

GUIA BRASILEIRO

de análise de dados

Claudio D. Shikida
Leonardo Monasterio
Pedro Fernando Nery



armadilhas & soluções



Brasília
-2021-



Guia Brasileiro de Análise de Dados: Armadilhas e Soluções

Editores:

Claudio D. Shikida
Leonardo Monasterio
Pedro Fernando Nery

Autores por ordem de aparição:

Guilherme Jardim Duarte
Pedro H. G. Ferreira de Souza
Alexandre Dias Porto Chiavegatto Filho
Fredí Alexander Diaz-Quijano
Ilona Becskeházy
Joana Monteiro
Bárbara Caballero
Guilherme Tinoco
Daniel Duque
Pedro Masson
Álvaro J. Pereira Filho



Expediente

Presidente

Diogo Costa

Diretor de Educação Executiva

Rodrigo Torres

Diretor de Desenvolvimento Profissional

Paulo Marques

Diretora de Altos Estudos

Diana Coutinho

Diretora de Inovação

Bruna Santos

Diretora de Gestão Interna

Alana Regina Biagi Silva Lisboa

Revisão

Renata Fernandes Mourão

Luiz Augusto Barros de Matos

Editoração e idealização gráfica

Samyra Lima

A Escola Nacional de Administração Pública (Enap) é uma escola de governo vinculada ao Ministério da Economia (ME).

Tem como principal atribuição a formação e o desenvolvimento permanente dos servidores públicos. Atua na oferta de cursos de mestrados profissionais, especialização lato sensu, cursos de aperfeiçoamento para carreiras no setor público, educação executiva e educação continuada.

A instituição também estimula a produção e disseminação de conhecimentos sobre administração pública, gestão governamental e políticas públicas, além de promover o desenvolvimento e a aplicação de tecnologias de gestão que aumentem a eficácia e a qualidade permanente dos serviços prestados pelo Estado aos cidadãos. Para tanto, desenvolve pesquisa aplicada e ações de inovação voltadas à melhoria do serviço público.

O público preferencial da Escola são servidores públicos federais, estaduais e municipais. Sediada em Brasília, a Enap é uma escola de governo com abrangência nacional e suas ações incidem sobre o conjunto de todos os servidores públicos, em cada uma das esferas de governo.

Ficha catalográfica elaborada pela equipe da Biblioteca Graciliano Ramos da Enap

G9431 Guia brasileiro de análise de dados: armadilhas & soluções / Editores
 Claudio D. Shikida, Leonardo Monasterio, Pedro Fernando Nery.
 -- Brasília: Enap, 2021.
 251 p. : il. --

Inclui bibliografia.

ISBN: 978-65-87791-25-8

1. Análise de Dados – Guia - Brasil. 2. Causalidade. 3. Pobreza. 4.
Desigualdade Social. 5. Saúde. 6. Educação. 7. Crime. 8. Macroeconomia.
9. Mercado de Trabalho. 10. Opinião Pública. I. Título. II. Claudio, Shiki-
da D. III. Monasterio, Leonardo. IV. Nery, Pedro Fernando.

CDU 519.25(036)(81)

Bibliotecária: Tatiane de Oliveira Dias – CRB1/2230



Enap, 2021

Este trabalho está sob a Licença Creative Commons – Atribuição: Não Comercial – Compartilha Igual 4.0 Internacional.

As informações e opiniões emitidas nesta publicação são de exclusiva e inteira responsabilidade do(s) autor(es), não exprimindo, necessariamente, o ponto de vista da Escola Nacional de Administração Pública (Enap). É permitida a reprodução deste texto e dos dados nele contidos, desde que citada a fonte. Reproduções para fins comerciais são proibidas.

Escola Nacional de Administração Pública (Enap) Diretoria de
Pesquisa e Pós-Graduação Coordenação-Geral de Pesquisa
SAIS – Área 2-A – 70610-900 — Brasília-DF, Brasil

SU MÁ RIO

cap 01

APRESENTAÇÃO

10

CAUSALIDADE

por Guilherme Jardim Duarte

• Introdução	15
• Causas e efeitos e identificação	17
• Experimentos controlados	20
• Grafos diretos e acíclicos	23
• Confusores	24
• Colisores	26
• Mediadores	28
• Dados observacionais	28
• Experimentos naturais	29
• Diferenças em diferenças	29
• Matching	31
• Regressão descontínua	33
• Variável instrumental	34
• Controle Sintético	34
• Considerações finais	35
• Referências bibliográficas	36

cap 02

POBREZA E DESIGUALDADE

por Pedro H. G. Ferreira de Souza

• Introdução	39
• Distribuição de renda	40
• Pobreza	48
• Desigualdade	55
• Fontes de dados para o Brasil	68
• Perguntas frequentes	73
• Referências bibliográficas	79

cap 03

ANÁLISE DE DADOS EM SAÚDE

*por Alexandre Dias Porto Chiavegatto Filho e
Fredy Alexander Diaz-Quijano*

• Introdução	83
• Fontes de dados de saúde	89
• Considerações finais	96
• Referências bibliográficas	97

cap 04

EDUCAÇÃO

por Ilona Becskeházy

• Introdução	99
• Proficiência e nível socioeconômico (NSE)	103
• Proficiência e nível de complexidade do ambiente educacional	107
• A Matemática do Ideb	110
• Ideb 6,0 não é o mesmo que nível 3 no PISA!	114
• Habilidades socioemocionais não passam à frente do conhecimento relacionado às disciplinas escolares	118
• Considerações finais	122
• Referências bibliográficas	123

cap 05

CRIMES E VIOLÊNCIA

por Joana Monteiro e Bárbara Caballero

• Introdução	127
• Dados de Segurança Pública no Brasil	129
• Armadilhas no cômputo dos dados	137
• Armadilhas na interpretação dos dados	147
• Considerações finais	164
• Referências bibliográficas e fonte de dados	166
• Anexo 1	167

cap 06

MACROECONOMIA

por Guilherme Tinoco

• Introdução	171
• Medidas usuais na conjuntura econômica	173
• Os dados no Brasil	180
• Armadilhas dos dados e na interpretação dos dados	188
• Perguntas frequentes	195
• Considerações finais	199
• Referências bibliográficas e leituras adicionais	200

cap 07

MERCADO E TRABALHO

por Daniel Duque

• Introdução	203
• Principais indicadores	204
• Dados no Brasil	214
• Armadilhas dos dados: interpretação, uso e questões recentes	221
• Referências bibliográficas e leituras adicionais	228

cap 08

OPINIÃO PÚBLICA

por Pedro Masson e Álvaro J. Pereira Filho

• Introdução	231
• Elaboração de survey	233
• Aplicação de surveys	239
• Análise de dados	245
• Considerações finais	250
• Referências bibliográficas	251

SU
MÁ
RIO

Dados não são dados. Ou seja, não caem do céu, prontos e perfeitos para quem os consome. São o resultado, por vezes, de longos processos de construção que envolvem várias decisões metodológicas. Além disso, o consumo dos dados não é imediato. Quem lê ou analisa os dados precisa também estar capacitado para compreendê-los. Mesmo pesquisadores experimentados podem cair nas diversas armadilhas que uma nova e desconhecida base de dados geralmente apresenta.

Nunca houve tantos dados disponíveis. E nunca houve, portanto, tantos problemas na sua interpretação. Chamamos este livro de “Guia” justamente por pretendemos que o livro mostre rumos, sugira roteiros, aponte perigos e indique atalhos. A criação de um material didático como este foi um projeto que entusiasmou os editores e os 11 especialistas que assinam os oito capítulos deste Guia. Além da excelência acadêmica de cada um, registre-se a generosidade de cada autor em ajudar no esclarecimento de conceitos que são muitas vezes citados, mas mal compreendidos, no debate público.

O público-alvo deste Guia? Pensamos que qualquer um que já tenha questionado alguma manchete de jornal, aquele que escreve uma notícia ou mesmo aquele pesquisador qualificado que se depara com um tema novo. Isto é, buscamos contemplar aquele leitor interessado que precisa dar seus primeiros passos em um território novo.

Os temas do Guia se dividem em oito capítulos que podem ser lidos separadamente.

No primeiro deles, Guilherme Duarte (Princeton University) fala sobre as dificuldades que temos para definir relações causais entre variáveis. São problemas que surgem quando, por exemplo, lemos que a reforma trabalhista aumentou o desemprego ou que armas aumentam os homicídios. Será mesmo? O entendimento científico acerca dos problemas que envolvem a causalidade tem avançado bastante nos últimos anos e este capítulo ajuda o leitor a se situar neste importante debate.

O capítulo de Pedro Souza (Ipea) trata de dois temas caros ao debate econômico: pobreza e desigualdade. Como podemos afirmar que, por exemplo, uma sociedade é mais desigual que outra? Ou quantas pessoas são pobres? O popular índice de Gini é a mais conhecida, há outras medidas que podem revelar características da distribuição de renda em uma sociedade.

Com a pandemia, termos técnicos da área da saúde se disseminaram nas notícias. Entendê-los passou a ser uma tarefa básica para quem deseja se informar sobre a Covid-19. Fala-se muito, por exemplo, em co-morbidades. Será que você realmente sabe a diferença entre mortalidade e morbidade? Alexandre Chiavegatto e Fredi Dias-Quijano (USP), no terceiro capítulo, ajudam-nos a compreender melhor estes e outros conceitos da área.

A participação dos alunos brasileiros no exame internacional PISA é sempre notícia e nos leva a pensar sobre a relação de seu desempenho com o que se mede com testes aplicados pelo governo brasileiro, como o IDEB. Há alguma relação entre estes testes? Ilona

Becskeházy ajuda-nos a entender estes e outros problema em educação no quarto capítulo.

Outro tema presente nos debates é o problema da criminalidade no Brasil e esta é uma área em que o leigo se depara com uma gama notável de variáveis. Quando se diz que “as mortes violentas aumentaram x% no último ano”, poucos notam a complicação existente já que a contabilização destas mortes pelo sistema de saúde e o de justiça não segue uma mesma metodologia. O quinto capítulo, de Joana Monteiro (FGV/EBAPE e Ministério Público do RJ) e Bárbara Caballero (Instituto de Segurança Pública - RJ) ajuda o leitor a desvendar o complexo mundo das estatísticas criminais.

Qual a diferença entre o PIB e o IBC-BR? Não é novidade que uma das maiores vítimas da má interpretação dos dados é a macroeconomia. O sexto capítulo, de Guilherme Tinoco (BNDES), explica de forma didática os principais conceitos macroeconômicos, as suas principais medidas no Brasil e explica que medidas de variação você deve usar em cada situação.

Outro tema econômico muito presente em noticiários é o mercado de trabalho, alvo de análise do sétimo capítulo, de Daniel Duque (Norwegian School of Economics) . O problema de se medir o emprego e o desemprego não é trivial e há várias medidas diferentes produzidas no Brasil. Caso você queira discutir o desemprego nos últimos anos, que fonte de dados usar? O CAGED? A PNAD Contínua? Pistas para a resposta neste capítulo.

O oitavo e último capítulo deste Guia tem como tema a opinião pública. Apesar de aparentemente mais fáceis de serem compreendidas, Pedro Masson (Enap) e Álvaro Pereira Filho (Western University of Ontario) mostram que há muitas armadilhas nessa área. Quem quiser entender melhor como elaborar um survey ou como analisar suas respostas encontrará neste ótimo capítulo um texto sucinto e didático.

Este guia não cobre todas as questões relacionadas com a análise de dados. Não só isso é inviável, como alguns autores, por motivos compreensíveis, foram forçados a desistir do projeto no seu decorrer. Por essa razão, algumas áreas relevantes ficaram a descoberto. Além disso, tal como os guias de viagens, este já sai da editora desatualizado. Novas bases surgem, questões metodológicas são revistas e, com isso, outros problemas surgem. Prudência segue sendo necessária.

Mesmo assim, nós - editores e autores - esperamos ter dado nossa contribuição para a tão necessária melhoria do debate sobre dados, evidências e políticas públicas no Brasil.

Somos muito gratos a Ana Rita Cunha, Mônica Waldvogel, Gustavo Maultasch e Pedro Burgos. Eles foram essenciais para pensarmos na estrutura e objetivos gerais do livro.

Por fim, agradecemos acima de tudo aos autores. Todos são especialistas reconhecidos em suas áreas que usaram seu precioso tempo, sem qualquer remuneração, para que contribuíssem com esta obra.

Os Editores

GAP

CAUSALIDADE

Guilherme Jardim Duarte

01

Introdução

O objetivo deste capítulo é expor maneiras simples de verificar causa e efeito em análise de dados. Apresentaremos qual é a noção de causa e efeito para a ciência atualmente e a concepção da hierarquia causal de Pearl. Em seguida, mostraremos como representar causa e efeito utilizando modelos estruturais, e quais questões de inferência causal são passíveis de serem respondidas (problema de identificação). Apresentaremos a representação de grafos acíclicos e diretos, muito utilizada para desenvolver argumentações de natureza causal. Por fim, mostraremos as técnicas de estimação mais utilizadas para inferência causal no dia a dia e quais as pressuposições por trás delas.

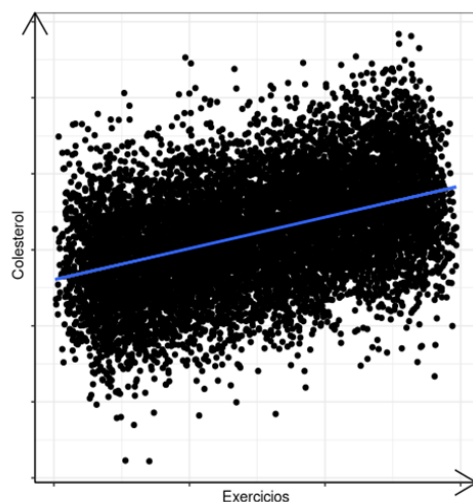
Considere as seguintes perguntas. Diminuir velocidade em vias também diminui o número de acidentes? A reforma trabalhista melhorou o desemprego? Armas aumentam o número de homicídios? Ser contra o crime aumentou a probabilidade de voto no Bolsonaro nas eleições presidenciais de 2018? Deve uma companhia enviar um certo anúncio de marketing para seus clientes? Cigarro causa câncer?

Todas essas questões guardam uma característica em comum. Elas pedem uma resposta que indique causa e efeito. O problema é que indicar causa e efeito é muito mais difícil que apenas descrever um fenômeno ou realizar uma predição.

Descrever e prever é apenas parte do trabalho de um analista. Por mais que esse tipo de tarefa pareça trivial para a população em geral, há dificuldades bastante sérias. Considere a Figura 1¹. Nela, temos um gráfico de dispersão com dados da quantidade de exercício praticada por pessoas no eixo X e medidas de colesterol no eixo Y. Apenas olhando esses dados, alguém poderia concluir que exercícios aumentam os índices de colesterol. Assim, se essa pessoa fosse chamada a dar uma recomendação de saúde, ela provavelmente diria, “com base nos dados, o melhor seria se tornar sedentário e ter a saúde milagrosamente melhorada”.

¹ Exemplo retirado de Pearl, Glymour e Jewell (2016).

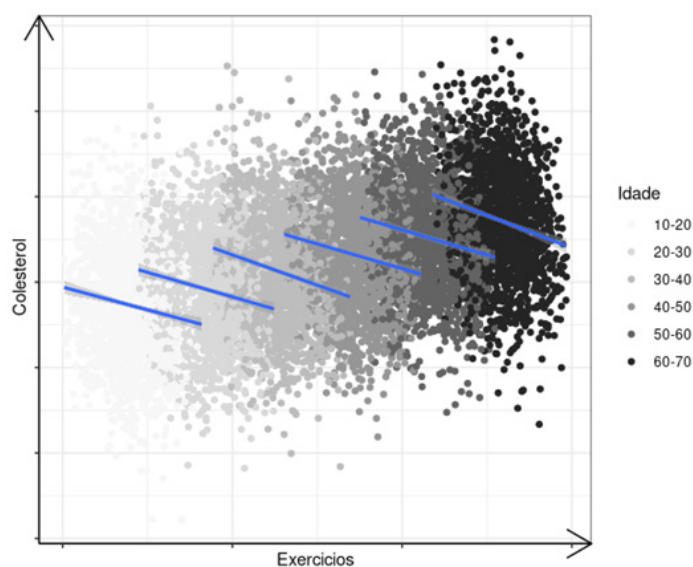
Figura 1 – Quantidade de exercício x medidas de colesterol



Fonte: autor com base em Perl, Glymour e Jewell (2016)

Essa conclusão errônea é um exemplo do problema de tentar extrapolar resultados estatísticos como causais. Como se diz comumente, “correlação não implica causalidade”. O problema, no caso, é que a quantidade praticada de exercícios e o nível de colesterol estão sendo impactados por uma terceira causa, a idade das pessoas participantes. Quando plotamos o mesmo gráfico, mas utilizando cores para representar faixas de idade, como na Figura 2, nossa conclusão pode mudar. Isto é, agora conseguimos enxergar melhor que fazer exercício reduz colesterol e é muito bom para sua saúde.

Figura 2 – Quantidade de exercício x medidas de colesterol agrupadas por faixas de idade



Fonte: autor com base em Perl, Glymour e Jewell (2016).

Esse exemplo também é suficiente para mostrar as diferenças entre predição comum e predição causal. Se pegarmos alguém aleatório, apenas com a sua quantidade de exercícios, conseguimos prever a sua quantidade de colesterol de maneira acurada², usando apenas dados. Porém, se fôssemos questionados sobre a quantidade de colesterol que esse alguém perderia com a quantidade de exercícios, teríamos que considerar também faixa de idade. É fácil perceber que as duas tarefas são bem diferentes. Uma pede apenas uma predição simples. A outra, uma predição após uma intervenção.

Se estamos interessados, então, em predição causal, algumas perguntas permanecem. Quando devemos condicionar por idade? Em quais outras variáveis devemos condicionar? O que devemos fazer para achar causa e efeito? A ciência hoje tem procedimentos claros para lidar com essas questões. Há boas referências para o público em geral³. Neste capítulo, vamos apresentar alguns desses métodos. Mostraremos como identificar situações chave. Em primeiro lugar, definiremos o que significa causalidade. Em seguida, trataremos de experimentos, o método clássico de identificar causas. Por fim, apresentaremos vieses relativos a confusores, colisores e mediadores.

Causas e efeitos e identificação

O que significa uma coisa ser causa ou efeito?

Essa é uma questão que vem desafiando filósofos pelo menos desde a Antiguidade. Uma resposta comum tinha a ver com eventos que se seguiam. Se, após presenciarmos um evento, um outro ocorre, estamos diante de uma causa e um efeito. Por exemplo, se uma bola de bilhar toca outra em movimento e a segunda se move, temos o movimento da primeira causando o movimento da segunda. Essa ideia é atraente e encontra apelo, porém ela enfrenta óbvias dificuldades.

A primeira dificuldade diz respeito ao problema de se perceber que nem todo evento que antecede outro seria aceito como causa. Imagine que após sofrer um acidente não fatal de carro, Joaquim ganhe um prêmio na mega-sena. Ou então que todo cliente de um supermercado que comprou refrigerante também comprou farinha de trigo.

² Embora não de maneira muito precisa. Site: http://fisbio.biof.ufrj.br/restrito/bioEstatistica/02_conc_basicos/acuracia_e_precisao.htm

³ Pearl and Mackenzie (201Ca8) é boa indicação para quem quer se aprofundar no tema.

Descrições desse tipo são sempre possíveis e utilizadas no dia a dia de empresas para tomada de decisão. Esse tipo de análise é conhecida como associacional, porque só indica associação entre variáveis. Poucas pessoas aceitariam que o acidente causou o prêmio de Joaquim ou, no segundo caso, que refrigerante levou as pessoas a comprar farinha de trigo. Não existe conexão entre os eventos. Desta forma, associação não é suficiente para causalidade. Precisamos de algo a mais.

Uma solução é separar os eventos que resultam de uma intervenção e os resultados que ocorrem naturalmente. Por exemplo, alguém poderia verificar se pessoas que tomam um determinado produto natural X possuem menores chances de ataque cardíaco. Todavia, esse resultado é diferente das chances de ataque cardíaco, dado que fomos lá e introduzimos o produto a uma pessoa diretamente. Essa diferença decorre da existência de fatores que podem levar pessoas a tomar o produto e a ter ataque cardíaco ao mesmo tempo (confusores). Por exemplo, um gene qualquer pode ser uma variável confusora no caso. Quando fazemos uma intervenção e introduzimos diretamente o produto a uma pessoa, estamos eliminando a influência de confusores. Assim conseguimos medir o efeito causal do produto.

Filósofos e estatísticos no século 20 passaram a acrescentar ao conceito de causalidade a ideia de contrafactual. É metafísico, mas suficiente para o esclarecimento. Suponha que Joaquim tenha sofrido o acidente e ganhado na loteria. O que ocorreria se voltássemos no tempo e o salvássemos do acidente? Será que ele venceria na loteria nesse caso? Se a resposta pra essa pergunta é sim, temos uma relação de causa e efeito. Assim, desvendar causa e efeito depende de uma máquina do tempo, com a qual mediríamos a diferença entre um evento ocorrer e não ocorrer.

Fácil utilizar esse mecanismo para pensar. Infelizmente, para além da especulação, máquinas do tempo não existem. Questões contrafactuais assim são em tese impossíveis de serem respondidas sem pressuposições adicionais. Isso ocorre também em menor grau com as questões de intervenção. É fácil medir o efeito de uma intervenção, em que damos pílulas para grupos de pessoas. Não é fácil, no entanto, saber o efeito de se introduzir parlamentarismo em um determinado país. São coisas que não podemos fazer, por serem fisicamente impossíveis, impraticáveis ou antiéticas.

A questão que se faz então é: quando é possível responder uma questão associacional, intervencional ou contrafactual? Esse problema é conhecido também como problema de identificação causal. Uma questão é não identificável quando uma resposta única não é possível. Basicamente alguém introduz dados (observacionais e experimentais) e perguntas num hipotético computador e esse computador indica quais

respostas são possíveis.

O conjunto de proposições de associação (o que?), intervenção (e se?) e contrafactual (por quê?) formam uma hierarquia, conhecida como hierarquia causal de Pearl (HCP). É uma hierarquia, porque a possibilidade de responder uma questão contrafactual implica a possibilidade de responder uma questão intervencional, que por si implica a possibilidade de responder uma questão associacional. O oposto não é verdadeiro, porém. Mas quais as condições pra responder cada uma?

A primeira condição para se responder a uma questão de causal é ter um modelo de como as variáveis se relacionam. Esse modelo, também conhecido como modelo estrutural causal, é composto de quatro conjuntos: variáveis não observadas, uma distribuição de probabilidade em relação a essas, variáveis observadas e uma relação entre todas as variáveis que indique qual causa qual. Com a ajuda desses modelos, podemos checar quais tipos de questões de qual nível da hierarquia podem ser respondidas. A partir de então, existe um método completo para responder se essas questões são passíveis de serem respondidas ou não (se são identificáveis) baseado em regras conhecidas como do calculus (SHPITSER; PEARL, 2006). Uma exposição dessas regras foge do escopo deste texto, mas é importante saber da sua existência.

Uma colher de chá. Para identificar se acidentes causam prêmios da loteria, podemos tentar analisar dados para vários indivíduos e casos, e comparar aqueles que sofreram um acidente e aqueles que não sofreram. Para fazer isso, só precisamos exigir que esses casos ou indivíduos sejam suficientemente parecidos (em termos de confundidores, como veremos abaixo). Essa exigência não é trivial e está por trás de críticas a trabalhos científicos. Um crítico sempre tentará mostrar que o pesquisador falhou em obter casos semelhantes. Por exemplo, como identificar se hospital melhora a saúde de indivíduos?⁴ Não conseguimos comparar a condição de um indivíduo que foi a um hospital com o contrafactual em que ele não foi. Um cientista x surge com a seguinte sugestão: basta comparar indicadores de saúde daqueles que vão a hospitais com aqueles que não vão. Ele obtém os dados, estima alguns modelos estatísticos e traz a seguinte conclusão: ir ao hospital é ruim para a saúde de um indivíduo, uma vez que os dados mostram que aqueles que foram ao hospital estão piores que os que não foram. Um cientista y, então, resolve criticar o trabalho. No ponto de vista de y, o trabalho do cientista x está incorreto, uma vez que ele comparou grupos de indivíduos diferentes. Aqueles que vão a hospitais fazem isso porque já se encontravam com a

4 Exemplo retirado de Angrist e Pischke (2008).

saúde debilitada antes. Nesse exemplo, é óbvio que o cientista y está correto.

Sem malabarismos, é fácil perceber que nosso sistema cognitivo é programado com um sistema para realizar inferência causal. Fazemos isso no dia a dia. Voltando ao exemplo da bola de bilhar, sabemos que a segunda vai se mover, porque no passado nos deparamos com a mesma situação e esperamos que ela vá ocorrer novamente. Entretanto, há uma série de casos que exigem um cuidado de raciocínio e metodologia.

Um último ponto de atenção. Esse raciocínio usando grupos não precisa necessariamente ser determinístico. É bastante óbvio que isso funciona para exemplos da física clássica, mas não para ciências sociais ou biológicas. A diferença entre grupos pode ser meramente probabilística. No debate público, é comum encontrar argumentos que fazem essa confusão. Por exemplo, na discussão sobre proibição de armas, é comum ver o argumento “armas não causam mais homicídios, porque nem todo mundo que possui arma, mata”. Esse raciocínio é errôneo, porque foge da questão central que é se um maior número de armas aumenta a probabilidade de mais homicídios. Outro exemplos é “não é porque é pobre que vai cometer um crime”. É importante evitar esse viés nos debates.

Experimentos controlados

Na seção anterior, falamos que, para identificar causa e efeito, basta comparar grupos semelhantes que diferem apenas pela presença de uma causa. Há uma técnica para obter grupos semelhantes. Chamamos essa técnica de experimentos controlados ou RCT (Randomized Controlled Trial). Essa técnica foi popularizada pelo famoso estatístico Ronald Fisher (ver, por exemplo, ARMITAGE, 2003). Ele escreveu um livro para pesquisadores em agricultura em 1925, mostrando os detalhes de como distinguir o que era causal do que não era.

Experimentos controlados consistem em dividir aleatoriamente casos em dois ou mais grupos. Cada caso possui, assim, chances pré determinadas de cair em cada grupo. É garantido estatisticamente que, fazendo isso, você obterá grupos muito parecidos. Em seguida, a cada um dos grupos apresentamos um evento a ser investigado. Esses grupos são chamados de grupos de tratamento. A um dos grupos, não se apresenta evento nenhum. Esse último grupo é chamado de grupo de controle. É importante ressaltar que experimentos são diretamente intervencionais. Então a identificação nesse caso é imediata.

