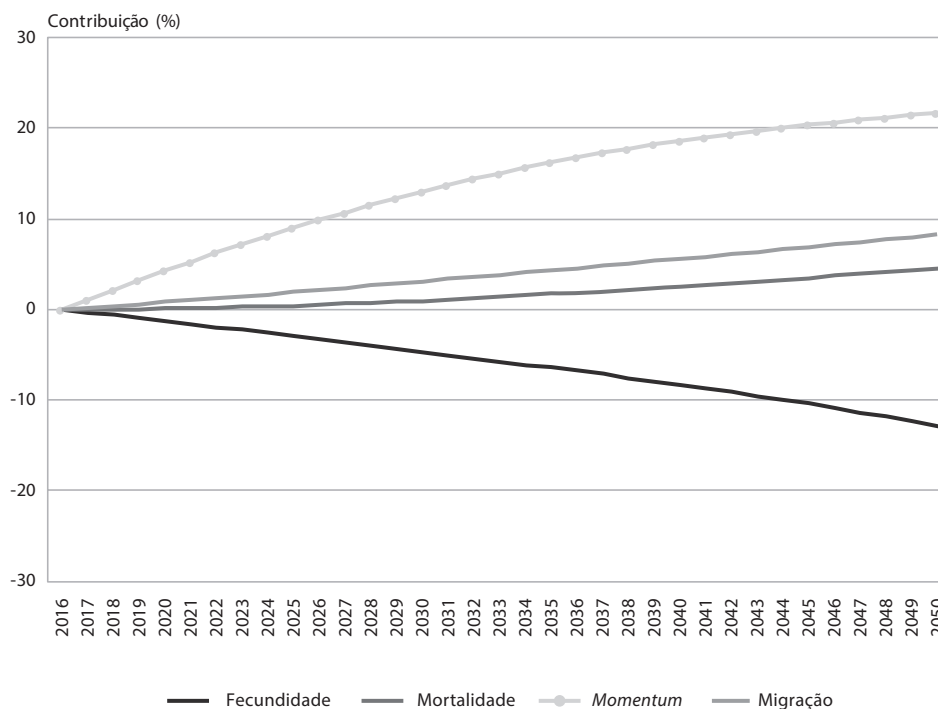


Fonte: Elaboração própria.

Nota: População inicial da projeção em 2010 (IBGE, 2013).

Por fim, o Gráfico 5 corresponde ao Cenário TLM da Alemanha, com resultados que se assemelham aos encontrados no Cenário Atenuação da RDI.

A principal diferença entre ambos reside no fato de que a contribuição da componente migratória cresce gradativamente ao longo do período analisado e não somente a partir da década de 2030, como no cenário anterior. Sob esse cenário, a estrutura etária mostra-se como principal contribuinte, seguida da migração, mortalidade e fecundidade. Enquanto a estrutura etária é responsável por 21% do crescimento, a migração o é por 7. Entre o ano inicial e final de projeção, o crescimento populacional nesse cenário foi de 22%.

Gráfico 5 - Mudança nas componentes demográficas ao longo do tempo no Cenário "TLM da Alemanha", como proporção da população total em 2016 - Brasil - 2016-2050

Fonte: Elaboração própria.

Nota: População inicial da projeção em 2010 (IBGE, 2013).

Para todos os cenários, as hipóteses acerca da fecundidade e mortalidade são as mesmas, e, portanto, suas contribuições são mantidas. Nos contextos analisados, a fecundidade seria a única força agindo sobre a população que levaria ao seu decréscimo, já que a TFT encontrasse abaixo do nível de reposição. Ao final do período, essa componente isolada é responsável por um decréscimo da população de 10%. De maneira contrária, o aumento da expectativa de vida levaria ao aumento populacional, principalmente na segunda metade do período considerado. As informações obtidas com os cálculos encontram-se resumidas na Tabela 1.

Tabela 1 - Contribuição percentual dos cenários analisados, segundo as componentes demográficas - Brasil - 2030/2050

Componente demográfica	Contribuição percentual dos cenários analisados (%)					
	RDI 12		Atenuação da RDI		TLM da Alemanha	
	2030	2050	2030	2050	2030	2050
Fecundidade	(-) 5,0	(-) 13,0	(-) 5,0	(-) 13,0	(-) 5,0	(-) 13,0
Mortalidade	1,0	4,0	1,0	4,0	1,0	4,0
Migração	76,0	657,0	1,0	13,0	3,0	8,0
Momentum	13,0	22,0	13,0	22,0	13,0	22,0

Fonte: Elaboração própria.

Notas: 1. População inicial da projeção em 2010 – IBGE (2013).

2. RDI = Razão de dependência de idosos.

3. TLM = Taxas líquidas de migração.

Nos cenários com saldo migratório menores, ou seja, nos casos Atenuação da RDI e TLM da Alemanha, nota-se que a principal contribuição ao crescimento populacional será da estrutura etária, enquanto para o decréscimo será a fecundidade. Isso apenas se altera no Cenário RDI 12, já que a migração é a principal responsável pelo aumento populacional sob essas hipóteses.

De acordo com os resultados apresentados, nota-se que a componente migratória pode ter papel fundamental no crescimento populacional, dependendo do número de imigrantes considerado. Cabe ressaltar, porém, que a estrutura etária da populacional também continuará exercendo forte influência nesse processo.

Considerações finais

A interação entre os diferentes cenários de migração, mortalidade, fecundidade e estrutura etária resultará em diferentes totais e composições populacionais no Brasil.

Analisando os cenários propostos, sob a hipótese de interrupção do envelhecimento, a migração faz com que a população seja quase 500% maior em relação ao ano de 2016. Esse resultado mostra que, ainda que fosse desejável, impedir de maneira definitiva o avanço do envelhecimento via migração levaria realmente a uma explosão populacional, o que traria diversas consequências econômicas e sociais, além do fato de não ser factível.

Por outro lado, o Cenário Atenuação da RDI mostrou que a suavização dos efeitos dessas transformações seria menos impactante. Ainda que o fluxo migratório aumentasse, a chegada de imigrantes não seria a principal força impulsionando o crescimento populacional nesse cenário. Esse papel seria assumido pelo *momentum* populacional, que seria responsável por 21% do aumento da população entre 2016 e 2050.

Do mesmo modo, a alteração no fluxo migratório de maneira considerável, considerando dinâmicas de países receptores de imigrantes (como assume o cenário TLM da Alemanha), não implica que a migração seja a principal contribuinte ao crescimento. Mais uma vez, a estrutura etária seria o principal componente nesse processo.

Assim, percebe-se que apesar da migração exercer influência sobre o tamanho populacional, por vezes, sua contribuição é superada pela estrutura de idades da população. Isso mostra que o uso da migração com intuito demográfico, como suavizar o envelhecimento populacional, conforme sugerido por Bacci (2012), traz os benefícios nesse sentido e ainda que leve ao aumento do contingente populacional, é o *momentum* populacional que trará a principal contribuição ao crescimento no caso brasileiro. Diante de um cenário de migração internacional com aumento do fluxo migratório sem alcançar valores absurdos, o foco potencial das políticas pode reunir políticas relacionadas à estrutura etária e à migração, de modo a considerar a demografia não como destino, mas como dinâmica que pode influenciar determinados processos sociais e econômicos. Cabe ressaltar que o crescimento populacional (inclusive de grupos etários específicos, como dos idosos), não precisa ser visto como um problema. Como amplamente tratado na literatura econômica e da demografia econômica, o crescimento da população é um fator importante no desenvolvimento econômico e social, e ainda mais, a distribuição etária da população também seria um fator com relevância e que merece a devida atenção (OLIVEIRA, 2014).

Por fim, ressalta-se que as análises aqui realizadas são apenas exercícios que buscam retratar o futuro diante de certas condições. A dinâmica demográfica é um processo complexo, que não pode ser previsto com exatidão. Ainda que seja possível estimar a contribuição de cada componente e que isso tenha se mostrado relevante, é importante ressaltar que esses resultados referem-se às hipóteses realizadas. Mesmo diante dessa limitação, a preparação para os processos em curso são fundamentais para que haja o direcionamento das políticas públicas de modo que os impactos dessas transformações tragam o mínimo de consequências possíveis, tanto sociais, quanto econômicas.

Referências

- ALVES, J. E. D. Transição demográfica, transição da estrutura etária e envelhecimento. *Revista Portal de Divulgação*, Ano IV, n.40, p. 8–15, 2014
- BACCI, M.L. *Breve historia de las migraciones*. Traducción: Marco Aurelio Galmarini. Madrid: Alianza Editorial, 2012
- BONGAARTS, J.; BULATAO, R. A. *Completing the demographic transition*. *Population and Development Review* vol. 25, No. 3, pp. 515–529., 1999
- BORGES, A. D. S., MARQUES, C. S., BRITO, L. P. G., SILVA, V. R. L., & JANNUZZI, P. D. M. *Projeções populacionais no Brasil: subsídios para seu aprimoramento*. Caxambu: Associação Brasileira de Estudos Populacionais (ABEP), 2006
- 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, p. 177-210, 2014
- COLEMAN, D. A. Replacement migration, or why everyone is going to have to live in Korea: a fable for our times from the United Nations. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, v. 357, n. 1420, p. 583-598, 2001.
- FARIA, V. E; POTTER, J. E. 2002. Televisão, telenovelas e queda de fecundidade no Nordeste. *Novos Estudos*, no 62, pp. 21-39.
- IBGE, *Projeções da população: Brasil e Unidades da Federação*. Rio de Janeiro: IBGE, 2013.
- _____. *Projeções da população: Brasil e Unidades da Federação*. Rio de Janeiro: IBGE, 2018.
- OLIVEIRA, A. T. R. Transição demográfica e migrações internacionais: O Brasil na rota dos países desenvolvidos? *Plataforma Política Social*. Texto para discussão, n. 2, 2014
- ONU Replacement migration: is it a solution to declining and ageing populations? New York: United Nations, 2000
- ONU Demographic components of future population growth. Technical Paper 3, 2013
- ONU World Population Prospects: The 2015 Revision, custom data acquired via website, 2015
- PEREIRA, J. B.; PORSSE, A. Transição demográfica, acumulação de capital e progresso tecnológico: desafios para o crescimento brasileiro. *Revista Economia & Tecnologia*, v. 9, n. 1, p. 49-60, 2013.
- ROSA, M. J. V. Imigração e envelhecimento: ligações perigosas. *Revista Migrações*, n. 10, p. 183-184, 2012.
- SIMÕES, C. C. S. *A transição da fecundidade no Brasil: análise de seus determinantes e as novas questões demográficas*. São Paulo: Arbeit Factory Editora e Comunicação, 2006.
- WONG, L. L. R.; CARVALHO, J. A. O rápido processo de envelhecimento populacional do Brasil: sérios desafios para as políticas públicas. *Revista Brasileira de Estudos de População*, 23(1), p. 5–26. 2006.

Morte e vida dos barnabés do IBGE*

André Bruno de Oliveira**

Kaizô Iwakami Beltrão***

Resumo

Objetivo: Verificar as diferenças no perfil de mortalidade dos funcionários do IBGE para as carreiras do nível médio e do nível superior, entre 2000 e 2018. **Metodologia:** Foram concatenados os dados administrativos dos funcionários falecidos aos dados do sistema de mortalidade do Ministério da Saúde. Adotou-se a lista de causas de morte evitáveis para o Brasil, ajustando a idade limite para 79 anos. Como método de medição utilizou-se o cálculo dos anos de vida perdidos extraídos de tábuas de mortalidade construídas para os funcionários. **Resultados:** O valor dos anos perdidos por neoplasias é maior para o nível superior do que para o nível médio e no caso das doenças do aparelho circulatório a situação é a inversa. A comparação feita com os paulistas mostra valores por neoplasias maiores para os funcionários. Os anos perdidos por mortes evitáveis são de 4,9 anos por funcionários da carreira do nível superior e 5,8 anos por funcionários da carreira de nível médio, contra 9,1 anos por habitantes paulistas. A situação apresentada demanda ações para melhorar a qualidade de vida e possivelmente direcionadas a diagnóstico precoce, imunização e mudança de hábitos.

Palavras-chave: Mortalidade por causa. Política de saúde. População. Ocupação.

* O capítulo é resultado de um estudo iniciado na dissertação de mestrado concluída em 2012 por André Bruno de Oliveira, intitulada "Tábua de Mortalidade por grande grupo de causas para os servidores do IBGE", sob orientação de Kaizô Iwakami Beltrão. Estes dois pesquisadores, em um segundo momento, desenvolveram este presente estudo sobre mortalidade dos funcionários do IBGE para um período de tempo mais recente. O objetivo principal foi identificar as diferenças existentes na mortalidade entre os funcionários das carreiras/escolaridade do nível médio e do nível superior, como forma de melhor direcionar os futuros exames médios periódicos que vierem a ser realizados na instituição.

** Tecnologista do IBGE, concluiu o mestrado em Estudos Populacionais e Pesquisas Sociais na ENCE em 2012.

*** Pesquisador/professor da EBAPE/FGV-RJ, atuou como Professor/pesquisador da ENCE/IBGE de 1987 a 2010.

Introdução

Informações relacionadas às principais causas de morte nas populações são insumos fundamentais para os debates sobre políticas de saúde (LOZANO, 2012). Nesta situação, os registros de óbito são imprescindíveis para a avaliação da saúde de uma população. As principais características do falecido como, idade, sexo, escolaridade, ocupação, etnia, local de residência e de morte, e a própria causa da morte, servem para identificar fatores de risco associados às características da população exposta e propor ações para mitigar os riscos de morte (ARAGÓN, 2008).

Os dados administrativos de funcionários públicos têm sido usados para a construção de tábuas de mortalidade (TM) e tábuas de múltiplos decrementos (TMD). Em todos esses estudos o risco de morte do pessoal de nível médio está acima do risco do pessoal de nível superior para ambos os sexos (BELTRÃO & SUGARA, 2002; SILVA, 2008; BORGES, 2010; ARTEAGA, 2009; BELTRÃO & SUGARAH, 2017).

A recuperação da causa de morte para avaliação da mortalidade por causas é uma solução que complementa o alcance de dados administrativos de funcionários públicos. O trabalho de Oliveira (2010) apresenta uma metodologia que concatena os dados dos funcionários do Instituto Brasileiro de Geografia e Estatística (IBGE) com os dados do Sistema de Mortalidade (SIM) do Ministério da Saúde (MS). A pesquisa aponta o número de anos de vida perdidos (AVP) por grupo de causas de morte não evitáveis e evitáveis associadas às ações de diagnóstico precoce e mudança de hábitos.

Um modo de aferir o impacto das causas de morte na população é calcular o número de AVP. Por ser um indicador que pondera as mortes pelo tempo de vida perdido, diversos pesquisadores usam o AVP para avaliar as mortes prematuras. Uma parte dessas mortes incluem casos de óbitos evitáveis por algum tipo de ação de prevenção (SILVA, 2001; BELTRÃO & DELLASOPPA, 2011; MESALLES-NARANJO *et al.*, 2018; FRINGS *et al.*, 2018; EDNEY *et al.*, 2018). A adoção deste indicador combinado às causas de morte permite conhecer o perfil epidemiológico da população e colabora para evidenciar demandas por ações de saúde (LEITE, 2015).

O objetivo deste trabalho é identificar as diferenças no perfil de mortalidade dos funcionários do IBGE de 2000 a 2018 segundo a escolaridade. Os grupos populacionais selecionados são os da carreira do nível médio e os do nível superior. Para identificar essas diferenças, primeiramente o trabalho faz uma revisão bibliográfica sobre causas evitáveis, AVP e ações de prevenção. A seção de metodologia descreve as bases de dados, metodologia para recuperação da causa de morte, os grupos de causas de morte, método para calcular AVP, probabilidade de morte e TM. A seção de resultados inclui a identificação das diferenças. A seção de considerações finais resume os resultados.

Revisão bibliográfica

Causas evitáveis

A implementação de um método para medir a qualidade dos cuidados médico-hospitalares deve identificar casos desnecessários de doenças, incapacidades e mortes prematuras desnecessárias: primeiramente é importante avaliar em que condições esses casos ocorreram (por que aconteceu?); depois identificar as condições nas quais aumentos críticos nas taxas de doença, invalidez e morte prematura podem servir como índices da qualidade do atendimento médico-hospitalar; por fim, é importante identificar categorias amplas de doenças nas quais o estudo possa revelar características que sirvam como índices de saúde (Rutstein *et al.*, 1976).

O conceito de morte evitável vem sendo utilizado por diferentes pesquisadores desde a década de 1970. A identificação e classificação das causas básicas de morte em grupos segundo as ações para evitabilidade colabora para aferir a qualidade e a diversidade da atenção à saúde.

Este tipo de avaliação permite estabelecer uma associação entre os casos de mortes prematuras e os fatores de risco reduzíveis por ações de intervenção (MALTA *et al.*, 2007 e 2018).

Para o Brasil, Malta *et al.* (2007 e 2018) propõem duas listas de causas de morte evitáveis: uma para menores de 5 anos e outra para as pessoas com 5 até 75 anos de idade. As duas listas utilizam como referência a tecnologia disponível no Sistema Único de Saúde (SUS) para definir grupos de ações de causas evitáveis. A idade limite da lista de 75 anos possui carácter parcialmente arbitrário, sendo motivada em parte pela expectativa de vida da população e pela maior dificuldade em se definir a causa básica, por oposição as causas contribuintes. As causas definidas como evitáveis são aquelas que total ou parcialmente são preveníveis pela ação efetiva dos serviços de saúde disponíveis (ou acessíveis) em um determinado local e momento histórico. O estudo está fundamentado em uma revisão da literatura referente à base conceitual e empírica das listas de causas de morte evitáveis, publicadas entre 1975 e 2004, e nas reflexões de um grupo de trabalho organizado pelo MS do Brasil.

Anos de vida perdidos

Há registros de que Makeham (1874) publicou um dos primeiros estudos envolvendo as perdas de anos de vida devido à causa de morte, neste caso a varíola. O pesquisador, com base nos estudos de Bernouilli, D'Alembert e Laplace, usa a composição das forças decrementais de saída por morte para explicar a relação entre a extinção da varíola e o aumento da população. Desde então diversos estudos têm avaliado o impacto dos riscos competitivos na população.

No Brasil em 2008, o cálculo dos AVP ajustados por incapacidade evidencia as causas não transmissíveis mais proeminentes: Doenças cardiovasculares; transtornos mentais, com destaque para a depressão; o diabetes; e a doença pulmonar obstrutiva crônica. Com base nos resultados, os pesquisadores apontam demandas por ações transversais, que extrapolem as políticas específicas que envolvem a área da saúde (LEITE *et al.*, 2015).

Um estudo longitudinal com dados de 1990 a 2015 descreve o uso de AVP para aferir aumentos de expectativa de vida no Brasil (FRANÇA *et al.*, 2017). A redução das mortes por doenças cardiovasculares trouxe importante contribuição para a expectativa de vida nas regiões Sul, Sudeste e Centro-Oeste; e na região Norte, o aumento da expectativa de vida está mais associado à redução da mortalidade por diarreia, infecções respiratórias inferiores e outras doenças infecciosas.

O trabalho conduzido por Kanson (2013) estima o impacto das causas de morte evitáveis sobre a expectativa de vida dos idosos em São Paulo. A eliminação das causas evitáveis principais indica ganhos de 20% na expectativa de vida aos 60 anos. A avaliação por sexo e causas apresenta *grossa modo* sobremortalidade masculina, o que sugere maior exposição a fatores de risco e menor utilização dos serviços de saúde pelos homens.

Para os funcionários do IBGE de 1996 a 2010, a avaliação de AVP por causas evitáveis sinaliza a existência de fatores de risco compartilhados: sal em excesso, uso de tabaco, alimentação pouco saudável, sedentarismo, ingestão nociva de álcool, falta de diagnóstico tempestivo. Estes fatores são mais agravados nos funcionários do sexo masculino do que no feminino (OLIVEIRA, 2010).

Estudos internacionais têm usado os AVP para identificar fatores de risco reduzíveis por diferentes ações de intervenção e medir ganhos de expectativa de vida: na Escócia de 2000 a 2015, o estudo aponta redução das perdas por morte em anos de vida e desigualdades desacerbadas para algumas causas específicas (MESALLES-NARANJO *et al.*, 2018); na Austrália as análises indicam que os aumentos nos gastos com saúde promovem resultados mais evidentes nas áreas mais pobres (EDNEY *et al.*, 2018); na região de Friuli-Venezia Giulia, Itália, a utilização do AVP evidencia as neoplasias malignas e as doenças cardiovasculares (VALENT & LORIS ZAINER, 2015).

Ações de intervenção

Esta seção contém uma revisão das doenças, agravos e ações de intervenção em saúde que podem ser sugeridas para assegurar à saúde dos funcionários do IBGE. Um exemplo é a imuniza-

ção para evitar a tuberculose e outras doenças bacterianas como a hepatite B. Uma outra ação é o uso de preservativo para evitar o contágio pelo vírus da imunodeficiência humana (REDE, 2008). Essas doenças estão incluídas no capítulo I das doenças infecciosas e parasitárias (WHO, 2018).

As neoplasias (capítulo II) mais comuns nos homens brasileiros estão localizadas na próstata e no pulmão e nas mulheres na pele, mama e colo do útero. Essas neoplasias possuem fatores de riscos reduzíveis e compartilhados com outras causas de morte: i) no pulmão, traqueia e brônquios estão associadas ao tabagismo; ii) no esôfago, ao consumo de álcool e dieta pobre em fibras; iii) no estômago, ao consumo de sal e alimentos defumados; iv) no cólon do reto, às dietas ricas em gorduras; v) na pele, à radiação solar; vi) na mama feminina, ao comportamento hormonal e reprodutivo; vii) no colo do útero, à higiene precária e à exposição ao vírus do papiloma humano; viii) na boca, à irritação mecânica crônica (WHO, 2007).

As doenças endócrinas, nutricionais e metabólicas (capítulo IV) incluem a diabetes mellitus, desnutrição, obesidade e transtornos de glândulas (e.g. tireoide) (WHO, 2018). A obesidade, por exemplo, está vinculada ao excesso de calorias e ao uso de medicamentos (WHO, 2007). Em relação à diabetes, Barbosa, Barceló e Machado (2001) descrevem que as consequências no aparelho circulatório podem levar ao infarto agudo do miocárdio, ao acidente vascular cerebral, insuficiência renal crônica, amputações de pés e pernas, cegueira definitiva, abortos e mortes perinatais.