Um exemplo de experimento controlado aparece na medicina. Vamos supor que queremos testar um novo hipotético medicamento para diminuição de peso⁵. Para fazer isso, recebemos um grupo de pessoas voluntárias ($n = 1000$) e, jogando dados, atribuímos cada uma para um grupo diferente. O procedimento é composto de dois passos:

- Definimos quatro diferentes grupos:
 - a) tratamento 1: novo medicamento;
 - b) tratamento 2: medicamento tradicional;
 - c) tratamento 3: placebo;
 - d) controle: sem medicamento.
- Para cada indivíduo, jogamos um dado: se cair 1, o indivíduo deve ir para o grupo a; se cair 2, grupo b; se cair 3, grupo c; se cair 4, grupo d; e, por fim, se cair 5 ou 6, jogamos o dado novamente.

Tabela 1 – Dados de grupos em um teste para um medicamento de redução de peso

Tamanho da amostra	Grupo	% Homens	Idade média	Peso médio (antes do trat)	Peso médio (após o trat)	Diferença
243	Trat 1	48%	35	66	55	11
248	Trat 2	49%	36	67	59	8
254	Trat 3	47%	34	67	67	0
255	Control	48%	36	67	67	0

A Tabela 1 mostra os resultados do experimento. Vale a pena notar alguns detalhes. Em primeiro lugar, os grupos não possuem o mesmo número de indivíduos, mas um número aproximado. Isso ocorre porque o procedimento é probabilístico. Além disso, podemos notar pelas colunas que indicam as porcentagens de homens, a média de idade e a média de peso antes do tratamento, que os grupos são bastante semelhantes, garantindo nossa premissa de comparação. Por fim, podemos verificar a diferença do peso médio após e antes do tratamento. Como podemos ver, o medicamento novo é mais efetivo que o medicamento tradicional, e os dois geram redução do peso.

5 Este medicamento é hipotético. Os dados foram simulados.

Os grupos de controle e de tratamento com placebo não apresentaram nenhuma diferença⁶. Para conclusão, o experimento atestou a efetividade do medicamento. Podemos dizer que, em média, ele reduz o peso de indivíduos quando administrado.

Demos um exemplo hipotético da medicina, mas hoje os experimentos controlados são bastante utilizados em economia, ciência política e para análise de políticas públicas. Como é um método claro para inferência causal, é bastante comum que políticas públicas no mundo inteiro sejam avaliadas dessa forma⁷. Um exemplo é o **Progreso-Oportunidades**, programa de transferência de renda mexicano, que inspirou o brasileiro **Bolsa Família**. Iniciado em 1997 como um programa nacional, o **Progreso** distribuiu renda e estabeleceu incentivos em saúde, educação, e nutrição, para cerca de 2,6 milhões de famílias, em 50 mil comunidades rurais. (ver, por exemplo, GERTLER, 2004). Inicialmente, foram criados grupos aleatórios de tratamento e controle para 505 comunidades (320 para o tratamento e 185 para o controle). O impacto do programa foi avaliado e ele foi considerado um sucesso em termos de seus impactos.

Dessa forma, pela sua garantia de obter causa e efeito, experimentos controlados são favorecidos dentro da pesquisa científica. Porém, ele também apresenta limitações. A primeira delas tem a ver com viabilidade. Alguns experimentos são naturalmente impossíveis. Por exemplo, para verificar o efeitos da Segunda Guerra Mundial, precisaríamos criar grupos de tratamento de planetas com a mesma história humana que o nosso. A não ser que sejamos deuses, isso é impossível. Experimentos também podem ser inviáveis de um ponto de vista ético. Se tomarmos novamente o exemplo do hospital, seria eticamente reprovável recusar entrada para pacientes terminais por meio de sorteio. Dificilmente também um experimento para investigar compra de votos e que efetivamente comprasse votos seria aprovado por uma comissão de ética.

Outra limitação dos experimentos tem a ver com validade externa ou transportabilidade. Isso ocorre quando desejamos extrapolar conclusões em relação a um grupo experimentado para um mais geral. Em relação ao impacto de determinadas drogas, é muito comum que sejam experimentadas em animais em vez de humanos. Porém, não é garantido que as conclusões atingidas sejam generalizáveis de um grupo para o outro. Em termos de políticas públicas locais, é sempre necessário justificar que os resultados de um experimento aplicado em um país (por exemplo, Cingapura ou o Quênia) possam ser extrapolados para outro país (Brasil).

6 Na prática, apresentaríamos também o erro padrão e o nível de significância da diferença.

7 No Brasil, infelizmente, isso ainda não é muito comum.

Grafos diretos e acíclicos

Como vimos acima, é impossível responder questões de natureza causal sem um modelo estrutural e outras pressuposições. Para identificação, precisamos de uma linguagem que descreva como causas e efeitos podem ser estruturados. Uma forma de representar esses modelos é através de grafos.

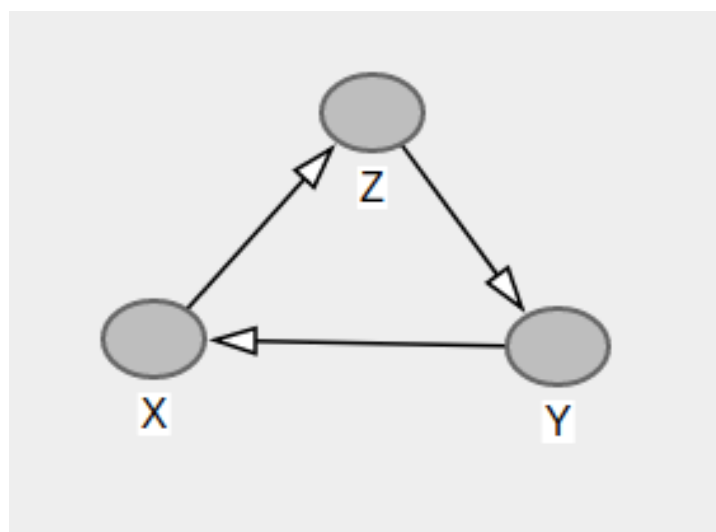
Grafos são estruturas matemáticas utilizadas para modelar relações. São como setas que utilizaremos para identificar causa e efeito. Sejam duas variáveis quaisquer X , representando democracia, e y , representando crescimento econômico. Se queremos afirmar que democracia causa crescimento econômico, podemos utilizar o diagrama:

Democracia $\rightarrow$ Crescimento Econômico

Podemos representar toda uma estrutura complexa usando esses grafos. Eles precisam apenas de duas restrições. A primeira delas é que as arestas nunca são bidirecionais. Em outras palavras, não é possível representar o ponto de que democracia esteja causando crescimento econômico e crescimento econômico esteja causando democracia desta forma:

Democracia $\leftarrow \rightarrow$ Crescimento Econômico

Por fim, os grafos não podem ter ciclos. Em outras palavras, é impossível que haja uma representação da seguinte forma:



Neste caso, X está causando Y, que está causando Z, que está causando X novamente. Há um ciclo que não será permitido dentro dos modelos causais que usaremos.

É importante também trazer a ideia de condicionar. Probabilidade condicionada tem um sentido preciso em estatística, mas para o escopo deste texto, vale dar uma definição mais simples. Falamos em condicionamento numa variável Z, quando estamos considerando alguns de seus valores em separado. Por exemplo, se Z é uma variável que representa calvície de um indivíduo, condicionar em Z significa verificar o que ocorre quando Z é calvo ou não. Se Z, entretanto, é uma variável quantitativa e não categórica – por exemplo, salário recebido – nesse caso, vamos considerar intervalos de salário.

Utilizaremos a ideia de condicionamento sobretudo para analisar quando existe de fato uma causa. Podemos falar em causa apenas quando todo caminho não direto entre duas variáveis é bloqueado ou por condicionamento ou por uma variável colísona⁸. Veremos algumas dessas situações abaixo.

Não pretendemos obviamente estender o tema, pois há trabalhos e trabalhos resolvendo problemas utilizando os grafos. Apenas descrevemos três situações comuns. Para uma abordagem completa, o leitor pode consultar Pearl (2000) e Spirtes et al. (2000), que são livros bastante extensivos sobre grafos e causalidade.

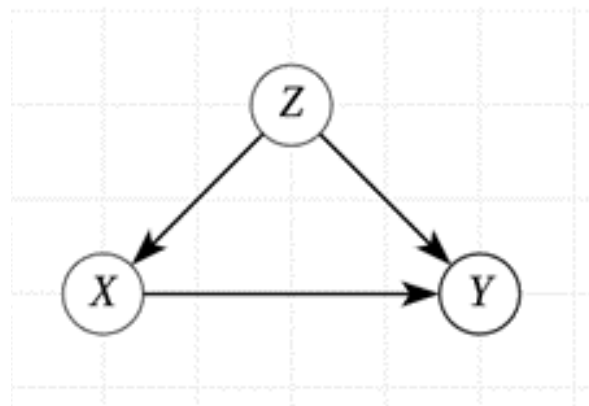
Confusores

Suponha o exemplo inicial de colesterol e exercício. A princípio, parece que fazer exercício está aumentando nosso colesterol. Por que isso ocorre? Na verdade, o que está ocorrendo é que a idade das pessoas entrevistadas está ao mesmo tempo impactando a quantidade de exercícios e a quantidade de colesterol no corpo. Assim, quando isso ocorre, estamos diante de variáveis confusoras (ou *confounding variables* em inglês). Esses tipos de exemplos são geralmente conhecidos na literatura como paradoxo de Simpson e são muito comuns.

Voltemos aos grafos. Confusores aparecem quando temos X causando Y, mas am-

8 Pode parecer confuso agora, mas veja mais exemplos a seguir. Interessados podem consultar Pearl, Glymour e Jewell (2016) pela ideia de d-separação. Na verdade, em relação às ideias exploradas aqui e nesse livro, os leitores podem encontrar métodos mais precisos em Pearl e Mackenzie (2018).

bas sendo causadas por Z. Por exemplo, o seguinte modelo:



O que fazer quando temos situações deste tipo então? Afinal, se existe uma confusora Z causando X e Y ao mesmo tempo, não basta olhar apenas para X e Y, pois aparecerá uma associação entre os dados. Nesse caso, basta condicionar nas variáveis confusoras. Isto é, no exemplo do colesterol e exercício, se olharmos a relação para cada grupo de idade, vamos constatar que a correlação é negativa e não positiva, como ocorria antes.

Um outro exemplo de que já tratamos. Vamos supor que queremos identificar o efeito de anos de estudo no rendimento futuro. O problema desse caso é que existe uma variável confusora clara, habilidade natural. Isto é, existe para cada indivíduo uma habilidade que está provavelmente causando mais anos de estudo e rendimento futuro e nesse caso é difícil condicionar nessa variável (uma vez que ela não é observada). Aqui também aparece a dificuldade de técnicas como matching, uma vez que elas dependem de confusoras observadas.

Um caso clássico. Muito conhecida historicamente é a correlação entre uso de tabaco e câncer. Por muito tempo, foi investigado se essa correlação seria causal ou não⁹. Ronald Fisher, um dos fundadores da estatística moderna, era um fumante inveterado. Ele aceitava que os dados mostravam uma associação, mas apenas porque haveria um terceiro fator – talvez de natureza genética – causando os dois ao mesmo tempo. Nesse caso, o que ele estava dizendo é que haveria uma variável confusora na qual não se poderia condicionar. Como se sabe, Fisher estava enganado.

Atenção, alegar haver fatores escondidos é a melhor forma para se destruir argumentos causais. É óbvio que nunca conseguiremos excluir todos os fatores escondidos que poderiam aparecer na realidade. Então é impossível se determinar causalidade?

Por honestidade intelectual, devemos sempre exigir do crítico que aponte qual a natureza do terceiro fator. Para defender o pesquisador, também há métodos de análise de sensibilidade que permitem avaliar qual deveria ser a força de um terceiro fator para anular a relação encontrada (por exemplo, CINELLI; HAZLETT, 2020).

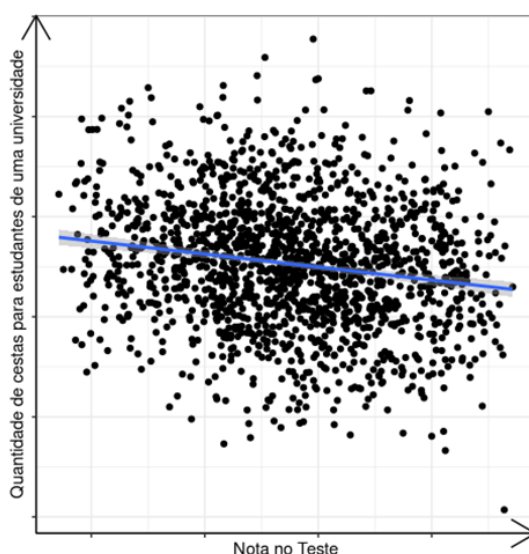
De qualquer forma, em muitos casos, a variável confusora é observada, e conseguimos fazer o condicionamento correto para extrair o efeito pesquisado. Porém, contrariamente ao que geralmente se faz em artigos, não se pode condicionar em quaisquer variáveis, mas sim naquelas que podem ser consideradas confusoras. Veremos casos em que condicionar em variáveis pode ser uma ideia ruim.

Colisores

Colisores são variáveis causadas ao mesmo tempo por X e Y. O problema de se condicionar em uma variável colidora é que isso acaba gerando uma relação de dependência espúria entre as duas.

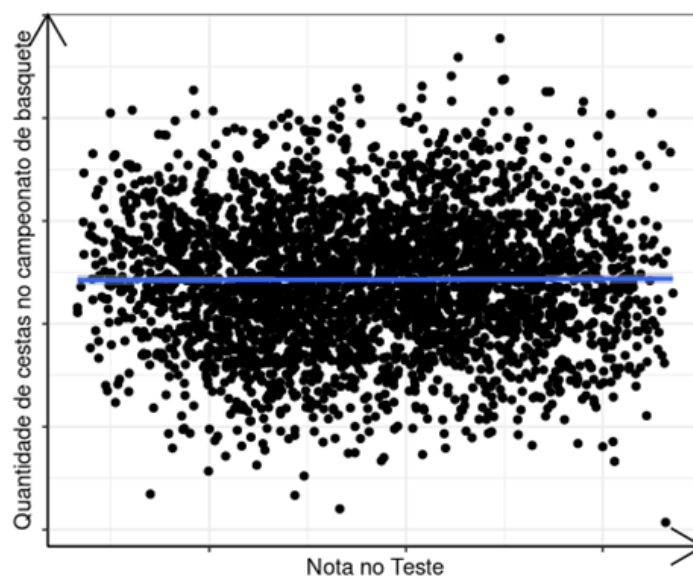
O exemplo clássico é o seguinte. Vamos supor que um pesquisador queira saber se existe relação causal entre habilidade intelectual e habilidade física. Para isso, ele obtém dados de uma determinada universidade, que concede bolsas para estudantes que vão muito bem em um determinado teste (X) ou que são grandes atletas (Y). A Figura 3 parece indicar que ser um atleta significa não ter um bom desempenho intelectual e vice-versa.

Figura 3 – Notas em um teste final x quantidade de cestas no campeonato de basquete



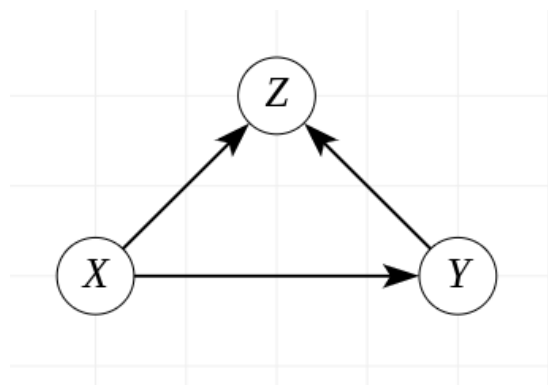
Na verdade, trata-se de um grande exemplo de viés por condicionar no collider. Esse viés também é bastante conhecido na literatura como paradoxo de Berkson. No exemplo anterior, tanto notas quanto desempenho no basquete estão causando a entrada na universidade. Assim, entrar ou não na universidade (Z) é uma variável colisor. Condicionar na colisor cria uma dependência estatística entre as outras duas variáveis, gerando viés. Quando colocamos todos os estudantes (não só os que passaram na universidade), a figura muda, a correlação se torna 0, como na Figura 4.

Figura 4 – Notas x cestas para amostra geral de estudantes



Há vários exemplos de problemas causados por condicionar em colisoras. Um bom texto abordando problemas em ciências sociais é o paper de Elwert e Winship (2014), que discute alguns exemplos. Um exemplo mais complexo é o correspondente ao Viés-M, abordado por Judea Pearl (2009) e que mostra os problemas de se condicionar em qualquer variável.

Representamos Z como um colisor de X e Y no seguinte DAG.



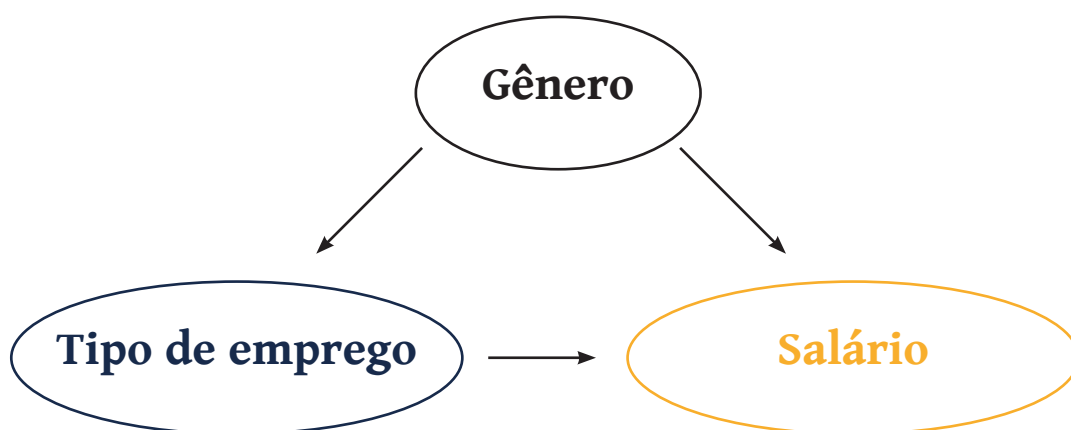
Mediadores

Mediadores geralmente aparecem quando a variável Z está entre o caminho causal de X e Y, como no grafo seguinte:

$$X \rightarrow Z \rightarrow Y$$

Nesse caso, ao se condicionar em Z, acabamos por transformar X e Y em independentes, o que é absurdo, uma vez que, pelo nosso grafo, X está causando Y. Geralmente conhecem esse viés como regressão Everest, termo cunhado pelo economista Garrett Jones: controlando pela altitude, a temperatura no Everest é amena.

É um caso comum de estudos de gênero. Muitos argumentam que, condicionando no tipo de emprego, muito do efeito das diferenças salariais entre homens e mulheres some. Isso é óbvio, uma vez que gender gap pode ser descrito seguindo o seguinte grafo:



X aqui representa gênero; Y, salário; e Z, tipo de emprego. É óbvio aqui que condicionando em Z, remove-se parte da dependência entre X e Y, o que não elimina o problema.

Dados observacionais

Todos os dados que não derivam de intervenção direta, como no caso dos experimentos controlados, são dados observacionais, isto é, apenas podemos observá-los. É possível responder questões intervencionais e contrafactuais aqui? Sim, desde que façamos certas pressuposições.

Experimentos naturais

Os experimentos naturais partem da mesma ideia que os experimentos controlados, mas há uma diferença fundamental: a divisão dos grupos não é feita por intervenção direta do pesquisador, mas sim realizada pela natureza. A partir disso, os cientistas tentam explorar situações em que surgem grupos de tratamento e controle para tentar identificar causalidade.

O exemplo clássico de experimento natural é a história de como se descobriu que cólera era causada por água contaminada¹⁰. John Snow era um médico inglês do século 19. Anteriormente, dizia-se que a cólera era trazida pelo ar ou por miasma. Snow tentou derrubar essa teoria e para isso utilizou o que chamaríamos hoje de experimento natural. Ele catalogou taxas de morte por cólera em locais servidos por duas companhias de água, Southwark and Vauxhall Company e Lambeth Company. As duas obtinham água de regiões sujas do Tâmesa. Em 1852, a segunda companhia passou a coletar sua água em outra região, pouco contaminada. Então John Snow comparou a mudança na mortalidade e pôde constatar que possivelmente a diferença de abastecimento de água estava causando aquelas mortes.

O resumo do caso da cólera mostra como a separação de grupos não gerada pela intervenção do pesquisador pode trazer resultados positivos em relação à inferência causal. A ideia aqui é que naturalmente não existem confusores ou esses confusores podem ser controlados de alguma forma. Há vários métodos de estimação usados por pesquisadores em relação aos experimentos naturais que alternam algumas dessas suposições. Veremos de maneira resumida o que cada um faz.

Diferenças em diferenças

A técnica utilizada por Snow é um exemplo do que os cientistas chamam de diferenças em diferenças. A ideia é de que casos semelhantes tendem a se comportar de maneira similar no tempo. Esta pressuposição é fundamental. Também é necessário que confusores afetem as variáveis de maneira aditiva. O procedimento é o seguinte. Se temos um conjunto de dados, em que uma causa X afetou apenas parte deles, consideramos a parte afetada o grupo de tratamento, e a não afetada, o grupo de controle. Para o grupo de tratamento, vemos qual a diferença na variável a ser ex-

10 História contada em Angrist e Pischke (2008, p. 227) e Johnson (2008).

plicada, antes e após a intervenção. Em seguida, também verificamos essa diferença no grupo de controle. Por fim, estimamos o efeito causal como a diferença entre essas diferenças.

Alguns exemplos nas ciências sociais mostram como avaliar impactos de programas usando a técnica. Por exemplo, Firpo e De Pieri (2012) usaram diferenças em diferenças para avaliar os efeitos da introdução de computadores em escolas públicas brasileiras. Eles verificaram diferenças na escolaridade média dos professores (para saber se as escolas atraíram professores mais qualificados) e na existência de outros laboratórios. O grupo de tratamento é composto por escolas públicas que foram admitidas no programa, e o grupo de controle, por escolas similares, mas não admitidas. O efeito foi próximo a 0 para a escolaridade dos professores, mas positivo para a existência de laboratórios.

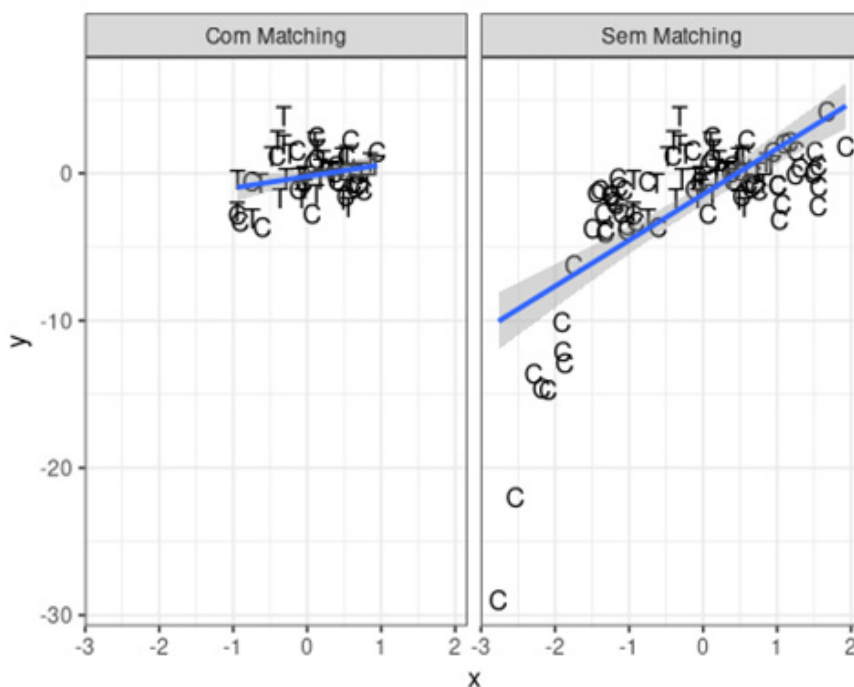
Outro exemplo interessante é o trabalho de Neri e Soares (2007), que avaliou o impacto de um programa de transferência de renda no estado de saúde de idosos. Os autores compararam grupos de idosos, beneficiários e não de programas, antes e depois da intervenção. Chegaram à conclusão que o programa trouxe efeitos positivos em relação à saúde dos idosos beneficiários.

Matching

Matching é também uma técnica bastante utilizada. O princípio dessa técnica é agrupar cada elemento do grupo de tratamento com os membros do grupo de controle mais similares. Com isso, obtemos o efeito médio do tratamento nos tratados (ETT)¹¹.

Na Figura 5, por exemplo, mostramos dados simulados antes e pós matching. Anteriormente ao matching, uma regressão de x em y indica um efeito positivo. Após o procedimento, porém, o efeito passa a ser nulo. Dessa forma, a técnica permite que se removam possíveis vieses decorrentes do desbalanceamento dos dados.

Figura 5 - Dados com e sem procedimento de matching¹²



Fonte: elaboração própria a partir de Ho et al. (2007).

Para realizar o procedimento, precisamos de uma técnica que indique se dois casos são similares ou não. Há muitas técnicas para isso, uma das mais utilizadas é o propensity scores. Trata-se do cálculo da probabilidade de se pertencer ao grupo de

11 Nada impede de se fazer o mesmo para o grupo de controle. No caso, obteríamos o efeito médio para os controles (ETC).

12 Gráfico baseado em Ho et al. (2007).

tratamento. Esse cálculo é feito com base em outras variáveis além das de causa e efeito (covariáveis).

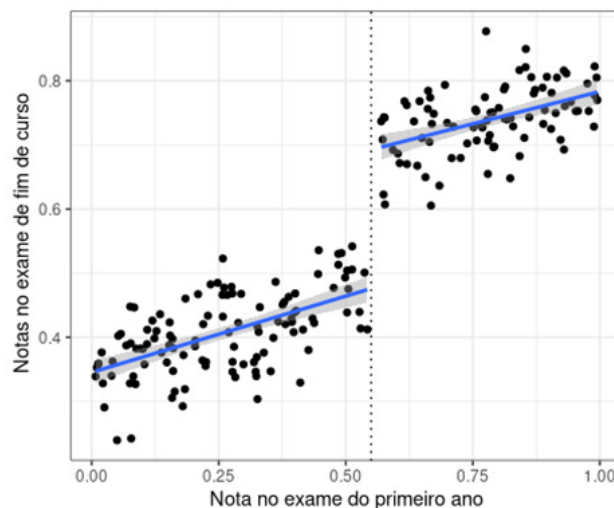
Há uma premissa importante do matching. Precisamos que variáveis a serem condicionadas (confusoras; *confounders*, em inglês) sejam observadas. Suponhamos que gostaríamos de investigar o efeito de estudar em uma universidade sobre os rendimentos futuros. O problema nesse caso é que há uma possível variável confusora (habilidade natural) que não é medida e que provavelmente está causando tanto a presença na universidade quanto os rendimentos. Nesse caso, usar matching não possibilita remover o viés. Porém, em muitas aplicações, o uso de matching é adequado.