As doenças do aparelho circulatório (capítulo IX), também conhecidas como doenças cardiovasculares, incluem as doenças hipertensivas, isquêmicas do coração, cerebrovasculares, reumáticas, outras doenças cardíacas e transtornos do aparelho circulatório. Essas doenças são responsáveis pela principal causa de incapacidade e de morte prematura em todo o mundo e contribuem substancialmente para os custos de cuidados com a saúde (WHO, 2007).

A redução do risco das doenças cardiovasculares está associada aos seguintes fatores: i) fazer escolhas alimentares saudáveis; ii) ser fisicamente ativo; iii) reduzir índice de massa corporal; iv) controlar pressão arterial; v) controlar colesterol no sangue; vi) controlar glicemia; vii) não usar tabaco; e viii) bem-estar mental. Estas medidas comportamentais não têm efeitos prejudiciais à saúde e desempenham um papel etiológico importante em outras doenças não transmissíveis, como: as neoplasias; doenças respiratórias; diabetes; osteoporose; e doenças do fígado. Além disso, a mudança comportamental é mais econômica do que o tratamento à base de medicamentos que podem ter efeitos adversos (WHO, 2007).

Uma medida importante para reduzir os riscos das doenças cardiovasculares é a alimentação equilibrada que exige o controle e ingestão de determinados alimentos importantes à saúde: i) a gordura saturada está associada ao aumento do colesterol LDL que é prejudicial à saúde; ii) a ingestão de peixe e certos nutrientes como canola e soja, ricos em Omega-3, podem agir como protetores em pessoas portadoras de doenças cardíacas coronarianas; iii) o consumo excessivo de sal está associado ao aumento da pressão sanguínea; iv) a ingestão excessiva de álcool está associada ao aumento das mortes súbitas e das doenças cerebrovasculares; e v) o consumo de frutas e vegetais ricos em potássio e fibras podem produzir protetores contra as doenças cardíacas e acidente vascular cerebral (WHO, 2007).

Segundo a WHO (2004), os fatores de risco mais comuns associados às doenças respiratórias (capítulo X) são: i) exposição ao fumo do tabaco, como fator primário, e as causas associadas a outras doenças como as cardiovasculares, neoplasia e diabetes; ii) outros fatores associados à poluição atmosférica; iii) desnutrição e baixo peso ao nascer; e iv) múltiplas infecções pulmonares nas idades mais jovens. Além destes fatores citados, regiões com baixa condição socioeconômica e falta de acesso aos serviços de saúde contribuem para a prevalência e incidência das doenças respiratórias. Um exemplo de ação preventiva é imunizar a população com alto risco de complicações relacionada à gripe (influenza e pneumonia), por exemplo, os idosos.

As doenças do aparelho digestivo (capítulo XI) incluem as doenças do fígado associadas ao álcool, pancreatite e hemorragias digestivas (WHO, 2018). Um estudo evidencia que a cirrose hepática tem alto nível de morte entre os militares ativos e aposentados da marinha do Brasil. Os dados de mortalidade desse estudo são extraídos de atestados de óbitos referente ao período

de 1991 a 1995. Os resultados indicam a existência de possíveis fatores de risco ocupacionais associados ao estilo de vida deste grupo (SILVA & SANTANA, 2004).

Uma doença muito frequente entre as doenças do aparelho geniturinário (capítulo XIV) é a infecção do trato urinário que na idade adulta acomete preferencialmente as pessoas do sexo feminino e está relacionada à atividade sexual. A maior susceptibilidade feminina está associada à proximidade dos órgãos genitais uretra e ânus. No homem, o maior comprimento uretral atua como fator protetor, no entanto quando ocorre pode estar associada à doença da próstata (carcinoma de próstata) e ao estreitamento da uretra. Em pessoas idosas, a doença é mais comum em ambos os sexos. Além destes fatores, a infecção urinária na presença de diabetes mellitus pode comprometer o controle glicêmico do paciente (HEILBERG e SCHOR, 2003).

O capítulo XVIII dos Sintomas, sinais e achados anormais de exames clínicos e de laboratório, não classificados em outra parte inclui aqueles sintomas e afecções mal definidos, sem que o devido estudo tenha sido feito para se estabelecer um diagnóstico final. As causas de morte deste capítulo geralmente estão vinculadas a etiologias desconhecidas, não especificadas ou transitórias. A etiologia mal definida está associada aos problemas de diagnóstico por precariedade de assistência, de infraestrutura ou de acesso ao serviço de saúde (REDE, 2008; WHO, 2018).

O capítulo XX das causas externas de morbidade e mortalidade possibilita a classificação de ocorrências e circunstâncias ambientais como a causa de lesões, envenenamento e outros efeitos adversos. Os acidentes (de trânsito e quedas), lesões autoprovocadas intencionalmente (suicídios) e agressões (homicídios) fazem parte do grupo de doenças deste capítulo (WHO, 2018).

Os acidentes de trânsito são considerados causas passíveis de intervenção mediante as práticas comportamentais seguras e saudáveis. As pessoas mais susceptíveis a estes acidentes são portadoras de deficiência, consumidores de álcool e outras drogas. As ações de intervenção de natureza educativa geralmente visam a maior conscientização da população sobre o risco de acidente de trânsito (SOUZA, MINAYO e FRANCO, 2007). As mortes por causas externas em homens no Brasil são mais acentuadas do que nas mulheres, com proeminência nas idades jovens (BELTRÃO e DELLASOPA, 2011).

Metodologia

Para atender o objetivo proposto no final da parte introdutória deste trabalho, alguns passos precisam ser resolvidos: tabulação dos dados dos funcionários por grupos etários, escolaridade e carreira – nível médio e nível superior; construir tábua de vida para a população de funcionários do IBGE de 2000 a 2018; recuperar as causas de morte dos anos de 2000 a 2016; classificar as causas em grupos de evitáveis; calcular AVP por grupos de causas; e analisar os resultados, apontando as diferenças por carreira e escolaridade.

Base de dados

Este trabalho usa duas bases de dados: o Sistema Integrado de Administração de Recursos Humanos (SIAPE) que é um sistema de banco de dados com abrangência nacional, instituído pelo Decreto 99.328, de 19 de junho de 1990 para gerir a folha de pagamento e dar manutenção aos dados cadastrais dos servidores da administração pública civil federal (MPOG, 2017); e o Sistema de Mortalidade do MS (SIM), que coleta e codifica informações de óbitos dos estabelecimentos de saúde e cartórios, tem abrangência em nível nacional desde 1979 e o registro da causa de morte desde 1996 está alicerçado na Classificação Internacional de doenças, 10ª revisão - CID-10 (MS, 2011).

É notório o subregistro do SIM, porém este subregistro é diferenciado por nível sócio econômico, muito menor para as classes mais afluentes. Concomitante com a melhor cobertura vem uma melhor qualidade dos dados reportados (KANSO *et al.*, 2011).

O acesso aos dados dos funcionários do IBGE, no período de janeiro de 2000 até agosto de 2018, se deu através de ofício enviado ao departamento de recursos humanos da instituição, que ao tomar conhecimento da pesquisa prontamente concedeu o acesso.

Segundo (BELTRÃO & SUGAHARA, 2002), a vantagem em utilizar dados administrativos, como os do SIAPE, é que os numeradores e denominadores vêm da mesma fonte e os dados são coletados diretamente de documentos oficiais, evitando assim problemas de cobertura e erro de dígito preferencial. Somado a isso, como o uso principal desses registros é o pagamento dos funcionários, existe uma boa base para se crer na fidedignidade dos registros.

Recuperação da causa básica

A metodologia para recuperar as causas básicas de morte para os funcionários do IBGE adota um procedimento programado para concatenar os registros de óbitos do SIAPE aos do SIM. O procedimento usa um conjunto de variáveis para identificação dos registros comuns entre as duas bases: data de nascimento, data de óbito, sexo, unidade federativa de residência, município de residência e escolaridade. A solução tem por base o algoritmo usado por Oliveira (2012).

O procedimento informatizado resolve essas concatenações por etapas: primeiramente as concatenações feitas com todas as variáveis recuperam boa parte das causas; depois com os registros restantes do SIAPE não concatenados, novas buscas são feitas no SIM sem usar o dia da data de óbito; o dia e mês da data de óbito; o dia de nascimento; o dia e mês de nascimento; o município de residência; e a unidade federativa de residência. Os percentuais de causas recuperados por este procedimento atingem 92%.

As etapas subsequentes, feitas sem o procedimento informatizado, procuram tratar as idiosincrasias inerentes às duas bases de dados para ampliar os percentuais de causas recuperadas. A parcela de registros sem concatenação, se deu basicamente devido as duplicidades de registros encontrados no SIM. Parte dessas duplicatas (2%) é resolvida pela comparação da variável escolaridade. O ano limite selecionado para recuperar as causas, 2016, está relacionado aos microdados de causas de morte disponíveis para download no site do MS.

Classificação das mortes evitáveis

A classificação das causas básicas de morte recuperadas em grupos de evitáveis usa a Lista de causas de morte evitáveis (LCME) por intervenções do SUS do Brasil nas idades de 5 a 74 anos (MALTA *et al.*, 2007 e 2018). A idade limite usada neste trabalho, 79 anos, considera a expectativa de vida dos funcionários do IBGE aos 20 anos de idade, em torno de 81 anos, o que eleva em 5 anos a idade limite dessa lista.

A LCME possui 5 grupos de ações de intervenção para evitar a morte: 1.1 Reduzíveis por ações de imunoprevenção; 1.2 Reduzíveis por ações adequadas de promoção à saúde, prevenção, controle e atenção às doenças de causas infecciosas; 1.3 Reduzíveis por ações adequadas de promoção à saúde, prevenção, controle e atenção às doenças não transmissíveis; 1.4 Reduzíveis por ações adequadas de prevenção, controle e atenção às causas de morte materna; 1.5 Reduzíveis por ações intersetoriais adequadas de promoção à saúde, prevenção e atenção às causas externas (acidentais e violências). As não evitáveis estão em dois grupos: 2.0 das causas mal definidas; e 3.0 das não claramente evitáveis (MALTA *et al.*, 2007). MALTA *et al.* (2018), revisaram esta classificação e redistribuíram as causas mal definidas (2.0) proporcionalmente entre as evitáveis. Entre os funcionários públicos e a população paulista estas causas constituem uma proporção muito pequena e não foram distribuídas neste estudo.

Cálculo dos anos de vida perdidos

Uma das soluções possíveis para calcular AVP por causas é usar a técnica da TM para determinar quantos anos de vida uma população perde ou pode ganhar respectivamente na presença ou na ausência de certos fatores de risco. Segundo, Ortega (1987), para estimar o ganho é fundamental considerar que as pessoas salvas de morrer por uma causa determinada tenham a mesma probabilidade de morrer por outras causas que qualquer outra pessoa na população.

Para calcular as perdas ou ganhos é preciso elaborar a TMD. Sua construção diverge da TM comum somente pela probabilidade de morte líquida na idade, ou seja, aceita que uma determinada causa ou conjunto de causas tenham sido excluídos da população (BORGES, 2009). Segundo Chang (1968), o valor da probabilidade líquida pode ser estimado com base na causa j e idade i . Na equação (1) D_{ij} é o número de óbitos ocorridos na idade i pela causa ou grupo de causas j e D_i é o número total de óbitos ocorridos na idade i .

$$q_i = 1 - p_i \left[\frac{(D_i - D_{ij})}{D_i} \right] \quad (1)$$

A partir da probabilidade líquida (q_i sem a atuação do grupo de causas) é possível construir a TMD, adotando um procedimento semelhante ao que é feito para a TM comum. A expectativa de vida líquida (e_{ij}) é extraída da TMD e corresponde aos anos que os indivíduos esperam viver se determinado grupo de causas for eliminado. A diferença entre as esperanças de vida da TMD e da TM resulta nos AVP por grupo de causas.

Tábua de mortalidade

Este trabalho adota o modelo de coorte sintética para construir a TM e a TMD. A técnica adotada para construção está baseada na metodologia descrita por Beltrão e Sugaraha (2017). A etapa inicial de cálculo da probabilidade de morte consiste na graduação da probabilidade bruta e para isso, o modelo proposto por Heligman e Pollard (1980) é fundamental. O modelo (2) possui 3 componentes. A primeira refere-se às primeiras idades e não é usada para funcionários públicos, a segunda referente às causas externas, e a terceira refere-se às idades adultas e idosas. D, E, F, G, H e K são os parâmetros que devem ser estimados iterativamente através de uma técnica de regressão não linear.

$$q_x = A^{(x+B)^C} + D e^{E(\ln x - \ln F)^2} + \frac{GH^x}{1+KGH} \quad (2)$$

A probabilidade bruta de morte geralmente pode apresentar valores com mudanças bruscas em idades contíguas. Essa situação está associada ao tamanho reduzido da população para certas idades. A suavização desses valores deve usar uma técnica de graduação que procure o melhor ajuste possível entre a probabilidade predita e a bruta (Beltrão e Sugaraha, 2017). Uma boa prática é visualizar as probabilidades brutas num gráfico, antes de selecionar o modelo polinomial para aplicar a graduação. A equação a seguir exemplifica o cálculo da probabilidade bruta como função dos óbitos (Obt_x) e da população exposta (N_x).

$$Q_x = \frac{Obt_x}{N_x + 0.5 * Obt_x} \quad (3)$$

Conforme mencionado, o processo de ajuste da probabilidade bruta (Q_x) é iterativo. Usa-se um procedimento de regressão não linear de um pacote estatístico para o qual a escolha direta da distribuição das mortes na população não está disponível. Para isso, utiliza o conceito de desvio padrão da distribuição binomial para aplicar um peso que auxilie nos ajustes da curva. Este peso é inversamente proporcional à variância da distribuição binomial $B(N_x, q_x)$, onde N_x é a população em determinado grupo de idade x e q_x é a probabilidade de morte que deve ser estimada.

O processo de ajuste usa primeiramente a população na idade como peso para estimar a curva. A partir da probabilidade de morte obtida no primeiro passo, o peso passa a ser calculado através do inverso da variância da distribuição binomial, usando para isso a probabilidade de morte do passo anterior e a população. O processo segue até que seja encontrada uma convergência, quando as diferenças de sucessivas estimativas forem muito próximas

a zero (10⁻¹⁰). Este processo de convergência equivale a calcular o estimador de máxima verossimilhança para uma distribuição binomial, procedimento este que é o utilizado pelos pacotes estatísticos cujos estimadores não apresentam uma forma fechada.

Resultados

População exposta

Os dados da Tabela 1 seguem a proposta de estudo para analisar a população de funcionários por carreira e nível educacional: funcionários do nível superior de escolaridade – S; do nível médio - M, que compreende as escolaridades até o nível superior incompleto; funcionários da carreira do nível superior – CS; e da carreira do nível médio – CM. A população exposta no período de 2000-2018 é de 240.332 e o número de óbitos registrados é de 2.652.

Tabela 1 – Distribuição de funcionários do IBGE e óbitos dos mesmos, por carreira, escolaridade e sexo entre 2000 e 2018, com indicação da população de São Paulo/SP, segundo faixas etárias em 2010

Faixas etárias	Distribuição de funcionários do IBGE							Faixas etárias	População São Paulo/SP
	Total	Percentual (%)							
		Carreira		Escolaridade		Sexo			
		Nível superior	Nível médio	Nível superior	Nível médio	Homem	Mulher		
18-29 anos	7 934	24,3	75,7	72,9	27,1	72,4	27,6	20-29 anos	7 427 477
30-39 anos	14 368	48,9	51,1	83,2	16,8	66,4	33,6	30-39 anos	6 740 568
40-49 anos	42 033	21,9	78,1	56,8	43,2	57,4	42,6	40-49 anos	5 733 797
50-59 anos	81 880	19,2	80,8	51,4	48,6	57,7	42,3	50-59 anos	4 424 095
60-69 anos	60 509	22,9	77,1	49,3	50,7	58,9	41,1	60-69 anos	2 646 095
70-79 anos	23 757	27,0	73,0	45,3	54,7	62,2	37,8	70-79 anos	1 457 533
80-89 anos	8 403	34,9	65,1	47,2	52,8	62,6	37,4	80-89 anos	581 153
90-99 anos	1 409	43,0	57,0	49,5	50,5	57,1	42,9	90 anos e mais	86 655
100-03 anos	39	48,7	51,3	48,7	51,3	51,3	48,7		

Faixas etárias	Distribuição de óbitos de funcionários do IBGE							Faixas etárias	Óbitos São Paulo/SP
	Total	Percentual (%)							
		Carreira		Escolaridade		Sexo			
		Nível superior	Nível médio	Nível superior	Nível médio	Homem	Mulher		
18-29 anos	5	20,0	80,0	20,0	80,0	80,0	20,0	20-29 anos	7 488
30-39 anos	15	40,0	60,0	66,7	33,3	53,3	46,7	30-39 anos	10 617
40-49 anos	106	8,5	91,5	47,2	52,8	70,8	29,2	40-49 anos	19 609
50-59 anos	426	14,1	85,9	42,0	58,0	75,1	24,9	50-59 anos	33 163
60-69 anos	684	17,5	82,5	39,9	60,1	74,3	25,7	60-69 anos	41 866
70-79 anos	627	21,5	78,5	39,1	60,9	76,4	23,6	70-79 anos	57 264
80-89 anos	581	34,4	65,6	45,6	54,4	72,1	27,9	80-89 anos	50 555
90-99 anos	197	36,5	63,5	43,7	56,3	62,9	37,1	90 anos e mais	16 577
100-03 anos	11	54,5	45,5	54,5	45,5	72,7	27,3		

Fonte: SIAPE (2018) e Censo Demográfico 2010.

A população paulista de 2010 na mesma tabela é usada neste trabalho como referência para as comparações. São Paulo em 2010 tem o segundo melhor índice de desenvolvimento humano do país e a maior receita orçamentária (IBGE, 2018). O ano de 2010 está próximo do ponto médio do período selecionado para este estudo, 2000-2018.

Para a população de funcionários como um todo, a moda está no grupo etário de 50 a 59 anos. Nos grupos etários extremos a população está em menor número. Entre as mortes, a classe modal está no grupo etário de 60 a 69 anos. As ocorrências de óbitos nas idades idosas são em maior número do que nas idades jovens. No caso dos paulistas, a moda está no grupo etário de 20 a 29 anos e a classe modal das mortes, entre 70 e 79 anos.

Entre os funcionários, o quantitativo de homens supera o de mulheres em todos os grupos etários, inclusive nas mortes. Mais de 70% da população exposta nos grupos etários de 18 a 29 e 30 a 39 tem escolaridade de nível superior. Nos grupos etários seguintes, os valores são maiores para o pessoal da escolaridade de nível médio. Os valores da população exposta por carreira e grupos etários são maiores para o nível médio do que para o superior. Essas comparações apontam que nas idades jovens há uma proeminência de funcionários da escolaridade de nível superior na carreira de nível médio.

Os óbitos dos funcionários do nível superior de escolaridade superam os do nível médio somente no grupo etário de 30 a 39 anos e entre a carreira de nível superior e a de nível médio a situação é inversa (TABELA 1). Esta comparação aponta que há um percentual significativo de óbitos de funcionário da escolaridade de nível superior na carreira de nível médio.

Causas evitáveis

A Tabela 2 contém as causas de morte dos funcionários do IBGE recuperadas e classificadas em grupos de causas (2.212 causas): evitáveis (11 a 15); mal definidas (20); e não claramente evitáveis (30). O período selecionado inclui os anos 2000 a 2016 (último ano com dados disponíveis no site do MS). O percentual de causas não recuperadas (6%) está redistribuído entre as causas identificadas. As causas de morte da população paulista de 2010 estão agrupadas seguindo este mesmo critério.

A distribuição proporcional dos grupos de causas de morte em cada grupo etário corresponde a um ambiente em que a população esteja exposta à múltiplos riscos competitivos de morte. Os valores percentuais servem de base para a estimação da probabilidade de morte líquida usada na construção da TMD, conforme já explicado na seção metodológica.

Entre os funcionários, as proporções de óbitos por causas mal definidas (2.0) têm respectivamente valores menores entre os funcionários com escolaridade de nível superior do que entre os de nível médio. Essas causas podem ser usadas como sentinela para notificar ocorrências de mortes devido à dificuldade de acesso aos serviços de saúde médico-hospitalar de qualidade (RUTSTEIN *et al.*, 1976). Na população de São Paulo, esses casos não ultrapassam os 7% de óbitos nos grupos etários.

Entre os funcionários, as causas externas evitáveis estão mais proeminentes no grupo etário jovem de 30 a 39 anos do pessoal da carreira de nível médio sem o nível superior. Entre os paulistas, as causas externas apresentam proporções proeminentes nos grupos etários jovens (decenais de 20 a 49 anos).

O grupo das mortes evitáveis por doenças não transmissíveis (grupo 1.3), entre os funcionários como um todo, tem os valores mais altos nos grupos etários decenais de 40 a 79 anos. O nível de importâncias dessas causas nos grupos etários de 50 a 59 e de 60 a 69 anos está mais alto para as pessoas com escolaridade de nível superior do que para os de nível médio. As causas não transmissíveis entre os paulistas assumem a maior importância nos grupos etários decenais de 50 a 79 anos. Estas mortes evitáveis estão ocorrendo nos grupos etários de maior número absoluto de funcionários.

As mortes por doenças infecciosas evitáveis (grupo 1.2), entre os funcionários, têm valores mais proeminentes nos grupos etários decenais de 40 a 59 anos e uma situação próxima

é observada para a população paulista. As causas evitáveis por imunoprevenção (1.1) e morte materna (1.4) possuem valores muito baixos nas populações observadas.