Há exemplos de análises que empregam matching para corrigir desbalanceamentos. Zucco Jr (2013), por exemplo, analisou o impacto eleitoral de programas de transferência de renda, como o Bolsa Família. Para isso, o autor usou propensity scores para agregar municípios de acordo com variáveis controle de nível municipal ou estadual, como resultados eleitorais em anos anteriores, o crescimento do PIB per capita no ano anterior às eleições, entre outras. Por fim, foram regredidos a porcentagem de voto no candidato incumbente sobre o valor do gasto para os programas em cada município. O autor mostrou que os gastos corresponderam a um significativo aumento de votos para três anos: 2002, 2006 e 2010. Em seguida, ele também utilizou dados individuais de survey para recebedores ou não de programa, para mostrar que havia um efeito individual para voto no incumbente.

Regressão descontínua

Regressão descontínua é uma técnica bastante utilizada quando queremos medir o efeito de uma intervenção que ocorre após uma escala atingir certo valor de corte. Exemplos podem ilustrar o ponto. Numa determinada universidade, alunos obtiveram uma determinada bolsa, após atingir 55 pontos (de 100) num exame no primeiro ano¹³. Essa bolsa permite que alunos se dediquem mais e tenham acesso a mais recursos que os outros estudantes. Suponhamos que um pesquisador queira medir o efeito da bolsa em relação às notas no final de todo o curso. Ele poderia verificar como a relação entre notas se alteraria dado o corte.

Figura 6 – Notas no exame final vs rendimento escolar no último ano



Como podemos ver na Figura 6, a regressão descontínua permite que calculemos o efeito ao comparar a diferença entre a regressão estimada para valores inferiores ao valor de corte em relação aos valores acima dele.

Essa técnica vem sendo muito utilizada nos últimos anos. Cepaluni e Hidalgo (2016), por exemplo, utilizaram a descontinuidade que existe na obrigatoriedade do voto no Brasil (limite de idade: maior de 17 anos e menor de 70) para estimar o efeito negativo do voto compulsório em relação à desigualdade de participação entre ricos e pobres. Bruce e Lima (2019), usando a mesma estratégia, mediram o efeito positivo da obrigatoriedade do voto em relação à probabilidade de se informar pela televisão. Brambor e Ceneviva (2012) empregaram a descontinuidade que existe em eleições majoritárias (se a margem de voto for maior que 0%) para estimar o efeito negativo de incumbência sobre o desempenho eleitoral futuro.

13 Exemplo adaptado de Angrist e Pischke (2008).

Variável instrumental

Vamos supor que desejamos analisar o impacto de crescimento econômico em relação à incidência de guerra civil. Essa é uma questão de inferência causal bastante difícil, porque é bastante difícil separar causalidade reversa ou outros terceiros fatores causando os dois ao mesmo tempo. Quando isso ocorre, é possível encontrar variáveis causando o efeito estudado, apenas através da causa estudada. Justamente foi essa situação enfrentada por Miguel, Satyanath e Sergenti (2004). Eles obtiveram dados relativos à incidência chuva, que tem impacto causal sobre crescimento econômico. O argumento é que não deve haver outra forma pela qual a incidência de chuva esteja afetando guerra civil a não ser por meio do crescimento econômico. Uma variável que atenda essa condição é conhecida na literatura como instrumento¹⁴. Uma pressuposição chave aqui é a linearidade.

Outro uso comum de variáveis instrumentais é no caso de experimentos com compliance imperfeita. Por exemplo, num experimento com pílulas, isso ocorre, quando parte do grupo do tratamento não cumpre as determinações e deixa de tomar o remédio. Ou então quando parte do grupo de controle passa a tomar o remédio independentemente da intervenção. Nesse caso, a variável instrumental de atribuição ao grupo pode ser utilizada pra avaliar o efeito do remédio.

Controle sintético

Controle sintético é um método de estimação bastante utilizado que combina outras técnicas, como matching e diferenças em diferenças, para gerar um grupo de controle dos dados existentes. A ideia é que misturando características de outros casos, consigamos criar um caso de controle que seja comparável a um prévio caso de tratamento. Em seguida, analisa-se como esses casos estão se comportando no tempo.

Há dois bons exemplos na literatura que estuda Brasil. Entre outros, Freire (2018) utilizou controle sintético para avaliar o efeito de medidas de prevenção à violência no Estado de São Paulo. Para isso, ele construiu um Estado de São Paulo sintético com dados de todos os outros estados e Distrito Federal. Com isso, ele conseguiu estimar o efeito das políticas de segurança do estado em relação ao índice de homicídios.

14 Na verdade, há versões generalizadas de variáveis instrumentais, muito mais complexas que o exemplo dado, mas que fogem ao escopo do capítulo. Ver Brito e Pearl (2002).

Considerações finais

Neste texto, tentamos introduzir o tema de inferência causal e como extrair conclusões de causa e efeito dos dados. O objetivo, é claro, não é exaurir o tema, mas apenas apresentar os principais resultados para o leitor. Não é um tema novo, mas também não é antigo. Muitos dos resultados de inferência causal surgiram apenas após a década de 1980 e foram popularizados apenas a partir de 2000. Espera-se, assim, que pelo menos duas conclusões fiquem presentes. A primeira é de que causalidade é algo diferente de estatística e exige uma linguagem nova. A segunda é que há técnicas boas e framework adequado para apresentar resultados causais e também para criticá-los.

Referências bibliográficas

- ANGRIST, Joshua D.; PISCHKE, Jörn-Steffen. Mostly harmless econometrics: an empiricist's companion. Princeton: Princeton University Press, 2008.
- _____. Mastering'metrics: the path from cause to effect. Princeton: Princeton University Press, 2014.
- ARMITAGE, Peter. Fisher, Bradford Hill, and randomization. *International Journal of Epidemiology*, v. 32, n. 6, p. 925-28, 2003.
- BAREINBOIM, Elias; PEARL, Judea. Causal inference and the data-fusion problem. *Proceedings of the National Academy of Sciences*, v. 113, n. 27, p. 7345-52, 2016..
- BRAMBOR, Thomas; CENEVIVA, Ricardo. Reeleição e continuísmo nos municípios brasileiros. *Novos Estudos CEBRAP*, n. 93, p. 09-21, 2012.
- BRITO, Carlos; PEARL, Judea. Generalized instrumental variables. In: *Proceedings of the Eighteenth Conference on Uncertainty in Artificial Intelligence*. Burlington: Morgan Kaufmann Publishers Inc, 2002. p. 85-93
- BRUCE, Raphael; LIMA, Rafael Costa. Compulsory voting and Tv news consumption. *Journal of Development Economics*, v. 138, p. 165-79, 2019.
- CEPALUNI, Gabriel; HIDALGO, F. Daniel. Compulsory voting can increase political inequality: evidence from Brazil. *Political Analysis*, v. 24, n. 2, p. 273-80, 2016.
- CINELLI, Carlos; HAZLETT, Chad. Making sense of sensitivity: Extending omitted variable bias. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, v. 82, n. 1, p. 39-67, 2020.
- ELWERT, Felix; WINSHIP, Christopher. Endogenous selection bias: the problem of conditioning on a collider variable. *Annual Review of Sociology*, v. 40, p. 31-53, 2014.
- FIRPO, Sergio Pinheiro; DE PIERI, Renan Gomes. Avaliando os efeitos da introdução de computadores em escolas públicas brasileiras. *Revista Brasileira de Inovação*, v. 11, p. 153-90, 2012.
- FREIRE, Danilo. Evaluating the effect of homicide prevention strategies in São Paulo, Brazil: a synthetic control approach. *Latin American Research Review*, v. 53, n. 2, p. 231-249, 2018.
- GERTLER, Paul. Do conditional cash transfers improve child health? Evidence from Progres's control randomized experiment. *American Economic Review*, v. 94 n. 2, p. 336-41, 2004.
- HO, Daniel E.; IMAI, Kosuke; KING, Gary; STUART, Elizabeth A. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis*, v. 15, n. 3, p. 199-236, 2007.
- HOLLAND, Paul W. Statistics and causal inference. *Journal of the American Statistical Association*, v. 81, n. 396, p. 945-60, 1986.
- JOHNSON, Steven. O mapa fantasma. Rio de Janeiro: Zahar, 2008.

- MIGUEL, Edward; SATYANATH, Shanker; SERGENTI, Ernest. Economic shocks and civil conflict: an instrumental variables approach. *Journal of Political Economy*, v. 112, n. 4, p. 725-53, 2004.
- NERI, Marcelo C.; SOARES, Wagner Lopes. Estimando o impacto da renda na saúde através de programas de transferência de renda aos idosos de baixa renda No Bra-sil. *Cadernos de Saúde Pública*, v. 23, p. 1845-56, 2007.
- PEARL, Judea. *Causality: models, reasoning and inference*. Cambridge: Cambridge University Press, 2000.
- _____. Remarks on the method of propensity score. *Statistics in medicine*, v. 28, n. 9, p. 1415, 2009.
- PEARL, Judea; GLYMOUR, Madelyn; JEWELL, Nicholas P. *Causal inference in statistics: a primer*. Hoboken: John Wiley & Sons, 2016.
- PEARL, Judea. The seven tools of causal inference, with reflections on machine learning. *Communications of the ACM*, v. 62, n. 3, p. 54-60, 2019.
- PEARL, Judea; MACKENZIE, Dana. *The book of why: the new science of cause and effect*. New York: Basic Books, 2018.
- SHPITSER, Ilya; PEARL, Judea. Identification of joint interventional distributions in recursive semi-Markovian causal models. (Proceedings of the National Conference on Artificial Intelligence, v. 21, n. 2). Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2006.
- SPIRTES, Peter; GLYMOUR, Clark N.; SCHEINES, Richard; HECKERMAN, David; MEEK, Christopher; COOPER, Gregory; Richardson, Thomas. *Causation, prediction, and search*. Cambridge: MIT press, 2000.
- ZUCCO JR., Cesar. When payouts pay off: conditional cash transfers and voting behavior in Brazil 2002–10. *American Journal of Political Science*, v. 57, n. 4, p. 810-22, 2013.

GAP

POBREZA E DESIGUALDADE

Pedro H. G. Ferreira de Souza¹

02

¹ Técnico de Planejamento e Pesquisa do Ipea. Email: pedro.ferreira@ipea.gov.br. O autor agradece os comentários e sugestões de Ana Paula Soares Carvalho.

Introdução

O objetivo deste capítulo é apresentar os conceitos, medidas e fontes de dados mais importantes nas análises distributivas brasileiras, bem como esclarecer dúvidas recorrentes e sugerir boas práticas para auxiliar leitores e pesquisadores iniciantes. A intenção não é esgotar o assunto, e sim oferecer uma introdução acessível ao tema. Dessa forma, passo ao largo dos intermináveis debates filosóficos e disputas conceituais sobre pobreza e desigualdade, ao mesmo tempo em que dispenso, sempre que possível, formalizações matemáticas que assustam leitores leigos.

Embora diversos resultados empíricos sejam mencionados ao longo do texto, meu foco é essencialmente metodológico. Portanto, também evito entrar em controvérsias sobre resultados substantivos e/ou explicações.

“Pobreza” e “desigualdade” só ganham conteúdo quando se referem a uma distribuição específica, no sentido descritivo ou estatístico do termo, o que implica definir o atributo (distribuição do quê?) e a unidade e o universo de análise (entre quem?). Para simplificar, privilegio aqui a distribuição de renda entre pessoas, ainda que boa parte da discussão também se aplique a outras dimensões. Como veremos adiante, essa definição ainda é muito imprecisa, mas serve com ponto de partida.

A distribuição de renda

Qual distribuição?

Qualquer investigação sobre pobreza ou desigualdade precisa, antes de tudo, definir qual seu objeto de interesse, isto é, qual distribuição de renda está sendo analisada. Rendimentos são fluxos recebidos ao longo do tempo de fontes muito heterogêneas e com periodicidade distinta, que podem ser compartilhados ou não entre os membros de uma família. Três definições são fundamentais:

- a) Qual o **conceito de renda** utilizado?
- b) Qual a **unidade de análise**?
- c) Qual o **universo de referência**?

Idealmente, todo pesquisador deveria partir de uma definição prévia do conceito de renda para em seguida buscar os dados mais adequados. Na prática, toda fonte de informações implícita ou explicitamente define a renda de modo mais ou menos restrito para tornar a coleta viável, e os pesquisadores acabam adaptando-se à disponibilidade de dados. De um modo ou de outro, a definição de um determinado conceito de renda envolve decisões em dimensões como:

i. **Forma dos rendimentos:** somente rendimentos monetários ou também incluímos rendimentos não monetários?

ii. **Período de referência:** qual a data de referência para recebimento dos fluxos de renda?

iii. **Tributação:** rendimentos brutos ou líquidos, isto é, antes ou depois do pagamento de tributos?

iv. **Composição dos rendimentos:** quais as fontes de rendimentos contempladas?

As escolhas da unidade de análise e do universo de referência costumam andar juntas. Este capítulo prioriza a distribuição de renda entre pessoas, mas, ainda assim, os detalhes fazem toda a diferença. Por exemplo, podemos considerar apenas rendimentos individuais ou levar em conta o papel redistributivo das famílias, do mesmo modo que podemos ter como universo de referência a população como um todo ou apenas algum subgrupo, como a população adulta, os trabalhadores ocupados, entre

outros.

O papel redistributivo das famílias, em particular, merece atenção. Estudos sobre a distribuição de rendimentos individuais pressupõem que não há qualquer efeito das famílias, enquanto estudos que usam a renda domiciliar per capita pressupõem que os rendimentos são divididos igualmente entre os moradores e que não há economias de escala relacionadas ao tamanho das famílias. Por sua vez, análises que empregam escalas de equivalência – pouco comuns no Brasil – também pressupõem que a renda é compartilhada, mas tentam levar em conta essas economias de escala, ou seja, o fato de que há despesas fixas e outras que não crescem linearmente conforme aumenta o tamanho das famílias.

Dependendo dos objetivos, todas essas escolhas são razoáveis, ainda que, em última instância, não sejam completamente realistas. Como sói acontecer, são simplificações úteis dada a impossibilidade de coletar informações adicionais sobre a forma como cada família combina e reparte os rendimentos de seus membros. Definições diferentes podem gerar resultados distintos e até contraditórios. Entender essas diferenças ajuda a desfazer muitas confusões.

Desafios na captação da renda

Rendimentos são tabu para a maioria das pessoas, não só no Brasil, mas em todo mundo. Poucos falam abertamente sobre o tema com amigos e parentes, quanto mais com desconhecidos. No entanto, a própria existência de análises distributivas exige que essas informações sejam obtidas de alguma forma. Grosso modo, podemos dividir as fontes de dados em dois grandes grupos – levantamentos amostrais e registros administrativos –, que possuem vantagens e desvantagens específicas.

Levantamentos amostrais

As vantagens dessas fontes de dados são óbvias. Em primeiro lugar, essas pesquisas domiciliares são relativamente baratas, por serem amostrais e de fácil acesso, já que preservam o anonimato dos respondentes. Além disso, a coleta simultânea de informações sobre temas como demografia, educação, migração, mercado de trabalho e afins permite relacionar rendimentos a diversas outras características de indivíduos e famílias. Infelizmente, pesquisas amostrais também possuem desvantagens intrínsecas potencialmente graves para análises distributivas, como:

1. **Erros amostrais:** a maioria dos levantamentos têm amostras de tamanho insatisfatório para a análise de grupos com características raras. Indivíduos muito ricos, em especial, quase nunca são sorteados para entrevistas.

2. **Limitações do questionário:** como questionários longos aumentam os custos e prejudicam a qualidade das informações, o rol de rendimentos pesquisados costuma ser restrito às principais fontes regulares (salários, aposentadorias etc.) apenas em um único mês de referência.

3. **Taxas diferenciais de não participação:** pesquisas domiciliares mundo afora quase sempre registram maiores taxas de não participação – isto é, recusa ou indisponibilidade para ser entrevistado – entre os mais ricos, seja por dificuldades de acesso às moradias, preocupações com segurança, falta de tempo ou outros motivos.

4. **Taxas diferenciais de não resposta:** quando de fato concordam em responder às entrevistas, os mais ricos tendem a apresentar maiores taxas de não resposta às perguntas sobre rendimentos, seja por preocupações com segurança ou privacidade, constrangimento ou mesmo desconhecimento total ou parcial da própria renda.

5. **Taxas diferenciais de subdeclaração:** mesmo quando respondem, os mais ricos tendem a subestimar ou subdeclarar seus rendimentos, de forma intencional ou não, pelos motivos citados nos pontos anteriores. Embora a subdeclaração seja particularmente forte no topo da distribuição de renda, ela também afeta outros estratos sociais. Rendimentos informais, irregulares ou de cálculo complexo (como rendimentos de capital, de trabalhadores por conta própria e/ou de empregadores) são propensos à subdeclaração, mas nem mesmo rendimentos formais regulares estão isentos de problemas, já que nem todos informantes entendem a distinção entre rendimentos líquidos e brutos. Por fim, quase sempre transferências decorrentes de programas sociais são subdeclaradas, seja por confusão dos informantes, estigma social ou medo de perder o benefício.

Essas limitações prejudicam a captação dos rendimentos ao longo de toda a distribuição de renda, mas, em geral, atingem de forma mais forte os mais ricos. Dependendo do desenho amostral, questionário e outros fatores, alguns surveys são mais afetados, outros são menos, ainda que nenhum esteja completamente imune.

Parte dessas questões foi identificada já nos anos 1930 e a discussão sobre o assunto foi incorporada há tempos nos manuais internacionais. A novidade é que em muitos países a qualidade das informações coletadas parece estar em queda. Nos Esta-

dos Unidos, em especial, as taxas de não participação vêm aumentando rapidamente, bem como a subdeclaração de benefícios sociais, entre outros problemas.

Registros administrativos

Registros administrativos são bases de dados operadas pelos governos para alinhar o funcionamento de políticas públicas e do sistema tributário. Antes do primado das pesquisas amostrais, os estudos pioneiros sobre rendimentos dependiam deste tipo de informação. Depois de um longo hiato, seu uso voltou a ser frequente devido à disponibilidade de capacidade computacional para o processamento de volumes gigantescos de dados e à crescente consciência das limitações intrínsecas às pesquisas domiciliares.

Em particular, as informações reportadas nas declarações anuais de imposto de renda têm sido cada vez mais usadas para o estudo da distribuição de renda. Em relação às pesquisas domiciliares amostrais, os dados tributários se destacam sobretudo pela melhor captação dos rendimentos dos mais ricos. O caráter compulsório do imposto de renda, as sanções legais associadas, o escopo de rendimentos coletados e o período de referência anual reduzem tanto os problemas de participação quanto os de declaração. Outros registros administrativos, por sua vez, são atraentes porque contêm informações impossíveis de serem coletadas em pesquisas domiciliares – por exemplo, informações sobre empresas, exames educacionais e assim por diante.

Evidentemente, nem tudo são flores. Informações administrativas também possuem desvantagens, como:

1. **Dificuldade de acesso aos dados:** nem sempre os órgãos responsáveis concedem acesso a microdados; quando o fazem, o processo é burocrático, pois envolve no mínimo um requerimento formal e a assinatura de termo de responsabilidade pelo sigilo das informações.
2. **Cobertura populacional restrita:** registros administrativos não são desenhados para representar a população como um todo, mas sim o público-alvo de cada programa.
3. **Menor abrangência de informações:** bases administrativas costumam ser muito detalhadas sobre temas específicos, mas não se comparam em abrangência às pesquisas domiciliares.

4. **Evasão, elisão e comportamento estratégico:** mesmo nos casos em que há sanções legais contra o fornecimento de informações falsas, como no caso do IRPF, nunca é possível eliminar totalmente a evasão (ilegal), a elisão (legal) de rendimentos e outros tipos de comportamento estratégico. Em casos de difícil verificação (como para rendimentos informais), pode haver até incentivos para a subdeclaração de rendimentos.

5. **Mudanças legais, operacionais e na fiscalização:** a qualidade dos dados de registros administrativos pode flutuar ao longo do tempo em função de mudanças na legislação, nos sistemas de gestão e na fiscalização.

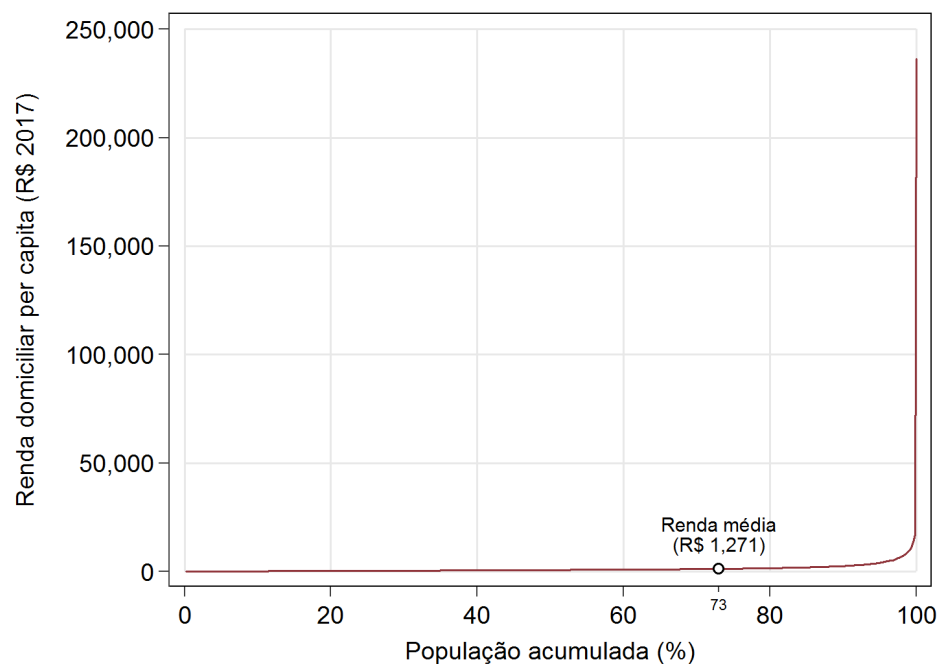
Representações visuais

Parada de Pen

A Parada de Pen, também conhecida como Parada de Anões, é apenas uma curva de quantis. O nome pitoresco se decorre da metáfora do economista Jan Pen, que, para tornar o conceito de “distribuição de renda” intuitivo, sugeriu aos leitores imaginar um desfile ordenado por renda, envolvendo toda a população, do mais pobre ao mais rico, em um mundo em que a altura das pessoas é proporcional à sua renda. Pen então descreveu o que o público do desfile veria: uma procissão com uma longa fila de anões seguida por alguns indivíduos de estatura relativamente normal e, finalmente, por um punhado de gigantes tão altos que ninguém seria capaz de enxergar além dos seus calcanhares.

Para ilustrar, a Figura 1, na página seguinte, constrói a Parada de Pen, ou curva de quantis, para os rendimentos domiciliares per capita reportados pela Pesquisa Nacional por Amostra de Domicílios Contínua (PNADC) para 2017. O gráfico destaca a presença de rendimentos extremamente elevados – os gigantes de Jan Pen. A curva é totalmente achatada para mais de 90% da população, sendo que quase três quartos está abaixo da renda média (R\$ 1,271 na PNADC 2017). Não se trata de um artefato dos dados. Com efeito, a PNADC, assim como outras pesquisas domiciliares, subestima os rendimentos dos mais ricos e a concentração de renda no topo.

Figura 1 – Parada de Pen da renda domiciliar per capita no Brasil, 2017



Fonte: Pesquisa Nacional por Amostra de Domicílios Contínua (PNADC) 2017. Nota: rendimentos brutos efetivos recebidos por moradores (exceto pensionistas, empregados domésticos e afins), registrados nas variáveis vd4020 e vd4048. Ver seção 5 para mais informações sobre a PNADC.

Para fins analíticos, a Parada de Pen é mais usada para comparar duas distribuições diferentes, como, por exemplo, um mesmo país em dois momentos do tempo ou dois países distintos no mesmo momento. Nesse caso, costuma ser mais fácil visualizar o gráfico da diferença entre as duas curvas: com isso, o eixo vertical passa a representar a diferença absoluta de renda entre as distribuições em cada percentil p da população, podendo assumir valores positivos ou negativos. Naturalmente, a comparação entre duas distribuições só faz sentido se os rendimentos estiverem expressos na mesma unidade monetária e com a mesma data de referência.

Por representar níveis absolutos de renda, a curva de quantis – bem como sua inversa, a função de distribuição acumulada – pode ser usada para retratar tanto a pobreza em relação a um patamar absoluto como a dispersão ou desigualdade dos rendimentos. No primeiro caso, para melhorar a visualização, costuma-se truncar o eixo horizontal, de modo a considerar apenas, por exemplo, os 20% ou 40% mais pobres, em vez da população inteira.

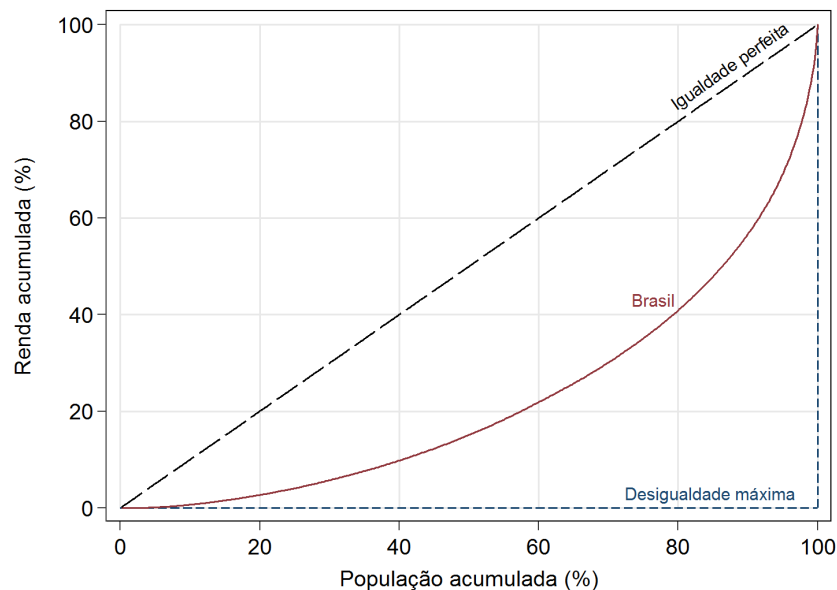
Curva de Lorenz

Outra representação visual comum das distribuições de renda é a chamada Curva de Lorenz, proposta em 1905 pelo economista americano Max Lorenz. A curva de Lorenz também ordena toda a população de acordo com a renda: mais uma vez, o eixo horizontal representa a população acumulada, começando do zero, pelos mais pobres, até chegar aos mais ricos entre os ricos (100% da população). A diferença em relação à curva dos quantis está no eixo vertical, que agora representa a renda relativa acumulada até determinado percentil p , e não mais os rendimentos absolutos de cada percentil. Em outras palavras, cada ponto da curva de Lorenz mostra qual o percentual da renda total (eixo vertical) recebida pela população acumulada até aquele ponto (eixo horizontal).

Como sugeriu o economista Frank Cowell, a curva de Lorenz pode ser compreendida a partir de uma metáfora parecida com a da Parada de Anões de Jan Pen. Suponha que toda a renda de uma sociedade foi usada para preparar um bolo gigante, e que chegou a hora de distribuir as fatias. Mais uma vez, formamos uma fila com toda a população ordenada dos mais pobres até os mais ricos, e cada um recebe uma fatia proporcional à sua contribuição financeira para o bolo. De início, são servidos apenas farelos, mas, conforme chegamos aos mais ricos, as fatias vão se tornando gigantes. A cada fatia entregue, nós anotamos o percentual da população que já foi servida e o percentual do bolo que já foi consumido. A curva de Lorenz é a representação gráfica dessas anotações.

A Figura 2 apresenta a curva de Lorenz para o Brasil em 2017, calculada a partir dos mesmos dados utilizados para a Figura 1. A linha tracejada diagonal em preto é a chamada “reta da igualdade perfeita”, isto é, a curva de Lorenz em um cenário extremo em que todos os indivíduos têm exatamente a mesma renda per capita. Dessa forma, os 20% mais pobres teriam 20% da renda total, os 40% mais pobres teriam 40% da renda, e assim por diante. A reta da igualdade perfeita quase sempre é incluída em gráficos da curva de Lorenz porque serve como referência visual: quanto mais próxima dela estiver a distribuição empírica, menor a desigualdade.

Figura 2 – Curva de Lorenz da renda domiciliar per capita no Brasil, 2017



Fonte: Pesquisa Nacional por Amostra de Domicílios Contínua (PNADC) 2017. Nota: rendimentos brutos efetivos recebidos por moradores (exceto pensionistas, empregados domésticos e afins), registrados nas variáveis vd4020 e vd4048. Ver seção 5 para mais informações sobre a PNADC. A linha pontilhada indica a reta de igualdade perfeita.

A curva azul escura em “L” representa o polo oposto da reta da igualdade perfeita, correspondendo à maior desigualdade possível, quando apenas um único indivíduo detém toda a renda em uma população numerosa.

Evidentemente, os extremos nunca são vistos na prática. Em situações normais, as curvas de Lorenz reais estão sempre entre esses dois limites, como ocorre na Figura 2. Nela, vemos que os 20% mais pobres ficam com menos de 3% da renda total. Quando avançamos para o percentil 60, que demarca os 60% mais pobres, a renda acumulada é pouco maior do que 20%. No percentil 80, esse percentual está em 40%, o que significa que os 20% mais ricos detêm os outros 60% da renda total.

Por representar apenas frações acumuladas da população e da renda total, a curva de Lorenz é indiferente ao nível absoluto de renda em uma dada distribuição, ao contrário da curva de quantis. Em outras palavras, a curva de Lorenz não serve para o estudo da pobreza, mas para investigações sobre desigualdade. Por não levar em conta os valores absolutos, a curva de Lorenz permite a comparação direta entre distribuições mesmo que a unidade monetária e/ou o período de referência sejam distintos, sem necessidade de conversão e/ou deflacionamento.

Pobreza

Definições

Pobreza pode carregar dezenas de definições conceituais diferentes. Para ficar apenas nos exemplos óbvios: a pobreza deve ser entendida como privação de direitos, de capacidades, de recursos econômicos ou em qual outra dimensão? Essa privação é absoluta, ligada a padrões mínimos de consumo, ou relativa, já que esses padrões podem variar entre sociedades? Ou pobreza é algo puramente subjetivo?

É seguro afirmar que jamais haverá consenso acerca das respostas certas. Por isso, as pesquisas empíricas costumam passar ao largo desses temas, privilegiando definições operacionais com conteúdo preciso, escolhido em função do contexto e da disponibilidade de dados.

No Brasil, a produção acadêmica recente define operacionalmente a pobreza como insuficiência de renda monetária em relação a uma linha absoluta de pobreza, expressa em termos de renda domiciliar per capita mensal bruta (isto é, antes do pagamento de tributos). Essa escolha não significa que os pesquisadores brasileiros considerem que a pobreza é apenas e tão somente falta de renda monetária, ou seja, que ela pode ser reduzida à sua definição operacional. Pelo contrário, a maioria não teria problemas em reconhecer que, no plano conceitual, a pobreza de fato é melhor compreendida como fenômeno multidimensional.

A opção por tratar a pobreza como insuficiência de renda se justifica por dois motivos. Em primeiro lugar, independentemente das definições conceituais adotadas, as operacionalizações da pobreza como um fenômeno multidimensional costumam incluir um rol de dimensões – acesso a saneamento básico, educação, saúde etc. – que, em última instância, são altamente correlacionadas com a renda.

Em segundo lugar, análises verdadeiramente multidimensionais da pobreza sofrem com dificuldades técnicas que não afetam análises baseadas na renda. A escolha e comparação de variáveis heterogêneas não é simples, exigindo procedimentos arbitrários de normalização e ponderação. A agregação dos resultados é mais complicada ainda, porque ou impõe o achatamento da multidimensionalidade em um único indicador sintético cujo valor absoluto não possui nenhum sentido intrínseco ou, no outro extremo, limita-se à apresentação de dezenas de indicadores que inevitavelmente serão analisados separadamente.

Linhas de pobreza

A definição operacional da pobreza como insuficiência absoluta de renda dribla esses obstáculos, mas ainda exige um passo adicional: a escolha de uma linha de pobreza. Para isso, o método calórico sempre foi uma abordagem comum na América Latina, possuindo variantes que buscam, de diferentes formas, estipular como linha de pobreza a renda necessária para o consumo dos nutrientes mínimos para sobrevivência.

Entretanto, poucas pesquisas recentes seguem esse caminho. As dificuldades para o método calórico são muitas: falta de informações, extrema sensibilidade dos resultados a pequenas mudanças metodológicas, incongruências teóricas com o conceito de bem-estar, definição de cestas alimentares intragáveis, foco restrito a calorias, e assim por diante.

Por isso, a posição predominante no Brasil nos últimos 15-20 anos tem sido o uso de linhas administrativas de pobreza – ou seja, linhas definidas de modo mais ou menos arbitrário por governos ou organismos internacionais para fins de elegibilidade ou monitoramento de políticas públicas. Essa abordagem abdica da definição científica de uma linha de pobreza. Para seus defensores, contudo, esse último ponto é uma qualidade, não um problema, pois o verniz de cientificidade dado pelos outros métodos disponíveis esconde uma série de decisões arbitrárias, enquanto o uso de linhas administrativas torna essa arbitrariedade explícita e, presumivelmente, traduz uma solução de compromisso entre os recursos fiscais disponíveis e a vontade dos eleitores.

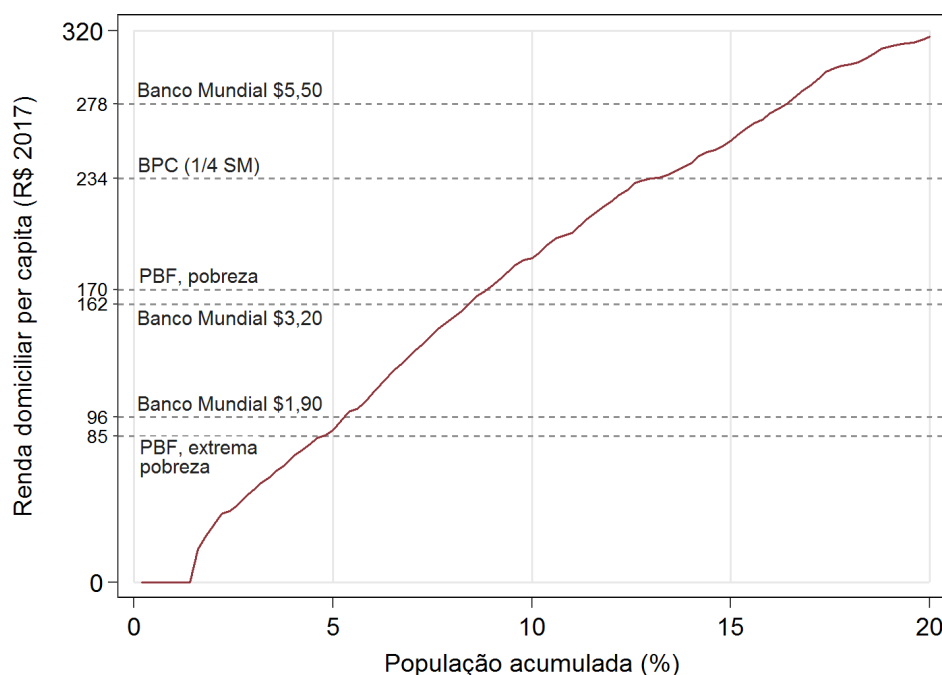
Como nosso país não possui uma linha de pobreza oficial, os trabalhos empíricos tendem a usar como referência nossos principais programas sociais. As linhas de elegibilidade ao Programa Bolsa Família (PBF) ou ao Benefício de Prestação Continuada (BPC) têm sido amplamente utilizadas. Outra opção igualmente comum é a adoção das linhas de pobreza calculadas pelo Banco Mundial, que foram encampadas na Meta 1 dos Objetivos de Desenvolvimento Sustentável da Organização das Nações Unidas (ONU).

A Figura 3 recorre à Parada de Pen para comparar diferentes linhas de pobreza. No gráfico, a mesma distribuição de renda da PNADC 2017 apresentada na Figura 1 está truncada no percentil 20, enquanto as linhas de pobreza mais comuns são representadas por linhas horizontais em cinza.

O percentual de pobres gerado por cada linha corresponde ao ponto no eixo horizontal em que a linha cruza com a Parada de Pen. Assim, cerca de 5% dos brasileiros são pobres de acordo com as linhas mais baixas, percentual que sobe para quase 17% para a linha mais alta. Ou seja, o percentual de pobres observado com a linha mais elevada é mais de três vezes maior do que o visto com a menor linha – e não dá para dizer que uma opção é a “certa” e a outra é “errada”.

Em alguns casos, por pura coincidência, há grande semelhança de valores entre linhas calculadas de forma completamente diferente. As duas linhas mais baixas do Banco Mundial são muito próximas das linhas do Programa Bolsa Família, ao passo que a linha mais elevada é razoavelmente parecida com a linha do BPC. Quando diferentes linhas estão muito próximas, as conclusões empíricas tendem a ser as mesmas.

Figura 3 – Parada de Pen truncada da renda domiciliar per capita no Brasil e linhas de pobreza mais comuns, 2017.



Fonte: Pesquisa Nacional por Amostra de Domicílios Contínua (PNADC) 2017. Nota: rendimentos brutos efetivos recebidos por moradores (exceto pensionistas, empregados domésticos e afins), registrados nas variáveis vd4020 e vd4048. Ver seção 5 para mais informações sobre a PNADC. As linhas de pobreza do Banco Mundial, expressas em dólares internacionais (PPP\$) por dia, foram transformadas em reais com fator de conversão de 1.66 e multiplicadas por 30,42 para obtenção de valores mensais.

Seja como for, é indispensável que a linha de pobreza tenha um valor real fixo. No Brasil, por uma série de motivos, diferenças de custo de vida entre regiões costumam

ser ignoradas. Dado nosso histórico inflacionário, a prioridade sempre foi manter o valor real em comparações ao longo do tempo.

Por exemplo, suponha que um estudo baseado na PNADC 2017 usou a linha inferior (“extrema pobreza”) do Programa Bolsa Família naquele ano (R\$ 85) e que, posteriormente, os autores decidiram comparar os resultados com os de 2016. Como a inflação entre os dois períodos foi de aproximadamente 6,5%, há duas opções: ou multiplicar todos os rendimentos na PNADC 2016 por 1.065 e aplicar a mesma linha de R\$ 85 ou manter os rendimentos nominais e aplicar a linha deflacionada de R\$ 79,81 em 2016. Seria errado simplesmente utilizar o valor nominal da linha inferior de elegibilidade do programa em 2016: como os valores do PBF não são indexados e desfrutam apenas de reajustes irregulares, seu valor real flutua muito ao longo do tempo. No caso, em 2016 a linha inferior já era de R\$ 85, valor que só foi atualizado para R\$ 89 em 2018, ficando aquém da inflação acumulada no período.

Medidas de pobreza

A medida de pobreza mais utilizada é o percentual de pobres na população, vale dizer, a fração da população cuja renda é menor do que a linha de pobreza. Adicionalmente, o número absoluto de pobres também desperta interesse, embora a incidência relativa seja o padrão para comparações no tempo e no espaço por motivos óbvios.

O percentual de pobres ignora a intensidade da pobreza e a desigualdade entre os pobres. Pensemos em três sociedades hipotéticas (A, B e C), todas com taxa de pobreza de 10% da população, para uma linha de pobreza de \$100. Na sociedade A todos os pobres têm renda de \$99, enquanto na sociedade B todos os pobres têm renda de \$50. Na sociedade C, por sua vez, metade dos pobres têm renda igual a \$10 e a outra metade tem renda de \$90.

Embora a proporção de pobres seja idêntica nos três casos, a situação na sociedade A é muito melhor do que nas outras, pois os pobres são muito menos pobres, isto é, a intensidade da pobreza é menor. Ao mesmo tempo, as sociedades B e C têm a mesma intensidade média da pobreza, já que a renda média dos pobres é de \$50 nas duas, só que metade dos pobres em C está em privação relativa muito mais intensa do que os pobres em B. Logo, se estivermos preocupados com os mais pobres entre os pobres, a situação em B pode ser considerada melhor do que em C.

Para captar essas diferenças, o economista indiano Amartya Sen propôs uma abordagem axiomática para a mensuração da pobreza, depois estendida por muitos outros

autores. Nesse enfoque, uma medida de pobreza só é satisfatória se atender às condições especificadas em uma série de axiomas ou princípios. O próprio Sen construiu um índice de pobreza apropriado segundo esses critérios, mas, na prática, a abordagem unificada desenvolvida por James Foster, Joel Greer e Erik Thorbecke (FGT) nos anos 1980 é ainda hoje a mais usada.

Suponha uma população de tamanho N , em que cada indivíduo i tem renda y_i e a linha de pobreza escolhida é representada por z . Com isso, para todos os q indivíduos que estão abaixo da linha de pobreza, a forma geral da família de medidas FGT é dada por:

$$P(\alpha) = \frac{1}{N} \sum_{i=1}^q \left(\frac{z - y_i}{z} \right)^\alpha, \text{ com } \alpha \geq 0$$

O parâmetro α representa a aversão à desigualdade: quanto mais alto seu valor, maior peso damos para os indivíduos mais pobres entre os pobres. Tipicamente, os valores mais usados são 0, 1 e 2. Quando $\alpha = 0$, a fórmula se reduz para a proporção de pobres na população. Quando $\alpha = 1$, a fórmula equivale à multiplicação da proporção de pobres pela intensidade média da pobreza. Quando $\alpha = 2$, leva-se em conta também a desigualdade entre os pobres, medida pelo quadrado do coeficiente de variação.

Em termos axiomáticos, a medida $P(2)$ é a mais satisfatória – por exemplo, ela é a única que diferencia entre as sociedades B e C no cenário hipotético descrito acima. O paradoxo é que, apesar disso, essa medida é a menos utilizada das três.

O motivo é que o preço a pagar pela informação adicional é a perda de transparência. Todo mundo é capaz de entender o que significa um percentual de pobres igual a 10%, mas não há nenhuma interpretação intuitiva possível para a medida $P(2)$. A medida $P(1)$, por sua vez, é só um pouco menos opaca, podendo ser concebida como o custo mínimo de eliminação da pobreza, supondo focalização perfeita, como proporção da linha de pobreza.

Essa opção pragmática pela simplicidade não deve obscurecer o fato de que há julgamentos de valor implícitos no uso da taxa de pobreza. Com efeito, não é só a escolha da linha que carrega um substrato normativo: todas as medidas de pobreza implicam a adoção de algum princípio de justiça, que, no caso da família FGT, torna-se explícito com a escolha do parâmetro α . O próprio foco na pobreza, por sinal, carrega em si

mesmo uma determinada concepção de bem-estar social que ignora tudo o que acontece entre os não pobres.

Seja como for, no mundo real, privilegiar a proporção de pobres em detrimento das outras medidas não é um grande problema porque são raras as situações como a do exemplo hipotético acima. Comparações no tempo ou no espaço tendem a produzir as mesmas conclusões substantivas para as medidas $P(0)$, $P(1)$ e $P(2)$.

Dominância de primeira ordem

Suponha que queremos saber se a pobreza na PNADC 2017 era menor no Pará ou em Pernambuco. Se adotarmos como referência a linha inferior do PBF (R\$ 84 *per capita*, em valores de 2017), o Pará se sai melhor, com 6,8% de pobres, contra 8,3% de Pernambuco. Porém, se preferirmos a linha de pobreza mais elevada do Banco Mundial (R\$ 278 *per capita*), a relação se inverte: Pernambuco aparece com 27,5% de pobres, contra 31% do Pará.

Nesse exemplo, não podemos dizer inequivocamente onde a pobreza é menor, pois os resultados dependem da linha de pobreza escolhida. O problema maior é que, mesmo que obtivéssemos o mesmo *ranking* para essas duas linhas, nada garante que o resultado seria válido para outras escolhas. Em outras palavras, conclusões obtidas para uma determinada linha de pobreza nem sempre são válidas para outras.

Como proceder então se quisermos extrair conclusões mais robustas? Felizmente, não precisamos repetir o cálculo para todas as linhas de pobreza imagináveis. Há um jeito mais fácil, baseado na Parada de Pen, que consiste na busca por relações de *dominância de primeira ordem* entre distribuições de renda.

Para duas distribuições de renda expressas na mesma unidade monetária, dizemos que há dominância de primeira ordem quando as curvas de quantis não se cruzam: a renda no percentil p na distribuição dominante é sempre maior do que a renda no mesmo percentil na distribuição dominada. Ou seja, todos na distribuição dominante possuem rendimentos mais elevados do que seus equivalentes na distribuição dominada.

Caso ocorra dominância de primeira ordem, podemos concluir que o percentual de pobres e intensidade da pobreza são sempre menores na distribuição dominante, seja qual for a linha absoluta de pobreza escolhida. Não podemos chegar a conclusões semelhantes quando não há dominância, isto é, quando as curvas de quantis se cruzam, pois é sempre teoricamente possível achar linhas de pobreza que reflitam

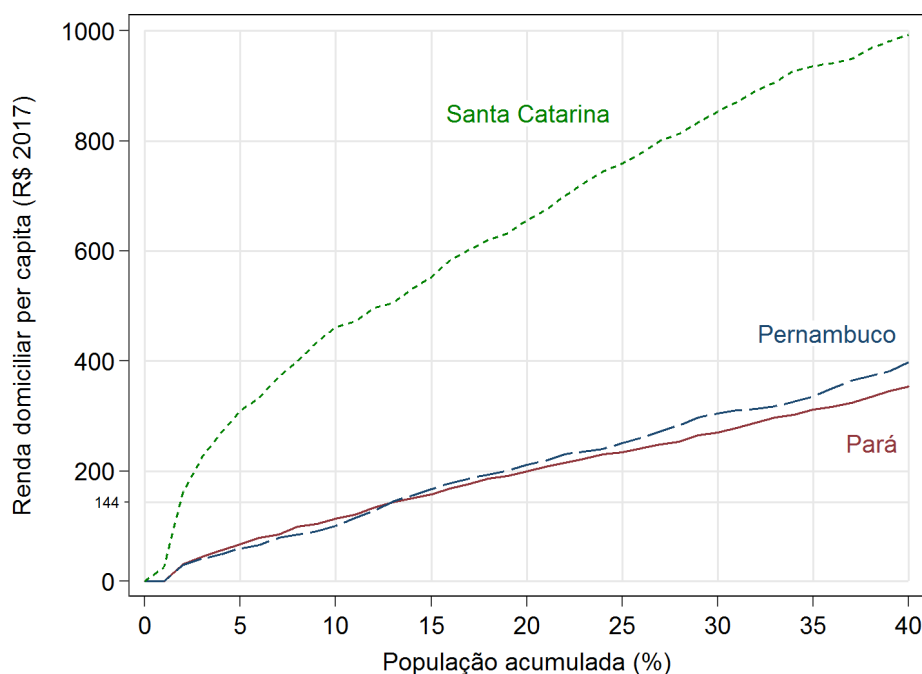
ordenamentos diferentes.

A Figura 4 traz exemplos concretos, comparando as curvas de quantis de Pará, Pernambuco e também Santa Catarina. Para facilitar a visualização, o eixo horizontal foi truncado no percentil 40, ou seja, a figura traz os rendimentos absolutos apenas dos 40% mais pobres de cada UF.

O gráfico explica por que obtivemos ordenamentos contraditórios entre Pará e Pernambuco: não há dominância de primeira ordem entre as duas distribuições, que se cruzam em torno do percentil 13. Dessa forma, o percentual de pobres no Pará será sempre menor do que o de Pernambuco para linhas de pobreza de até R\$ 144 *per capita*; porém, Pernambuco terá resultados melhores para todas as linhas de pobreza com valor mais alto do que isso.

Por outro lado, Santa Catarina possui nítida dominância de primeira ordem sobre Pará e Pernambuco: seja qual for a linha de pobreza, tanto a proporção de pobres quanto a intensidade da pobreza serão menores na UF sulista.

Figura 4 – Análise de dominância de primeira ordem – Pará, Pernambuco e Santa Catarina, 2017



Fonte: Pesquisa Nacional por Amostra de Domicílios Contínua (PNADC) 2017. Nota: rendimentos brutos efetivos recebidos por moradores (exceto pensionistas, empregados domésticos e afins), registrados nas variáveis vd4020 e vd4048. Ver seção 5 para mais informações sobre a PNADC.

Desigualdade

Definições

Na vida cotidiana, “pobreza” imediatamente evoca um estado concreto de carência ou privação, seja lá como esse estado for definido. “Desigualdade”, por sua vez, não se trata de um estado vivido por indivíduos, mas sim de uma característica da *relação* entre indivíduos. Em termos mais técnicos, a desigualdade não é uma característica das unidades de análise, mas da dispersão na distribuição de uma variável entre as unidades de análise.

A desigualdade ou dispersão na distribuição de renda pode ser entendida em termos absolutos ou relativos. A diferença está em como medimos a distância entre indivíduos. Suponha, por exemplo, que a pessoa A tem renda de R\$ 1.000, enquanto a pessoa B tem renda de R\$ 200.000. Em termos absolutos, a distância entre elas é de R\$ 199.000, enquanto em termos relativos a métrica relevante é que a renda de B é 200 vezes maior do que a de A. Se dobramos a renda dos dois indivíduos, a desigualdade absoluta aumenta, mas a desigualdade relativa permanece idêntica.

Quase todas as pesquisas acadêmicas definem a desigualdade em termos relativos, por motivos teóricos e empíricos. Por isso, para a análise da desigualdade, a curva de Lorenz é uma forma de representação visual da distribuição de renda mais adequada do que a Parada de Pen, pois é construída a partir de rendimentos relativos ao longo de toda a distribuição de renda.

Essa definição em termos de diferenças relativas implica que a desigualdade pode variar de modo independente do padrão de vida ou da renda média. Sociedades com a mesma renda média podem ter níveis de desigualdade radicalmente diferentes e vice-versa. Ao longo do tempo podemos ter todas as combinações possíveis entre mudanças na renda *per capita* e variações na desigualdade.

Raciocínio análogo vale para a relação entre desigualdade e pobreza, que permanecem dois conceitos distintos e que também podem variar de forma independente. Não obstante, os dois conceitos guardam laços entre si. Partindo dos aspectos mais práticos em direção aos mais teóricos, podemos citar pelo menos quatro formas de conexão entre eles. Primeiro, algumas medidas axiomáticas de pobreza incorporam a desigualdade, ao menos entre os mais pobres. Segundo, para um determinado nível médio de renda, quanto menor a desigualdade, menor costuma ser a taxa de pobreza. Terceiro, mesmo linhas absolutas de pobreza são definidas, em última instância,

conforme as condições de vida e a desigualdade em algum grupo de referência: como lembra Amartya Sen, a satisfação absoluta de algumas necessidades depende da posição relativa da pessoa em seu meio social, algo já reconhecido por Adam Smith no século 18. Quarto, como as medidas de pobreza, todas as medidas de desigualdade têm um substrato normativo: o modo como quantificamos os desvios em relação à igualdade perfeita reflete um julgamento de valor prévio.

Esse último *insight* data do trabalho pioneiro de Hugh Dalton, no início do século 20. Nos anos 1970, Anthony Atkinson foi além e mostrou que todas as medidas usuais de desigualdade – derivadas da curva de Lorenz – baseiam-se em funções de bem-estar social côncavas e simétricas com formas funcionais específicas. Isso faz com que algumas sejam mais sensíveis aos extremos da distribuição enquanto outras apresentam maior sensibilidade a mudanças nos estratos intermediários.

Medidas axiomáticas

Uma medida sintética de desigualdade resume todas as diferenças relativas observadas ao longo da distribuição de renda em um único número, assim como uma medida de pobreza é uma forma sintética de resumir informações sobre a distribuição de renda entre indivíduos ou famílias abaixo da linha de pobreza. Ao contrário da discussão sobre pobreza, contudo, não há para a desigualdade uma medida imperfeita, porém incrivelmente simples e intuitiva, que tenha apelo imediato, como é o caso do percentual de pobres.

Por isso, o debate técnico sobre mensuração da desigualdade orienta-se a partir de axiomas, isto é, propriedades desejáveis por motivos autoevidentes. Esses axiomas não determinam a forma da função de bem-estar subjacente a cada medida ou família de medidas de desigualdade, mas apenas garantem que elas obedecem a critérios mínimos.