Tabela 2 - Distribuição percentual das mortes evitáveis e não claramente evitáveis dos funcionários do IBGE, por escolaridade, carreira e grupos da LCME, segundo as faixas etárias entre 2000 e 2016, em comparação com os dados de São Paulo/SP em 2010

Faixas etárias	Distribuição percentual das mortes evitáveis e não claramente evitáveis dos funcionários do IBGE, por grupos da LCME (%)													
	Funcionários do IBGE							São Paulo/SP						
	1.1	1.2	1.3	1.4	1.5	2.0	3.0	1.1	1.2	1.3	1.4	1.5	2.0	3.0
20-29 anos	-	-	-	-	-	-	-	0,1	5,4	6,9	0,7	67,2	4,7	15,0
30-39 anos	-	-	-	-	50,0	25,0	25,0	0,2	7,8	10,6	0,8	61,3	4,6	14,7
40-49 anos	-	16,7	33,3	-	25,0	-	25,0	0,1	14,6	22,6	0,5	36,2	6,9	19,0
50-59 anos	-	17,0	43,0	-	12,0	2,0	26,0	0,1	13,6	40,8	0,1	16,3	7,0	22,2
60-69 anos	0,3	9,1	51,2	-	8,8	4,3	26,4	0,1	9,6	52,3	-	7,7	6,1	24,3
70-79 anos	0,2	7,6	54,8	-	3,2	3,4	30,9	0,0	8,9	55,7	-	3,9	5,6	25,9

Faixas etárias	Distribuição percentual das mortes evitáveis e não claramente evitáveis dos funcionários do IBGE, por grupos da LCME (%)													
	Escolaridade													
	Nível superior							Nível médio						
	1.1	1.2	1.3	1.4	1.5	2.0	3.0	1.1	1.2	1.3	1.4	1.5	2.0	3.0
20-29 anos	-	-	-	-	-	-	-	-	-	-	-	-	-	-
30-39 anos	-	-	-	-	-	-	100,0	-	-	-	-	66,7	33,3	-
40-49 anos	-	28,6	28,6	-	28,6	-	14,3	-	-	40,0	-	20,0	-	40,0
50-59 anos	-	12,5	47,9	-	10,4	-	29,2	-	21,2	38,5	-	13,5	3,8	23,1
60-69 anos	-	11,3	53,8	-	6,9	3,1	25,0	0,5	7,4	49,3	-	10,2	5,1	27,4
70-79 anos	-	6,6	55,8	-	3,1	1,3	33,2	0,3	8,3	54,1	-	3,3	4,7	29,3

Faixas etárias	Distribuição percentual das mortes evitáveis e não claramente evitáveis dos funcionários do IBGE, por grupos da LCME (%)													
	Carreira													
	Nível superior							Nível médio						
	1.1	1.2	1.3	1.4	1.5	2.0	3.0	1.1	1.2	1.3	1.4	1.5	2.0	3.0
20-29 anos	-	-	-	-	-	-	-	-	-	-	-	-	-	-
30-39 anos	-	-	-	-	-	-	100,0	-	-	-	-	66,7	33,3	-
40-49 anos	-	25,0	50,0	-	25,0	-	-	-	12,5	25,0	-	25,0	-	37,5
50-59 anos	-	-	57,1	-	14,3	-	28,6	-	18,3	41,9	-	11,8	2,2	25,8
60-69 anos	-	7,1	57,1	-	7,1	1,8	26,8	0,3	9,4	50,2	-	9,1	4,7	26,3
70-79 anos	-	7,6	54,3	-	2,9	1,9	33,3	0,2	7,6	54,9	-	3,3	3,7	30,3

Fonte: SIAPE (2018), Censo Demográfico 2010 e Ministério da Saúde (2018).

Notas: 1. Os grupos de mortes evitáveis vão de 1.1 a 2.0.

2. O grupo 3.0 é de mortes não claramente evitáveis.

3. LCME = Lista de causas de mortes evitáveis

Probabilidade de morte

Na Figura 1, para o período de 2000-2018, podem ser visualizadas as comparações dos resultados relativos aos ajustes do modelo escolhido: dados brutos (pontos); e ajustados para carreira de nível superior (linha fina); carreira de nível médio (linha grossa); escolaridade do nível superior (linha mais fina segmentada); e escolaridade do nível médio (linha mais grossa segmentada). Os parâmetros ajustados no polinômio seguem as explicações da seção metodológica. Tipicamente o nível de mortalidade aumenta com a idade.

Para as idades abaixo de 40 anos, os dados brutos parecem mais afastados dos valores ajustados e estão rarefeitos, indicando uma variância maior. Para as demais idades, a aderência dos dados ao modelo parece melhor.

O parâmetro K negativo, para o nível médio de escolaridade e a carreira de nível médio, aponta uma mortalidade com aceleração no crescimento nas altas idades.

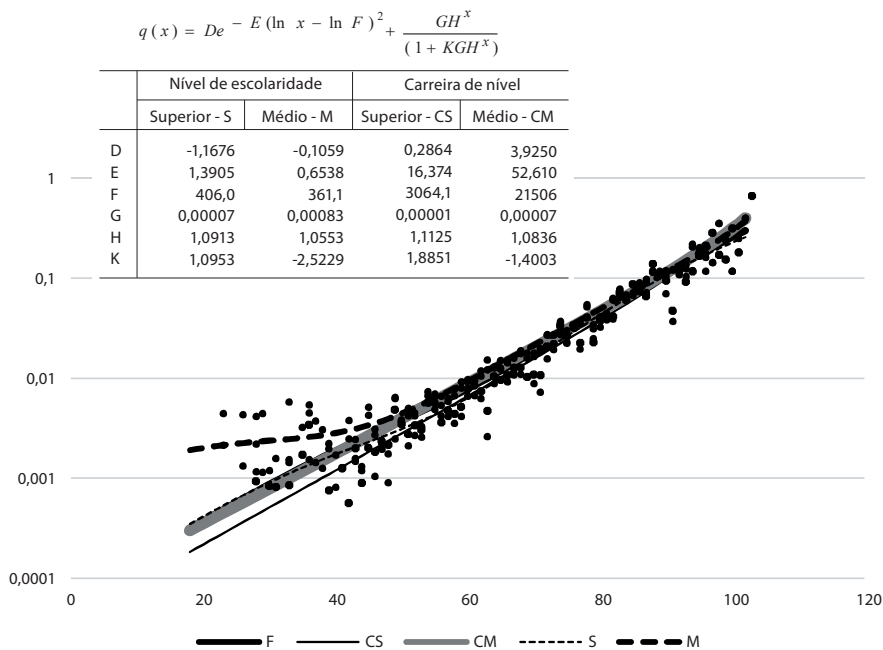
Para o pessoal com nível superior o valor de K positivo indica uma mortalidade com desaceleração no crescimento das altas idades. O parâmetro D maior para o nível médio aponta a mortalidade nas causas externas. Estas causas afetam a população jovem e geralmente os homens (BELTRÃO & SUGARA, 2017).

A probabilidade de morte nas idades abaixo de 45 anos está maior para o pessoal da escolaridade de nível médio, depois vem o segundo grupo - pessoas da escolaridade de nível superior e carreira de nível médio - e em terceiro a carreira de nível superior.

Próximo a 50 anos, o pessoal da escolaridade de nível superior fica na segunda posição acima da carreira de nível superior. A sobremortalidade entre a escolaridade de nível médio e a carreira de nível médio fica próximo a 1 e parece se manter assim até a idade limite adotada, 103 anos. Próximo a 60 anos, uma situação semelhante de sobremortalidade ocorre entre a escolaridade de nível superior e a carreira de nível superior.

As situações observadas apontam que a carreira e a escolaridade são variáveis que afetam positivamente as condições de vida. Abaixo de 50 anos, a carreira combinada à escolaridade exerce uma atuação mais importante nas condições de sobrevivência do que só a escolaridade.

Figura 1 - Probabilidades de morte dos funcionários do IBGE por nível de escolaridade e por carreiras de nível superior ou médio - 2000-2018



Fonte: SLAPE (2018).

AVP por causas

A situação das principais causas de morte dos funcionários do IBGE está apresentada na Tabela 3. As causas de morte do período de 2000-2016 por grupos etários são utilizadas como *proxy* para estimar os AVP no período de 2000-2018 nos grupos etários decenais de 20 a 79 anos. A mortalidade proporcional inferior a 1% para São Paulo em 2010 define o fator de corte para os grupos de causas. Os AVP dos funcionários por capítulos são obtidos através da metodologia já descrita. Para os AVP de São Paulo, buscou-se a tábua de vida de 2010 publicada pelo IBGE (2018).

Tabela 3 - Distribuição percentual das principais causas mortis e anos de vida perdidos no grupo etário de 20 a 79 anos, dos funcionários do IBGE entre 2000 e 2018 e da população de São Paulo/SP em 2010, segundo capítulos do CID-10 e grandes grupos de causas mortis

Capítu- los CID-10	Grandes grupos de causas mortis CID-10	Distribuição percentual das principais causas mortis (%)		Anos de vida perdidos no grupo etário de 20 a 79 anos				
		Funcio- nários do IBGE	São Paulo/ SP	Funcionários do IBGE				São Paulo/ SP
				Carreira		Escolaridade		
				Nível superior	Nível médio	Nível superior	Nível médio	
Total		100,0	100,0	6,9	8,5	7,7	10,1	9,4
I	Outras doenças bacterianas	1,5	1,0					
	Doença pelo vírus da imunodeficiência humana [HIV]	0,5	1,2	0,3	0,5	0,5	0,5	0,4
II	Neoplasias malignas	25,5	17,6	2,7	2,1	2,3	2,1	2,0
IV	Diabetes mellitus	5,0	3,9	0,3	0,5	0,3	0,6	0,4
VI	Outras doenças degenerativas do sistema nervoso	2,2	1,4	0,1	0,1	0,1	0,2	0,2
IX	Doenças hipertensivas	1,8	3,2					
	Doenças isquêmicas do coração	12,1	10,7					
	Outras formas de doença do coração	4,1	5,8	1,7	2,3	2,0	2,3	2,9
	Doenças cerebrovasculares	7,4	8,5					
	Doenças das artérias, arteríolas e capilares	2,0	1,6					
X	Influenza e pneumonia (microorganismo)	5,6	7,0	0,5	0,7	0,5	0,8	0,9
	Doenças crônicas das vias aéreas inferiores	3,2	3,8					
XI	Doenças do fígado	3,3	2,7	0,6	0,7	0,7	0,9	0,6
XIV	Insuficiência renal	1,1	1,0	0,1	0,2	0,1	0,2	0,2
	Outras doenças do aparelho urinário	1,0	1,1					
XVIII	Causas mal definidas e desconhecidas	3,1	5,0	0,1	0,4	0,2	0,7	0,5
XX	Acidentes	2,8	4,8	0,4	0,7	0,5	1,4	1,1
	Agressões	0,5	2,2					
Outros	Outras causas	17,4	17,3	0,1	0,2	0,4	0,4	0,2

Fonte: SIAPE (2018), Tábua de Vida 2010 e Ministério da Saúde (2018).

Nota: CID-10 = Classificação Internacional de Doenças e Problemas Relacionados à Saúde.

O grupo das neoplasias (cânceres) no IBGE e em São Paulo possuem os valores mais altos, em 2º lugar estão as doenças isquêmicas do coração e em 3º, as doenças cerebrovasculares. Os grupos de causas selecionados ultrapassam levemente os 80% do total de mortes.

Os valores de AVP por capítulos nas populações observadas estão maiores nas neoplasias (II) e nas doenças do aparelho circulatório (IX). Os AVP por neoplasias estão mais elevados nos funcionários com escolaridade de nível superior e os valores dos funcionários do IBGE superam os AVP dos paulistas. Entre as doenças do aparelho circulatório (IX), esta situação está invertida.

Os valores de AVP para o pessoal de nível médio, seja por escolaridade ou carreira, superam os dos paulistas nos seguintes capítulos: algumas doenças infecciosas (I); neoplasias (II); doenças endócrinas e nutricionais (IV); doenças do aparelho digestivo (XI); doenças do aparelho geniturinário (XIV); Sintomas, sinais e achados anormais de exames clínicos e de laboratório, não classificados em outra parte (XVIII); causas externas XX. A mesma comparação feita para o pessoal de nível superior aponta os seguintes capítulos: capítulo I; capítulo II; capítulo IV; capítulo XI. Estas causas observadas apontam que o pessoal de nível médio na escolaridade e na carreira estão expostos aos mesmos riscos que seus pares na escolaridade e carreira, mas o pessoal do nível superior tem uma lista menor de causas preocupantes. Esta situação ajuda a reforçar o fato de que a expectativa de vida da carreira de nível médio (61 anos aos 20 anos de idade) seja menor do que a do nível superior (63 anos aos 20 anos de idade).

AVP por causas evitáveis

A Figura 2 apresenta os AVP por causas evitáveis e as causas mal definidas. As causas de morte por doenças não transmissíveis (1.3) são as mais importantes e incluem os capítulos II (e.g. neoplasias de traqueia, brônquios, pulmões, mama, cólon, anus, reto, fígado e estômago), IV (e.g. diabetes), IX (e.g. doenças do coração, cerebrovasculares, hipertensivas), X (e.g. doenças pulmonares) e XI (e.g. doença alcoólica do fígado).

Na segunda posição (1.5), as causas externas do capítulo XX incluem os acidentes de transporte e agressões. As causas infecciosas (1.2) incluem os capítulos I (e.g. HIV), X (e.g. pneumonia e influenza) e VI (e.g. estado de mal epilético) (MALTA *et al.*, 2007; MALTA *et al.*, 2018).

Para os funcionários, a magnitude dos anos perdidos por causas de morte evitáveis está menor para o pessoal de nível superior (carreira 4,6 anos e escolaridade 4,9) do que para o pessoal de nível médio (carreira de nível médio 5,8 anos e escolaridade de nível médio 6,2 anos). Para os paulistas, essas perdas são de 6,6 anos.

Os valores de AVP por causas evitáveis estão menores do que os valores vistos na Tabela 3 por capítulos. Essa diferença alerta para a necessidade de avanços na área da tecnologia médica do Brasil e na qualidade do atendimento médico-hospitalar (MALTA *et al.*, 2018).

Os AVP relativa aos casos de doenças mal definidas estão associados à dificuldade de acesso ao atendimento médico-hospitalar de qualidade. A resolução deste problema nas áreas que essas pessoas residem e trabalham é essencial.

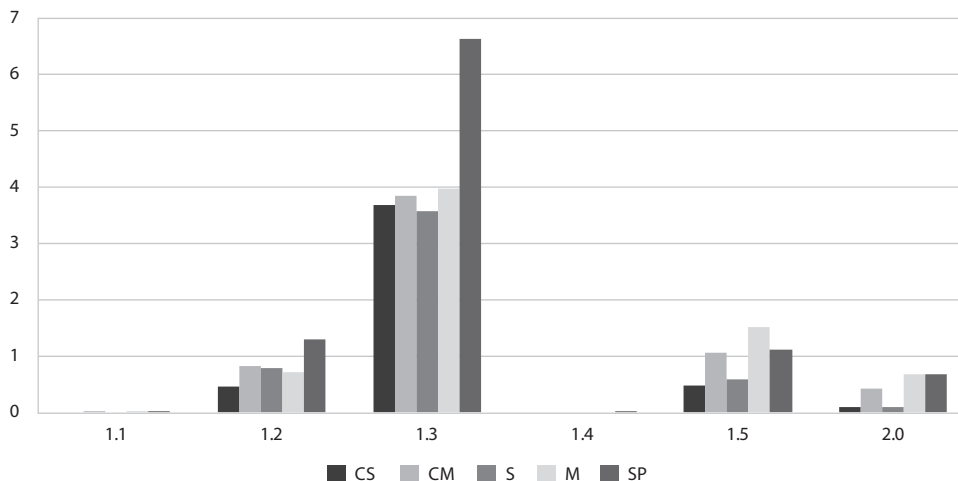
Os AVP por doenças não transmissíveis, entre os funcionários da carreira de nível superior, constituem 78% dos AVP evitáveis, superando a proporção entre os paulistas, 73%. Entre os de escolaridade de nível superior esta proporção é de 72%, entre os de carreira de nível médio é de 67% e entre os de escolaridade de nível médio, 64%.

Os AVP por causas externas evitáveis constituem 1,5 ano para os funcionários da escolaridade de nível médio, representando 24%, mais altos do que para os paulistas, com 1,1 ano, representando 12%. Para os funcionários da escolaridade de nível superior e da carreira de nível superior os valores são menores, respectivamente 0,57 e 0,55 para os AVP, 12% e 11% para as proporções.

A redução dessas causas externas demanda ações conjuntas entre vários setores, como saúde, educação, judiciário, serviço de trânsito e serviço social (MALTA *et al.*, 2018).

O valor de AVP por doenças infecciosas evitáveis é o mais proeminente entre os paulistas, com 1,3 ano, e menor entre os funcionários da carreira de nível superior, com 0,5 anos. Os outros recortes populacionais de funcionários apresentam valores intermediários.

Figura 2 - AVP por grupo de causas de morte evitáveis por escolaridade (S – superior; M – médio) e por carreira (C) para o IBGE de 2000/2018 e São Paulo em 2010



Fonte: SIAPE (2018) e MS (2018).

Considerações finais

O estudo pode verificar que a probabilidade de morte dos funcionários do IBGE da carreira de nível médio é mais alta do que para a carreira do nível superior para todas as idades entre 20 e 80 anos. No caso dos funcionários públicos federais, uma situação parecida é encontrada em todas as idades por um período de 10 anos (BELTRÃO & SUGAHARA, 2017).

Nas idades jovens as causas externas são as mais proeminentes para o pessoal da escolaridade de nível médio e menores para a carreira de nível superior, conforme verificado em outras pesquisas (BELTRÃO & SUGAHARA, 2017; BELTRÃO e DELLASOPA, 2011).

Para idades mais avançadas, as doenças não transmissíveis são as mais frequentes. Os AVP por neoplasias são maiores para o nível superior do que para o nível médio. No caso das doenças do aparelho circulatório, esta situação está invertida. Diversos estudos nacionais e internacionais têm relatado a alta incidência de mortes por doenças não transmissíveis nas idades adultas e idosas (LEITE, 2015; FRANÇA *et al.*, 2017; KANSO, 2013; VALENT & LORIS ZAINER, 2015).

A eliminação dos grupos de causas evitáveis relacionadas às ações de intervenções pelo SUS e ações conjunta de outros órgãos aponta perdas de 4,6 anos para a carreira do nível superior e 5,8 anos para a carreira de nível médio.

A situação verificada demanda a implantação de políticas para diminuir o risco de morte das doenças e agravos evitáveis apresentados. O hiato de mortalidade associadas as condições socioeconômicas, escolaridade e carreira, deve ser eliminado. Este perfil de mortalidade aponta a falta do diagnóstico tempestivo de doenças e dos hábitos protetores à saúde.

A recuperação das causas básicas de morte enriqueceu os resultados. Como sugestão para trabalhos futuros, é interessante fazer uma análise comparativa da mortalidade por causas dos funcionários do IBGE com a morbidade dos funcionários ativos.

Referências

- ARAGÓN, T. J. *et al.* Calculating expected years of life lost for assessing local ethnic disparities in causes of premature death, *BMC public health*, v. 8, n. 1, p. 116, 2008.
- ARTEAGA, T. C. S. *Tábuas Seletas de mortalidade: comportamento da mortalidade dos aposentados por invalidez do regime próprio de previdência social.* 2009. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2009.
- BARBOSA, R. B.; BARCELÓ, A.; MACHADO, C. A. Campanha nacional de detecção de casos suspeitos de diabetes mellitus no Brasil: relatório preliminar, *Revista Panamericana de Salud Pública*, v. 10, p. 324–327, 2001.
- BELTRÃO, K. I.; DELLASOPPA, E. E. El designio de los hombres. Años de vida perdidos en Brasil y en sus grandes regiones, 1980 a 2005, *Estudios demográficos y urbanos*, v. 26, n. 2, p. 299–343, 2011.
- BELTRÃO, K. I.; SUGAHARA, S. Mortalidade dos funcionários públicos civis do Executivo por sexo e escolaridade - 1993/2014, *Revista Contabilidade & Finanças*, v. 28, n. 75, p. 445–464, 2017.
- BELTRÃO, K. I.; SUGAHARA, S. *Tábua de mortalidade para os funcionários públicos civis federais do poder executivo por sexo e escolaridade: comparação com tábuas do mercado.* Rio de Janeiro: IBGE, 2002.
- BORGES, G. M.; BELTRÃO, K. I. Tábua de múltiplos decrementos para a saída da atividade no funcionalismo público federal, *19 SINAPE-Simpósio Nacional de Probabilidade e Estatística*, 2010.
- CHANG, C. L. *Introduction to stochastic processes in biostatistics.* New York: John Wiley And Sons, Inc, 1968.
- EDNEY, L. C. *et al.* Mortality reductions from marginal increases in public spending on health, *Health Policy (Amsterdam, Netherlands)*, v. 122, n. 8, p. 892–899, 2018.
- FRANÇA, E. B. *et al.* Cause-specific mortality for 249 causes in Brazil and states during 1990-2015: a systematic analysis for the global burden of disease study 2015, *Population Health Metrics*, v. 15, n. 1, p. 39, 2017.
- FRINGS, M. *et al.*, Modeling and mapping the burden of disease in Kenya, *Scientific Reports*, v. 8, n. 1, p. 9826, 2018.
- HEILBERG, I. P.; SCHOR, N., Abordagem diagnóstica e terapêutica na infecção do trato urinário: ITU, *Revista da Associação Médica Brasileira*, v. 49, n. 1, p. 109–116, 2003.
- HELIGMAN, L.; POLLARD, J. H., The age pattern of mortality, *Journal of the Institute of Actuaries*, v. 107, n. 1, p. 49–80, 1980.
- IBGE, INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA, **Downloads | IBGE: Instituto Brasileiro de Geografia e Estatística.** Disponível em: <https://www.ibge.gov.br/estatisticas-novoportal/downloads-estatisticas.html>. Acesso em: 23 set. 2018.
- IBGE, INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA, **IBGE | Brasil em Síntese | São Paulo | Panorama.** Disponível em: <https://cidades.ibge.gov.br/brasil/sp/panorama>. Acesso em: 20 set. 2018.
- KANSO *et al.* Diferenciais geográficos, socioeconômicos e demográficos da qualidade da informação da causa básica de morte dos idosos no Brasil. *Cadernos de Saúde Pública*, Rio de Janeiro, 27(7):1323-1339, jul, 2011
- KANSO, S. *et al.* A inevitabilidade de óbitos entre idosos em São Paulo, Brasil: análise das principais causas de morte, *Cadernos de Saúde Pública*, v. 29, n. 4, p. 735–748, 2013.