A lista exata de axiomas é objeto de disputa entre diferentes autores, ainda que seja possível identificar um núcleo duro com quatro axiomas quase sempre aceitos: o princípio de transferências de Pigou-Dalton; o princípio de independência de escala; o princípio da população de Dalton; e o princípio do anonimato.

Quadro 1 – Principais axiomas para medidas de desigualdade

1. Princípio (fraco) de transferências de Pigou-Dalton

Transferências progressivas de renda de um indivíduo mais rico para outro mais pobre que não alterem as posições relativas dos dois devem reduzir a desigualdade.

2. Princípio de independência de escala

Mudanças proporcionais nas rendas de todos nunca alteram a desigualdade, isto é, a desigualdade é independente da escala de mensuração, pois o que interessa são as diferenças relativas.

3. Princípio de população de Dalton

A desigualdade de uma distribuição independe do tamanho da população, pois é invariante a replicações da população, ou seja, combinações de duas ou mais distribuições de renda idênticas nunca alteram o nível de desigualdade.

4. Princípio do anonimato

A desigualdade é simétrica, dependendo apenas da distribuição de frequência dos rendimentos, de modo que nenhuma permutação aleatória de rendimentos em uma população afeta o nível de desigualdade. A renda é a única característica relevante dos indivíduos para a mensuração da desigualdade.

Juntos, esses axiomas impõem limitações brandas, porém importantes. Por exemplo, o princípio de transferências exige que a desigualdade caia quando ocorrem transferências progressivas, mas não especifica qual deve ser o tamanho da queda, que pode variar entre medidas. Por si só, essa exigência já derruba medidas como a variância dos logaritmos e impõe que medidas satisfatórias de desigualdade precisam levar em conta toda a distribuição de renda, o que exclui o uso de indicadores como a razão de Palma ou a fração da renda recebida pelo 1% mais rico, que serão discutidos na próxima seção. O axioma de independência de escala, por sua vez, preconiza que só devemos nos importar com a desigualdade relativa, de modo que nossas medidas não devem mudar se, por exemplo, multiplicarmos os rendimentos de todos por mil ou convertermos as rendas de reais para dólares. Embora pouco exigente, esse axioma significa que a variância e o desvio-padrão, duas medidas de dispersão fundamentais em estatística, não são boas medidas de desigualdade.

As principais medidas que satisfazem a esses quatro axiomas são discutidas abaixo. Como elas se baseiam em diferentes princípios normativos, nem sempre seus resultados concordam entre si. Por exemplo, ao compararmos duas sociedades A e B , não é incomum que o coeficiente de Gini aponte A como a mais igualitária, enquanto o $GE(0)$, *que será analisada adiante*, diga que a desigualdade é menor em B . Por isso, a seção 4.4 introduz o conceito de dominância de Lorenz, que serve em alguns casos para dirimir dúvidas de modo análogo ao que fizemos acima com a dominância de primeira ordem.

Coefficiente de Gini

A medida de desigualdade mais conhecida que atende aos quatro axiomas é o coeficiente de Gini, introduzido em 1912 pelo estatístico italiano Corrado Gini, que posteriormente foi colaborador de Mussolini e em 1926 se tornou o primeiro presidente do instituto oficial de estatística da Itália.

O coeficiente de Gini varia entre 0 e 1, sendo às vezes multiplicado por 100 apenas para facilitar a apresentação. O valor mínimo representa a igualdade perfeita, isto é, o cenário em que todos os indivíduos possuem exatamente a mesma renda. No outro extremo, o valor máximo sinaliza a maior desigualdade possível, correspondente a uma situação hipotética em que um único indivíduo recebe toda a renda, enquanto todo o resto da população tem renda igual a zero.

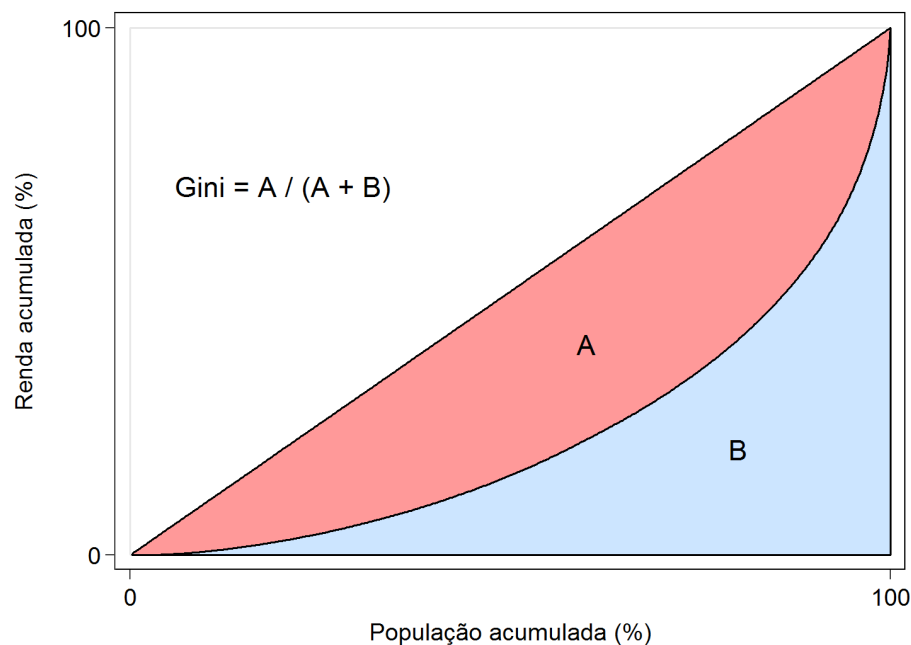
Na prática, o Gini da renda domiciliar *per capita* costuma variar entre 0,25 e 0,40 em países mais igualitários e entre 0,40 e 0,65 em países muito desiguais. Dada a definição acima, esses valores podem ser interpretados como o percentual da desigualdade máxima possível, ao menos no sentido estatístico do termo.

Além dessa, há mais de uma dúzia de interpretações para o coeficiente de Gini. A mais intuitiva é a que mostra como essa medida pode ser calculada diretamente a partir da curva de Lorenz. Matematicamente, considere uma variável aleatória não negativa y com função de distribuição acumulada $F = F(y)$, e seja $L(\cdot)$ a função que calcula a curva de Lorenz. Então:

$$G = 1 - 2 \int_0^1 L(F) dF$$

Visualmente, essa equação significa que o coeficiente de Gini corresponde à razão entre duas áreas: o numerador é a área entre a reta da igualdade perfeita e a curva de Lorenz, ao passo que o denominador é a área total sob a reta da igualdade perfeita. Para facilitar o entendimento, a Figura 5 reapresenta os dados da Figura 2, deixando claro como computar o Gini a partir da curva de Lorenz.

Figura 5 – Cálculo do coeficiente de Gini a partir da Curva de Lorenz da renda domiciliar per capita no Brasil, 2017



Fonte: Pesquisa Nacional por Amostra de Domicílios Contínua (PNADC) 2017. Nota: rendimentos brutos efetivos recebidos por moradores (exceto pensionistas, empregados domésticos e afins), registrados nas variáveis vd4020 e vd4048. Ver seção 5 para mais informações sobre a PNADC.

A representação gráfica explicita por que os valores extremos do Gini são 0 e 1. Caso todos tenham a mesma renda, a área A será igual a zero, pois a curva de Lorenz estará completamente sobreposta à reta de igualdade perfeita. Logo, o Gini será zero. Em uma sociedade suficientemente numerosa, caso apenas um indivíduo concentre toda a renda, a área B será igual a zero, o que implica um coeficiente de Gini igual a 1. A curva de Lorenz de fato representada nas Figuras 2 e 5 corresponde a um coeficiente de Gini de 0,549.

Medidas de entropia generalizada

Além dos quatro axiomas listados acima, muitos autores consideram que há um princípio adicional que também deveria ser levado em conta: trata-se do axioma de decomponibilidade aditiva. Uma medida de desigualdade atende a esse axioma se, para uma dada partição da população em subgrupos mutuamente exclu-

dentes, a desigualdade total puder ser escrita como a soma da desigualdade entre grupos e intragrupos. Em termos formais, decompor aditivamente uma medida I significa escrevê-la de forma semelhante a $I = I_{entre} + I_{intra}$.

Por esse critério, uma boa medida de desigualdade deve permitir, por exemplo, que a desigualdade total no Brasil seja decomposta por regiões e expressa como a soma de dois componentes: a desigualdade entre regiões (I_{entre}) e a desigualdade dentro das regiões (I_{intra}).

O axioma de decomponibilidade exige coerência entre a desigualdade na sociedade como um todo e a desigualdade em suas partes constitutivas: *ceteris paribus*, se a desigualdade dentro de cada grupo cair e não houver mudanças nas rendas médias relativas entre os grupos, então a desigualdade total também deve cair (e vice-versa). No entanto, nem sempre isso ocorre: medidas como o coeficiente de Gini e a variância de logaritmos não obedecem a este axioma.

A única família de medidas que atende simultaneamente ao princípio de decomponibilidade e aos quatro axiomas anteriores é a família de medidas de entropia generalizada, cujas raízes remetem à teoria da informação. Suponha uma população de tamanho N , com renda média igual a μ , em que cada indivíduo i tem renda y_i . Então:

$$GE(\alpha) = \begin{cases} \frac{1}{N(\alpha^2 - \alpha)} \sum_{i=1}^N \left[\left(\frac{y_i}{\mu} \right)^\alpha - 1 \right], & \text{para } \alpha \neq 0, 1 \\ -\frac{1}{N} \sum_{i=1}^N \ln \left(\frac{y_i}{\mu} \right), & \text{para } \alpha = 0 \\ \frac{1}{N} \sum_{i=1}^N \left(\frac{y_i}{\mu} \right) \ln \left(\frac{y_i}{\mu} \right), & \text{para } \alpha = 1 \end{cases}$$

O parâmetro α determina o peso dado às distâncias entre rendimentos em diferentes partes da distribuição. Valores mais baixos significam medidas mais sensíveis a mudanças nos estratos mais pobres da distribuição de renda, e valores mais altos causam o efeito oposto. Se $\alpha = 1$, a medida dá o mesmo peso para mudanças ao longo de toda a distribuição.

Os valores mais usados para α são 0, 1 e 2. A medida $GE(0)$ é também chamada de L de Theil ou MLD (*mean log deviation*); a medida $GE(1)$ é o T de Theil ou índice de Theil; e a medida $GE(2)$ corresponde à metade do quadrado do coeficiente de variação ($GE(2) = \frac{1}{2} CV^2$).

A Tabela 1 ilustra as vantagens do axioma de decomponibilidade, trazendo a decomposição da renda domiciliar *per capita* na PNADC 2017 por Unidades da Federação (UFs). Quando o parâmetro α é 0 ou 1, a desigualdade entre UFs representa entre 9% e 10% da desigualdade total no Brasil, mas quando $\alpha=2$, esse valor cai para apenas 3%. Com efeito, quanto maior o parâmetro α – ou seja, quanto mais importância damos aos muito ricos – menor o peso da desigualdade de renda média entre UFs na desigualdade total.

Tabela 1 – Decomposição da desigualdade da renda domiciliar per capita por Unidades da Federação (UF) para as principais medidas de entropia generalizada – Brasil, 2017

Medidas	Desigualdade total	Desigualdade entre UFs	Desigualdade intra-UFs
GE(0)	0,556	0,054	0,502
(%)	100	10	90
GE(1)	0,600	0,051	0,549
(%)	100	9	91
GE(2)	1,569	0,051	1,518
(%)	100	3	97

Fonte: Pesquisa Nacional por Amostra de Domicílios Contínua (PNADC) 2017.

Nota: rendimentos brutos efetivos recebidos por moradores (exceto pensionistas, empregados domésticos e afins), registrados nas variáveis vd4020 e vd4048. Ver seção 5 para mais informações sobre a PNADC. As medidas GE(0) e GE(1) excluem pessoas com renda domiciliar per capita igual a zero.

Cabe observar que o componente entre grupos só pode ser interpretado de forma contrafactual para a medida GE(0) ou L de Theil. Por essa métrica, se a renda média em todas as UFs fosse rigorosamente a mesma, a desigualdade total recuaria 0,054, ou 10%. Não podemos afirmar nada semelhante para as outras duas medidas.

Não é fácil explicar por que as medidas de entropia generalizada são menos populares do que o coeficiente de Gini. Uma hipótese plausível é que elas são ainda menos intuitivas do que o bom e velho Gini, com algumas inconveniências adicionais. Por exemplo, embora todas as medidas da família GE(α) sejam iguais a zero quando há igualdade perfeita, elas não assumem o valor máximo de 1 quando há a maior desigualdade possível, ao contrário do Gini. No caso do GE(0) sequer há um limite

superior definido. Além disso, como as medidas $GE(0)$ e $GE(1)$ são computadas a partir de transformações logarítmicas, elas não são definidas caso existam indivíduos com rendimentos iguais a zero. Nessas situações, ou esses casos são removidos da análise ou os zeros são substituídos por algum valor infinitesimal positivo - duas soluções por elegantes, por assim dizer.

Medidas não axiomáticas

Os axiomas para mensuração da desigualdade servem para guiar a construção de medidas capazes de resumir toda a distribuição de renda em um único número. Como se depreende da discussão acima, tal rigor técnico tem seu preço: medidas axiomáticas são demasiado abstratas.

Mesmo o coeficiente de Gini, talvez a medida axiomática mais intuitiva, é indesejavelmente misterioso. O Gini pode indicar um grau alto ou baixo de desigualdade, mas nada nos diz sobre sua forma concreta. A desigualdade pode ser alta porque os pobres são muito mais pobres do que o resto da população, ou porque os ricos são incrivelmente mais ricos, e assim por diante.

Por esses motivos, muitos pesquisadores argumentam que análises baseadas em medidas não axiomáticas são insubstituíveis e, em muitos casos, até preferíveis. Por essa visão, valeria a pena sacrificar a aderência estrita a alguns axiomas em nome de medidas mais transparentes, facilmente compreensíveis até por leigos e que, de quebra, ainda deixam explícito seu conteúdo normativo.

Fração da renda recebida pelos mais ricos e similares

A medida não axiomática mais usada é a fração da renda recebida pelos mais ricos, definidos em termos de estratos com tamanho relativo fixo. A fração da renda recebida pelo 1% mais rico da população, em particular, ganhou enorme popularidade na esteira dos trabalhos do francês Thomas Piketty e seus coautores. Por definição, essa medida só registra um tipo de desigualdade, a que separa os $x\%$ mais ricos do resto da população. Portanto, ela viola o princípio das transferências, já que é imune a quaisquer mudanças distributivas que ocorram entre os mais pobres e/ou entre os mais ricos. Os outros axiomas, no entanto, são obedecidos.

Em compensação, a fração dos mais ricos exemplifica com louvor as virtudes das medidas não axiomáticas. Por um lado, qualquer leigo é capaz de entender seu sig-

nificado e relacioná-la à experiência cotidiana. Por outro, é evidente que, em termos normativos, o foco dessa medida é a concentração de renda no topo e suas possíveis consequências.

A fração recebida pelos mais ricos pode ser reconciliada com o coeficiente de Gini. Como “ricos” e “não ricos” constituem dois grupos sem sobreposição de rendimentos, temos um dos poucos casos em que o Gini se torna aditivamente decomponível em um componente intra-grupos e outro componente entre grupos. O último componente é dado por $G_{entre} = S_{ricos} - P_{ricos}$, ou seja, a contribuição para o Gini da desigualdade entre grupos é igual à fração da renda recebida pelos mais ricos (S_{ricos}) menos a fração da população escolhida para delimitar os mais ricos (P_{ricos}). Por exemplo, se os 10% mais ricos recebem 60% da renda total, então o componente entre grupos é igual a $G_{entre} = 0,60 - 0,10 = 0,50$.

Por sinal, nada impede que outras medidas não axiomáticas sejam construídas de modo análogo. Por exemplo, pode-se calcular a fração recebida pelos x% mais pobres ou por qualquer outro grupo. Os próprios trabalhos recentes de Piketty e associados têm trilhado esse caminho, apresentando resultados também para a fração da renda recebida pelos 50% mais pobres e para os 40% que ocupam posição intermediária entre a mediana e o percentil 90.

Razão de Palma e similares

Medidas como a fração da renda recebida pelos mais ricos impõem uma classificação dicotômica à distribuição de renda, contrastando o topo da distribuição com todo o resto da população. Porém, o conflito distributivo mais relevante para a análise pode ser distinto. Afinal, podemos querer analisar a desigualdade entre os mais ricos e a “classe média” (seja lá como for definida), ou apenas entre ricos e pobres, entre outros.

Dessa forma, outro tipo de medida não axiomática frequentemente utilizada é a razão entre rendimentos observados em diferentes pontos da distribuição. Grosso modo, as possibilidades de escolhas são infinitas. Graças aos trabalhos do economista chileno José Gabriel Palma, a medida mais usada na literatura recente é a razão entre a renda dos 10% mais ricos e a dos 40% mais pobres. O cálculo pode ser feito tanto para a renda total quanto para a fração da renda ou mesmo para a renda média dos dois grupos – as duas primeiras opções produzem sempre os mesmos resultados, en-

quanto a terceira implica um nível quatro vezes maior, mas com variações relativas idênticas.

A razão de Palma viola o princípio de transferências porque ignora tudo o que acontece entre a os percentis 40 e 90, bem como a desigualdade interna entre os 40% mais pobres e os 10% mais ricos. Seu foco recai exclusivamente sobre a polarização entre ricos e pobres. Sua motivação decorre de um padrão observado por Palma em dados transversais internacionais. Segundo ele, a metade da população situada entre os percentis 40 e 90 na distribuição de renda quase sempre recebe cerca de metade da renda nacional, de modo que as diferenças entre países resultam principalmente da divisão da outra metade da renda entre os 40% mais pobres e os 10% mais ricos.

Não há consenso na literatura especializada acerca da validade dessa generalização empírica, em parte porque ela só é falsificável se definirmos rigorosamente o que constitui “razoável estabilidade” da fatia da classe média. Além disso, alguns pesquisadores consideram que a delimitação dos grupos é arbitrária e esconde muita heterogeneidade.

Um resultado curioso é que a razão de Palma está altamente correlacionada com o coeficiente de Gini, “explicando” quase toda a variação do Gini entre países. Para seus defensores, isso significa que, para fins de políticas públicas, a razão de Palma deve ser usada em detrimento do Gini, pois é muito mais transparente e explícita em termos de preferências. Para seus detratores, trata-se apenas de uma forma tecnicamente imperfeita de obter as mesmas informações.

Para fins deste capítulo, a validade dos fatos estilizados reportados por Palma é bem menos importante do que a estrutura da medida não axiomática que ele ajudou a popularizar. Nada impede, por exemplo, o uso de medidas como a razão entre a renda do 1% mais rico e dos 50% mais pobres, ou qualquer outra do tipo. Pode-se, inclusive, usar a renda observada em diferentes percentis da distribuição no lugar da renda média ou total de diferentes estratos populacionais.

Pesquisadores que estudam o mercado de trabalho têm esse hábito. As mais comuns são as razões $P90/P10$, $P90/P50$ e $P50/P10$. Seu cálculo é simples: basta ordenar a população de acordo com a renda, do mais pobre para o mais rico, anotar os valores observados nos percentis de interesse e calcular a razão entre eles. Para a razão $P90/P10$, por exemplo, é só dividir a renda do indivíduo que está no percentil 90 (isto é, o indivíduo que é mais rico que 90% da população) pela renda do indivíduo no percentil 10 (isto é, o indivíduo que só é mais rico que 10% da população).

Dominância de Lorenz

Assim como no caso da pobreza, análises da desigualdade em geral desejam obter respostas inequívocas para comparações entre diferentes sociedades ou entre diferentes momentos no tempo. O problema, aqui, é que temos uma profusão de medidas que usam pesos específicos para quantificar desvios em relação à igualdade perfeita.

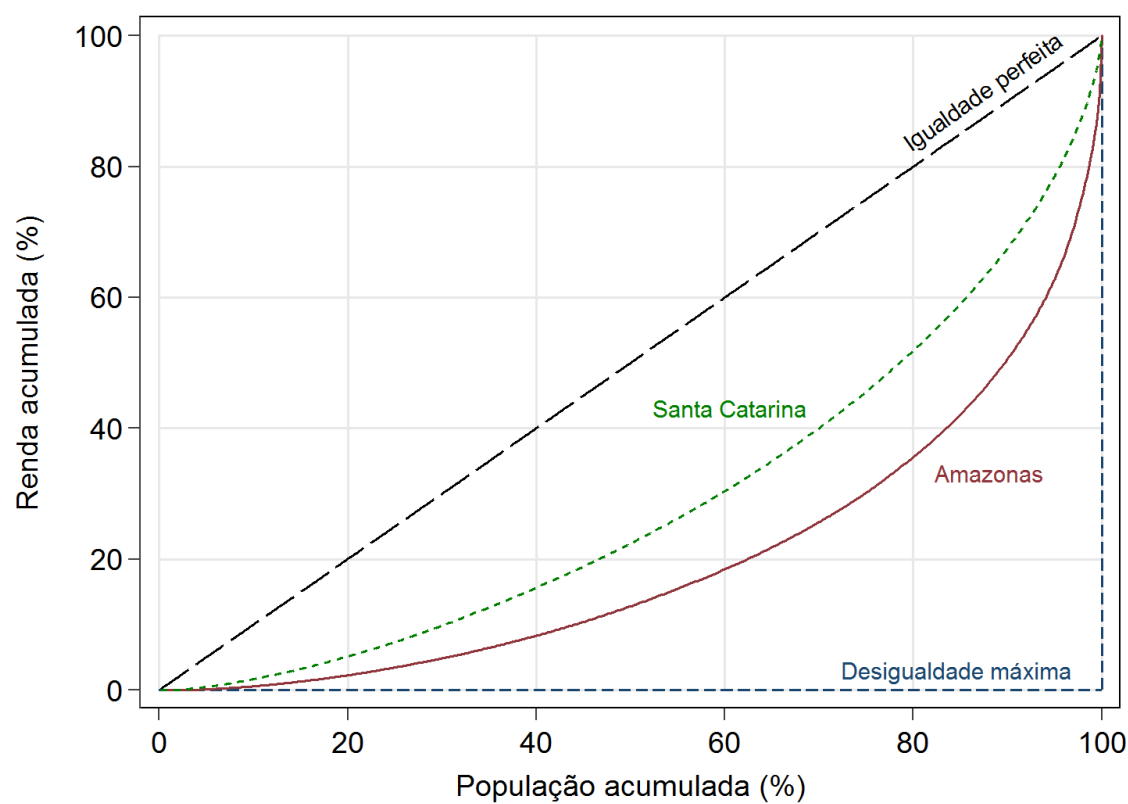
Como saber se o ordenamento obtido para uma medida é robusto? Mais uma vez, não é preciso recorrer à força bruta e calcular os resultados para todas as medidas possíveis e imagináveis para ter certeza de que é possível ordenar distribuições de modo não ambíguo. Se na análise da pobreza resolvemos o problema averiguando a existência ou não de dominância de primeira ordem, para a desigualdade a solução é o teste para outro tipo de dominância – a **dominância de Lorenz**.

Há dominância de Lorenz de uma distribuição sobre outra quando as duas curvas de Lorenz não se cruzam, ou seja, a fração da renda acumulada até o percentil p é sempre maior ou igual na distribuição dominante do que na dominada, para qualquer p . Por consequência, a fração da renda total recebida pelos $p\%$ mais pobres na distribuição dominante é sempre maior do que na distribuição dominada.

Se a dominância de Lorenz for confirmada, podemos afirmar que a desigualdade é menor na distribuição dominante: todas as medidas que obedecem aos quatro primeiros axiomas registrarão valores superiores de desigualdade para a distribuição dominada.

A Figura 6, na página seguinte, traz um exemplo de dominância de Lorenz, comparando a curva de Santa Catarina com a do Amazonas – respectivamente, as Unidades da Federação com o menor e o maior Gini na PNADC 2017. Ao longo de toda a distribuição, a curva de Santa Catarina está sempre mais perto da linha de igualdade perfeita do que a do Amazonas, e elas não se cruzam em nenhum momento. Portanto, Santa Catarina possui dominância de Lorenz sobre o Amazonas, sendo inequivocamente menos desigual.

Figura 6 – Análise de dominância Lorenz para renda domiciliar per capita – Amazonas e Santa Catarina, 2017



Fonte: Pesquisa Nacional por Amostra de Domicílios Contínua (PNADC) 2017. Nota: rendimentos brutos efetivos recebidos por moradores (exceto pensionistas, empregados domésticos e afins), registrados nas variáveis vd4020 e vd4048. Ver seção 5 para mais informações sobre a PNADC.

Fontes de dados para o Brasil

Pesquisas domiciliares

Censo demográfico

Desde 1960, o Censo utiliza dois instrumentos de coleta: um questionário básico, respondido pela maior parte da população, e um questionário mais detalhado, aplicado apenas a uma amostra. Para preservar o anonimato dos respondentes, o IBGE só disponibiliza microdados para as amostras, publicando os dados do universo de modo agregado.

Informações sobre rendimentos foram divulgadas pela primeira vez para a amostra do Censo de 1960, que contou com uma única pergunta sobre o rendimento total dos indivíduos, codificada de forma categórica. O Censo 1970 também contou com um único item, mas as respostas foram registradas em uma variável contínua, o que se repetiu nas edições posteriores. Em 1980, o questionário da amostra foi expandido para captar seis fontes de rendimentos, número que caiu para quatro em 1991, subiu para sete em 2000 e foi novamente reduzido no Censo de 2010, que discriminou apenas três fontes de rendimentos.

O questionário básico incorporou rendimentos somente em 1991, sempre em um único item. Em 1991 e 2000, apenas o rendimento da pessoa de referência dos domicílios foi registrado; em 2010, a coleta foi estendida aos demais moradores.

Apesar de todas essas mudanças, a qualidade e a comparabilidade dos dados sobre rendimentos nos censos é boa, pelo menos para o questionário da amostra. Com efeito, os censos são mais compatíveis entre si do que com as outras pesquisas do IBGE, apresentando níveis de renda, desigualdade e pobreza maiores do que nas PNADs mais próximas. Pelo menos para os Censos 2000 e 2010 a evolução temporal também é distinta das PNADs, embora em menor grau. As evidências para os censos mais recentes indicam que essas discrepâncias são causadas por questões amostrais e de coleta e tratamento dos dados. Afinal, as amostras dos censos são desenhadas para garantir representatividade municipal e, por isso, são mais de 50 vezes maiores do que as amostras das antigas PNADs. Dessa forma, os censos conseguem ser muito mais representativos para os muito ricos do que as PNADs. No outro extremo, também há nos censos uma quantidade maior de famílias com rendimentos muito baixos, o que reflete tanto a concentração geográfica dos miseráveis quanto questões específicas de processamento dos dados do censo, especialmente em 2010.