- LEITE, I. da C. *et al.*, Carga de doença no Brasil e suas regiões, 2008, *Cadernos de Saúde Pública*, v. 31, n. 7, p. 1551–1564, 2015.
- LOZANO, R. *et al.*, Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010, *The Lancet*, v. 380, n. 9859, p. 2095–2128, 2012.
- MAKEHAM, W. M., On an application of the Theory of the Composition of Decremental Forces, *Journal of the Institute of Actuaries*, v. 18, n. 5, p. 317–322, 1874.
- MALTA, D. C. *et al.*, Mortes evitáveis no Sistema Único de Saúde na população brasileira, entre 5 e 69 anos, 2000 - 2013, *Revista Brasileira de Epidemiologia*, v. 21, 2018.
- MALTA, D. C.; DUARTE, E. C., Causes of avoidable mortality through effective healthcare services: a review of the literature, *Ciência & Saúde Coletiva*, v. 12, n. 3, p. 765–776, 2007.
- MESALLES-NARANJO, O. *et al.* Trends and inequalities in the burden of mortality in Scotland 2000–2015, *PloS One*, v. 13, n. 8, p. e0196906, 2018.
- MPOG, MINISTÉRIO DO PLANEJAMENTO, DESENVOLVIMENTO E GESTÃO, **Boletim Estatístico de Pessoal e Informações Organizacionais - Ministério do Planejamento, Desenvolvimento e Gestão**, Brasília: MINISTÉRIO DO PLANEJAMENTO, DESENVOLVIMENTO E GESTÃO, 2017.
- MS, MINISTÉRIO DA SAÚDE. Secretaria de Vigilância em Saúde. Departamento de Análise de Situação de Saúde, **Manual de Instruções para o preenchimento da Declaração de Óbito**, Brasília – DF: Ministério da Saúde, 2011.
- MS, MINISTÉRIO DA SAÚDE. Secretaria de Vigilância em Saúde. Departamento de Análise de Situação de Saúde, **DATASUS, Departamento de Informática do SUS**, Portal da Saúde. Disponível em: <http://datasus.saude.gov.br>. Acesso em: set. 2018.
- OLIVEIRA, A. B. *Tábuas de mortalidade por grande grupo de causas para os servidores do IBGE*. 2012. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2012.
- ORTEGA, A. Tablas de mortalidad, *Tablas de mortalidad*, p. 295, 1987.
- REDE INTERAGENCIAL DE INFORMAÇÕES PARA A SAÚDE (REDE). *Indicadores básicos para a saúde no Brasil: conceitos e aplicações*, 2. ed. Brasília: Organização Pan-Americana da Saúde, Escritório Regional para as Américas da Organização Mundial da Saúde, 2008.
- RUTSTEIN, D. D. *et al.* Measuring the quality of medical care. A clinical method, *The New England Journal of Medicine*, v. 294, n. 11, p. 582–588, 1976.
- SILVA, L. G. C. *Estudo da Mortalidade dos Servidores Públicos Civil do Estado de São Paulo: Tábua de Mortalidade Destinada aos Regimes Próprios de Previdência Social*. Disponível em: http://www.ence.ibge.gov.br/images/ence/doc/mestrado/dissertacoes/2008/dissertacao2008-luciano_goncalves.pdf. Acesso em: set. 2018.
- SILVA, M.; SANTANA, V. S. Occupation and mortality in the Brazilian Navy. *Revista de Saúde Pública*, v. 38, n. 5, p. 709–715, 2004.
- SOUZA, E. R.; MINAYO, M. C. S.; FRANCO, L. G. Avaliação do processo de implantação e implementação do Programa de Redução da Morbimortalidade por Acidentes de Trânsito, *Epidemiologia e Serviços de Saúde*, v. 16, n. 1, p. 19–31, 2007.
- VALENT, F.; ZANIER, L. A population-based study of the years of life lost in the Friuli Venezia Giulia region, Italy, *Annali dell'Istituto Superiore Di Sanita*, v. 51, n. 4, p. 291–297, 2015.

WHO (World Health Organization). *International Statistical Classification of Diseases and Related Health Problems 10th Revision, ICD-10 Version: 2016*. Disponível em: <http://apps.who.int/classifications/icd10/browse/2010/en>. Acesso em: set. 2018.

WHO (World Health Organization). *Prevention and Control of Chronic Respiratory Diseases at Country Level*. Geneva: World Health Organization, 2004.

WHO (World Health Organization). *Prevention of cardiovascular disease: guidelines for assessment and management of cardiovascular risk*. Geneva: World Health Organization, 2007.

Determinantes do tempo de aleitamento materno exclusivo no Brasil em 2006*

Letícia Batista de Paula Barros**

Angelita Alves de Carvalho***

Resumo

Este estudo investigou os principais determinantes da interrupção precoce do aleitamento materno exclusivo (AME) a partir dos dados levantados pela Pesquisa Nacional de Demografia e Saúde realizada no ano de 2006. Os resultados foram obtidos a partir da aplicação do método de análise de sobrevivência, em que as probabilidades do tempo de aleitamento materno foram analisadas a partir de curvas de Kaplan-Meier referentes às características maternas, à fatores regionais e socioeconômicos e demográficos, à prestação dos serviços de saúde e características relativas à criança. Posteriormente, foram ajustados modelos de Cox, os quais geraram estimativas dos riscos de interrupção do AME associado à estas variáveis. Os resultados indicam que o aleitamento materno exclusivo no Brasil apresenta números aquém do recomendado pelos principais organismos de saúde, o percentual de crianças menores de seis meses sendo amamentadas exclusivamente foi de apenas 34,3%. Dentre os determinantes encontrados para a interrupção dessa prática foi observado que características maternas, como idade, características sociodemográficas, como a situação do domicílio e estrato socioeconômico, bem como características associadas à criança, como uso da chupeta e número de consultas 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: Nutrição infantil. Interrupção precoce do AME. Aleitamento Materno Exclusivo. Pesquisa Nacional de Demografia e Saúde.

* Este capítulo é resultado da dissertação de mestrado concluída em 2017 por Letícia Baptista de Paula Barros, intitulada "Determinantes do tempo de aleitamento materno exclusivo – Brasil (1986-2006)", sob orientação de Angelita Alves de Carvalho e Gustavo da Silva Ferreira.

** Pesquisadora no Instituto Nacional de Saúde da Mulher, da Criança e do Adolescente Fernandes Figueira (IFF-FIOCRUZ). Concluiu o mestrado em População, Território e Estatísticas Públicas em 2017.

*** Doutora em Demografia e Docente no programa de pós-graduação da Escola Nacional de Ciências Estatísticas desde 2014.

Introdução

À prática natural do aleitamento materno atribuem-se inúmeros benefícios registrados na literatura, uma vez que se constitui em uma medida importante para estabelecer uma boa condição de saúde para a criança (DE MOURA *et al.* 2015). De acordo com Bueno (2013), o leite materno é o alimento ideal para o crescimento e desenvolvimento da criança, além de auxiliar no combate à desnutrição e doenças comuns na infância, como a diarreia e infecções respiratórias.

Além disso, evidências científicas mostram que o aleitamento materno exclusivo gera benefícios para as mães, pois auxilia no processo de retorno do útero ao seu tamanho normal, além de diminuir os casos de anemia por sangramentos após o parto e auxiliar na perda de peso mais rápida (OLIVEIRA, 2010). Diversos estudos também mostram que a amamentação funciona como proteção contra a concepção de alguns tipos de câncer, como o de mama e de ovário.

Há ainda benefícios de natureza econômica para a família, sociedade e o Estado. A criança que é amamentada de forma adequada fica menos doente, e conseqüentemente, necessita de menor atendimento médico e remédios, além disso, diminui a necessidade dos seus pais faltarem ao trabalho (MUNIZ, 2010). Giugliani (2002) ressalta que em um país como o Brasil, onde muitas famílias sofrem com dificuldades financeiras, o benefício econômico que o aleitamento materno traz é valioso, pois se economiza na compra de outros alimentos, mamadeiras, gás e ainda nas eventuais despesas com medicação.

Sendo assim, a Organização Mundial da Saúde (OMS) e o Ministério da Saúde (MS) recomendam, com base em evidências científicas, que o aleitamento seja realizado de forma exclusiva por seis meses e de forma complementada até 2 anos ou mais. No entanto, seu abandono precoce ainda é bastante frequente (BRASIL, 2009).

A prática do aleitamento materno sofreu uma grande desvalorização em meados do século XX em todo o mundo, em decorrência de fatores relacionados, principalmente, à ascensão do capitalismo. Dentre esses fatores pode-se citar: o estímulo ao consumo de alimentos artificiais, inclusive na fase neonatal; inserção da mulher no mercado de trabalho; e, de um modo geral, alterações no estilo de vida da maior parte das sociedades (TRINDADE *et al.*, 2008).

No Brasil, a interrupção precoce do aleitamento materno exclusivo (AME) passou a ser considerado um grave problema de saúde no final da década de 1970. Nesse período, dados do Ministério da Saúde apontavam que a mortalidade infantil era de 88 por 1000 nascidos vivos. Diante deste quadro, reverter essa situação passou a ser uma das mais importantes estratégias para redução dos índices de morbimortalidade infantil no país. Esse contexto levou organismos de saúde a criarem recomendações de incentivo à prática do aleitamento materno. Sendo assim, em 1981, aconteceu a implantação do Programa Nacional de Incentivo ao Aleitamento Materno (ALMEIDA, 1999).

A partir de 1990 o Ministério da Saúde passou a intensificar o incentivo a amamentação com maiores investimentos nessa área. Atualmente existem diversos programas e políticas que buscam aumentar os índices do aleitamento materno em todo o Brasil, como os Bancos de Leite Humano e o programa Iniciativa Hospital Amigo da Criança (PARIZOTTO e ZORZI, 2008).

Segundo Boccolini, Carvalho e Oliveira (2017), a amamentação exclusiva no Brasil de crianças menores de 6 meses era próximo de 2,9% em 1980, passou para e passou para mais de 37,1% em 2013, o que representa um aumento de 34,2 pontos percentuais. Essa tendência do aleitamento materno exclusivo observada entre os anos de 1980 e 2000 é ascendente, com crescimento estatisticamente significativo em cada década. Além disso, o mais recente inquérito que abrangeu questões ligadas ao aleitamento materno exclusivo no Brasil, a Pesquisa Nacional de Saúde, realizada em 2013, mostrou que a prevalência de AME entre as crianças menores de seis meses de idade foi de 36,6% e trouxe evidências de que os percentuais de prevalência do aleitamento materno exclusivos de 2006 até 2013 sofreram estabilização.

Para além da prevalência do AME é importante compreender quais fatores levam as mulheres a desmamarem precocemente seus filhos, pois esta compreensão tem grande relevância para a saúde pública (BOCCOLINI, CARVALHO e OLIVEIRA, 2015). Entre os aspectos ligados ao binômio mãe-filho, que mais comumente são associados à interrupção precoce do AME na li-

teratura, estão: idade materna (OLIVEIRA *et al.*, 2005; CHAVES *et al.*, 2007), escolaridade materna (VENÂNCIO *et al.*, 2002; DAMIÃO, 2008), cor ou raça (KUMMER *et al.*, 2000), trabalho materno fora do lar (VIEIRA *et al.*, 2004; SALUSTIANO *et al.*, 2012; QUELUZ *et al.*, 2012), estado conjugal da mãe (CARRASCOZA *et al.*, 2006; BERNARDI *et al.*, 2009), situação do domicílio (BERSOT, 2011), número de consultas pré-natais (CALDEIRA e GOULART, 2000; CAMINHA *et al.*, 2010), tipo de parto (AUDI *et al.*, 2003), local do parto, peso ao nascer (VENANCIO *et al.*, 2002; DAMIÃO, 2008), tempo decorrido entre o nascimento e a primeira mamada (REA *et al.*, 1997; BECHE, HALPERN e STEIN, 2009), uso de chupeta (SOARES *et al.*, 2003; PARIZOTO *et al.*, 2009), sexo da criança (VIEIRA *et al.*, 2004; TABAI *et al.*, 1998).

Visto a importância desta temática e a quantidade de fatores que influenciam a duração do aleitamento materno, é de extrema relevância entender quais e como os determinantes maternos, regionais, socioeconômicos e da própria criança afetam o AME, seja agindo de forma a estender a sua duração, seja influenciado sua interrupção. Estudos nesta área são importantes, pois podem auxiliar a formulação de políticas públicas que objetivam elevar os índices de aleitamento materno no Brasil e suas regiões.

Desta forma, este artigo possui como objetivo geral analisar os determinantes da duração do aleitamento materno exclusivo no Brasil a partir da Pesquisa de Demografia e Saúde da Mulher e da Criança (PNDS) realizada no Brasil em 2006. Pretende-se ainda relacionar a importância das variáveis maternas, regionais e socioeconômicas, dos serviços de saúde e das crianças para a determinação do tempo de aleitamento materno exclusivo.

Metodologia

Fonte dos dados

Foram utilizados os dados da Pesquisa Nacional de Demografia e Saúde da Criança e da Mulher (PNDS) do ano de 2006, a qual foi realizada pelo Ministério da Saúde, coordenada pelo Cebrap e o trabalho de campo ficou a cargo do Ibope. Esta é uma pesquisa amostral com representatividade nacional, que permite a inferência para as macrorregiões, incluindo o contexto urbano e rural. A PNDS teve como objetivo contribuir para os conhecimentos acerca saúde da criança e da mulher, e para isso foram investigados temas relativos ao domicílio, reprodução, anticoncepção, conjugalidade, planejamento da fecundidade, entre outros. Seu público alvo envolve mulheres entre 15 a 49 anos totalizando 15575 e informações sobre seus filhos com até 5 anos (PNDS, 2006).

As entrevistadas foram selecionadas por amostragem probabilística complexa, realizada em dois estágios. No primeiro estágio foram selecionados os setores censitários e no segundo estágio os domicílios. A cada domicílio sorteado aplicou-se um questionário identificado como “ficha do domicílio” para conhecimento das características sociodemográficas de cada morador e para cada mulher elegível, em idade reprodutiva, moradora do domicílio, foi aplicado o questionário básico (BRASIL, 2008). A temática do aleitamento materno fez parte da seção chamada de Amamentação e Nutrição contida no Questionário da Mulher, onde foram respondidas perguntas em relação aos filhos nascidos vivos que tinham menos de 5 anos na data do levantamento. A base de dados da PNDS 2006 pode ser acessada no site do Ministério da Saúde do Brasil (<http://bvsms.saude.gov.br/bvs/pnds>).

Variáveis utilizadas e plano amostral

A escolha das variáveis explicativas analisadas e incorporadas na análise descritiva e na modelagem baseou-se em fatores citados na literatura como determinantes para a duração do aleitamento materno e na disponibilidade dos dados na pesquisa. Dessa forma, foram selecionadas para o estudo variáveis relativas às características maternas (idade, escolaridade, situação conjugal, raça/cor, religião e trabalho fora do lar) e da criança (sexo, peso ao nascer, uso de chupeta,

primeira mamada logo após o parto e alimentação nos primeiros 3 dias), à fatores regionais e socioeconômicos (situação de residência, região geográfica e estrato econômico) e relativos à prestação de serviços de saúde (nº de consultas pré-natal, tipo de parto e local do parto).

Para verificar se as diferenças entre estas categorias foram estatisticamente significativas na prevalência de aleitamento materno foi realizado o Teste Qui-Quadrado que tem seu valor-p também apresentado nas tabelas seguintes. Para valores p abaixo de 0,1 rejeita-se a hipótese nula de igualdade entre as categorias de determinada variável e conclui-se que há diferenças significativas entre elas.

Considerando que a pesquisa analisada apresenta um desenho amostral complexo, foi utilizado o prefixo `svy` no pacote `survey` (`svy.design`) do software estatístico R para o cálculo das estimativas, visto que este comando incorpora a complexidade da amostra. Seguindo as recomendações dos relatórios técnicos da pesquisa, para incorporação das informações do plano amostral foram utilizadas as variáveis que correspondem à informação dos clusters, à informação do estrato que é o código das áreas de ponderação e o peso amostral de cada indivíduo.

Método e variável dependente

Foram ajustados modelos de regressão de Cox para a análise do impacto das variáveis independentes na duração do aleitamento exclusivo. A partir dessas análises foi possível medir o impacto de cada covariável no risco de abandono da prática de aleitamento materno exclusivo. Esse tipo de metodologia faz parte da análise de sobrevivência em que a variável resposta, geralmente, é o tempo até que o evento de interesse ocorre. A principal característica deste tipo de estudo é a presença de censura que significa que no tempo observado, ou seja, na data da pesquisa, o evento de interesse não ocorreu, qual seja a interrupção do aleitamento materno exclusivo (COLOSIMO; GIOLO, 2006).

A censura pode ser de várias naturezas, como à esquerda, quando o tempo registrado foi maior que o tempo de ocorrência do evento, censura intervalar, em que o evento ocorreu em um intervalo de tempo e por último, e mais típica, censura à direita quando o evento ocorre apenas após o fim do período observado no estudo. Nesse estudo o único tipo de censura encontrado foi o de censura à direita, a qual ocorreu nos casos em que a criança deixou de receber aleitamento exclusivo apenas depois dos 6 meses.

A variável resposta nesse estudo foi o tempo em dias até a interrupção do AME. Para sua construção foram consideradas todas as crianças que tinham até 6 meses de idade na data da pesquisa. A pergunta chave para construção desta variável foi a que questionava o tempo que a mãe deu somente o leite do peito à criança, ou seja, o tempo de interrupção do AME, o qual foi explicitamente mencionado. Nos casos em que a mãe respondeu que a criança ainda estava em amamentação exclusiva, foi feita uma conferência com as variáveis que indicavam o que a criança consumia nas 24 horas antecedentes a pesquisa, e procedeu-se da seguinte forma: em situações que a criança não recebeu qualquer tipo de alimento, nem mesmo água, a observação foi considerada como uma censura a direita, pois a interrupção do aleitamento não havia ocorrido até a data do inquérito. Nas situações em que a criança havia feito o consumo de algo (inclusive água), entendeu-se que o AME já havia sido interrompido, sendo considerada esse tipo de observação como tempo de duração do AME a idade da criança em dias, naquele momento.

Função de sobrevivência, função taxa de falha e estimador não paramétrico de Kaplan Meier

A presença de censuras implica em limitações nas técnicas mais usuais de análise descritiva. Isso acontece devido à falta de conhecimento das frequências exatas para cada uma das observações. Em casos deste tipo a principal componente descritiva é a função de sobrevivência (COLOSIMO; GIOLO, 2006).

Para estimar a função de sobrevivência muitas vezes recorre-se ao estimador de Kaplan Meier, que também é conhecido como estimador limite-produto (COLOSIMO; GIOLO, 2006).

Este estimador considera que a sobrevivência até cada tempo é independente da sobrevivência até outros tempos, além disso, considera que a probabilidade de se chegar até o tempo t é o produto das probabilidades de se chegar até cada um dos tempos anteriores (CARVALHO, 2011).

O método de Kaplan-Meier foi utilizado para comparar curvas de sobrevivência de diferentes grupos. Essa estratégia recebe o nome de estratificação. A estratificação consiste em dividir o conjunto de todas as observações em diferentes grupos, de acordo com as covariáveis de interesse para então estimar as funções de sobrevivência para cada uma delas (CARVALHO, 2011).

Modelo de regressão de Cox

O modelo de regressão de Cox é útil na análise de dados de sobrevivência. Este modelo é caracterizado pelos coeficientes β 's, que medem os efeitos das covariáveis sobre a função de taxa de falha (COLOSIMO; GIOLO, 2006). Este modelo é dito semiparamétrico, pois não assume distribuição estatística para função de risco (CARVALHO, 2011).

O modelo de Cox também é conhecido como modelo de riscos proporcionais. Isso acontece porque a razão das taxas de falhas para dois indivíduos diferentes é constante no tempo, ou seja, a razão entre elas não depende do tempo. Por exemplo, se uma criança do sexo feminino tem um risco de ser desmamada precocemente igual a duas vezes o risco de uma criança do sexo masculino, essa razão de riscos será a mesma para todo o período de acompanhamento. Dessa forma o modelo de Cox tem como suposição básica que as taxas de falha sejam proporcionais.

A partir do ajuste do modelo de Cox pode-se verificar quais os fatores de risco têm significância na presença dos demais fatores. Após o ajuste do modelo é preciso avaliar a qualidade do ajuste. Essa avaliação pode ser feita recorrendo-se a diversos métodos. O método adotado neste artigo foi a análise do coeficiente de correlação de Person entre os resíduos padronizados de Schoenfeld e $g(t)^*$ para cada uma das covariáveis. Essa metodologia indica que valores próximos de zero significam que não há evidências para que se rejeite a hipótese de riscos proporcionais (COLOSIMO; GIOLO, 2006).

Comentários metodológicos

As análises da duração do AME tiveram início com o ajuste de modelos de Cox para cada uma das variáveis individualmente. Isso foi feito para se testar a significância destas variáveis de forma individual, a partir do teste de Wald, na duração do aleitamento materno exclusivo. Foi utilizado nível de significância de 0,1 e para as variáveis que apresentaram diferenças significativas entre suas categorias foram apresentados os gráficos de Kaplan Meier. As variáveis que não foram significativas, não foram consideradas na elaboração dos modelos finais. As variáveis que demonstram ser significativas na duração do AME, foram a idade materna, escolaridade materna, situação conjugal, situação do domicílio, região de residência, estrato, nº de consultas pré-natal, local do parto, uso de chupeta, alimentação nos primeiros dias e as classes econômicas, as quais foram reagrupadas em apenas dois estratos (ABC e DE) e a fim de melhorar os ajustes.

As variáveis que foram estatisticamente significativas neste procedimento inicial foram posteriormente colocadas em um modelo completo e retiradas uma a uma conforme critério de maior p-valor. As variáveis que tiveram p-valores maiores que 0,1 para o teste de Wald foram retiradas do modelo. Após os diversos ajustes, o modelo final ajustado incluiu as variáveis: idade materna, situação do domicílio, estrato socioeconômico, nº de consultas pré-natal e uso de chupeta.

Depois de avaliadas as significâncias das covariáveis, avaliou-se a qualidade do modelo final observado, verificando se o pressuposto de riscos proporcionais estava sendo satisfeito por meio dos resíduos de Schoenfeld. É importante ainda mencionar, que para verificar a adequação do modelo não se levou em conta o desenho amostral da pesquisa, pois, segundo RINALDI (2015), quando este é considerado há violação do pressuposto de proporcionalidade dos riscos. Este pressuposto não foi violado, observando-se um p-valor global maior que 0,01.