Pesquisa Nacional por Amostra de Domicílios (PNAD)

Por décadas, a antiga Pesquisa Nacional por Amostra de Domicílios (PNAD) foi a principal pesquisa domiciliar produzida pelo IBGE. Sua primeira edição, ainda como um survey trimestral e de cobertura limitada, foi em 1967. Entre interrupções e reformulações, seu formato definitivo só foi adotado em 1976. A partir de então, a PNAD foi realizada anualmente até 2015, exceto em anos censitários (1980, 1991, 2000 e 2010) e em 1994, em função de cortes orçamentários. Depois de um longo período de transição, a antiga PNAD foi a campo pela última vez em 2015, sendo substituída pela nova PNAD Contínua.

O tamanho das amostras variou ao longo do tempo, oscilando em torno de 150 mil domicílios por ano nas últimas edições. Até 1979, a pesquisa não contemplava as áreas rurais das regiões Norte e Centro-Oeste e, entre 1981 e 2003, ficaram de fora as áreas rurais da região Norte, exceto Tocantins. Somente a partir de 2004 a PNAD passou ter representatividade nacional de fato. Embora isso tenha que ser levado em conta na construção de séries históricas, não se trata de um grande problema, já que as regiões excluídas nas primeiras décadas eram até então esparsamente povoadas.

O questionário expandiu-se ao longo do tempo, sem grande prejuízo para a comparabilidade da série histórica. A PNAD sempre foi uma pesquisa com propósitos múltiplos, contando com módulos detalhados para demografia, educação, mercado de trabalho e outros temas. As edições de 1976 a 1979 apresentaram mudanças constantes no módulo de rendimentos, mas as PNADs de 1981 a 1990 adotaram o mesmo padrão, que só foi um pouco mais desagregado em 1992, permanecendo idêntico até 2015.

Tamanha estabilidade não é inteiramente positiva, uma vez que a PNAD continuou pesquisando separadamente até 2015 algumas fontes de renda irrelevantes (como o abono de permanência dos servidores que podem se aposentar), ao mesmo tempo em que relegou para a categoria residual de “outros rendimentos” as transferências sociais criadas nas últimas décadas (como o Programa Bolsa Família e o Benefício de Prestação Continuada).

O problema potencialmente mais grave da antiga PNAD é que, em comparação com outras fontes de dados, ela tende a “achatar” a distribuição de renda ao subestimar consideravelmente os rendimentos dos mais ricos e, em menor grau, superestimar os rendimentos dos mais pobres.

PNAD Contínua

A PNAD Contínua (PNADC) foi a campo em caráter experimental em 2011, sendo implantada em definitivo a partir de 2012 para substituir tanto a antiga PNAD quanto a Pesquisa Mensal do Emprego (PME), que fornecia estatísticas mensais de mercado de trabalho para seis regiões metropolitanas.

A maior mudança da PNADC é o formato de painel rotativo, em que cada domicílio responde a cinco entrevistas, ao longo de cinco trimestres consecutivos, o que permite grande diversificação do questionário. Os rendimentos do trabalho são sempre pesquisados, mas os rendimentos de outras fontes são coletados somente na primeira e na última visita.

Atualmente, os microdados são divulgados em dois formatos: dados trimestrais, sobretudo com informações sobre mercado de trabalho, e versões anualizadas apenas com as primeiras e quintas entrevistas, com todos os rendimentos.

Em relação à antiga PNAD, a nova PNADC possui amostras maiores, com cerca de 210 mil domicílios entrevistados por trimestre, e mais espalhadas ao longo do território brasileiro, além de novidades em conceitos, no questionário e no próprio processo de coleta e processamento dos dados. Por isso, os resultados para várias variáveis não são diretamente comparáveis com os da série histórica da PNAD.

Infelizmente, uma das maiores quebras metodológicas ocorreu justamente com os dados sobre rendimentos. Em comparação com a PNAD, a PNADC reporta níveis mais elevados de renda e desigualdade. Até o momento não há consenso sobre procedimentos de compatibilização que permitam o encadeamento das duas fontes de dados.

Pesquisa de Orçamentos Familiares (POF)

A POF é a única pesquisa regular do IBGE que captura informações sobre consumo e despesas das famílias, além de variáveis sobre rendimentos. Suas origens remontam ao Estudo Nacional de Despesas Familiares (Endef), conduzido em 1974-1975, que por seu alto custo e complexidade, só voltou a campo, já com nome POF, em 1987-1988. Tanto essa segunda edição quanto a terceira, em 1995-1996, abrangeram somente um conjunto reduzido de regiões metropolitanas. As edições subsequentes, em 2002-2003 e 2008-2009, recuperaram a cobertura nacional.

Nenhuma pesquisa domiciliar brasileira coleta dados de rendimentos de modo mais detalhado do que a POF, que contém informações sobre os 12 meses anteriores à entrevista, e não apenas sobre um único mês de referência. A POF 2002-2003 discriminou mais de 80 fontes de rendimentos, enquanto a POF 2008-2009 ampliou esse número para cerca de 110. Rendimentos não monetários, 13º salário, e outros fluxos ignorados parcial ou completamente em outras fontes de dados são bem contemplados pela POF.

Com isso, a renda média *per capita* nas POF dos anos 2000 são entre 24% e 36% maiores do que nas PNAD mais próximas. A desigualdade também é mais elevada, mas a diferença é menor, com os números da POF ultrapassando os da antiga PNAD em 3%-4%.

Depois de muito atraso, uma nova rodada da POF foi a campo em 2017-2018, mas os microdados só devem ser divulgados em 2020.

Registros administrativos

Cadastro Único (CadÚnico)

O CadÚnico para programas sociais foi criado em 2001 com o objetivo de unificar e consolidar as informações sobre famílias de baixa renda no país, racionalizando o funcionamento dos programas sociais do Governo Federal. Por servir como instrumento de seleção e inclusão de famílias no Programa Bolsa Família, o CadÚnico deslanchou em meados dos anos 2000, representando hoje a via de acesso obrigatória para diversos programas sociais.

O CadÚnico é gerido em nível federal pelo Ministério da Cidadania e operado pela Caixa Econômica Federal. O cadastramento das famílias, por sua vez, é de responsabilidade dos municípios. São elegíveis para o CadÚnico famílias com renda *per capita* de até 1/2 salário mínimo ou renda familiar total de até três salários mínimos. As famílias cadastradas devem atualizar suas informações a cada dois anos, no máximo.

Com algo entre 70 e 90 milhões de pessoas cadastradas, o CadÚnico praticamente possui cobertura censitária da população de baixa renda no país. As informações coletadas seguem, de modo geral, os módulos da antiga PNAD, contemplando temas como mercado de trabalho, educação e rendimentos.

O CadÚnico é uma excelente fonte de dados para o estudo da pobreza, desde que

se leve em conta que *a)* as datas de referência variam entre famílias, *b)* há inúmeros incentivos para que famílias subestimem ou escondam rendimentos ao realizar seu cadastramento.

Imposto de Renda de Pessoas Físicas (IRPF)

Não há fonte de dados melhor para capturar os rendimentos dos mais ricos do que as informações do IRPF. No Brasil, o IRPF existe desde o início dos anos 1920, com tabulações publicadas regularmente, ainda que de forma pouco sistemática, pelo menos até o ano-calendário 2007, quando a Receita Federal passou a produzir estatísticas mais detalhadas sobre as declarações.

Como discutido acima, as informações do IRPF não estão isentas de problemas ou limitações. Para estudos sobre desigualdade, os dois principais pontos são a ausência de microdados e a cobertura limitada. Ao contrário de muitos outros países, os órgãos oficiais no Brasil sempre somente divulgaram tabulações agregadas por faixas de rendimentos, sem jamais permitir o acesso direto de pesquisadores externos a microdados. Com isso, o uso do IRPF para fins de pesquisa exige a realização de procedimentos de interpolação, dado que o número de declarantes nas faixas reportadas varia de ano para ano.

No segundo caso, a cobertura é limitada pelo fato de que nem todos os brasileiros são obrigados a declarar anualmente o IRPF. Hoje, os declarantes correspondem a algo em torno de 25% da população adulta. Mais ainda, devido às regras do IRPF, não necessariamente esses 25% são os mais ricos da população, de modo que os estudos sobre o tema tendem a considerar apenas os 10% ou, no máximo, os 15% mais ricos. Para tratar da distribuição de renda como um todo, portanto, é preciso combinar o IRPF com pesquisas domiciliares, o que vem sendo feito com frequência crescente.

Perguntas frequentes

O que significa volatilidade de renda e por que isso importa?

Volatilidade de renda é o termo usado para descrever as variações, muitas vezes imprevisíveis, que ocorrem nos rendimentos individuais e/ou familiares no curto e no médio prazo. O fenômeno tende a não afetar muito servidores públicos, aposentados e pensionistas, cujos rendimentos (ou ao menos parte deles) são garantidos todos os meses. Contudo, muitas outras categorias estão expostas a doses consideráveis de riscos, recebendo rendimentos que podem flutuar radicalmente de um mês para o outro. Essas flutuações podem ser positivas ou negativas.

Basta pensar no que acontece com lavadores de carros em períodos de chuva, com pequenos produtores agrícolas quando há quebras de safra e/ou com trabalhadores informais como um todo quando adoecem ou são demitidos. Evidentemente, trabalhadores formais também podem ser demitidos e empresários e profissionais liberais, entre outros, podem sofrer ou se beneficiar de choques repentinos. Porém, no Brasil, um país cuja proteção social ainda é muito vinculada ao trabalho, a volatilidade de renda se expressa sobretudo como risco e insegurança para os mais pobres.

Dada essa volatilidade, na prática, há indivíduos e famílias constantemente entrando e saindo da pobreza, de modo que a pobreza “transversal” é sempre inferior à pobreza “longitudinal”. Ou seja, se medirmos a pobreza em um dado momento do tempo, invariavelmente vamos obter um percentual inferior ao que seria observado se acompanhássemos as mesmas famílias por um intervalo mais longo e calculássemos quantas foram pobres em algum momento.

Por isso, a pobreza não deve ser pensada como uma característica fixa. Famílias pobres são muito semelhantes àquelas que estão pouco acima da linha de pobreza, havendo frequente troca de posições. Pequenas flutuações em termos absolutos podem mudar muito as posições relativas na metade de baixo da distribuição de renda.

O desenho das políticas sociais precisa levar isso em conta, como mostraram os trabalhos do economista Sergei Soares e outros pesquisadores, que influenciaram inclusive os critérios de permanência no PBF. Apesar disso, infelizmente, a volatilidade de renda ainda é pouco estudada no Brasil, em boa parte por falta de informações, problema parcialmente remediado com a chegada da PNADC.

O que é o Índice de Desenvolvimento Humano (IDH)?

O IDH é uma medida sintética de bem-estar apresentada pela primeira vez em 1990 no primeiro relatório anual de desenvolvimento humano publicado pelo Programa das Nações Unidas para o Desenvolvimento (PNUD). Inspirado pela teoria das capacidades de Amartya Sen, o índice foi pensado para permitir comparações internacionais para além da dimensão meramente econômica.

Sua fórmula é periodicamente revisada, mas desde o início o IDH contemplou três dimensões principais: saúde, educação e padrão de vida. A última revisão, em 2010, optou por operacionalizar essas dimensões a partir de indicadores como a expectativa de vida, escolaridade média e renda *per capita*. O IDH propriamente dito equivale à média geométrica de índices normalizados (que variam entre zero e 1) para cada uma das três dimensões.

Como tal, o IDH não é nem uma medida de pobreza nem de desigualdade – não obstante o lançamento oficial, em 2010, de uma versão do IDH “ajustada à desigualdade”. Além disso, o IDH compartilha os mesmos problemas de outros índices multidimensionais: falta de significado intuitivo (o que significa concretamente um IDH de 0,800?); opacidade (é impossível compreender variações sem desagregar o índice e examinar seus componentes); e arbitrariedade na escolha das dimensões, dos indicadores e dos parâmetros para normalização e agregação das dimensões.

Qual a diferença entre pobreza e extrema pobreza (ou miséria)?

A resposta depende da abordagem utilizada. Em estudos baseados em abordagens calóricas ou alimentares, a linha de extrema pobreza costuma ser definida como a renda mínima para suprir as necessidades calóricas dos indivíduos; já a linha de pobreza leva em conta também a renda necessária para cobrir despesas não alimentares.

No Brasil, a maior parte dos estudos das últimas décadas vem optando por usar linhas administrativas de pobreza. Nesse contexto, uma estratégia comum para reforçar a robustez das análises é apresentar resultados para pelo menos duas linhas de pobreza diferentes, de modo que a distinção entre “pobreza” e “extrema pobreza” não possui nenhum significado substantivo, apenas prático: a linha mais elevada é chamada de “linha de pobreza”, enquanto a linha mais baixa vira a “linha de extrema pobreza”.

As regras de elegibilidade de alguns programas sociais replicam essa lógica. A linha inferior do PBF, por exemplo, é muitas vezes chamada de “linha de extrema pobreza”, garantindo o acesso potencial a todos os benefícios do programa. Já a linha superior é considerada uma “linha de pobreza”, qualificando famílias apenas para os benefícios variáveis para crianças e adolescentes.

Como são calculadas as linhas de pobreza do Banco Mundial?

A primeira linha de pobreza internacional do Banco Mundial foi calculada em 1990 tomando como referência as linhas nacionais de alguns dos países mais pobres do mundo, convertidas em dólares internacionais usando taxas de Paridade de Poder de Compra (PPC) para levar em conta diferenças de custo de vida. Os pesquisadores observaram que, nos seis países mais pobres de sua amostra, a linha de pobreza nacional estava muito próxima de PPC\$ 31 por pessoa por mês, o que deu origem à linha de pobreza global de “um dólar por dia”.

Desde então, o Banco Mundial atualiza periodicamente o valor da linha, incorporando informações para mais países e novos fatores PPC estimados pelo *International Comparison Program*, que coleta regularmente dados sobre os preços de bens e serviços mundo afora.

A última revisão, em 2015, obteve uma linha de pobreza de PPC\$ 1.90 por pessoa por dia, calculada como valor médio arredondado das linhas de pobreza nacionais observadas em 15 países pobres, quase todos africanos, tomando como referência os fatores PPC de 2011.

Mais recentemente, duas novas linhas de pobreza complementares foram criadas, de PPC\$ 3.20 e PPC\$ 5.50 por pessoa por dia. A primeira reflete o valor mediano das linhas de pobreza nacionais em 32 países de renda média-baixa, conforme a classificação em vigor do Banco Mundial, o que inclui países como Bolívia, Índia, Indonésia e Paquistão. Já a linha de PPC\$ 5.50 segue o valor mediano das linhas de pobreza nacionais em outros 32 países considerados de renda média-alta, como África do Sul, Brasil, Chile e Turquia.

Nem todas as linhas nacionais tomadas como referência são linhas oficiais. A linha de pobreza atribuída ao Brasil, por exemplo, foi estimada com base na POF 2002-2003 pelo método de necessidades básicas, que leva em conta custos calóricos e despesas não alimentares. Como o próprio relatório original aponta, o valor estimado para o

Brasil (R\$ 103, em valores de 2003) era quase idêntico à linha superior do PBF vigente naquele ano (R\$ 100).

O que significa “erradicar a pobreza”?

O primeiro dos Objetivos de Desenvolvimento Sustentável (ODS) prevê até 2030 a erradicação da extrema pobreza, definida pela linha do Banco Mundial de PPC\$ 1.90 por pessoa por dia. Em sentido literal, o objetivo só será cumprido se não houver mais ninguém no mundo com renda inferior à linha de pobreza em qualquer momento do tempo. Na prática, as coisas são diferentes, por dois motivos principais: problemas de captação dos rendimentos (discutidos na seção 2) e problemas relacionados à volatilidade de renda (discutidos acima, nesta seção).

Para recapitular, no primeiro caso, coletar os rendimentos dos muito pobres é quase tão difícil quanto medir os dos extremamente ricos. A renda dos muito pobres em geral consiste de rendimentos informais e irregulares do trabalho e transferências de programas sociais, duas fontes normalmente subestimadas, seja pela fragilidade e incerteza dos vínculos, seja por medo ou estigma. Além disso, choques temporários na renda ou mesmo problemas de declaração muitas vezes fazem com que famílias não pobres apareçam nas pesquisas amostrais com rendimentos nulos ou muito baixos. Nos censos demográficos e nas antigas PNADs, por exemplo, a maior parte das famílias com renda domiciliar *per capita* igual a zero tem perfil socioeconômico típico de famílias mais privilegiadas, o que sugere se tratar ou de erros de declaração ou simplesmente de choques temporários (por exemplo, desemprego).

No segundo caso, como vimos, boa parte das famílias brasileiras convive com grandes variações na sua renda de um mês para o outro. A cada momento do tempo, novas famílias entram na pobreza, enquanto outras saem. Mesmo se os programas focalizados funcionarem perfeitamente, sempre há um intervalo entre o momento em que uma família se torna pobre e o momento em que ela começa a receber os benefícios a que tem direito. O retrato social apresentado pelas pesquisas domiciliares pode inflar a pobreza e não dá conta desse dinamismo do mundo.

Em última instância, o que se passa é muito semelhante ao que ocorre com outras estatísticas: por exemplo, falamos em “pleno emprego” mesmo quando a taxa de desemprego é maior do que zero, assim como melhorias em saúde nunca conseguem reduzir a mortalidade infantil estritamente para zero.

O próprio Banco Mundial, instituição que capitaneia os esforços de erradicação da pobreza nos ODS, reconhece que “erradicação da pobreza” significa a redução das taxas de pobreza para níveis residuais, próximos, mas não idênticos a zero. Mais especificamente, o Banco considera que o objetivo será atingido se a taxa global de pobreza em 2030 for inferior a 3%, “já que algum grau de pobreza friccional deve persistir¹”.

O limiar de 3% não é uma regra imutável: dependendo da qualidade dos dados de um país, ele pode ser mais alto ou mais baixo. O ponto central é que, para todos os fins, a pobreza pode ser considerada “erradicada” mesmo sem que as taxas oficiais sejam rigorosamente iguais a zero.

Qual a diferença entre desigualdade de renda e desigualdade de riqueza ou patrimônio?

Renda é um fluxo, riqueza é um estoque: enquanto a renda é o quanto uma pessoa ganha em determinado período de tempo, sua riqueza equivale ao estoque de bens e serviços acumulados ao longo de toda a vida. Mais ainda, é um estoque *líquido*, isto é, a soma dos bens e ativos menos a soma de dívidas e empréstimos contraídos.

Evidentemente, rendimentos podem ser poupados e convertidos em patrimônio, da mesma forma que a riqueza pode ser liquidada e transformada em renda para sustentar o consumo em tempos difíceis. Analogamente, quase todos os conceitos e medidas analisados neste capítulo podem ser aplicados tanto à distribuição de renda quanto à distribuição de riqueza.

Não obstante, essa diferença básica – fluxo vs. estoque – recomenda cautela na interpretação de dados de patrimônio por três motivos. Primeiro, exceto para os muito ricos, o que sustenta o padrão de vida e consumo das famílias é sua renda mensal, enquanto seu patrimônio acumulado serve para reduzir incertezas e, no máximo, produzir fluxos secundários e/ou indiretos de renda. Segundo, a correlação entre renda e riqueza está longe de ser perfeita, e os próprios preços dos ativos que compõem o patrimônio das famílias podem passar por flutuações radicais em função de condições macroeconômicas. Terceiro, o fato de ser um estoque acumulado ao longo de

1 Tradução livre do autor do trecho: “The World Bank Group’s goal to end extreme poverty within a generation has the specific target of decreasing the global extreme poverty rate to no more than 3 percent by 2030, since a small amount of frictional poverty is likely to persist.”. Disponível em <<https://www.worldbank.org/en/topic/poverty/overview#2>>. Acesso 9 set. 2019.

uma trajetória faz com que os efeitos do ciclo de vida e de políticas de crédito sejam amplificados quando se olha para a riqueza: em qualquer momento, mesmo que os jovens tenham as rendas mais elevadas, eles provavelmente vão ser mais pobres em média do que os mais velhos, pois não tiveram tempo suficiente para acumular ou herdar patrimônio e possivelmente contraíram dívidas para financiar estudos e/ou abrir negócios.

Por isso, estatísticas sobre a distribuição de riqueza são sempre mais controversas do que para renda. Por exemplo, todo ano a organização não governamental Oxfam divulga um relatório sobre desigualdade global de riqueza que invariavelmente resulta em manchetes escandalosas na imprensa. Assim, em 2018, alardeou-se que as 26 pessoas mais ricas do mundo têm patrimônio igual aos 50% mais pobres da população global, algo próximo de quatro bilhões de indivíduos.

Os críticos sempre reagem dizendo que esses números são duvidosos, uma vez que se baseiam em muitas estimativas e imputações. Mas as acusações mais graves refletem os problemas listados acima e podem ser resumidos em uma anedota: pela métrica da Oxfam, um bebê recém-nascido em uma família miserável em algum dos países mais pobres do mundo estaria melhor do que cerca de um terço da população global em termos de riqueza, uma vez que não possui dívidas. Ao mesmo tempo, um jovem casal recém-formado contratado por muitas centenas de milhares de dólares para trabalhar em Wall Street provavelmente estará entre os 10% mais pobres do mundo na distribuição de riqueza em virtude de dívidas com crédito estudantil e afins.

Nada disso invalida análises da distribuição de riqueza. Com efeito, a Oxfam em geral retruca que nenhuma conclusão substantiva se altera se americanos e europeus ricos em renda, mas pobres em patrimônio, forem retirados da análise. Seja como for, é evidente que é preciso cautela quando se olha para a riqueza, pois há dificuldades que inexistem quando o assunto é renda.

Referências bibliográficas

- ATKINSON, A. On the measurement of inequality. *Journal of Economic Theory*, v. 2, p. 244-263, 1970.
- ATKINSON, A.; PIKETTY, T.; Saez, E. Top incomes in the long run of history. *Journal of Economic Literature*, vol. 49, n. 1, p. 3-71, 2011.
- COBHAM, A.; SUMNER, A. Is it all about the tails? The Palma measure of income inequality. Center for Global Development, 2013. (Working Paper n. 343)
- COUDOUÉL, A.; HENTSHCEL, J.; WODON, Q. Poverty measurement and analysis. In: KLUGMAN, J. (Ed.) *A Sourcebook for Poverty Reduction Strategies*, vol. 1, Core Techniques and Cross-Cutting Issues. Washington, D.C.: Banco Mundial, 2002.
- COWELL, F. *Measuring Inequality*. Londres: Oxford University Press, 2011.
- COWELL, F. Income distribution and inequality. In: DAVIS, J.B.; DOLFSMA, W. (Eds.) *The Elgar Companion to Social Economics*. Northampton, Massachusetts: Elgar, 2015.
- FERREIRA, F.; et al. A global count of the extreme poor in 2012: data issues, methodology and initial results. *Journal of Economic Inequality*, v. 14, p. 141-172, 2016.
- FOSTER, J.; GREER, J.; THORBECKE, E. A class of decomposable poverty measures. *Econometrica*, v. 52, n. 3, p. 761-766, 1984.
- GRUPO DE CANBERRA. *Handbook on Household Income Statistics*. Geneva: Organização das Nações Unidas, 2011.
- HOFFMANN, R. *Distribuição de renda: medidas de desigualdade e pobreza*. São Paulo: EdUSP, 1998.
- HOFFMANN, R. Transferências de renda e redução da desigualdade no Brasil e em cinco regiões, entre 1997 e 2005. In: BARROS, R.P.; FOGUEL, M.; ULYSSEA, G. (Eds.) *Desigualdade de renda no Brasil: uma análise da queda recente* (vol. 2). Brasília: Ipea, 2007.
- KAKWANI, N. *Income inequality and poverty: methods of estimation and policy applications*. Nova York: Banco Mundial, Oxford University Press, 1980.
- LERMAN, R.; YITZHAKI, S. Income inequality effects by income source: a new approach and applications to the United States. *The Review of Economics and Statistics*, v. 67, n. 1, p. 151-156, 1985.
- MEDEIROS, M. *Medidas de Desigualdade e Pobreza*. Brasília: Editora UnB, 2012.
- ROCHA, S. A investigação da renda nas pesquisas domiciliares. *Economia e Sociedade*, v. 12, n. 2, p. 205-224, 2003.
- SALVERDA, W.; NOLAN, B.; SMEEDING, T. (Eds.) *The Oxford Handbook of Economic Inequality*. Oxford: Oxford University Press, 2009.
- SHORROCKS, A. The class of additively decomposable inequality measures. *Econometrica*, v. 48,

n. 3, p. 613-625, 1980.

- SHORROCKS, A. Inequality decomposition by factor componentes. *Econometrica*, v. 50, n. 1, p. 193-212, 1982.
- SOARES, S. Análise de bem-estar e decomposição por fatores da queda na desigualdade entre 1995 e 2004. *Econômica*, v. 8, n. 1, p. 83-115, 2006.
- SOARES, S. Metodologias para estabelecer a linha de pobreza: objetivas, subjetivas, relativas, multidimensionais. Brasília: Ipea, 2009. (Texto para Discussão n. 1381)
- SOUZA, P.H.G.F. A distribuição de renda nas pesquisas domiciliares brasileiras: harmonização e comparação entre Censos, PNADs e POFs. *Revista Brasileira de Estudos de População*, v. 32, n. 1, p. 165-188, 2015.
- YITZHAKI, S.; SCHECHTMAN, E. *The Gini methodology: a primer*. Nova York: Springer, 2013.



ANÁLISE DE DADOS EM SAÚDE¹

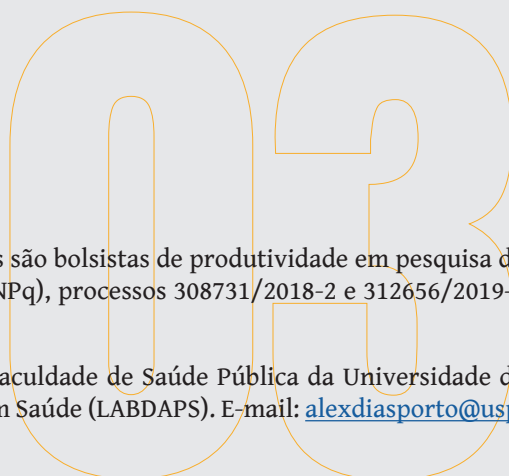
Alexandre Dias Porto Chiavegatto Filho²

Fredi Alexander Diaz-Quijano³

1 Esse trabalho não teve financiamento específico. Os autores são bolsistas de produtividade em pesquisa do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), processos 308731/2018-2 e 312656/2019-0 (ADPC e FADQ, respectivamente).