Resultados

Prevalência do AME e tempo de amamentação exclusiva

Primeiramente é importante relatar que entre as crianças de até 6 meses de idade presentes na amostra (um total de 510 crianças) estas apresentaram uma mediana de tempo de aleitamento exclusivo de 60 dias, ou seja, metade delas eram amamentadas exclusivamente por 2 meses. Sendo que 34% delas estavam em AME, ou seja, aproximadamente 173 crianças.

A Tabela 1 traz a prevalência do AME para um conjunto de covariáveis que tradicionalmente são apontadas na literatura como muito determinantes do aleitamento. O primeiro grupo delas, se relacionada com as características maternas, e pôde-se notar que mães de maior idade apresentaram o maior percentual de AME e o Teste Qui-Quadrado indicou que estas diferenças foram estatisticamente significativas. Quanto às demais variáveis deste grupo, os pequenos diferenciais observados não foram estatisticamente significativos indicando pouca relação entre a escolaridade, trabalho, religião e raça/cor da mãe para o tempo de amamentação exclusiva das crianças. O segundo grupo de variáveis diz respeito às características regionais e socioeconômicas e revela que a zona urbana apresenta as maiores prevalências de AME (37% contra 26% do rural). Nota-se ainda que a região Sudeste detém o maior percentual de crianças em AME (46,3%) e a região nordeste a menor (24,2%). Para os estratos socioeconômicos, pode-se observar que as maiores prevalências de AME estiveram concentradas nas categorias de estratos que representam pessoas com maior poder aquisitivo. Todas as diferenças neste conjunto de variáveis foram estatisticamente significativas ao nível de 0,05.

Ainda na Tabela 1 é possível verificar as prevalências de AME e sua relação com características dos serviços de saúde, em que somente a variável número de consultas pré-natal mostrou-se associado a grandes diferenciais no tempo da amamentação exclusiva. O grupo de crianças cujas mães frequentaram um número adequado de consultas tinham uma prevalência de AME de 41,4%, enquanto que entre aquelas em que a mãe fez o número reduzido de consultas apenas 16% estavam em AME. Já o tipo e local do parto, influenciou pouco os percentuais de AME e suas diferenças não foram estatisticamente significativas. Por fim, têm-se os percentuais do aleitamento exclusivo segundo variáveis relacionadas à criança, em que somente o uso de chupeta e sexo mostraram-se estatisticamente significativos. E, como era esperado, os grupos representados pelas crianças que não usavam chupeta apresentaram maior percentuais de AME. Em relação às variáveis sexo, peso ao nascer, primeira mamada após o parto e tipo de alimentação nos 3 primeiros dias, surpreendentemente, e apesar de algumas diferenças interessantes na prevalência de AME, não foram estatisticamente significativas.

Tabela 1 - Percentual de crianças menores de seis meses que estavam em aleitamento exclusivo segundo determinantes - Brasil - 2006

(continua)

Variáveis determinantes do AME	Percentual (%)	P-valor
Maternas		
Faixa etária		
Menos que 20 anos	36,2	<0,01
20 a 35 anos	31,8	
Maior que 35 anos	54,9	
Escolaridade		
0 a 3 anos	39,1	0,25
4 a 8 anos	27,7	
9 a 11 anos	40,1	
Maior de 12 anos	33,6	
Situação conjugal		
Casada/unida	40,4	0,72
Não casada/unida	34,4	
Raça/cor		
Branca	34,6	0,93
Preta/Parda	35,5	
Religião		
Católica	39,8	0,43
Evangélica	27,3	
Outra	21,1	
Nenhuma	32,2	
Trabalho fora do lar		
Sim	37,9	0,66
Não	34,2	
Regionais e socioeconômicas		
Situação de residência		
Urbano	37,1	0,02
Rural	26,1	
Grandes Regiões		
Norte	33,7	0,03
Nordeste	24,2	
Sudeste	46,3	
Sul	26,6	
Centro-Oeste	39,3	
Estratos econômicos		
A, B e C	26,4	0,05
D e E	37,6	
Serviços de saúde		
Tipo de parto		
Cesariana	31,1	0,39
Vaginal	37,8	
Consultas pré-natal		
Abaixo do recomendado	15,9	<0,01
Recomendado e mais	41,4	
Local do parto		
Hospital público	34,1	0,85
Hospital Privado	37,5	
Outro	37,8	

Tabela 1 - Percentual de crianças menores de seis meses que estavam em aleitamento exclusivo segundo determinantes - Brasil - 2006

(conclusão)

Variáveis determinantes do AME	Percentual (%)	P-valor
Criança		
Sexo		
Feminino	28,8	0,12
Masculino	41,2	
Peso ao nascer		
Menos de 2,5kg	33,3	0,25
Maior ou igual a 2,5 kg	46,3	
Uso de chupeta		
Usa	19,4	0,001
Não usa	46,9	
Primeira mamada após o parto		
Na primeira hora	35,1	0,38
Após a primeira hora	35,2	
Alimentação nos 3 primeiros dias		
Somente leite materno	42,9	0,39
Outros líquidos	34,6	

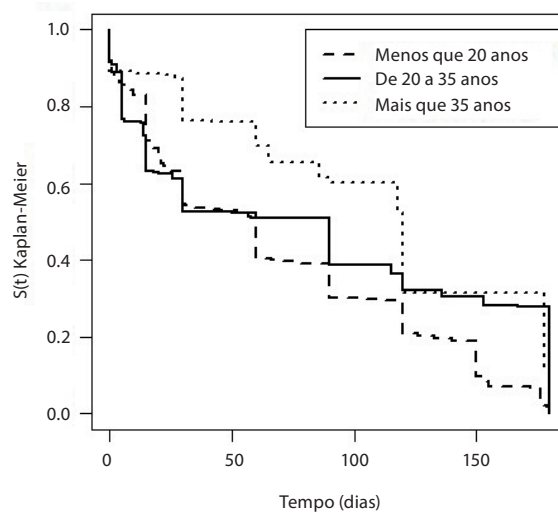
Fonte: PNDS 2006.

Nota: AME = Aleitamento materno exclusivo.

Com relação aos tempos de AME, tem-se abaixo somente as variáveis que foram estatisticamente significativas. Pode-se observar na Ilustração 1 que as mães mais jovens apresentaram maiores probabilidades de interromper o aleitamento materno exclusivo de seus filhos, para esta categoria o tempo mediano de AME encontrado foi de 60 dias. O maior tempo de amamentação exclusiva foi observado na categoria que concentra as mulheres mais velhas, mediana de 120 dias e isso possivelmente pode ser explicado pela maior experiência sobre aleitamento materno para esse grupo.

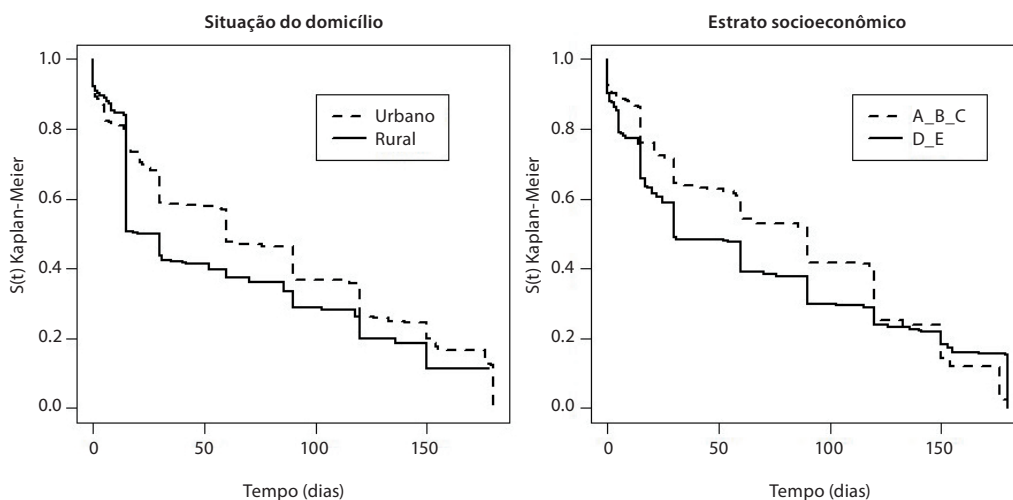
No grupo das características sociodemográficas, observa-se que visualmente nas curvas de Kaplan Meier apresentadas na Ilustração 2, que as mães que residem em áreas rurais amamentaram de forma exclusiva por mais tempo. O tempo mediano de AME para a categoria urbano foi de 60 dias contra 22 dias da categoria rural. As curvas referentes aos estratos econômicos mostram maiores probabilidades de AME para as crianças que fazem parte do estrato mais favorecido. Nota-se também que os estratos A_B_C apresentaram o maior tempo mediano de AME, que foi 90 dias. Enquanto que o tempo mediano de AME para o estrato socioeconômico de nível mais inferior foi de 30 dias. Logo a partir destas evidências é possível concluir que o fato de pertencer a grupos mais favorecidos está associado a maiores probabilidades de extensão da duração do aleitamento materno exclusivo.

Ilustração 1 - Curvas de Kaplan-Meier para a duração do AME em dias segundo idade materna - Brasil - 2006



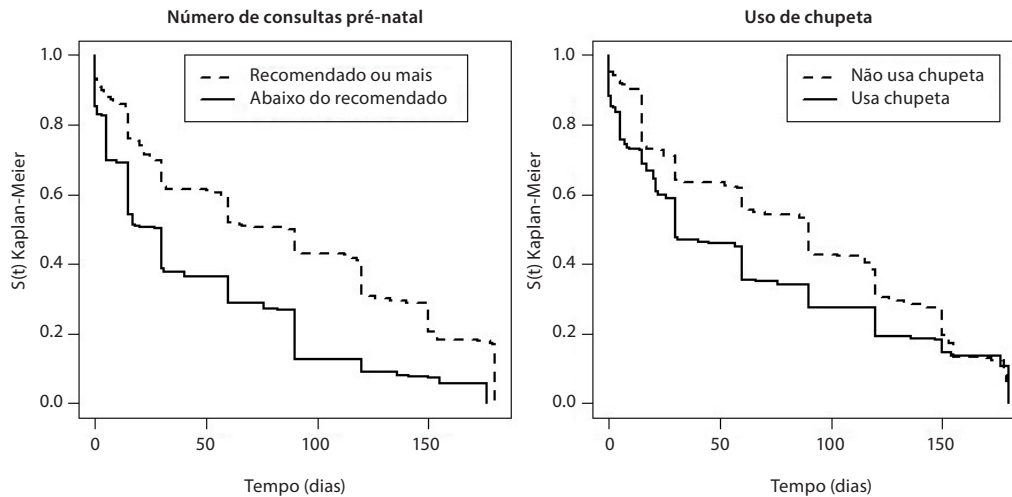
Fonte: PNDS-2006.

Ilustração 2 - Curvas de Kaplan-Meier para a duração do AME em dias, segundo características demográficas - Brasil - 2006



Fonte: PNDS-2006.

Quanto às variáveis relacionadas aos serviços de saúde (Ilustração3), somente o número de consultas pré-natal foi significativo e esta variável pode ser considerada fator determinante na duração do aleitamento materno exclusivo. Como pode ser visto na Ilustração, para as crianças cujas mães se submeteram a um número desejado de consultas pré-natal a duração do aleitamento materno exclusivo mostrou-se maior. Para estas crianças o tempo mediano de AME foi de 90 dias, o triplo do que foi observado para crianças cujas mães não frequentaram um número desejado de consultas. E na mesma ilustração tem-se a única variável do grupo das características da criança que se mostrou relevante, na qual o uso de chupeta está associado a um menor tempo de aleitamento materno exclusivo, em que a curva relacionada às crianças que não usaram chupeta tem comportamento superior durante todo o período analisado. As crianças que não usaram chupeta tinham um tempo mediano de AME de 90 dias, e para aquelas que usavam esse tempo foi de 60 dias.

Ilustração 3 - Curvas de Kaplan-Meier para a duração do AME em dias segundo fatores relacionados ao número de consultas pré-natal e uso de chupeta - Brasil - 2006

Fonte: PNDS-2006.

Modelos de Cox e discussões

A análise descritiva apresentada acima possibilitou um entendimento geral da maneira como cada covariável se relaciona com o tempo de amamentação exclusiva. Contudo, é necessário entender como estas variáveis agem em conjunto e por isso então a necessidade da modelagem dos dados. Os resultados estimados dos modelos de Cox encontram-se na Tabela 2, em que para a interpretação é preciso levar em conta a exponencial dos coeficientes que são os riscos relativos associados a interromper a amamentação exclusiva. Neste modelo, manteve-se somente as variáveis que foram estatisticamente significativas na etapa anterior.

Ao analisar os resultados nota-se que a idade materna foi a característica materna que apresentou associação significativa com o aleitamento materno, em que mães com 35 anos ou mais apresentaram um risco de interrupção do AME cerca de 53% menor em relação às mães que tinham entre 20 e 35 anos, esse resultado vai de encontro aos estudos de Oliveira *et al.* (2005) e Chaves *et al.* (2007). Segundo estes autores essa relação pode ser fruto da falta de informação sobre amamentação no grupo de mulheres mais novas. E por outro lado, acredita-se que as mulheres mais velhas amamentam por mais tempo porque possuem maior conhecimento e experiência em relação à vários fatores como por exemplo nascimentos anteriores.

Em relação à situação do domicílio, observa-se que mães que viviam em áreas rurais tiveram maior probabilidade de interromperem a amamentação exclusiva de forma precoce. O risco de interromper precocemente o AME para crianças que viviam em área rural foi aproximadamente 65% maior do que aquelas residentes em áreas urbanas. Esse resultado corrobora outros estudos (Bersot, 2011 e Wenzel; Souza, 2011) os quais também mostraram que mães que viviam em área rural tinham maior risco de abandonar o AME de forma precoce uma vez que haveria uma baixa qualidade dos serviços de saúde oferecidos nestas áreas bem, bem como devido à padrões culturais distintos, em que as estruturas de apoio familiar nas áreas rurais não permitem a manutenção do aleitamento materno por longos períodos de tempo. Acrescenta-se a isso, a alta informalidade no mercado de trabalho rural, em especial para as mulheres trabalhadoras do campo, em que muitas não possuem a licença maternidade garantida, o que leva ao desmame precoce.

O número de consultas pré-natal foi considerado fator determinante na duração do aleitamento materno exclusivo, em que mães que se submeteram a um número de consultas pré-natal abaixo do recomendado durante a gestação apresentaram um risco cerca de 90% maior de desmamarem precocemente seus filhos comparando-as com aquelas que seguiram as recomendações. Esses dados estão em concordância com estudos de Caldeira e Goulart (2000) e Caminha *et al.* (2010), os quais acredita que o maior contato das mães com a equipe de saúde durante o pré-natal aumenta a possibilidade de orientações sobre amamentação adequada, tornando esta mulher mais apta para permanecer no AME por mais tempo.

Tabela 2 - Resumo do Modelo de Cox referente a duração do AME e suas correspondentes razões de risco - Brasil - 2006

Covariáveis	Razão de risco	P-valor
Idade materna		<0,01
Menos de 20 anos	0,74	
20 a 35 anos	1	
Mais que 35 anos	0,47	
Situação do domicílio		0,08
Urbano	1	
Rural	1,65	
Consultas pré-natal		<0,01
Abaixo do recomendado	1,9	
Recomendado ou mais	1	
Uso de chupeta		<0,01
Usa chupeta	1,49	
Não usa chupeta	1	

Fonte: PNDS 2006.

Nota: AME = Aleitamento materno exclusivo.

Por fim, a variável uso de chupeta, mostrou-se significativamente associada a um menor tempo de aleitamento exclusivo, o risco de interrupção do AME para crianças que faziam seu uso foi cerca de 50% maior quando comparada com crianças que não a utilizavam. A associação entre o uso de chupeta e a interrupção do AME também foi observada nos estudos de Soares *et al.* (2003) e Parizoto *et al.* (2009). Entre as razões para esta relação está o fato de que o uso da chupeta diminui o número de mamadas durante o dia e isso leva a um menor estímulo da mama que desencadeia em uma diminuição da produção de leite. Além disso, pode ter relação com uma “confusão de sucção” que ocorre principalmente devido as diferenças na sucção da chupeta e do peito, acarretando em uma confusão que pode levar a interrupção do AME de forma precoce.

Considerações finais

Os resultados evidenciam que a prática do AME ainda está aquém do recomendado pelos principais organismos de saúde, uma vez que o percentual de crianças menores de seis meses sendo amamentadas exclusivamente foi de apenas 34,3%. Em comparação com outros estudos recentes, nota-se que tem ocorrido melhorias na prevalência de AME no Brasil. Boccolini, Carvalho e Oliveira, (2017) mostram que houve um crescimento importante desde os anos 80, contudo os autores chamam a atenção para o fato de que essa tendência ascendente teve seus principais ganhos entre 1986 e 2006 e que em seguida se observa uma certa estabilização, com pouco crescimento de 2006 para 2013.

Este avanço na prevalência do AME pode estar associada à campanhas e programas governamentais implementados neste período, em que se destacam: Iniciativa Hospital Amigo da Criança, a Rede Amamenta Brasil, A Rede de Banco de Leite Humano, extensão da licença maternidade para quatro meses e a criação do Código Internacional de Comercialização de Substitutos do Leite Materno. Tudo isso, direta ou indiretamente, podem ser consideradas medidas eficazes para o enfrentamento de barreiras que, muitas vezes, inviabilizam a extensão da amamentação exclusiva.

Infelizmente, para o período atual não se encontrou estimativas de prevalência de AME no Brasil. Isso, possivelmente, se deve à falta de disponibilidade de dados mais recentes que permitam atualizar as prevalências de AME para o país como um todo, ficando as iniciativas de estimação restrita a estudos de caso. Nesse sentido, faz-se urgente a necessidade de investimento em pesquisas que possibilitem a estimação da prevalência da amamentação exclusiva no Brasil, bem como permitam levantar os possíveis determinantes desse processo.

Com relação aos determinantes encontrados neste estudo para a interrupção da prática da amamentação exclusiva pode-se concluir que, entre os disponíveis no banco de dados, estes têm se tornado, cada vez mais, restritos a um pequeno número de variáveis. Entre as características maternas apenas a idade da mãe se mostrou relevante, a qual ainda é um determinante importante e que deve ser considerado quando se analisa os fatores relacionados ao desmame precoce, uma vez que mães com idade inferior tendem a introduzir alimentos mais precocemente na vida do bebê (Ferreira et al 2018).

O grupo das características regionais e socioeconômicas mostrou-se também relevante, uma vez que tanto a situação do domicílio quanto o estrato econômico foram significativos. Esses achados podem indicar a prevalência de contextos diferenciados tanto culturalmente quanto de acessos desiguais às políticas sociais, em especial de saúde, por mulheres rurais e urbanas e de diferentes estratos econômicos. Para reduzir estas diferenças faz-se necessário, sobretudo, combater as desigualdades regionais e sociais.

No que se refere às características associadas à criança, o destaque foi para o uso da chupeta, em que o seu uso esteve associado de forma negativa, ou seja, a maiores probabilidades de interrupção do AME. Este determinante tem sido apontado em diversos estudos como um dos mais importantes na determinação do AME. Apesar das recomendações contrárias sobre o uso da chupeta pelos órgãos de saúde, o uso da chupeta ainda é bastante comum no contexto brasileiro e combater a sua prática pode ser bastante difícil. Estudos mostram que muitas mães, por questões culturais, ainda veem na chupeta uma forma de consolar e acalmar o bebê e não a relacionam com o abandono precoce de AME (GALVÃO, et al 2018)

Entre as características dos serviços de saúde, apenas o número de consultas pré-natal se mostrou estatisticamente relevante para o AME, em que a maior frequência à consultas pré-natal esteve associada positivamente ao maior tempo de AME. Esta foi uma das variáveis de maior efeito para o tempo de AME e por isso, sugere-se que haja um fortalecimento das políticas de acesso ao pré-natal com maior fornecimento de orientações acerca do aleitamento materno, já que intervenções específicas de apoio profissional e acesso à informação adequada tem mostrado efetividade para melhorar as taxas dessa prática.

A redução do número de variáveis citadas na literatura que interferem no AME indicam que, embora o setor de saúde tenha papel fundamental na evolução desta prática, o AME possui forte associação com as mudanças ocorridas nos contextos sociais, culturais e econômicos vivenciados pelo país, como a urbanização, a expansão da escolaridade e inserção feminina no mercado de trabalho. Essas mudanças podem ser consideradas fundamentais na expansão do tempo de aleitamento materno ao longo do tempo, o que faz com que as características mais próximas desse processo, como aquelas relacionadas à mãe e à criança percam importância.

Além disso, não se considerou neste estudo, a influência dos fatores psicológicos que podem tanto contribuir prolongar ou desencorajar o AME. Se por um lado as crenças disfuncionais sobre a amamentação como prováveis dores, leite fraco ou problemas relacionados a distúrbios psicológicos como depressão, ansiedade, podem prejudicar AME. Por outro lado, também as

questões emocionais associadas ao afeto e laço emocional intenso que ocorre entre mãe e filho, em que se percebe a prática do AME como um facilitador do desenvolvimento da criança e que as mães veem as necessidades do bebê sendo supridas, estas sentem prazer e ocorre uma troca desse sentimento entre ambos, o que incentiva a continuação da amamentação exclusiva (Martins *et al.*, 2013).

Conclui-se, por fim, que ainda é necessária a intervenção do poder público no que diz respeito ao estímulo à amamentação exclusiva, sendo esta prática de fundamental importância para a saúde materno-infantil. A continuidade das políticas que atendem essa temática aliada à implementação de novas políticas de apoio e incentivo ao aleitamento materno possivelmente poderão ajudar o Brasil a alcançar a meta de redução da mortalidade infantil e desnutrição, dado o impacto positivo que esta prática proporciona.

Por fim, é fundamental que todas as pessoas incluindo a família, médicos e enfermeiros que estão envolvidas com o recém-nascido desde o seu nascimento tenham consciência da importância desta prática. As particularidades de cada mãe, principalmente quanto aos aspectos psicossociais, precisam ser compreendidas e os profissionais de saúde que lidam com as gestantes e recém-nascidos devem atentar-se para cada fator que prejudica o sucesso dessa prática, auxiliando no enfrentamento de todas essas dúvidas e dificuldades.