2 Professor Associado do Departamento de Epidemiologia, Faculdade de Saúde Pública da Universidade de São Paulo (FSP-USP). / Laboratório de Big Data e Análise Preditiva em Saúde (LABDAPS). E-mail: alexdiasporto@usp.br

3 Professor Associado do Departamento de Epidemiologia, Faculdade de Saúde Pública da Universidade de São Paulo (FSP-USP). / Laboratório de Inferência Causal em Epidemiologia da Universidade de São Paulo (LINCE-USP). E-mail: frediazq@usp.br



Introdução

A disponibilidade e análise de dados de qualidade em saúde é fundamental para o desenvolvimento, implementação e avaliação de intervenções clínicas e o estabelecimento de políticas públicas. A ciência da Epidemiologia envolve a análise da distribuição de eventos em saúde e de seus determinantes. O fundador honorário da área foi John Snow, que em 1854 começou a reparar que os casos de cólera de Londres estavam espacialmente concentrados em volta da rua Broad (1). Essa análise permitiu identificar que um poço de água estava no centro da epidemia. A descoberta não só fez uma contribuição ao entendimento do mecanismo causal da doença, mas também levou ao fechamento do poço diminuindo drasticamente o número de caso na cidade, salvando potencialmente milhares de vidas.

Boa parte das principais descobertas da área da saúde têm sido feitas pelo uso da análise de dados, como o clássico estudo de coorte prospectivo realizado por Doll & Hill, que em 1956 permitiu o primeiro resultado claro da ligação entre tabagismo e câncer de pulmão (2). Atualmente, epidemiologistas e cientistas de dados de saúde trabalham na linha de frente do combate a surtos de doenças infecciosas, na identificação de fatores de risco de doenças, em análises de custo-efetividades de novas tecnologias e medicamentos, e em avaliações do impacto de políticas públicas de saúde.

Nesta revisão apresentaremos os principais conceitos introdutórios para a análise de dados na área da saúde, incluindo algumas de suas particularidades e complexidades que devem ser consideradas em ciência de dados. O objetivo deste capítulo é auxiliar o público não científico como jornalistas e gestores públicos sobre como lidar com as bases de dados em saúde, fazendo um foco em dados frequentemente disponíveis pelos sistemas públicos de informação (dados secundários).

Fontes de dados de saúde

Segundo a fonte, podemos classificar os dados como primários ou secundários. **Dados primários** são aqueles que foram obtidos pelo pesquisador ou grupo de pesquisa que fará a análise, ou seja, por quem planejou a coleta e formulou as hipóteses do estudo. Os principais pontos negativos de utilizar dados primários são referentes ao alto custo de dinheiro e de tempo envolvidos devido à necessidade de treinamento e seleção da equipe de coleta, realização de uma pesquisa piloto e supervisão de todo o processo para evitar fraudes. No entanto, existem importantes pontos positivos, como o fato de poder controlar quais e como serão medidas as variáveis, podendo fazer um controle de qualidade das informações para diminuir vieses relacionados com a medição inadequada ou perdas de informação. Isso permite garantir a disponibilidade das variáveis relevantes para realizar as análises estatísticas.

Os **dados secundários** são aqueles coletados por outras pessoas que não são os pesquisadores que estão realizando a análise estatística. O contínuo aumento da coleta automática de grandes bancos de dados em saúde (o famoso *big data*) tem tornado esse tipo de dado cada mais popular e frequente. O principal ponto negativo, porém, é que os métodos e questionários de coleta já foram pré-estabelecidos e não podem ser alterados pelo pesquisador.

Em relação aos desfechos de interesse em saúde, as principais fontes podem ser divididas em três grupos: dados de nascimento, de mortalidade e de morbidade. A principal fonte de organização e distribuição desses dados no Brasil é o Departamento de Informática do Sistema Único de Saúde - SUS (DATASUS) criado em 1991. A maioria dos dados de abrangência nacional estão disponibilizados gratuitamente e sem necessidade de registro ou solicitação especial pelo site do DATASUS, por meio da plataforma TABNET, e em muitos casos também na forma de microdados (ou seja, com resultados detalhados para cada indivíduo).

Dados de nascimento

O primeiro desafio para utilizar dados de nascimento é definir precisamente o que é um **nascido vivo**. A Organização Mundial da Saúde utiliza uma definição que desde 1950 é adotada pela maioria dos países: “nascimento vivo é a expulsão ou extração completa, do corpo da mãe, independentemente da duração da gestação, de um produto de concepção, o qual, depois da separação, respire ou dê qualquer outro

sinal de vida, tal como batimentos do coração, pulsações do cordão umbilical estando ou não desprendida a placenta”.

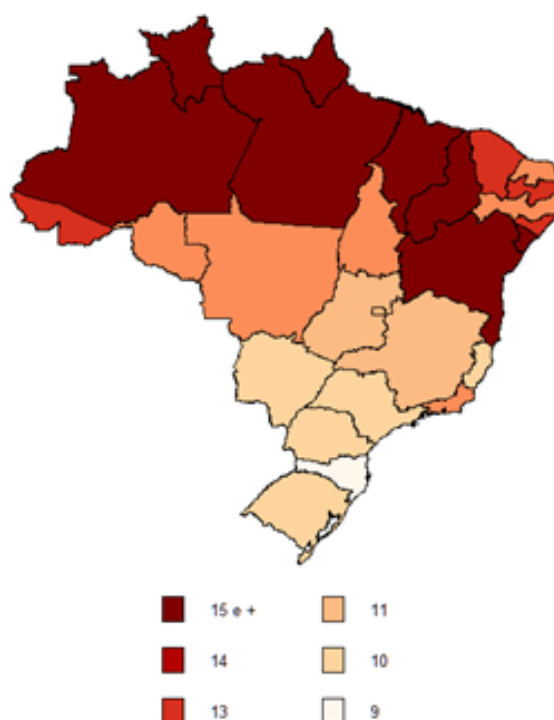
Essa definição é utilizada por alguns indicadores de saúde mais populares, entre eles o **coeficiente de mortalidade infantil** (CMI). Esse indicador é bastante utilizado para avaliar a situação de saúde por causa da vulnerabilidade de recém-nascidos a fatores externos. Em casos de grandes desastres e desestabilizações na atenção à saúde, como foi o terremoto do Haiti em 2010, as maiores vítimas são os recém-nascidos. Essa resposta imediata torna a mortalidade infantil um bom indicador para comparações locais e internacionais da atenção à saúde. O CMI é calculado da seguinte forma:

$$CMI = \frac{\text{número de óbitos menores de 1 ano, ano } X, \text{ local } Y}{\text{número de nascidos vivos, ano } X, \text{ local } Y} \times 1.000$$

Nota: a inclusão de ano X e local Y no numerador e denominador serve para ressaltar que os dois dados devem ser referentes ao mesmo período e local.

Os resultados para o CMI do Brasil segundo Estados em 2017 estão apresentados na Figura 1:

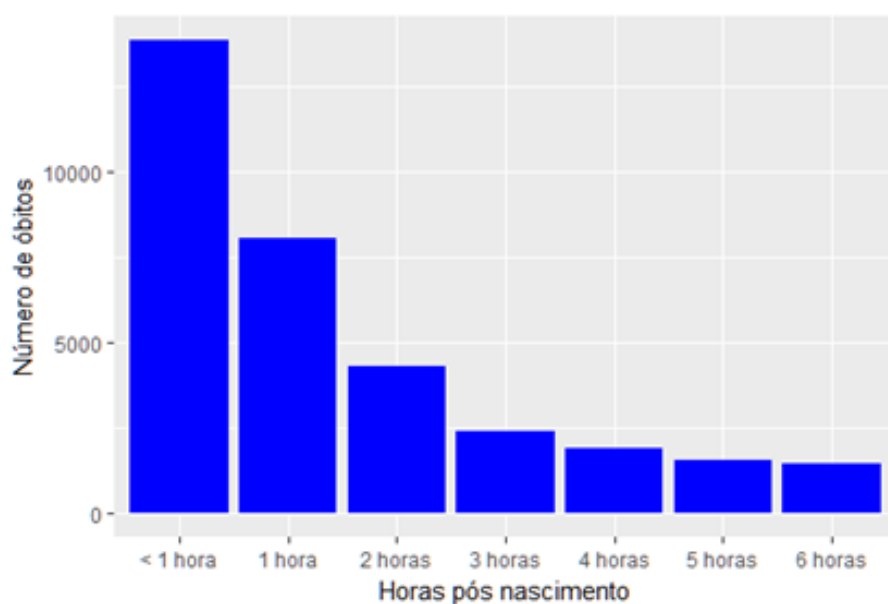
Figura 1 - Coeficiente de Mortalidade Infantil (por 1000 nascidos vivos) segundo Estados brasileiros, 2017.



Por ser muito utilizado como um indicador geral da atenção à saúde de um local, existem riscos de que esse resultado seja manipulado. Por exemplo, se alguns locais simplesmente alterarem a definição de nascido vivo e considerarem um recém-nascido que vai a óbito logo depois do parto como nascido morto, isso fará com que essa informação desapareça tanto do numerador quanto do denominador, e com isso diminuirá artificialmente o coeficiente (3,4).

A boa notícia é que essa manipulação, que pode ocorrer pontualmente em alguns hospitais ou mesmo países, é facilmente detectada por meio da análise da mortalidade segundo horas de vida. Como a maioria dos óbitos de recém-nascidos ocorre na primeira hora de vida, se esse valor estiver abaixo dos óbitos com 1-2h de vida é um possível alerta de que recém-nascidos que vão a óbito logo após o nascimento estão sendo considerados como nascidos mortos. Os dados de mortalidade nas primeiras horas de vida para o Brasil de 2013 a 2017 estão apresentados na Figura 2:

Figura 2 - Mortalidade nas primeiras horas de vida no Brasil, 2013-2017.



Os dados de nascidos vivos no Brasil são provenientes da **Declaração de Nascido Vivo** (DN), padronizada no Brasil desde a implantação em 1990 do Sistema de Informações sobre Nascidos Vivos (SINASC). A DN possui algumas informações relevantes sobre a saúde dos nascidos vivos no Brasil, como: peso ao nascer (recém-nascidos com menos do que 2.500g são considerados como **baixo peso**), idade gestacional (recém-nascidos com menos do que 37 semanas completas são considerados **pré-termo** e com 42 semanas ou mais são considerados **pós-termo**), número de consultas pré-na-

tal realizadas, tipo de parto (vaginal, cesáreo ou ignorado), entre outras.

Dados de mortalidade

Utilizar uma definição padronizada sobre **óbito** é também fundamental para permitir a comparação entre diferentes regiões de indicadores de mortalidade. Como no caso de nascimento vivo, a definição mais utilizada é a estabelecida pela OMS desde 1950: “óbito é o desaparecimento permanente de todo sinal de vida, em um momento qualquer depois do nascimento vivo, sem possibilidade de ressuscitação”.

Ao contrário da DN, que pode ser preenchida por qualquer profissional de saúde treinado para esse fim, a **Declaração de Óbito (DO)** é de responsabilidade do médico, principalmente devido ao fato de que nela consta uma das informações mais utilizadas da área da saúde: **a causa básica de óbito**. Sempre que aparecem notícias como “Número de homicídios caem no Brasil em 2019 em relação ao mesmo período de 2018”, ou “Santa Catarina já registra 457 mortes por câncer de mama em 2019”, a informação considerada é a causa básica.

A causa básica de óbito é definida como “doença ou lesão que iniciou a cadeia de acontecimentos patológicos que conduziram diretamente à morte, ou as circunstâncias do acidente ou violência que produziram a lesão fatal”. Ou seja, não se trata do evento terminal que levou à morte, ao contrário do que muita gente pensa, mas sim do agravo que iniciou todo o processo que eventualmente levou ao óbito. Do ponto de vista de saúde pública, é fundamental que o foco esteja no evento inicial, já que esse é mais fácil de ser evitado ou prevenido. Por exemplo, no caso de uma pessoa que tinha um câncer que evoluiu para falência múltipla dos órgãos, a causa básica nesse caso é o câncer e a falência múltipla de órgãos é incluída entre as **causas associadas de óbito**.

As causas básicas e associadas de óbito são listadas e referenciadas por meio dos códigos alfanuméricos da Classificação Internacional de Doenças da OMS, atualmente na sua 10ª Revisão (**CID-10**). Existem várias formas de agrupar as causas básicas de óbito. Utilizando-se os “óbitos segundo grupos” do Sistema de Informações sobre Mortalidade (SIM) do DATASUS as principais causas de óbito totais e para homens, mulheres, jovens (15-24 anos) e idosos (60 anos e mais) no Brasil em 2017 foram:

Total	Mulheres	Homens	Jovens	Idosos
1 Neoplasias malignas	Neoplasias malignas	Neoplasias malignas	Agressões	Neoplasias malignas
2 Doenças isquêmicas do coração	Doenças cerebrovasculares	Doenças isquêmicas do coração	Acidentes	Doenças isquêmicas do coração
3 Doenças cerebrovasculares	Doenças isquêmicas do coração	Agressões	Suicídios	Doenças cerebrovasculares
4 Gripe e pneumonia	Gripe e pneumonia	Acidentes	Neoplasias malignas	Gripe e pneumonia
5 Acidentes	Diabetes mellitus	Doenças cerebrovasculares	Gripe e pneumonia	Diabetes mellitus

Além da causa básica e causas associadas de óbito, a DO possui também outras informações relevantes disponibilizadas pelo DATASUS, como: idade, sexo, escolaridade, cor/raça, local de ocorrência (hospital, outro estabelecimento de saúde, domicílio, via pública e outros), estado civil, entre outras.

Uma questão importante para a análise de dados de nascimentos e de óbitos é estabelecer se os dados serão analisados segundo local de residência (no caso de nascimentos, residência da mãe) ou segundo local de ocorrência do evento. Por exemplo, o coeficiente de mortalidade por câncer (CMC) é calculado da seguinte forma:

$$CMC = \frac{\text{número de óbitos por câncer, ano } X, \text{ local } Y}{\text{população total, ano } X, \text{ local } Y}$$

Vamos supor que queremos calcular o CMC do município de Barretos, em São Paulo. Deve-se usar como numerador o número de óbitos de câncer que ocorreram em Barretos ou o número de óbitos de residentes de Barretos? Em saúde, em geral se analisam os dados de residência. Isso ocorre porque regiões com boas maternidades ou bons hospitais terão mais casos de nascimentos e de óbitos, respectivamente, porque pacientes de outras cidades se deslocam para serem atendidos nessas cidades de excelência. Se o CMC de Barreto for calculado segundo ocorrência, o resultado será alto e aparentará indicar que a cidade tem um sério problema na atenção ao câncer, quando na verdade o que ocorre é o contrário – as pessoas se deslocam para a cidade por ser um centro de referência no tratamento do câncer e algumas inevitavelmente acabam por falecer.

Dados de morbidade

A **morbidade** se refere ao conjunto de casos de doenças ou agravos à saúde em uma população. O Ministério da Saúde (MS) define doença como uma enfermidade ou estado clínico que represente ou possa representar um dano significativo aos seres humanos. Trata-se, portanto, de uma definição bem mais vaga daquela utilizada tanto para nascimentos vivos quanto para óbitos.

Medir doença é um desafio por várias razões. Em primeiro lugar, porque o significado de “dano significativo” (da definição do MS) pode depender de avaliações subjetivas. Em segundo lugar, algumas doenças reconhecidas internacionalmente têm mudado e podem ser consideradas um conceito histórico. Por exemplo, até a CID-9,

homossexualidade era uma doença incluída dentro da categoria “Desvios e Transtornos Sexuais” (a data do Dia Internacional Contra a Homofobia é 17 de maio em comemoração ao dia em que a CID-10 foi aprovada e homossexualidade foi retirada da lista de doenças). Por outro lado, existem também diversos casos de doenças que não eram reconhecidas como tal até recentemente, principalmente aquelas relacionadas à saúde mental. Em terceiro lugar, é difícil medir doença porque ela depende da existência de um diagnóstico, o que nem sempre é possível principalmente em regiões mais pobres.

Apesar da existência desses desafios, é importante fazer um esforço para medir a morbidade para auxiliar no controle de surtos de doenças infecciosas, planejar políticas públicas preventivas, estimar e controlar gastos futuros em saúde, conhecer a etiologia das patologias, identificar a influência de fatores sociais na distribuição de agravos, entre outros usos.

Ao contrário dos dados de nascimento e óbito que possuem um sistema unificado de coleta e divulgação de dados no Brasil, existem diversas fontes de dados de morbidade no Brasil, cada uma com suas limitações e pontos positivos. As principais fontes dessas informações no Brasil são a notificação compulsória de doenças, as estatísticas hospitalares, as estatísticas ambulatoriais e os inquéritos domiciliares.

A **notificação compulsória** de doenças é uma comunicação obrigatória aos órgãos de saúde de casos suspeitos ou confirmados de doenças específicas. O principal objetivo é evitar a proliferação de surtos, razão pela qual a maioria dos 48 grupos de doenças da lista são infecciosas e também pela qual casos suspeitos (ainda não confirmados) devem ser notificados. A lista pode sofrer alterações, sendo as mais frequentes a inclusão da notificação de doenças emergentes como ebola, Zika e doença de Chagas crônica.

Existem, entretanto, vários problemas referentes à subnotificação das doenças. Entre as causas de subnotificação, está o grande número de agravos presentes na lista, o que às vezes gera alguma confusão e desconhecimento. Adicionalmente, existe uma tendência a que apenas os casos mais graves das doenças seja notificados, o que dificulta o conhecimento sobre a carga global das doenças (5). Adicionalmente, eventos com elevada incidência ou emergentes, são suscetíveis a subregistro e a baixa confirmação principalmente quando o volume de casos ultrapassa as capacidades de atendimento e diagnóstico, como o que tem acontecido durante a pandemia de covid-19 (6).

O **Sistema de Informações Hospitalares (SIH)** recebe as informações de internações realizadas em hospitais participantes do SUS (públicos ou particulares conveniados), sendo consolidadas e disponibilizadas online pelo DATASUS. O objetivo dessa informação é financeiro (ressarcimento do custo de internação), mas são também disponibilizadas informações sobre os procedimentos realizados e a causa principal da internação, codificada por meio da CID-10.

O **Sistema de Informações Ambulatoriais (SIA)** também tem uma finalidade financeira, mas nesse caso referente a atendimentos ambulatoriais no SUS. Assim como os resultados do SIH, é também disponibilizada pelo sistema TABNET do DATASUS, com um grande número de variáveis como procedimentos realizados, forma de financiamento e esfera administrativa do atendimento.

Os **Inquéritos Domiciliares** normalmente selecionam uma amostra de domicílios em que um questionário validado e estruturado é aplicado por meio de perguntas ou por medições diretas feitas por profissionais de saúde treinados. O principal inquérito domiciliar de saúde hoje no Brasil é a Pesquisa Nacional de Saúde (PNS) cuja primeira versão foi feita em 2013 e a segunda em 2019. Trata-se de uma amostra representativa dos moradores de domicílio particulares do Brasil coordenada pelo IBGE, que realiza a coleta de dados de questionários, exame de sangue e medição de peso, altura, circunferência da cintura e pressão arterial. Os resultados da coleta de 2013 estão disponíveis no site do Sistema IBGE de Recuperação Automática (SIDRA).

Usos das informações

Entre os principais usos das informações em saúde está caracterizar uma população corretamente em relação a uma condição de saúde ou a um fator de risco para doença. Isso permite identificar prioridades, grupos vulneráveis e potenciais alvos de intervenção. Dessa forma, critérios usados para considerar doenças como prioritárias incluem a sua frequência, gravidade, efeito em algumas subpopulações como crianças, e a factibilidade de realizar alguma prevenção efetiva.

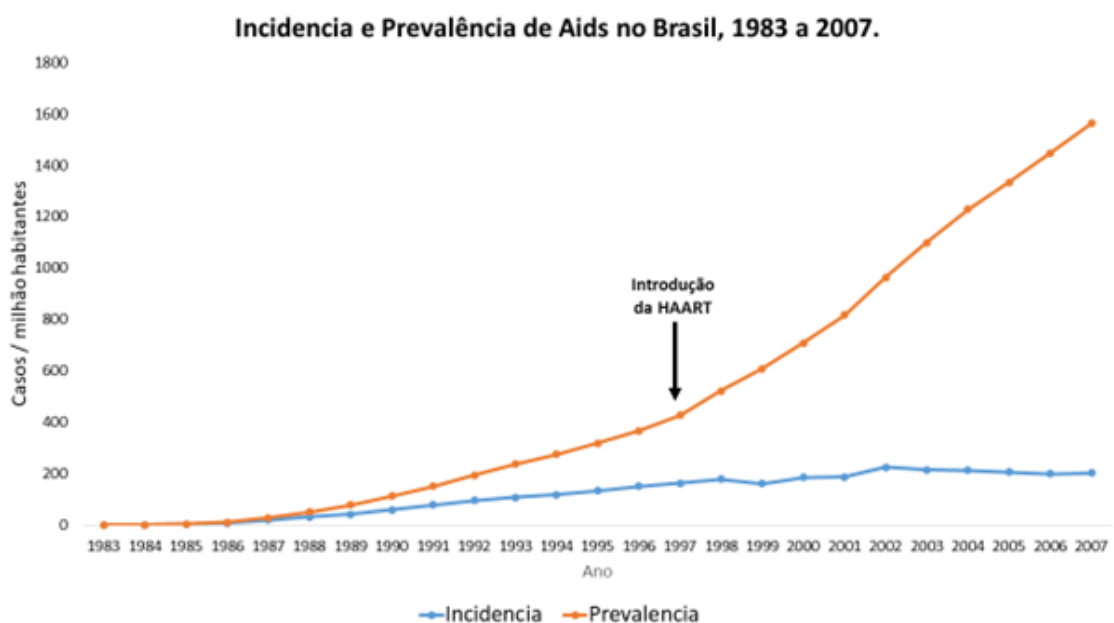
Usualmente as frequências dos problemas de saúde são descritas em função de variáveis relativas ao tempo, ao lugar, ou segundo características populacionais como as demográficas. Em cada caso, dependendo da natureza da variável podem ser escolhidas diferentes formas de resumir a informação de uma população. Assim, podemos usar proporções para descrever incidência ou a prevalência de eventos tais como uma

doença ou a morte. Dessa forma, denominamos **incidência acumulada** a proporção de pessoas que desenvolvem um evento novo durante um período de seguimento. Por outro lado, a **prevalência pontual** se refere à proporção de pessoas vivendo com problemas de saúde no momento específico.

A prevalência representa a extensão atual de um problema de saúde numa comunidade, independentemente do tempo que cada uma das pessoas apresenta esse agravo. Dessa forma, a prevalência depende diretamente tanto da incidência (aparição de casos novos) como da duração da doença. Sob condições relativamente estáveis, a prevalência é aproximadamente o produto da incidência e a duração da doença ($\text{Prevalência} \approx \text{Incidência} \times \text{Duração}$) (1).

Um aumento da prevalência não necessariamente representa um aumento do risco na comunidade, já que pode ser o resultado de um aumento da sobrevida no contexto de uma doença que não se cura facilmente. Assim, por exemplo, consideremos o caso da aids no Brasil. A incorporação da terapia antirretroviral altamente ativa (Highly Active Antiretroviral Therapy: HAART), no final de 1996, levou a uma queda acentuada da letalidade da doença. Como resultado, houve um aumento progressivo da sua prevalência, não por um aumento do risco, mas por um aumento da sobrevida dos pacientes (sem curá-los).

Figura 3 - Incidência e Prevalência de Aids no Brasil, 1983 a 2007.



Na Figura 3, é apresentada a evolução aproximada da prevalência de casos de aids no Brasil (linha laranja) de 1983 a 2007 (calculada a partir do acumulado de casos

novos menos a mortalidade registrada). Nos primeiros anos, o aumento progressivo da prevalência se explica essencialmente pela incidência crescente (linha azul). No entanto, após a introdução de HAART, a prevalência apresentou um aumento mais pronunciado devido ao aumento da sobrevida, apesar da incidência ter estabilizado nos anos seguintes.

Exemplos similares podem ser observados em outros tipos de doenças crônicas (geralmente incuráveis), onde a incorporação de tecnologias terapêuticas aumenta a sobrevida, como no caso da diabetes, doenças cardiovasculares, entre outras. Dessa forma, embora a prevalência seja uma medida da extensão do agravo numa comunidade e, por tanto, da necessidade de serviços específicos, seu aumento não necessariamente se explica por uma piora das condições de saúde.

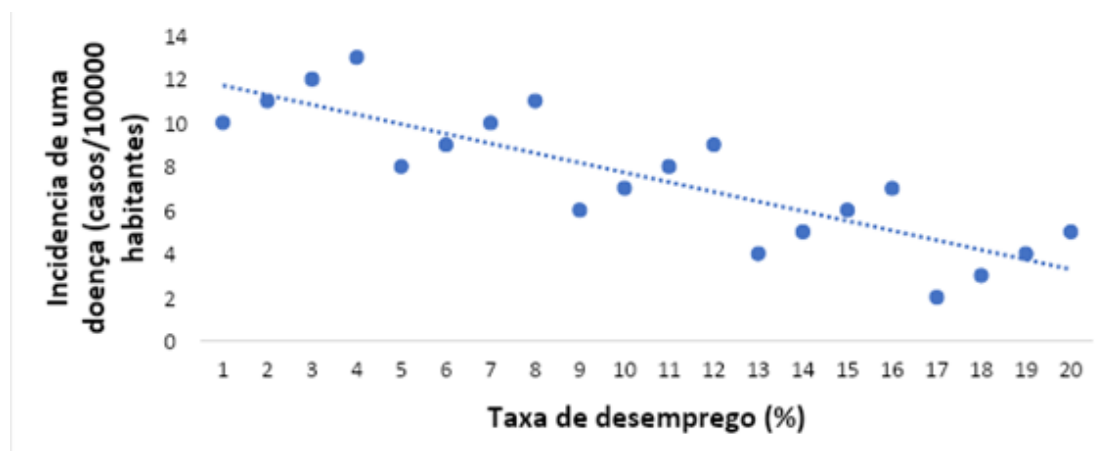
Outro cuidado a ser considerado na avaliação de tendências é a comparação dos valores observados em relação aos esperados para uma localidade e tempo específico. Muitas doenças têm padrões sazonais (mudanças previsíveis ao longo de um ano) ou seculares (tendências esperáveis através dos anos). Assim por exemplo, um aumento de casos não necessariamente representa um surto ou epidemia se a tendência faz parte do esperado para um momento específico do ano. Por outro lado, uma queda da incidência de casos de uma doença infecciosa nem sempre representa um sucesso das medidas de controle disponíveis. Em muitas ocasiões, após uma epidemia que alcançou seu auge, costuma haver uma diminuição de casos em decorrência da redução de pessoas suscetíveis (seja devido à imunidade adquirida ou morte das mesmas).