Referências

- AUDI, C. A. F.; CORRÊA, A. M. S.; LATORRE, M. R. D. Alimentos complementares e fatores associados ao aleitamento materno e ao aleitamento materno exclusivo em lactentes até 12 meses de vida em Itapira, São Paulo. *Revista Brasileira de Saúde Materno Infantil*, v. 3, n. 1, 2003.
- ALMEIDA, J. A. G. *Amamentação: um híbrido natureza-cultura*. Rio de Janeiro: Editora Fiocruz, 1999.
- BECHE, N.; HALPERN, R.; STEIN, A. T. Prevalência do aleitamento materno exclusivo em um município serrano do Rio Grande do Sul, Brasil. *Revista da Associação Médica do Rio Grande do Sul*, v. 53, n. 4, p. 345-353, 2009.
- BERNARDI, J. L. D.; JORDÃO, R. E.; BARROS FILHO, A. de A.. Fatores associados à duração mediana do aleitamento materno em lactentes nascidos em município do estado de São Paulo. *Revista de Nutrição*, v. 22, n. 6, p. 867-878, 2009.
- BERSOT, V. F. *Mudança temporal do aleitamento materno exclusivo na América Latina e Caribe: atualização de seus determinantes e da tendência secular*. Tese de Doutorado. Universidade de São Paulo. Faculdade de Saúde Pública. Departamento de Nutrição. 2011.
- BOCCOLINI, C. S.; CARVALHO, M. L. de; OLIVEIRA, M. I. C. de; VASCONCELLOS, A. G. G. Fatores associados à amamentação na primeira hora de vida. *Revista de Saúde Pública* [online], v. 45, n. 1, p. 69-78, 2011.
- BOCCOLINI, C. S.; CARVALHO, M. L. de; OLIVEIRA, M. I. C. de. Factors associated with exclusive breastfeeding in the first six months of life in Brazil: a systematic review. *Revista de Saúde Pública*, v. 49, 2015.
- BOCCOLINI, C.; BOCCOLINI, P. de M. M.; MONTEIRO, F. R.; VENÂNCIO, S. I.; GIUGLIANI, E. R. J. Tendência de indicadores do aleitamento materno no Brasil em três décadas. *Revista de Saúde Pública*, São Paulo, v. 51, n. 108, 2017.
- BUENO, K. de C. V. N. *A importância do aleitamento materno exclusivo até os seis meses de idade para a promoção da saúde da mãe e do bebê*. Universidade Federal de Minas Gerais, 2013.
- BRASIL, Ministério da Saúde. Pesquisa Nacional de Demografia e Saúde da Criança e da Mulher -PNDS 2006. *Relatório Final*. 2008.
- BRASIL, Ministério da Saúde. Secretaria de Atenção à Saúde. Departamento de Atenção Básica. *Saúde da criança: nutrição infantil: aleitamento materno e alimentação complementar*. 2009.
- CALDEIRA, A. P.; GOULART, E. M. A situação do aleitamento materno em Montes Claros, Minas Gerais: estudo de uma amostra representativa. *Jornal de Pediatria* (Rio J), v. 76, p. 65-72, 2000.
- CAMINHA, M. D. F. C., BATISTA FILHO, M., SERVA, V. B., ARRUDA, I. K. G. D., FIGUEIROA, J. N., LIRA, P. I. C. D. (2010). Time trends and factors associated with breastfeeding in the state of Pernambuco, Northeastern Brazil. *Revista de Saúde Pública*, v. 44, n. 2, p. 240-248, 2010.
- CARRASCOZA, K. C., POSSOBON, R. D. F., TOMITA, L. M., MORAES, A. B. A. D. Consequences of bottle-feeding to the oral facial development of initially breastfed children. *Jornal de Pediatria*, v. 82, n. 5, 395-397, 2006.
- CARVALHO, M. S.; ANDREOZZI, V. L.; CODEÇO, C. T.; CAMPOS, D. P.; BARBOSA, M. T. S. SHIMAKURA, S. E. *Análise de Sobrevida: teoria e aplicações em saúde*. SciELO-Editora Fiocruz, 2011.
- CHAVES, R. G.; LAMOUNIER, J. A.; CÉSAR, C. C. Fatores associados com a duração do aleitamento materno. *Jornal de Pediatria*, v. 83, n. 3, p. 241-246, 2007.
- COLOSIMO, E. A.; GIOLO, S. R. *Análise de sobrevivência aplicada*. São Paulo: Blucher, 2006.

DAMIÃO, J. de J. Influência da escolaridade e do trabalho maternos no aleitamento materno exclusivo. *Revista Brasileira de Epidemiologia*, vol. 11, n. 3, p. 442-452, 2008.

DE MOURA, E. R. B. B. *et al.* Investigação dos fatores sociais que interferem na duração do aleitamento materno exclusivo. *Revista Intertox de Toxicologia, Risco Ambiental e Sociedade*, v. 8, n. 2, 2015.

FERREIRA, H. L. O. C. *et al.* Fatores Associados à adesão ao aleitamento materno exclusivo. *Ciência & Saúde Coletiva*, v. 23, p. 683-690, 2018.

GALVÃO, D. M. P. G. *et al.* Conhecimento das mães acerca do uso da chupeta, amamentação e desenvolvimento da criança. *International Journal of Developmental and Educational Psychology*, v. 1, n. 1, p. 239-246, 2018.

GIUGLIANI, E. R. J. *Amamentação exclusiva e sua promoção*. Rio de Janeiro, 2002.

KUMMER, S. C.; GIUGLIANI, E. R.; SUSIN, L. O.; FOLLETO, J. L.; LERMEN, N. R.; WU, V. Y.; CAETANO, M. B. Evolução do padrão de aleitamento materno. *Revista de Saúde Pública*, v. 34, n. 2, p. 143-8, 2000.

MUNIZ, M. D. *Benefícios do aleitamento materno para a puérpera e o neonato: A atuação da equipe de saúde da família*. Universidade Federal de Minas Gerais. Faculdade de Medicina. Núcleo de Educação em Saúde Coletiva, 2010.

OLIVEIRA, K. A. de. *Aleitamento exclusivo até seis meses de vida do bebê: benefícios, dificuldade e intervenções na atenção primária de saúde*. Universidade Federal de Minas Gerais. Faculdade de Medicina. Núcleo de Educação em Saúde Coletiva, 2010.

OLIVEIRA, L. D., ASSIS, A. M. O., GOMES, G. D. S., PRADO, M. D. S., BARRETO, M. L. Duração do aleitamento materno, regime alimentar e fatores associados segundo condições de vida em Salvador, Bahia, Brasil. *Cadernos de Saúde Pública*, v. 21, n. 5, p. 1519-30, 2005.

PARIZOTTO, J.; ZORZI, N. T. Aleitamento Materno: fatores que levam ao desmame precoce no município de Passo Fundo, RS. *O Mundo da Saúde*, v. 32, n. 4, p. 466-74, 2008.

QUELUZ, M. C.; PEREIRA, M. J. B.; DOS SANTOS, C. B., LEITE, A. M.; RICCO, R. G. Prevalência e determinantes do aleitamento materno exclusivo no município de Serrana, São Paulo, Brasil. *Revista da Escola de Enfermagem da USP*, 2012.

REA, M. F. The Brazilian National Breastfeeding Program: A success story. *International Journal of Gynaecology and Obstetrics*, v. 31, p. 79-82, 1990.

REA, M. F.; VENÂNCIO, S. I.; BATISTA, L. E.; SANTOS, R. G.; GREINER, T. *et al.* Possibilidades e limitações da amamentação entre mulheres trabalhadoras formais. *Revista de Saúde Pública*, v. 31, p. 149-156, 1997.

RINALDI, A. E. M. *Tendência secular da alimentação de crianças brasileiras menores de cinco anos nas três últimas décadas*. Tese de Doutorado. Universidade de São Paulo, 2015.

SALES, C., CASTANHA, A & ALÉSSIO, R. Aleitamento Materno: representações sociais de mães em um distrito sanitário da cidade do Recife. *Arquivos Brasileiros de Psicologia*, v. 69, n. 1, p. 184-199, 2017.

SALUSTIANO, L. P. de Q.; DINIZ, A. L. D.; ABDALLAH, V. O. S.; PINTO, R. de M. C. Fatores associados à duração do aleitamento materno em crianças menores de seis meses. *Revista Brasileira de Ginecologia e Obstetrícia*, v. 34, n. 1, p. 28-33, 2012.

SOARES, M. E. M.; GIUGLIANI, E. R. J.; BRAUN, M. L.; SALGADO, A. C. N.; de OLIVEIRA, A. P.; de AGUIAR, P. R. *Uso de chupeta e sua relação com o desmame precoce em população de crianças nascidas em Hospital Amigo da Criança*, 2003.

TABAI, K. C., CARVALHO, J. F. D., SALAY, E. Aleitamento materno e a prática de desmame em duas comunidades rurais de Piracicaba-SP. *Revista de Nutrição*, v. 11, n. 2, p. 173-83, 1998.

TRINDADE, A. L. J.; LINHARES, E. F.; ARAÚJO, R. T. Aleitamento materno: conhecimentos das puérperas a respeito dessa prática. *Revista Saúde.Com*, v. 4, n. 2, p. 123-133, 2008.

VENANCIO, S. I.; ESCUDER, M. M. L.; KITOKO, P.; REA, M. F.; MONTEIRO, C. A. Frequência e determinantes do aleitamento materno em municípios do Estado de São Paulo. *Revista de Saúde Pública*, v. 36, n. 3, p. 313-8. 2002.

VIEIRA, G. O.; DE ALMEIDA, J. A. G.; SILVA, L. R., CABRAL, V. A.; NETTO, P. V. S. Fatores associados ao aleitamento materno e desmame em Feira de Santana, Bahia. *Revista Brasileira de Saúde Materno Infantil*, v. 4, n. 2, p.143-150, 2004.

Fatores que influenciam a medida socioeducativa aplicada ao adolescente autor de ato infracional, na comarca da capital do Rio de Janeiro*

Andrea Diniz da Silva**

Pedro Luis do Nascimento Silva***

Resumo

Com 30 anos de existência e algumas emendas, o Estatuto da Criança e do Adolescente, Lei 8.069/90, ainda é motivo de polêmica. Em especial, no que se refere à inimizabilidade penal de adolescentes, a qual é contestada pelos que preconizam a redução da maioridade penal de 18 para 16 anos. Legislação nacional e normativas internacionais de proteção dos direitos da criança e do adolescente seguem inócuas diante da superlotação de unidades de internação de adolescentes e da precariedade da Rede de Atendimento ao adolescente em conflito com a Lei. Para contribuir com o debate, este trabalho apresenta uma investigação da relação de fatores sociodemográficos e jurídico-processuais com a medida socioeducativa aplicada a adolescentes que cometeram ato infracional. Os dados foram coletados na Justiça da Infância e Juventude da Comarca da Capital do Rio de Janeiro, em uma amostra estratificada de 770 processos iniciados nos anos de 1996 a 1999. Usando um modelo multinomial logístico, foram identificados os seguintes fatores influenciadores da escolha da medida socioeducativa determinada pelo juiz: sexo, local de residência, tipo do ato infracional atribuído ao adolescente e reincidência na prática de ato infracional. O Estudo ajuda a entender como se perpetua o perfil do adolescente encarcerado e o porquê da precariedade do Sistema.

Palavras-chave: Estatuto da Criança e do Adolescente. Ato Infracional. Regressão Multinomial Logística.

* Este capítulo é resultado da dissertação de mestrado concluída em 2001 por Andrea Diniz da Silva, intitulada "Fatores que Influenciam a Medida Socioeducativa Aplicada ao Adolescente Autor de Ato Infracional, na Comarca da Capital do Rio de Janeiro", sob orientação de Pedro Luis do Nascimento Silva.

** Servidora do IBGE desde 2002, mestre em Estudos Populacionais e Pesquisas Sociais pela ENCE em 2001 e doutora em População, Território e Estatísticas Públicas pela ENCE em 2018. Atua como docente do programa de pós-graduação da ENCE desde 2019.

*** Pesquisador da ENCE e Docente do seu programa de pós-graduação desde sua criação em 1998.

Introdução

Desde os primeiros anos, a implantação do Estatuto da Criança e do Adolescente, Lei 8069/90, vem sendo marcada por divergências. Símbolos históricos do dissenso, a matéria veiculada na edição de 4 de setembro de 1996 da revista *Isto É* registrava os discursos do General Nilton Cerqueira e do Ouvidor da Polícia do Estado de São Paulo Benedito Domingos Mariano. Se de um lado o General mantinha uma fala pautada em sentenças como “Quem rouba os tênis das crianças que vão ao colégio? Quem assalta as crianças nos ônibus? São os menores infratores. A lei acaba deixando desprotegida a maioria, que são as vítimas. Os infratores ficam em liberdade por causa da impossibilidade de uma atuação serena e enérgica dos policiais. Com isso, aqueles elementos de alta periculosidade têm campo aberto para suas ações criminais.”, do outro, o Ouvidor lembrava que “No Brasil é muito comum que as questões sociais não resolvidas se transformem em questão de polícia. É o caso dos milhares de meninos e meninas de rua, marginalizados pela sociedade e pelo Estado. Usados muitas vezes pelos ‘pais’ de rua, por maus policiais e pelos narcotraficantes, alguns meninos e meninas de rua são utilizados para atos delituosos leves ou graves. E, quando isso acontece, a tese de penalizar o adolescente aos 14 ou 16 anos vem à tona...”.

Este debate se perpetuou e as questões relativas ao tratamento a ser dado ao jovem em conflito com a lei continuam sendo polêmicas. Uma extensa discussão sobre a maioridade penal, ou seja, a idade a partir da qual se pode aplicar o código penal e o sistema prisional de adultos, atualmente iniciada aos 18 anos completos, ainda é ponto de pauta em todos os segmentos sociais. Acaloradas discussões sobre propostas de rebaixar para dezesseis, quatorze e até mesmo doze anos a idade para a imputabilidade continuam frequentando muitas rodas de conversa.

A mídia segue sendo a principal colaboradora daqueles que defendem o rebaixamento da maioridade penal, despejando ostensivamente imagens que contribuem para a formação de um imaginário social voltado contra o “trombadinha”, o “menor bandido”, ou, em uma linguagem mais legalista, o “menor infrator”, termos frequentemente utilizados para designar a criança e o adolescente em situação de risco, mesmo quando não estão em conflito com a lei. Cenas veiculadas em séries de reportagens, mostrando crianças e adolescentes cometendo pequenos furtos, até mesmo de sanduíches, em locais como Praça da Sé, Cinelândia, Copacabana, substituem, rotineiramente, reportagens sobre outros assuntos de interesse nacional. A ação da mídia também contribui para degeneração da imagem destes adolescentes, com insistentes reportagens sobre adolescentes moradores de rua, usuários de drogas de todo o tipo (cola, maconha, cocaína, crack etc.), uma invasão de privacidade pouco comum em grupos de outras classes sociais, mesmo quando praticam os mesmos delitos.

Menos destacada na imprensa, a superlotação de unidades de internação de adolescentes não desperta solidariedades, apenas inspira normas de pouco efeito prático. Este é o caso da Lei 12.594, de 2012, que no inciso II do artigo 49 prevê que adolescente submetido ao cumprimento de medida socioeducativa tem o direito de “ser incluído em programa de meio aberto quando inexistir vaga para o cumprimento de medida de privação da liberdade...”. Passados sete anos e algumas rebeliões de adolescentes em razão das más condições das unidades de internação, o Ministro do Supremo Tribunal Federal ainda precisa lembrar a existência da norma aos magistrados de primeira instância e traz o assunto novamente para a pauta do dia em forma de novas controvérsias (STF, 2018; Bottari e Nunes, 2019).

Por outro lado, o cidadão comum, conservador de seu status quo, não mostra interesse em assumir uma postura crítica e acaba contribuindo para a consolidação da imagem do adolescente infrator que se protege atrás do Estatuto da Criança e do Adolescente. Frequentemente questiona o tratamento dado a estes jovens, ou mesmo uma política, nem sempre efetiva, de proteção aos direitos humanos, confundindo a prescrição de medidas socioeducativas, em princípio de caráter pedagógico e não punitivo, com impunidade.

Em meio a este cenário, juízes e técnicos que atuam nos processos judiciais de adolescentes a quem se atribui a autoria de ato infracional se sentem pressionados por uma parte

da sociedade e buscam sanar a “desordem social” com a aplicação de medidas punitivas o suficiente para restabelecer a “ordem”. Porém, diferentemente do que ocorre com o Código Penal Brasileiro (BRASIL, 1940), aplicável apenas a pessoas com 18 anos ou mais de idade que cometeram crime, no Estatuto da Criança e do Adolescente (BRASIL, 1990) não há prescrição que relacione o tipo do ato infracional cometido a uma medida socioeducativa específica. Ao contrário, este sistema estabelece apenas alguns princípios gerais para a aplicação de uma das medidas previstas, como o artigo 112 que prescreve que “A medida aplicada ao adolescente levará em conta sua capacidade de cumpri-la, as circunstâncias e a gravidade da infração” (BRASIL, 1990). Desta forma, fica o Juiz livre para, no curso do processo judicial, respeitando o devido processo legal, aplicar a remissão ou uma das medidas socioeducativas previstas no artigo 112 do Estatuto da Criança e do Adolescente. Tais medidas incluem: advertência; obrigação de reparar o dano; prestação de serviços à comunidade; liberdade assistida; inserção em regime de semiliberdade; internação em estabelecimento educacional. Neste caso, passam a ser de interesse geral os fatores ou critérios considerados pelo Juiz na escolha da medida socioeducativa que será determinada.

Para facilitar o estudo, os fatores foram classificados em dois grupos: subjetivos e objetivos. Quanto aos subjetivos, admite-se estarem ligados ao juiz e serem influenciados por sua história de vida, meio social onde vive, classe econômica a que pertence, origem étnica, idade, sexo etc. Já os objetivos podem ser denominados jurídico-processuais e são relativos a aspectos sociodemográficos do adolescente. Sendo assim, pode-se buscar entender o funcionamento do mecanismo de escolha e aplicação da Medida a jovens autores de ato infracional, buscando resposta à seguinte questão: Que fatores influenciam as Medidas aplicadas a adolescentes que cometeram ato infracional?

Considerando-se a volatilidade dos fatores subjetivos e a dificuldade de captá-los sem estudos aprofundados da personalidade e história de vida dos juizes, optou-se por estudar os fatores de natureza objetiva e sua relação de influência com a medida determinada pelo Juiz. Sendo assim, as hipóteses formuladas foram as seguintes: 1) A medida socioeducativa aplicada ao adolescente autor de ato infracional é influenciada por características sociodemográficas do adolescente; e 2) A medida socioeducativa aplicada ao adolescente autor de ato infracional é influenciada por fatores jurídico-processuais.

Para testar estas hipóteses, foram investigadas as seguintes variáveis sociodemográficas do adolescente: sexo, idade, número de anos de estudo, atividade profissional exercida, local de residência do adolescente, o fato de o adolescente ser ou não reincidente na prática de ato infracional. As variáveis jurídico-processuais consideradas incluíram: tipo de ato infracional atribuído ao adolescente, tipo de relatório ou laudo técnico existente no processo na data da audiência, além do conteúdo deste documento.

Foi escolhida a Comarca da Capital do Rio de Janeiro por ser a de maior fluxo de processos no Estado. O recorte temporal foi o período de 1996 a 1999, porque não havia quantidade significativa de prontuários para os anos anteriores, e em razão de o levantamento ter sido realizado no ano 2000, portanto não sendo este um ano completo. O levantamento foi feito coletando-se uma amostra estratificada de 800 processos para os quais foram levantadas informações sobre a medida aplicada ao adolescente pelo cometimento de ato infracional, juntamente com todas as variáveis ou fatores cuja influência se queria testar.

Para verificar a existência de relação entre a medida socioeducativa determinada pelo Juiz e os diversos fatores considerados, foi feita modelagem utilizando Regressão Logística Multinomial. O objetivo foi entender a relação entre a medida mais punitiva dentre as aplicadas com os fatores. O modelo de regressão foi ajustado levando em conta os efeitos do plano amostral estratificado aplicado para a coleta dos dados.

Assim, seguindo esta seção introdutória, são apresentados aspectos relacionados com a pesquisa realizada para obter os dados utilizados, como os que se referem à fonte, elaboração do cadastro, método de amostragem e coleta dos dados; a modelagem dos dados e uma análise dos resultados. Ao fim, alguns pontos para reflexão são propostos.

A pesquisa

Preliminarmente à coleta de dados, foi realizado um levantamento de possíveis fontes secundárias que pudessem fornecer as informações necessárias ao estudo. As duas principais fontes potenciais encontradas foram a pesquisa intitulada “Processos Judiciais”, realizada como parte de um convênio entre a Universidade do Estado do Rio de Janeiro e o Departamento de Ações Sócio-Educativas, Órgão do Governo do Estado do Rio de Janeiro responsável por gerir o sistema para o cumprimento das medidas socioeducativas, e o acervo documental do Centro de Documentação da Segunda Vara da Infância e Juventude.

Os dados da pesquisa, além de recentes estavam disponíveis em formato digital, o que facilitaria seu manuseio. Contudo algumas limitações para sua utilização foram encontradas. A primeira delas diz respeito ao conjunto de variáveis disponíveis. O objetivo da pesquisa era investigar o envolvimento de adolescentes com o uso e o tráfico de drogas e as variáveis utilizadas não coincidiam com aquelas elencadas para o presente estudo. A outra limitação diz respeito à população de pesquisa que, por razões estruturais, teve que ser limitada aos processos com mandado de busca e apreensão, ou seja, aos processos de adolescentes evadidos do sistema socioeducativo ou daqueles que não haviam sido localizados pela justiça. Estas diferenças impediriam o alcance dos objetivos fixados para o trabalho, e, portanto, optou-se por não utilizar esta fonte.

Quanto aos dados disponíveis no acervo documental do Juizado, apesar da diversidade de variáveis para as quais foram encontradas informações, estas se restringiam às frequências para algumas classes pré-definidas. Desta forma, não seria possível realizar o ajustamento dos modelos e a análise estatística necessários para a obtenção dos resultados pretendidos. Além disso, o conjunto total dos processos considerados na produção das estatísticas apresentadas nos relatórios inclui casos em que o adolescente autor do ato infracional está fora da faixa etária de interesse e não seria possível separá-los. Sendo assim, também esta fonte foi descartada como fonte principal, embora seus dados tenham sido utilizados para orientar o planejamento amostral aplicado.

Em razão de não ter sido encontrada fonte de dados secundária com as informações desejadas, optou-se pela realização de pesquisa no acervo dos processos da Segunda Vara da Infância e Juventude. Os procedimentos e métodos utilizados são descritos nas próximas seções.

Como o objetivo do trabalho era estudar os fatores que influenciam a escolha da medida socioeducativa aplicada ao adolescente a quem se atribui a autoria de ato infracional, buscou-se uma unidade de referência capaz de representar a singularidade de cada situação geradora de medida socioeducativa principal. Esta propriedade é considerada fundamental, pois se refere às características da situação que levou à aplicação de determinada medida, ou seja, às características pessoais do adolescente, as do ato infracional atribuído e as da intervenção de outros atores, como o defensor e o promotor, cuja análise é justamente o objeto do estudo. Preliminarmente, foram analisadas as características de duas possíveis alternativas: considerar o Adolescente ou o Processo Judicial como unidade de referência. A finalidade foi observar se uma ou ambas unidades possuem a propriedade de representar a singularidade mencionada.

Considerando ser o processo judicial o depositário da situação singular geradora da medida socioeducativa, pode-se reconhecer que o mais apropriado é admiti-lo como unidade de referência para o estudo e sobre ele obter informações das variáveis que serão estudadas, portanto sobre as características do adolescente acusado de cometer o ato infracional (sexo, idade, escolaridade, situação profissional e local de residência), sobre o ato infracional, sobre o tipo e a medida socioeducativa sugerida pelo defensor e pelo promotor e ainda sobre tipo e conteúdo de laudos e pareceres técnicos.

Definida a unidade de referência como sendo o processo judicial, foram então especificados os processos a serem considerados no estudo. Para esta tarefa, considerou-se que o que se deseja estudar são os “fatores que influenciam a medida socioeducativa aplicada a adolescentes aos quais se atribui a prática de ato infracional, na comarca da capital do Rio de Janeiro, nos anos de 1996 a 1999”. Assim, a população alvo foi formada pelos Processos Judiciais, referentes ao período de 1996 a 1999, onde se atribua a autoria de ato infracional a um adolescente cuja medida socioeducativa tenha sido aplicada por juiz da Comarca da Capital do Rio de Janeiro.

Antes da definição da população de pesquisa foi necessário fazer uma verificação preliminar do potencial informativo e da qualidade das fontes de dados disponíveis. Esta etapa foi essencial, pois seu resultado poderia limitar ou apresentar novas perspectivas em relação às unidades que poderiam fazer parte da população de pesquisa, portanto, da amostra.

Para um estudo dos procedimentos inerentes aos processos judiciais nos quais é atribuída a algum adolescente a autoria de ato infracional, foram encontradas, na Justiça da Infância e Juventude do Rio de Janeiro, três fontes distintas de dados: Processos Judiciais; Livros-Tombo; e Prontuários.

Como o Estatuto da Criança e do Adolescente prevê o prazo máximo de 45 dias para que o Estado mantenha um adolescente acautelado antes da primeira audiência e de seis meses para reavaliação da medida socioeducativa de Internação, a tramitação dos processos na Justiça da Infância e Juventude é muito rápida e dinâmica. Assim, antes de serem arquivados, há considerável dificuldade para encontrar e manusear os processos sem que sua tramitação seja comprometida. Por este motivo, apesar de se constituírem em fonte mais completa, os Processos Judiciais apresentam severas restrições quanto à possibilidade de acesso dos dados.

A segunda fonte é constituída de Livros-Tombo dos Cartórios. Esta fonte pode ser utilizada apenas de maneira auxiliar, pois não contém informações sobre os procedimentos ou sobre o adolescente, mas apenas a numeração sequencial e a data de início dos processos.

O Prontuário é um formulário contendo um resumo das informações de um ou mais processos judiciais referentes a um adolescente. Considerando a completa disponibilidade do arquivo de prontuários e que a maior parte das informações de interesse para o presente estudo poderia ser coletada pela consulta aos prontuários, estes se mostraram a fonte potencial mais apropriada para o desenvolvimento do trabalho. Além disso, não apresentaram nenhuma restrição de acesso ou em relação à qualidade dos dados de forma a comprometer o resultado do trabalho. Por estas razões decidiu-se que a fonte principal de dados para o presente estudo seria o arquivo de prontuários.

Assim, definiu-se a população de pesquisa da seguinte maneira: Processos Judiciais, referentes ao período de 1996 e 1999, onde se atribua a autoria de ato infracional a um adolescente, cuja medida socioeducativa tenha sido aplicada por juiz da Comarca da Capital do Rio de Janeiro, para os quais tenham sido confeccionados prontuários.

O cadastro utilizado para a seleção da amostra foi construído a partir dos prontuários. Além de informações sociodemográficas dos jovens acusados de cometer ato infracional, o cadastro incluiu informações sobre a medida socioeducativa aplicada e o número do processo judicial de referência, todas variáveis úteis para a especificação do plano amostral. Para sua constituição, foram consultados todos os prontuários elaborados até o dia 20 de abril de 2000, data do encerramento do cadastramento, totalizando 3.198 prontuários. Como em alguns casos o prontuário se refere a mais de um processo foi possível cadastrar 3.737 processos. Foram extraídas informações referentes a cinco variáveis: ano em que foi iniciado o processo; data de nascimento do adolescente; sexo do adolescente; medida socioeducativa aplicada; e data da aplicação da Medida.

Concluído o cadastramento inicial, procedeu-se à exclusão de unidades inelegíveis para o estudo. Tais unidades são aquelas que não pertencem à população de pesquisa definida. Portanto, foram declarados inelegíveis os processos com as seguintes características: processos criados apenas para acolher os casos oriundos de outras comarcas, ou seja, cuja medida socioeducativa foi aplicada por um Juiz de outra comarca e não da comarca da Capital do Rio de Janeiro; processos onde se observou que, no momento da confecção do prontuário, ainda não havia sido aplicada medida socioeducativa; processos referentes a jovens com idade inferior a 12 anos e superior a 17 anos, considerando-se que o estudo inclui apenas as Medidas aplicadas a adolescentes; processos cuja data de início era anterior ao ano de 1996 ou posterior ao ano de 1999, em observação ao corte temporal do estudo. Assim, o cadastro utilizado inclui 3.147 unidades.

Na especificação do plano amostral foram levadas em consideração todas as informações disponíveis sobre a população alvo do estudo, estas extraídas dos Relatórios Estatísticos da Segunda Vara da Infância e Juventude da Comarca da Capital e do Cadastro, construído es-

pecificamente para este trabalho. O conhecimento sobre a população permitiu aplicar um plano amostral estratificado simples. Este plano possibilita algum ganho de precisão sobre a aplicação de amostragem aleatória simples (COCHRAN, 1977). Contudo, mais importante é a possibilidade de assegurar que a amostra de processos inclua número suficiente de processos de algumas categorias menos frequentes, como é o caso de processos de adolescentes do sexo feminino.

Entre várias vantagens de se aplicar uma amostragem estratificada simples, tais como o aumento da precisão das estimativas e a possibilidade de estimar para subgrupos da população com eficiência e precisão controladas, uma em especial se traduziu em motivação para seu uso neste estudo: garantir maior representatividade de casos raros. Esta necessidade é observada em relação às adolescentes do sexo feminino, presentes como autoras de ato infracional em menos de 10% dos processos que constituem a população alvo e cerca de 6% daqueles que compõem o Cadastro (TJRJ, 1999). Sendo assim, a variável sexo foi considerada de fundamental importância para a definição dos estratos.

Também foi considerado o ganho de precisão possibilitado por uma estratificação em grupos homogêneos, e por este motivo foi utilizada a idade para a constituição dos estratos. Não foi encontrada nenhuma restrição para que a idade fosse utilizada como variável de estratificação, apenas a observação de que sua utilização implicaria em um aumento do número de estratos. A quantidade de estratos poderia passar de dois, com a estratificação apenas por sexo, para até quatorze estratos, com a utilização também da idade. Como o aumento da quantidade de estratos, neste caso, não implicaria em aumento do custo da coleta dos dados, e considerando o ganho de precisão, mesmo que pequeno, proporcionado pela inclusão da variável idade como variável de estratificação, optou-se por sua inclusão. Considerando-se a pequena quantidade de processos de adolescentes com idade entre 12 e 14 anos, optou-se por constituir uma classe que agrupasse as três idades. Sendo assim, a estratificação da população de pesquisa considerou o sexo (masculino e feminino) e a idade (12 a 14, 15, 16, 17 e Não Informada) do adolescente.

O tamanho da amostra foi estabelecido considerando a capacidade de trabalho e o tempo disponíveis para a coleta dos dados. Com base no trabalho já realizado por ocasião do cadastramento das unidades de pesquisa, estimou-se que o pesquisador seria capaz de coletar informações referentes a 20 processos por dia. Considerando-se 40 dias úteis, estimou-se, então em 800 o total de processos que poderiam ser consultados. Portanto, com base neste critério, decidiu-se extrair uma amostra de aproximadamente 800 processos.

A partir do tamanho estabelecido, avaliou-se a precisão esperada das estimativas, com a finalidade de verificar se seu valor era satisfatório, ou seja, compatível com os padrões de qualidade desejados. No cálculo dos valores esperados da precisão das estimativas foram utilizados os procedimentos referentes a um plano de amostragem aleatória simples (COCHRAN, 1977), pela simplicidade de sua aplicação e pelo fato de que, ao aplicar um plano estratificado simples pode-se obter algum ganho de precisão das estimativas em relação às obtidas sob amostragem aleatória simples. Assim, as estimativas de precisão aqui obtidas representam uma espécie de cota inferior para a precisão efetiva a ser alcançada com o plano amostral estratificado utilizado.

O risco de cometer um erro maior do que o especificado foi fixado em 0,05, considerando-se ser este o nível usual, e o valor da proporção a ser estimada foi aquele associado ao maior erro, portanto 50%. Assim, o erro máximo provável a ser cometido ao utilizarmos uma amostra de tamanho 800 para estimar uma proporção qualquer na população de processos da qual vai ser extraída a amostra foi de 0,035, ou seja, 3,5%.

Foi aplicada amostragem estratificada simples, com alocação desigual por sexo e proporcional por idade. A coleta dos dados foi feita por consulta direta aos prontuários selecionados a partir do cadastro construído especificamente para o presente estudo. As informações sobre as variáveis de interesse do estudo coletadas, foram atualizadas até o dia 20 de abril, data do encerramento da coleta de dados.

Em razão da indisponibilidade de alguns prontuários no momento da coleta dos dados, de alguns erros de cadastramento e ainda da decisão posterior de excluir os processos para os quais não havia qualquer informação sobre a idade do adolescente, o total de processos efetiva-

mente utilizado na análise foi reduzido para 770 (Tabela 1). Como a redução foi motivada por fatores que não poderiam ser controlados - falta de informações, não foi possível aplicar nenhum mecanismo que garantisse redução proporcional dentro dos estratos. Assim, a mudança no tamanho ocorreu em diferentes proporções em cada estrato.

Para os processos selecionados foram coletados dados referentes ao sexo, idade, escolaridade, atividade profissional e local de residência do adolescente; uma indicadora de reincidência, ato infracional atribuído, tipo de relatório ou laudo técnico, medida socioeducativa sugerida pelo técnico e medida socioeducativa determinada pelo juiz.

Tabela 1 - Processos na amostra, por sexo do adolescente, segundo os grupos de idade

Grupos de idade	Processos, por sexo		
	Total	Feminino	Masculino
Total	770	101	669
12 a 14 anos	130	26	104
15 anos	156	21	135
16 anos	224	30	194
17 anos	260	24	236

Fonte: Elaboração dos autores.

Modelagem e análise dos dados

As variáveis consideradas na modelagem foram aquelas coletadas no momento da pesquisa, já mencionadas na seção anterior. A variável resposta y foi a medida socioeducativa determinada pelo juiz. As variáveis explicativas inicialmente consideradas incluem fatores sociodemográficos como sexo, idade, escolaridade, atividade profissional, e local de residência, e fatores jurídico-processuais tais como indicador de reincidência, ato infracional atribuído, tipo de relatório ou laudo técnico, e medida socioeducativa sugerida pelo técnico.

Foi utilizado um modelo multinomial logístico (ver por exemplo HEERINGA, WEST e BERGLUND, 2010), que é uma generalização do modelo para variável resposta do tipo binária, que se aplica aos casos em que a variável resposta y possui mais que duas categorias. Nos casos em que as categorias de resposta são ordenadas (variável resposta do tipo ordinal) existem apenas $J-1$ *logits* não redundantes e a forma adequada de calcular estes *logits* é acumular as probabilidades até o nível escolhido e comparar com a probabilidade de categorias acima deste nível. Calculados os $J-1$ *logits*, o J -ésimo pode ser deduzido (POWERS e XIE, 2000, p.213). O *logit* para a j -ésima categoria de resposta para uma unidade i é dado por:

$$\text{logit } \Pr(y_i \leq j | x_i) = \log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_{0j} + \sum_{k=1}^K \beta_k x_{ik}$$

Onde:

y_i é o valor da resposta para a unidade i ;

x_i é o vetor de valores das covariáveis para a unidade i ;

$\pi_{ij} = \Pr(y_i \leq j | x_i)$ é a probabilidade condicional de a resposta da unidade i ser categoria menor ou igual a j , dado o valor das covariáveis x_i para essa unidade;

β_{0j} é o intercepto para resposta com valor igual a j ;

β_k é o efeito da k -ésima variável explicativa sobre o *logit*.

Para interpretar os resultados do modelo ajustado, basta comparar os *logits* considerando níveis distintos de cada uma das variáveis explicativas, para cada nível da variável resposta. Sendo x_k uma variável explicativa qualquer, é possível comparar a variação do *logit* da resposta y para dois níveis quaisquer desta variável calculando a **vantagem** a favor do nível 1 da k -ésima variável explicativa quando comparada com o nível base, para a ocorrência de uma das respostas acumuladas até o nível j , calculando a vantagem (razão de chances):

$$V(k,j) = \exp \left[\frac{\log \left(\frac{\pi_{1j}}{1-\pi_{1j}} \mid x_{ik}=1 \right)}{\log \left(\frac{\pi_{1j}}{1-\pi_{1j}} \mid x_{ik}=0 \right)} \right] = \exp(\beta_k)$$

Para o ajuste do modelo logístico multinomial foi utilizada como categoria base da variável resposta a medida socioeducativa mais punitiva dentre aquelas aplicáveis ao adolescente autor de ato infracional, ou seja, a medida de internação. O grau de punitividade das medidas foi estabelecido considerando a ordem estabelecida no Estatuto da Criança e do Adolescente. Esta ordenação foi validada considerando os valores estimados resultantes da aplicação de uma escala de Saaty a categorias atribuídas por um painel de especialistas – mais detalhes em SILVA (2001).

Inicialmente foi feito um ajuste de um modelo saturado contendo apenas os efeitos principais (sem interações entre as variáveis), incluindo as nove variáveis explicativas disponíveis: Sexo, Idade, Anos de Estudo, Local de Residência, Atividade Profissional, Ato Infracional, Reincidência, Tipo de Relatório Técnico, Medida sugerida no Relatório Técnico.

O procedimento utilizado consistiu em aplicar testes de hipótese acerca da contribuição de cada variável explicativa para o poder de explicação do modelo, a um nível de 5% de significância. Foram excluídas do modelo as variáveis cuja contribuição não foi considerada significativa, ao nível fixado, para explicar o *logit* da probabilidade de aplicação da Medida aplicada ao adolescente autor de ato infracional.

Considerou-se o tipo de amostragem utilizada e foram aplicados testes baseados na estatística de Wald, ao ajustar o modelo multinomial logístico. Com base nos testes de significância, concluiu-se que as variáveis idade, atividade profissional, anos de estudo, tipo de relatório e medida sugerida no relatório ou laudo técnico não contribuíam significativamente para a explicação do comportamento da medida socioeducativa (Tabela 2) ao nível de significância de 5%. Assim, tais variáveis foram excluídas, o modelo reduzido foi ajustado, e novos testes foram realizados com as variáveis restantes.

Tabela 2 - Estatísticas de teste de significância para manutenções das variáveis no modelo com efeitos principais

Variáveis	Graus de liberdade	Estatística de Wald	P-valor
Sexo	1	8,7	0,0032
Local de residência	8	20,9	0,0076
Ato infracional	7	233,1	0,0000
Reincidência	1	99,4	0,0000
Idade	5	1,8	0,8799
Atividade profissional	3	3,3	0,3443
Anos de estudo	5	10,2	0,0708
Tipo de relatório	2	4,9	0,0881
Medida sugerida no relatório	3	3,5	0,3157

Fonte> Elaboração dos autores.

O resultado dos novos testes de hipótese levou a concluir que o modelo contendo apenas quatro covariáveis explica o comportamento da medida determinada pelo juiz tão bem quanto os modelos anteriormente testados (Tabela 3).

Tabela 3 - Estatísticas de teste da significância para manutenção das variáveis do modelo reduzido contendo apenas as variáveis explicativas

Fonte	Graus de liberdade	Estatística de Wald	P-valor
Sexo	1	11,4	0,0007
Local de residência	8	26,1	0,0010
Ato infracional	7	249,9	0,0000
Reincidência	1	104,9	0,0000

Fonte: Elaboração dos autores.

Com a finalidade de obter um modelo alternativo ao modelo com efeitos principais ajustado, partiu-se para a inclusão de efeitos das interações duas a duas das quatro variáveis do modelo de efeitos principais considerado. Assim, buscou-se ajustar um modelo contendo as variáveis sexo, reincidência, ato infracional e local de residência, além dos seis cruzamentos possíveis entre pares destas variáveis. No entanto, ao processar os cruzamentos entre estas variáveis, muitas das frequências obtidas eram muito baixas ou mesmo nulas, o que dificultou o cálculo das estimativas dos parâmetros do modelo com interações com a precisão necessária para que se possa fazer alguma inferência.

Assim o modelo final considerando apenas os efeitos principais das quatro variáveis selecionadas (ver Tabela 3) foi ajustado, e as vantagens para os diversos níveis das variáveis explicativas foram calculadas (Tabela 4).

Tabela 4 - Estimativas dos parâmetros, erros padrões, vantagens e seus limites de confiança de 95%

Fator	Coefficiente estimado	Erro padrão	Vantagens	Limite inferior	Limite superior
Intercepto					
	1,6	1,0
	2,4	1,0
	2,8	1,0
	4,8	1,0
	7,0	1,0
Sexo					
Masculino	(-) 0,7	0,2	0,5	0,3	0,7
Feminino (base)	-	-	1,0
Local de residência					
Rua	(-) 1,7	0,7	0,2	0,1	0,7
Niterói ou Região dos Lagos	(-) 1,4	0,8	0,3	0,1	1,2
Zona Sul da cidade do Rio de Janeiro	(-) 0,0	0,7	1,0	0,3	3,6
Morros, favelas e Baixada Fluminense	(-) 0,9	0,6	0,4	0,1	1,3
Não informado	(-) 0,9	0,9	0,4	0,1	2,6
Outros estados	(-) 0,0	0,7	1,0	0,2	3,9
Outros municípios	0,0	1,0	1,0	0,1	7,1
Outros bairros do Rio (exceto Zona Sul)	0,6	0,6	0,4	0,1	1,4
Instituição de abrigo (base)	-	-	1,0
Ato infracional					
Uso de drogas	(-) 0,5	0,8	0,6	0,1	3,0
Não Informado	(-) 1,8	1,2	0,2	0,0	1,7
Contra a pessoa - grave	(-) 2,5	1,1	0,1	0,0	0,7
Contra o patrimônio	(-) 2,3	0,8	0,1	0,0	0,5
Outro tipo	(-) 4,3	0,8	0,2	0,0	1,1
Contra o patrimônio - grave	(-) 3,6	0,8	0,0	0,0	0,1
Tráfico de drogas	(-) 1,5	0,8	0,0	0,0	0,1
Contra a pessoa (base)	-	-	1,0
Reincidência					
Sim	(-) 2,0	0,2	0,1	0,1	0,2
Não (base)	-	-	1,0

Fonte: Elaboração dos autores.

Por meio da análise dos valores estimados dos parâmetros foi possível identificar o tipo de contribuição de cada um dos níveis dos fatores para a Medida aplicada ao adolescente a quem foi atribuída a prática de ato infracional. Realizando esta análise, pode-se observar que o grau de punitividade da medida determinada pelo juiz variou em razão das características do adolescente ou do ato praticado. Além destes valores, o modelo logístico multinomial utilizado permite fazer uma análise comparativa entre as chances de um adolescente que possui determinadas características nas variáveis sociodemográficas e jurídico-processuais incluídas no modelo, receber uma das medidas menos punitivas que a internação e as chances que tem um adolescente que não possui tais características. Esta comparação é feita através da vantagem (*odds ratio*) que têm os adolescentes de receber uma medida menos punitiva ao invés da internação.

Considerando-se então que a vantagem em favor de que um adolescente do sexo masculino receba Medidas menos punitivas é metade da verificada para adolescentes do sexo feminino, é possível afirmar que o grau de punitividade da Medida aumenta quando o sexo do adolescente é masculino. A vantagem em favor de que adolescentes reincidentes recebam uma Medida menos punitiva que a internação é menor que 1/6 daquela associada aos adolescentes que não são reincidentes. Logo, o grau de punitividade da Medida aplicada ao adolescente reincidente na prática de ato infracional, independentemente das demais características ou condições, é maior quando comparado com o dos não reincidentes.

Os casos em que o adolescente foi processado pelo cometimento de algum ato infracional contra a pessoa foram usados como categoria base de comparação para a análise. Desta forma, o grupo dos atos infracionais que diferem significativamente da categoria base é formado pelos atos contra o patrimônio, com ou sem uso de violência ou grave ameaça; análogos ao tráfico de drogas; e contra a pessoas, com uso de violência ou grave ameaça. Este grupo apresenta as menores vantagens em favor de que sejam aplicadas medidas menos punitivas. Tais vantagens variam de 0,01 a 0,10 daquela associada à categoria base. Para os casos em que o adolescente cometeu algum ato análogo ao uso de drogas; nos quais o ato praticado não foi informado ou mesmo quando foi praticado algum ato análogo a outro crime ou contravenção penal, na maioria composto por direção sem habilitação, pichação de monumentos públicos, estelionato, motins, porte ilegal de arma ou receptação, a vantagem em favor de que seja aplicada uma medida menos punitiva não se diferencia significativamente daquela associada à categoria base.

A categoria de local de residência utilizada como base de comparação para a análise foi a que agrupa os casos em que os adolescentes residem em alguma instituição de abrigo. Assim, observa-se que a vantagem em favor de que seja aplicada uma medida menos punitiva difere significativamente da base quando se trata de adolescente residente na rua. Nestes casos, a vantagem não chega a 1/5 daquela associada à categoria base. As vantagens associadas aos demais locais de residência não apresentam diferença significativa quando comparadas com a categoria base e variam de 1/4 para moradores de Niterói ou Região dos Lagos a 0,98 para moradores da Zona Sul da cidade do Rio de Janeiro.

Para reflexão

Os resultados apresentados foram obtidos em pesquisa realizada no ano de 2000. Vinte anos depois, duas coisas permanecessem iguais: o artigo 112 do Estatuto da Criança e do Adolescente, onde são elencadas as medidas socioeducativas aplicáveis ao adolescente que cometeu ato infracional, condicionada à existência de provas suficientes da autoria e da materialidade da infração; e a polêmica em torno da maioria penal “ideal”. Apesar dos debates havidos, não foi ainda possível fazer ver que:

- inimputabilidade penal não é sinônimo de impunidade;
- o elenco de medidas socioeducativas em muito se assemelha ao rol das penas previstas no Código Penal, diferindo destas apenas na prescrição de aplicação em instituição própria e separada de adultos;
- o grau de punitividade da medida escolhida varia de acordo com o perfil do adolescente a quem se atribui a autoria de ato infracional, tal qual ocorre, na prática, com os adultos a quem se atribua autoria de crime.

Por estes motivos, se faz necessário retomar a discussão e, mais uma vez, demonstrar que os atos infracionais são punidos e de maneira diferencial dependendo de quem o cometeu.

Para os casos estudados, apenas o sexo e o local de residência se mostraram significativamente influentes para o grau de punitividade da medida determinada pelo Juiz. Embora a influência do fator sexo no grau de punitividade da medida não seja tão grande quanto a do local de residência, observou-se que os adolescentes do sexo masculino têm maior chance de receber medidas mais punitivas pelo cometimento do mesmo ato infracional que adolescentes do sexo feminino, ainda que sob as mesmas condições. Embora a grande maioria dos adolescentes incluídos no Cadastro Nacional de Adolescentes em Conflito com a Lei, criado em fevereiro de 2009, e mantido pelo Conselho Nacional de Justiça, ser do sexo masculino (CNJ, 2011), não há qualquer prescrição legal que recomende maior grau de punitividade para grupos sociais mais representativos. Se há um esforço em coibir a prática de ato infracional pelos adolescentes do sexo masculino que inclui uma maior carga de punição para estes jovens, este procedimento não encontra amparo legal ou mesmo doutrinário.

Ao se tratar do local de residência do adolescente, houve maior variação das chances de que fosse determinada uma medida menos punitiva que a internação. Tomando-se como categoria base para a comparação o caso em que o adolescente é residente em alguma instituição de abrigo, cuja medida é em média menos punitiva que a aplicada aos demais, observou-se que os adolescentes que residem na rua têm menor chance de receber uma medida menos punitiva que a internação.

A influência dos fatores sociodemográficos mencionados sobre a medida socioeducativa imposta significa serem estes discriminadores na sua aplicação. Não obstante haja resultados apontando “relação entre delinquência e gênero, idade, escolaridade e atividade profissional” (ADORNO, 2000), deve ser levado em conta que não há previsão legal para qualquer tipo de discriminação em relação a estas características ou mesmo em relação ao local de residência do adolescente autor de ato infracional, principalmente no sentido de propor medidas com maior grau de punitividade para meninos quando comparados com meninas, tampouco para justificar variação no grau de punitividade da medida socioeducativa aplicada aos adolescentes moradores de rua. Neste caso, é inevitável concluir que os critérios para a determinação medida socioeducativa não têm origem em prescrição legal, mas em uma discriminação de grupos sociais específicos.

Dentre os fatores jurídico-processuais, são significativamente influentes na medida aplicada ao adolescente autor de ato infracional o tipo de ato atribuído, e o fato de ser ou não reincidente. Para o tipo de ato infracional atribuído ao adolescente processado, surpreende que as maiores chances de que seja aplicada uma medida menos punitiva que a internação esteja associadas aos atos praticados contra a pessoa. Quando comparados aos casos em que o ato praticado foi contra o patrimônio, tanto para o caso em que o ato foi praticado com uso de violência ou grave ameaça quanto no caso contrário, aqueles em que o ato foi contra a pessoa apresentam menor grau de punitividade. Além de socialmente condenável, esta inversão pela valorização do patrimônio em detrimento da pessoa não observa com rigor a prescrição legal, pois esta considera o uso de violência ou grave ameaça motivo suficiente para aumentar a gravidade do ato praticado. Dos atos análogos aos crimes previstos na lei de drogas o uso de substância entorpecente é o que apresenta associação com as maiores chances de que seja aplicada medida menos punitiva que a internação. Por outro lado, os atos análogos ao tráfico, estão associados a medidas, em média, mais punitivas.

Os jovens reincidentes na prática de ato infracional apresentaram menor chance de receber medida menos punitiva que a internação que aqueles que não reincidiram. Apesar de não ter sido possível fazer qualquer distinção em relação ao tipo de ato infracional cometido anteriormente, a grande quantidade de casos em que foram determinadas medidas de internação e semiliberdade quando o adolescente é reincidente revela sua utilização sem maiores restrições. Este procedimento pode estar associado a uma interpretação incorreta do artigo 122 do Estatuto da Criança e do Adolescente, que apenas autoriza a aplicação da medida de internação nos casos em que há reincidência na prática de ato infracional, mas não determina que a medida seja

aplicada. Além de discriminador, utilizar a reincidência como critério para aplicação da medida de internação contraria o artigo 121, que determina que a internação está sujeita ao princípio de excepcionalidade.

A despeito de não haver nenhuma prescrição legal que preestabeleça o tipo de medida a ser aplicada ao adolescente que cometeu ato infracional em razão de fatores sociodemográficos ou jurídico-processuais, houve um maior esforço em punir mais duramente os adolescentes do sexo masculino, moradores de rua, que cometeram atos análogos a crimes contra o patrimônio e tráfico de drogas, quando comparados com os demais. Nem mesmo as prescrições legais que autorizam a aplicação da medida de internação, considerada a mais punitiva para os fins deste estudo, quando do cometimento de atos com uso de violência ou grave ameaça (artigo 122 do Estatuto da Criança e do Adolescente) foram determinantes na escolha da medida a ser aplicada.

Por outro lado, quando tratados os casos em que o adolescente era reincidente, o inciso II do mesmo artigo, que estabelece como condição para a aplicação da internação a “reiteração no cometimento de outras infrações graves” não foi observado, uma vez que a medida foi mais punitiva quando observada reincidência de qualquer tipo, não necessariamente com atos “graves”.

Finalmente, cabe considerar que, apesar de haver sido observada a influência dos fatores acima mencionados na aplicação da Medida aos adolescentes autores de ato infracional, estes certamente não são os únicos fatores preponderantes no momento em que a medida é determinada pelo Juiz. Considera-se de grande importância investigar fatores associados à origem étnica do adolescente, ao tipo de defesa (se advogado particular, entidade de defesa de direitos humanos ou defensor público) recebida e ao perfil da pessoa ou instituição contra quem o ato infracional foi praticado. Contudo, não havia disponibilidade de informação para estas variáveis e seu estudo foi impossibilitado. Da mesma forma, acredita-se que o ajuste de modelos alternativos, incluindo interações entre as variáveis explicativas dos modelos considerados, pode trazer algum acréscimo na qualidade das análises feitas com base apenas no modelo com efeitos principais.

Não obstante as limitações apresentadas, vale destacar a importância dos resultados obtidos, uma vez que em 30 anos de vigência do Estatuto da Criança e do Adolescente, ainda não há uma posição de consenso em torno dos critérios utilizados na aplicação de medidas socioeducativas ao adolescente a quem se atribua a autoria de ato infracional. Vale também destacar a necessidade de se aprovar uma lei que preestabeleça a Medida a ser aplicada, dados o ato infracional praticado e as circunstâncias, uma vez que a ausência de legislação nesse sentido pode trazer benefícios ou prejuízos ao adolescente dependendo, em última instância, da vontade do Juiz. Embora os resultados deste trabalho não sejam definitivos e nem esgotem a discussão em torno do tema, eles podem trazer alguma contribuição para compreender que fatores influenciam a aplicação de Medidas Socioeducativas a adolescentes que cometeram ato infracional, na Comarca da Capital do Rio de Janeiro.

Referências

ADORNO, Sérgio (coord). *O Adolescente na Criminalidade Urbana em São Paulo*. São Paulo: Núcleo de Estudos da Violência a Universidade de São Paulo, 2000. Disponível em: <http://www.usp.gov.br>.

BOTTARI, E.; NUNES, M. *Com centros superlotados, crimes terão pontuação e jovens infratores só serão internados se estiverem no topo do ranking*. Publicado pelo O Globo em 31/05/2019.

BRASIL. *Código Penal*. Decreto-lei n. 2.848. Presidência da República, Casa Civil, Subchefia para Assuntos Jurídicos. 7 de dezembro de 1940. Disponível em: http://www.planalto.gov.br/ccivil_03/Decreto-Lei/Del2848.htm. Acesso em: nov. 2018.

BRASIL. *Estatuto da Criança e do Adolescente*. Lei Nº 8.069. Presidência da República, Casa Civil, Subchefia para Assuntos Jurídicos. 13 de julho de 1990. Disponível em: http://www.planalto.gov.br/ccivil_03/Leis/L8069.htm. Acesso em: nov. 2018.

COCHRAN, W. G. *Sampling Techniques*. New York: John Wiley & Sons, 1977.

CNJ. País tem mais de 86 mil adolescentes inscritos em cadastro de infratores. Notícias CNJ. Conselho Nacional de Justiça. 2011. Disponível em: <https://www.cnj.jus.br/pais-tem-mais-de-86-mil-adolescentes-inscritos-em-cadastro-de-infratores/>. Acesso em: out. 2020.

HEERINGA, S. G.; WEST, B. T.; BERGLUND, P. A. *Applied Survey Data Analysis*. 2nd edition. Boca Raton: Chapman & Hall/CRC, 2010. 467 p.

POWERS, D. A.; XIE, Y. *Methods for Categorical Data Analysis*. San Francisco: Academic Press, 2000.

SILVA, A. D. *Fatores que Influenciam a Medida Sócio-Educativa Aplicada ao Adolescente Autor de Ato Infracional, na Comarca da Capital do Rio de Janeiro*. Escola Nacional de Ciências Estatísticas, 2001. 184 p.

STF. Ministro impõe medidas contra superlotação em unidade de internação de adolescentes no ES. *Notícias STF*. Disponível em: <http://www.stf.jus.br/portal/cms/verNoticiaDetalhe.asp?idConteudo=387589>. Acesso em: jun. 2019.

TJRJ. Relatórios Estatísticos da Segunda Vara da Infância e Juventude, anos de 1996 a 1999. *Compêndio de Relatórios Estatísticos anuais*. Tribunal de Justiça do Estado do Rio de Janeiro. 1999.

Equipe técnica

Escola Nacional de Ciências Estatísticas

Maysa Sacramento de Magalhães

Planejamento e organização da publicação

César Marques

Colaboradores

Revisão dos capítulos

Ana Maria Nogales Vasconcelos (UnB)
André Junqueira Caetano (PUC Minas)
Andréa Branco Simão (CEDEPLAR/UFMG)
Carlos Henrique Conseuil (IPEA)
Carolina Guidotti (Universidad de la República Uruguay/URU)
Cassio Freitas Pereira de Almeida (ENCE/IBGE)
Douglas Sathler dos Reis (UFVJM)
Enrique Peláez (Universidad Nacional de Córdoba/ARG)
Flávia da Fonseca Feitosa (UFABC)
Gabriel Mendes Borges (IBGE)
Gustavo Junger da Silva (IBGE)
Ingrid Gomes Dias (UFV)
Joice de Melo Vieira (NEPO/UNICAMP)
Laetícia Rodrigues de Souza (NEPO/UNICAMP)
Leandro Mariano González (Universidad Nacional de Córdoba/ARG)
Luana Junqueira Dias Myhrra (UFRN)
Maira Covre Sussai Soares (UERJ)
Marcos Roberto Gonzaga (UFRN)
Marden Barbosa de Campos (UFMG)
Maria de Fátima de Lage Guerra (DIEESE)
Maria do Carmo Dias Bueno (IBGE)
Maria Tereza Serrano Barbosa (UNIRIO)
Mary Garcia Castro (FLACSO Brasil)
Pedro Luis do Nascimento Silva (ENCE/IBGE)
Ricardo de Sampaio Dagnino (UFRGS)
Sonia Albieri (IBGE) (aposentada)
Wasmalia Bivar (IBGE) (aposentada)

Projeto Editorial

Centro de Documentação e Disseminação de Informações

Coordenação de Produção

Marisa Sigolo

Gerência de Editoração

Estruturação textual

Leonardo Martins

Diagramação textual

Solange Maria Melo de Oliveira

Programação visual da publicação

Luiz Carlos Chagas Teixeira

Gerência de Gráfica

Ednalva Maia do Monte

Newton Malta de Souza Marques

Impressão e acabamento

Helvio Rodrigues Soares Filho

Série Documentos para Disseminação

ISSN 0103-6335

- 1 – O IBGE e o atendimento à sociedade: (prefácio ao projeto técnico CDDI), de Nelson de Castro Senra e Lídia Vales de Souza.
ISBN 85-240-0329-4. 1990. 43 p.
- 2 – Projetos de disseminação: contribuição ao estabelecimento de uma metodologia, de Cláudio Alex Fagundes da Silva.
ISBN 85-240-0355-3. 1990. 29 p.
- 3 – Pensando a disseminação de informações: (o caso do IBGE), de Nelson de Castro Senra.
ISBN 85-240-0459-2. 1993. 39 p.
- 4 – Memória institucional do IBGE: em busca de um referencial teórico, de Icléia Thiesen Magalhães Costa.
ISBN 85-240-0446-0. 1992. 40 p.

Subsérie Memória Institucional

ISSN 0103-6459

- 1 – Teixeira de Freitas: pensamento e ação, de Mario Augusto Teixeira de Freitas.
Organizado pelo Setor de Memória Institucional.
ISBN 85-240-0351-0. 1990. 140 p.
- 3 – Pró-censo: algumas notas sobre os recursos para o processamento de dados nos recenseamentos do Brasil, de Francisco Romero Feitosa Freire.
ISBN 85-240-0460-6. 1993. 53 p.
- 4 – A criação do IBGE no contexto da centralização política do Estado Novo, de Eli Alves Penha.
ISBN 85-240-0463-0. 1993. 123 p.
- 5 – IBGE: um retrato histórico, de Jayci de Mattos Madeira Gonçalves.
ISBN 85-240-0542-4. 1995. 61 p.
- 6 – Síntese histórica da formação dos Estados, Distrito Federal e Território da República Federativa dos Estados Unidos do Brasil e divisas inter-estaduais, de Ildefonso Escobar.
ISBN 85-240-0545-9. 1995. 144 p.
- 7 – O pensamento de Fábio de Macedo Soares Guimarães: uma seleção de textos.
Organizado por Nelson de Castro Senra.
ISBN 85-240-3868-3. 2006. 282 p.
- 8 – Isaac Kerstenetzky: legado e perfil.
Organizado por Nelson de Castro Senra.
ISBN 85-240-3900-0. 2006. 213 p.
- 9 – Giorgio Mortara: ampliando os horizontes da demografia brasileira.
Organizado por Nelson de Castro Senra.
ISBN 85-240-3937-9. 2007. 105 p.
- 10 – A estatística brasileira e o Esperanto: uma história centenária: 1907-2007.
Organizado por Nelson de Castro Senra.
ISBN 85-240-3944-7. 2007. 161 p.
- 11 – Bulhões Carvalho, um médico cuidando da estatística brasileira.
Organizado por Nelson de Castro Senra.
ISBN 978-85-240-3982-9. 2007. 433 p.
- 12 – Embaixador Macedo Soares: um príncipe da conciliação: recordando o primeiro presidente do IBGE.
Organizado por Nelson de Castro Senra.
ISBN 978-85-240-4008-5. 2008. 331 p.
- 13 – O IBGE na história do municipalismo e sua atuação nos municípios: o pensamento de Teixeira de Freitas e de Rafael Xavier.
Organizado por Nelson de Castro Senra.
ISBN 978-85-240-4017-7. 2008. 432 p.

- 14 – Lyra Madeira, um mestre da demografia brasileira.
Organizado por Nelson de Castro Senra.
ISBN 978-85-240-4032-0. 2008. 134 p.
 - 15 – Teixeira de Freitas, Um Cardeal da Educação Brasileira: sua atualidade intelectual.
Organizado por Nelson de Castro Senra.
ISBN 978-85-240-4052-8. 2008. 266 p.
 - 16 – Geografia e Geopolítica: a contribuição de Delgado de Carvalho e Therezinha de Castro.
Organizado por Marco Aurelio Martins Santos
ISBN 978-85-240-4084-9. 2009. 432 p.
 - 17 – Evolução da divisão territorial do Brasil 1872-2010.
ISBN 978-85-240-4208-9. 2011. 264 p.
 - 18 – Christovam Leite de Castro e a Geografia no Brasil.
Organizado por Leandro Malavota.
ISBN 978-85-240-4274-4. 2013. 340 p.
 - 19 – ENDEF
Organizado por Leandro Malavota.
ISBN 978-85-240-4330-7. 2014. 340 p.
 - 20 – PNAD
Organizado por Leandro Malavota, Luigi Bonafé e Vera Abrantes.
ISBN 978-85-240-4364-2. 2015. 202 p.
 - 21 – Indicadores Sociais: passado, presente e futuro
Organizado por André Simões e Antônio Carlos Alkmim.
ISBN 978-85-240-4424-3. 2017. 174 p.
 - 22 – 40 anos da unidade de métodos estatísticos do IBGE
Organizado por Sonia Albieri e Antonio José Ribeiro Dias.
ISBN 978-85-240-4430-4. 2017. 214 p.
 - 23 – O Centro Brasileiro de Estudos Demográficos: Uma trajetória dos estudos populacionais no IBGE
Organizado por Leila Regina Ervatti, Antônio Tadeu Ribeiro de Oliveira e Antonio de Ponte Jardim.
ISBN 978-85-240-4431-1. 64 p.
 - 24 – Desbravar, Conhecer, Mapear: Memórias do Projeto RADAM/RADAMBRASIL
Organizado por Leandro M. Malavota e Therence Paoliello de Sarti
ISBN 978-85-240-4471-7. 364 p.
 - 25 – Bibliotecas do IBGE: 80 anos de disseminação e democratização de informações sobre o Brasil
Organizado por Claudia Rodrigues do Prado, Gerlaine Braga e Leandro M. Malavota
ISBN 978-85-240-4501-1. 192 p.
 - 26 – População, Território e Estatísticas Públicas: 20 anos do Programa de Pós-Graduação da ENCE
Organizado por César Marques
ISBN 978-65-87201-34-4. 227 p.
- ### **Subsérie Fontes de Documentação**
- ISSN 0103-6459
- 1 – A indexação do banco de metadados do IBGE, de Philippe Jean Damian, Marília de Almeida March e Vera Lucia Cortes Abrantes.
ISBN 85-240-0475-4. 1993. 25 p.

Se o assunto é **Brasil**,
procure o **IBGE**.



/ibgecomunica



/ibgeoficial



/ibgeoficial



/ibgeoficial

www.ibge.gov.br 0800 721 8181

POPULAÇÃO, TERRITÓRIO E ESTATÍSTICAS PÚBLICAS

20 anos do Programa de Pós-Graduação da ENCE

O Programa de Pós-Graduação *stricto sensu* da Escola Nacional de Ciências Estatísticas - ENCE, do IBGE, foi criado em 1998 com o Mestrado em **Estudos Populacionais e Pesquisas Sociais**. A partir de 2014, o Programa mudou de denominação, para **Pós-Graduação em População, Território e Estatísticas Públicas**, alinhado com a criação do Curso de Doutorado, então aprovado pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES.

Nesses 20 anos, a expressiva produção de 350 trabalhos acadêmicos, entre teses e dissertações, concluídos até 2018, foi marcada por duas características singulares desse Programa. A primeira corresponde ao esforço de articulação entre diferentes cortes disciplinares e enfoques metodológicos na concepção e na prática da pesquisa social, sobretudo no que tange aos fenômenos territoriais e populacionais. A segunda é a sua vinculação direta ao IBGE, o que não só propicia uma pós-graduação integrada com a prática de produção e análise das estatísticas públicas, como também possibilita acesso e interação com parte do corpo de experientes pesquisadores na área de produção de estatísticas públicas do IBGE e de outras instituições, facilitando, assim, a troca de competências aliada a uma formação que preza a realização de pesquisas e levantamentos de elevado rigor metodológico.

Com o lançamento deste livro, que reúne, em 12 capítulos, estudos e metodologias elaborados por docentes, discentes e egressos de seu Programa de Pós-Graduação, a ENCE traz a público um recorte dos trabalhos oriundos das pesquisas desenvolvidas no âmbito de suas duas linhas de pesquisa. A primeira, *População, território e condições de vida*, tem como foco o estudo da dinâmica populacional e territorial e das condições de vida da população, englobando aspectos sociais, econômicos e ambientais; a segunda, *Produção de estatísticas públicas*, abarca questões ligadas ao planejamento e à condução de levantamentos, como censos, pesquisas amostrais e registros administrativos usados para a produção de estatísticas públicas que retratem as dinâmicas populacional e territorial e as condições de vida da população.

A publicação traz, ainda, reflexões sobre perspectivas e questões centrais para as áreas de pesquisa do Programa. Em um panorama geral, são colocadas as tendências e alguns dos desafios atuais e futuros, e informações adicionais sobre discentes, defesas de teses e dissertações e atuação docente, descrevendo parte dos resultados alcançados em suas duas décadas de existência acadêmica.