Importância de uma abordagem multinível

A partir das informações disponíveis, cada vez mais temos indicadores para representar os problemas de saúde e formular hipótese sobre seus determinantes. No entanto, a organização dos indivíduos em agregados coletivos é um aspecto que deve ser considerado na interpretação das análises de dados (7).

Um erro frequente, conhecido como falácia ecológica, ou viés de agregação, é decorrente de inferir a um nível individual um fenômeno observado em nível grupal. A situação mais comum da falácia ecológica ocorre quando uma variável de grupo é usada como substituta de dados desconhecidos no nível individual. Considere o exemplo hipotético representado na seguinte figura onde se observa uma correlação entre medidas agregadas de desemprego e incidência de uma doença.

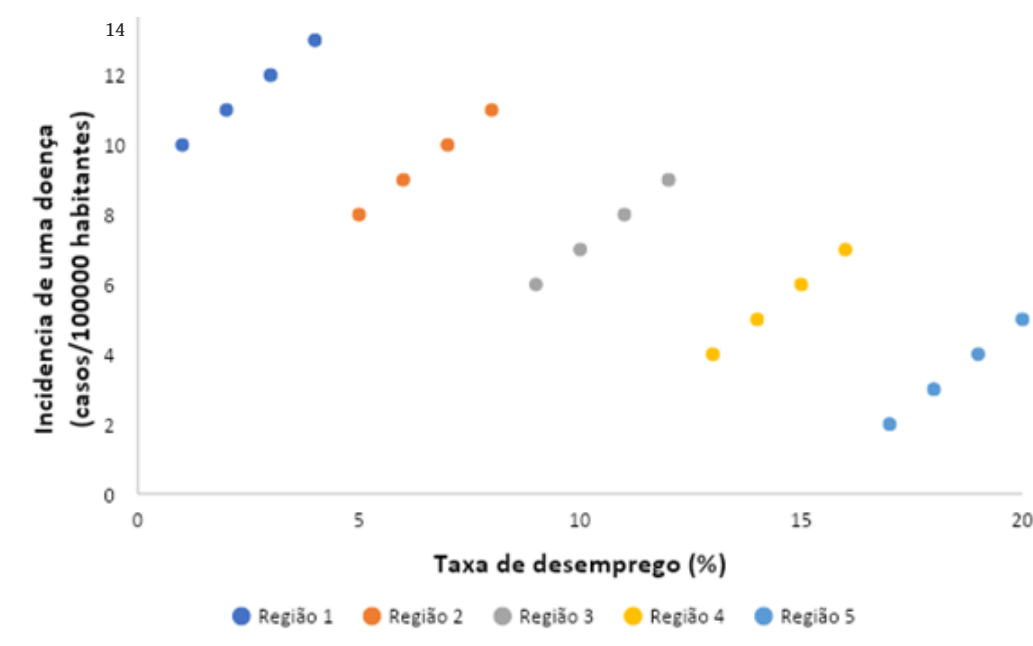
Figura 4 - Correlação hipotética entre a taxa de desemprego e a incidência de uma doença, em um conjunto de cidades.



Cada um dos pontos representa uma cidade e a linha cortada mostra a aparente tendência de uma correlação negativa, onde cada aumento em um ponto da taxa de desemprego diminui aproximadamente 0,44 casos da doença por 100.000 habitantes.

No entanto, essa interpretação desconsidera eventuais fenômenos de agregação. Agora considere que os dados da figura anterior correspondem a cidades de cinco regiões diferentes e considere avaliar as tendências em cada uma dessas regiões separadamente (Figura 5).

Figura 5 - Correlação hipotética entre a taxa de desemprego e a incidência de uma doença, em um conjunto de cidades agrupadas em regiões.



Como se observa, a taxa de desemprego se associou positivamente com a incidência do problema de saúde em cada uma das regiões separadamente. Esse resultado paradoxal⁴ é um exemplo extremo, mas ilustra a necessidade de considerar esses níveis de agregação para evitar o viés que levaria a conclusões erradas.

Inferência causal

Entre os principais usos dos dados em saúde está a identificação de grupos de risco e dos possíveis determinantes das doenças. Esse tipo de conhecimento permite orientar intervenções sobre alvos específicos. Nesse sentido, a evidência pertinente deve considerar relações de causalidade, seja entre determinantes e doenças, ou entre intervenções e resultados desejáveis.

Uma associação implica resultados diferenciais segundo gradientes de exposição. Assim, espera-se que pessoas mais expostas a um fator de risco (e.g. hipertensão) tenham uma maior frequência de doenças (e.g. cardiovasculares). No entanto, nem todas as associações são causais e nas pesquisas epidemiológicas devem ser levantados questionamentos tais como:

1. A associação poderia ser atribuível a algum viés, seja na seleção das observações ou na coleta das informações?

Em relação a isso, os pesquisadores devem fundamentar suas medições em estudos bem desenhados evitando erros sistemáticos relacionados à escolha de participantes ou perdas de seguimento, e também implementando métodos acurados para realizar as medições.

2. Mesmo na ausência de vieses, poderia existir alguma outra variável ou mecanismo alternativo que explique a associação?

Isso se refere ao fenômeno conhecido como “confusão⁵” em que variáveis não consideradas poderiam explicar uma associação. Os pesquisadores devem fazer um pla-

4 Fenômenos como esse, no qual a associação marginal tem uma direção oposta àquela condicionada (estratificada ou ajustada) por uma terceira variável, são frequentemente conhecidos como o paradoxo de Simpson (8,9).

5 Leituras recomendadas sobre confusão: capítulo 15 do livro Epidemiologia, Gordis L. (1); o artigo “Graphical presentation of confounding in directed acyclic graphs”, Suttrop MM (10)

nejamento adequado dos seus estudos para considerar todas as variáveis pertinentes e fazer os controles adequados para estimar os parâmetros.

3. Em ausência do anterior, a associação poderia ser por variabilidade aleatória?

Depois de ter descartado vieses e fenômenos de confusão, o pesquisador deve prestar atenção à possibilidade de que a diferença entre os grupos seja explicada por acaso. Para isso, os intervalos de confiança representam o erro amostral. Adicionalmente, o chamado “valor p” corresponde a uma probabilidade calculada de obter resultados como os observados ou mais extremos na situação hipotética de ausência de associação.

Para cada um desses questionamentos, os pesquisadores devem abordar as particularidades de cada problema específico de saúde, as características da população, a evidência prévia e os propósitos de estudo. Assim, a postulação de uma relação causal e as intervenções derivadas dela devem estar suportadas por um corpo de evidência consistente, baseado na melhor informação disponível e factível de ser obtida.

Considerações finais

Diversas fontes de dados podem ser consultadas para serem obtidas medidas de frequência de eventos de interesse em saúde. Essas informações são relevantes para quantificar e monitorar a carga causada pelas doenças. Além disso, sua análise tem o potencial de ajudar na identificação de seus determinantes e de grupos de risco, e na avaliação de intervenções. É essencial que as abordagens analíticas levem em consideração o conhecimento específico das patologias em estudo, assim como aspectos metodológicos para mitigar o risco de vieses. Em consequência, recomenda-se que os cientistas de dados se apoiem em especialistas da área com o conhecimento fisiopatológico, clínico e epidemiológico suficiente para os devidos planejamento e interpretação das análises.

Referências bibliográficas

- (1) Gordis L. Epidemiology. 5th ed. (Gordis L, ed.). Philadelphia: ELSEVIER; 2014.
- (2) Doll R, Hill AB. Lung Cancer And Other Causes Of Death In Relation To Smoking A Second Report On The Mortality Of British Doctors. Br Med J. 1956;2(5001):1071. doi:10.1136/bmj.2.5001.1071
- (3) Gonzalez RM, Gilleskie D. Infant Mortality Rate as a Measure of a Country's Health: A Robust Method to Improve Reliability and Comparability. Demography. 2017;54(2):701-720. doi:10.1007/s13524-017-0553-7
- (4) Berdine G, Geloso V, Powell B. Cuban infant mortality and longevity: health care or repression? doi:10.1093/heapol/czy033
- (5) Romero-Vega L, Pacheco I O, de la Hoz-Restrepo F, Díaz-Quijano FA. Evaluation of dengue fever reports during an epidemic, Colombia. Rev Saude Publica. 2014;48(6):899-905. doi:10.1590/S0034-8910.2014048005321
- (6) Diaz-Quijano FA, Silva JMN da, Ganem F, Oliveira S, Vesga-Varela AL, Croda J. A model to predict SARS-CoV-2 infection based on the first three-month surveillance data in Brazil. Trop Med Int Health. 2020 Nov;25(11):1385-1394. doi: 10.1111/tmi.13476.
- (7) Ana V. Diez Roux. La necesidad de un enfoque multinivel en epidemiología. Región y Soc. 2008;20(SPE2):77-91. http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1870-39252008000400004. Accessed October 16, 2019.
- (8) Hernán MA, Clayton D, Keiding N. The Simpson's paradox unraveled. doi:10.1093/ije/dyr041
- (9) Julious SA, Mullee MA. Confounding and Simpson's paradox. BMJ. 1994;309(6967):1480. doi:10.1136/bmj.309.6967.1480
- (10) Suttorp MM, Siegerink B, Jager KJ, Zoccali C, Dekker FW. Graphical presentation of confounding in directed acyclic graphs. Nephrol Dial Transpl. 2015;30:1418-1423. doi:10.1093/ndt/gfu325

GAP

EDUCAÇÃO

Ilona Becskeházy

04

Introdução

Uma das coisas que precisamos saber sobre o setor educacional brasileiro, particularmente o público, é que há pelo menos um aspecto que não nos faz passar tanta vergonha em comparações internacionais: a nossa capacidade de aplicar provas padronizadas em larga escala, de processar as informações obtidas e de transformá-las em enormes bases de dados que ficam à disposição on-line para quem por elas se interessar, como, por exemplo, autoridades, pesquisadores e imprensa.

Isso não quer dizer que as questões que aferem o desempenho dos alunos e os questionários contextuais sejam brilhantes. Contudo, ao longo do tempo, as autoridades educacionais em nível federal, estadual e, em alguns casos, até municipal, conseguiram desenvolver a capacidade de preparar e aplicar não apenas exames que podem atestar, em diferentes níveis de agregação, a proficiência dos alunos, mas também formulários capazes de captar informações sobre variados aspectos da vida escolar e pessoal deles e de suas famílias, escolas e condições de ensino (cada um desses instrumentos podendo ser preenchido tanto por alunos, quanto por seus pais, professores, diretores e até secretários de educação).

Essa política de usar dados em larga escala para compreender o que acontece dentro e no entorno das escolas (principalmente) públicas no Brasil começou no início da década de 1990. Um dos maiores defensores do uso de bons dados para melhorar a qualidade e a equidade das escolas no país foi Sérgio Costa Ribeiro. Esse brilhante pesquisador, falecido precocemente, escreveu, entre inúmeras produções importantíssimas¹, o icônico artigo “A pedagogia da repetência²”, em que cunhou o termo que resume a prática naturalizada, apesar de equivocada e abjeta, de selecionar os alunos “mais capazes” deixando quase todo o resto “na chuva” em termos de atendimento pedagógico.

Mostra-se que as taxas de evasão escolar, ao contrário do que indicam os dados do MEC, são muito baixas nas primeiras séries do 1º Grau. Este fato faz com que a visão de que as crianças no Brasil abandonam precocemente a Escola por motivos de ordem social ou cultural está completamente equivocada e que as famílias fazem um

1 <http://www.sergiocostaribeiro.ifcs.ufrj.br/artigos.html>

2 <http://www.sergiocostaribeiro.ifcs.ufrj.br/artigos/1991%20Pedagogia%20da%20repetencia.pdf>

esforço enorme para que seus filhos permaneçam na escola e que é a péssima qualidade da escola que impede, através da repetência, a universalização da Educação Fundamental da população brasileira (RIBEIRO, 1991, p. 20).

Essa contextualização é importante para esclarecer a motivação institucional e política de se investir em sistemas de coleta de informação educacional de larga escala: informar os agentes sobre a evolução de políticas públicas que são de interesse geral, para melhor informar as tomadas de decisão.

O histórico e as informações sobre o mais importante sistema de coleta de informações de desempenho estudantil e de contexto do Brasil podem ser encontrados no site do Ministério da Educação. O Sistema de Avaliação da Educação Básica (Saeb³), que começou em 1990 e vem sendo expandido, aprofundado e aprimorado desde então, tem abrangência nacional e é censitário para as séries participantes desde 2005. Além do Governo Federal, todos os estados contam com sistemas locais para o mesmo fim, implementados, obviamente, com algumas diferenças (uma vez que estão sob o comando de governadores e secretários de cada unidade da Federação), mesmo que uma boa parte deles seja operada pela mesma empresa e use a mesma escala de proficiência do Saeb. O mais longo e contínuo deles é o do Estado do Ceará⁴, que começou a fazer avaliações em larga escala em âmbito estadual em 1992.

Além do sistema nacional e dos subnacionais que capturam informações sobre o desempenho dos alunos e fatores associados, existem alguns sistemas internacionais correlatos de que o Brasil participa. O mais famoso deles é o Programa Internacional de Avaliação de Estudantes (Pisa), que avalia uma amostra de alunos brasileiros de 15 anos desde sua primeira edição, em 2000. Mas ainda há a Pesquisa Internacional sobre Ensino e Aprendizagem (Talis), desde 2008 e, mais recentemente, foi anunciada a participação dos alunos brasileiros cursando o 4º ano do ensino fundamental no Progress in International Reading Literacy Study (PIRLS) em sua próxima edição, de 2021. Tanto as avaliações nacionais, quanto as internacionais estão sob responsabilidade do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), uma autarquia federal vinculada ao Ministério da Educação (MEC).

Que aspectos devemos levar em conta quando lemos matérias na imprensa e até

3 <http://portal.inep.gov.br/web/guest/educacao-basica/saeb>

4 Sistema Permanente de Avaliação da Educação Básica do Ceará (SPAECCE): <http://www.spaece.caedufjf.net/>

artigos acadêmicos que utilizam esse tipo de dados? Os principais estão listados abaixo e serão detalhados a seguir:

1) Alunos, turmas, escolas, municípios e estados não são iguais – existem diferenças importantes, por exemplo, de nível socioeconômico (NSE) dentro de cada um desses níveis de desagregação de dados. Reconhecidamente, essa é a categoria de características que mais interfere no desempenho escolar dos alunos, de sua turma, escola, município ou estado (infelizmente, pois uma das funções republicanas da instituição escola – filosoficamente elaborada no espírito de igualdade – é promover a equidade entre alunos de diferentes origens). A relação costuma ser diretamente proporcional: quanto mais baixo o NSE, mais baixos os parâmetros de proficiência e vice-versa.

2) Dentro desses mesmos níveis de desagregação, pode haver outras diferenças substanciais, como o nível de complexidade de gestão: tamanho da clientela atendida, distribuição territorial das unidades escolares, existência de comunidades com particularidades educacionais etc., que adicionam percalços a professores e gestores, para além do desafio usual que é ensinar. Ainda mais quando há a associação desses fatores ao item anterior.

3) O Índice de Desenvolvimento da Educação Básica (Ideb) é um índice que agrega desempenho dos alunos com o fluxo deles (passar de um ano para o outro, sem repetir ou sair da escola) em cada nível de desagregação já citado. Nesses níveis, há autoridades pedagógicas que ainda lutam para vencer as mazelas decorrentes da já citada “pedagogia da repetência” – turmas, escolas, municípios e estados que apresentam altos níveis de evasão e repetência – e há os que, meritoriamente, já venceram essa etapa e partem para melhorar a aprendizagem. Em relação ao Ideb, isso quer dizer que, quando o fluxo tende a ter 100% de alunos passando de um ano escolar para o seguinte, o Ideb tende a refletir simplesmente a média das proficiências nas provas (por enquanto, Língua Portuguesa e Matemática). Mas quando um Ideb é inicialmente baixo e as notas de proficiência não estão subindo, o truque de “passar todo mundo” pode dar a impressão de que a qualidade da educação naquele determinado recorte está melhorando – matematicamente, em vez de qualitativamente – porque, ao se multiplicar uma média padronizada de notas por um número mais próximo de 100% (que representa percentual de fluxo), o desconto do valor em cima da nota de proficiência é menor e seu Ideb “aumenta” de uma medição para a seguinte, mesmo que os alunos não estejam realmente aprendendo mais.

4) Ideb = 6,0 no Brasil não é o mesmo que o desempenho médio dos alunos de 15 anos dos países da Organização para a Cooperação e Desenvolvimento Econômico

(OCDE) que fazem o já mencionado Pisa. Isso se explica porque as perguntas que se faz na Prova Brasil nas provas de leitura e de Matemática – a partir da edição de 2019, de Ciências também – são muito diferentes daquelas que formam o exame do Pisa. O nível das perguntas das provas padronizadas brasileiras é muito menos complexo que o das provas do exame gringo.

5) É muito comum ler e ver comentários, debates, inserções em novelas, conversas no almoço de domingo ou no bar etc. que sugerem que as “competências do séc. 21” sejam mais importantes que “apenas” saber o que consta das disciplinas tradicionais da escola. É preciso tomar conhecimento do que está especificado nessas disciplinas em países desenvolvidos para compreender que a maior parte do que se espera que os alunos aprendam para viver bem neste século já está previsto em suas normativas curriculares, o que é complementado por aprender a conviver em um ambiente institucional, por meio de exemplos e de normativas que, em geral, pressupõem regras diferentes das que vigem nas famílias e outros espaços que os alunos frequentam.

Proficiência e nível socioeconômico (NSE)

Há muita literatura sobre a associação entre o NSE de alunos e seu desempenho acadêmico. É importante saber que o que deveria estar sendo perseguido obsessivamente em cada uma das escolas no Brasil é a equidade decorrente da eficácia escolar, ou seja, a neutralização dessa associação deletéria entre origem social e proficiência escolar dentro das escolas. Ainda estamos longe de consegui-lo, embora já comecemos a aparecer casos de sucesso nesse aspecto.

Assim, quando se lê um ranking de proficiência, de Ideb ou afins, é possível que se esteja lendo, sem perceber, um ranking de NSE. É preciso ficar atento. Os dados deveriam ser divulgados de forma a evidenciar essa relação, até para valorizar as turmas, escolas, municípios e estados que já superaram o desafio da equidade por meio da eficácia escolar. Para dar acesso a pesquisadores e jornalistas a essa associação e facilitar a sua identificação, o Inep criou em 2015 o Indicador de Nível Socioeconômico das Escolas de Educação Básica (Inse). A Tabela 1 apresenta a descrição dos três níveis de Inse nos quais se localizam a quase totalidade dos alunos brasileiros.

Tabela 1 – Níveis selecionados de INSE

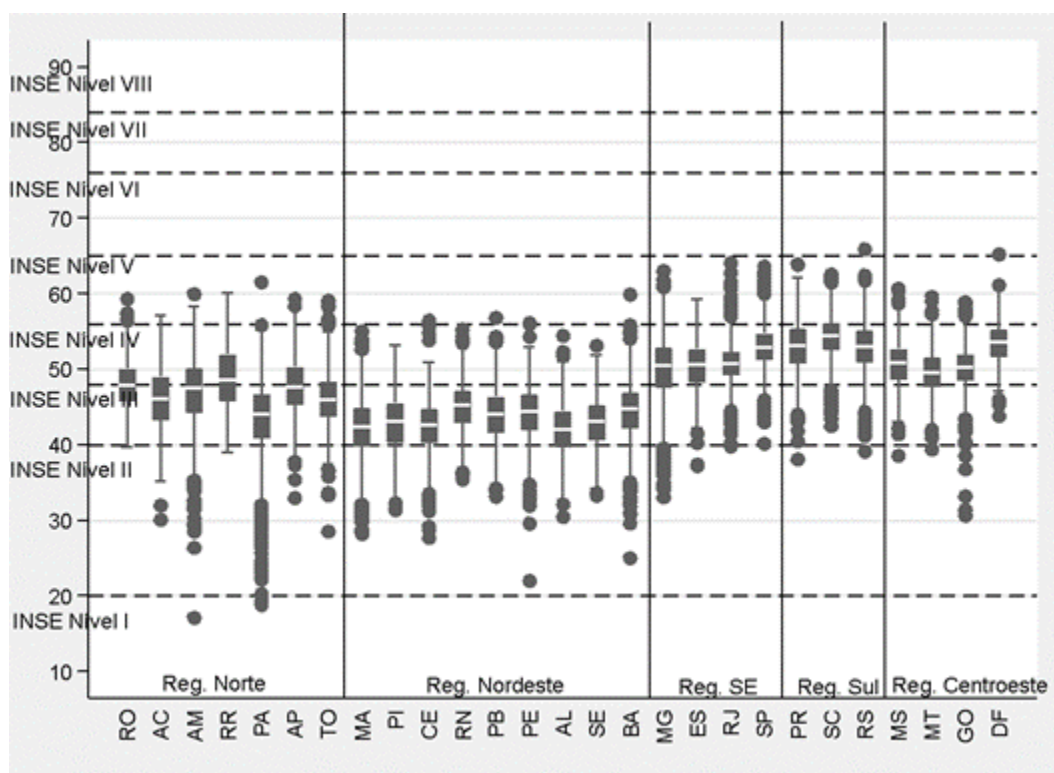
Inse Nível II	Inse Nível III	Inse Nível IV
Nível II - (20;40]: Neste, os alunos, de modo geral, indicaram que há em sua casa bens elementares, como uma televisão, uma geladeira, um ou dois telefones celulares, um banheiro e até dois quartos para dormir. Não possuem máquina de lavar roupa ou computador entre seus bens. A renda familiar mensal é de até 1 salário mínimo; e seu pai e sua mãe (ou responsáveis) sabem ler e escrever tendo ingressado no ensino fundamental, completando ou não o 5º ano de estudo.	Nível III - (40;48]: Neste, os alunos, de modo geral, indicaram que há em sua casa bens elementares, como banheiro e até dois quartos para dormir, possuem televisão, geladeira, dois ou três telefones celulares; bens complementares como máquina de lavar roupas e computador (com ou sem internet); a renda familiar mensal é entre 1 e 1,5 salários mínimos; e seus responsáveis completaram o ensino fundamental ou o ensino médio.	Nível IV - (48;56]: Já neste nível, os alunos, de modo geral, indicaram que há em sua casa bens elementares, como dois ou três quartos para dormir, um banheiro, uma geladeira, três ou mais telefones celulares, e um ou dois televisores; e bens complementares, como máquina de lavar roupas, micro-ondas, computador (com ou sem internet), um telefone fixo e um carro; bens suplementares, como freezer; a renda familiar mensal está entre 1,5 e 3 salários mínimos; e seus responsáveis completaram o ensino médio ou a faculdade.

Fonte: Nota Técnica – Indicador de Nível Socioeconômico das Escolas de Educação Básica (Inse). Disponível em: http://download.inep.gov.br/informacoes_estatisticas/indicadores_educacionais/2015/nota_tecnica/nota_tecnica_inep_inse_2015.pdf. Destaques autora.

Veja abaixo como é importante levar essa interdependência em conta. O Gráfico 1 apresenta a distribuição da classificação, por escola, em relação ao Inse por UF, as

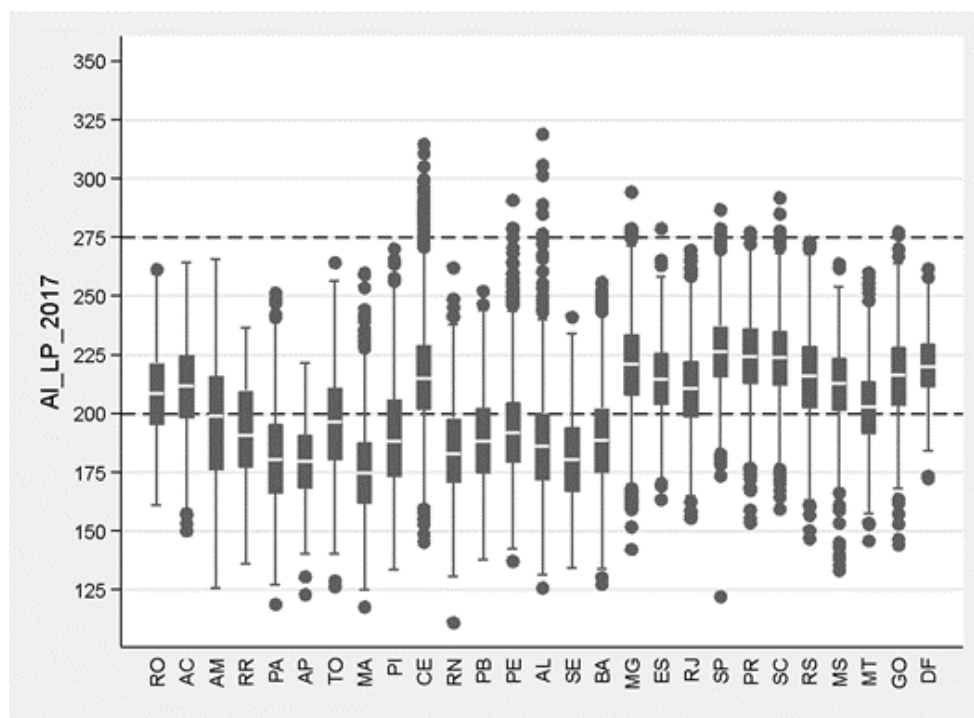
quais foram agrupadas por Região. É fácil notar que a distribuição é mais homogênea dentro de uma mesma Região e que as escolas do Nordeste atendem a uma clientela mais concentrada no nível III – mas com um contingente não desprezível no nível II –, enquanto as do Sudeste, Sul e Centro-Oeste servem mais frequentemente alunos do nível IV.

Gráfico 1 - Distribuição do ISNE por escola no SAEB de 2017 para o 5º ano do EF I - UF's e Grandes Regiões



Fonte: <http://portal.inep.gov.br/indicadores-educacionais>. Inep (2015). Nota Técnica—Indicador de Nível Socioeconômico das Escolas de Educação Básica (Inse). Tabulação autora. Recuperado de Inep website: http://download.inep.gov.br/informacoes_estatisticas/indicadores_educacionais/2015/nota_tecnica/nota_tecnica_inep_inse_2015.pdf.

Gráfico 2 - Distribuição das notas de proficiência por escola no SAEB de 2017 para o 5º ano do EF I por UF's

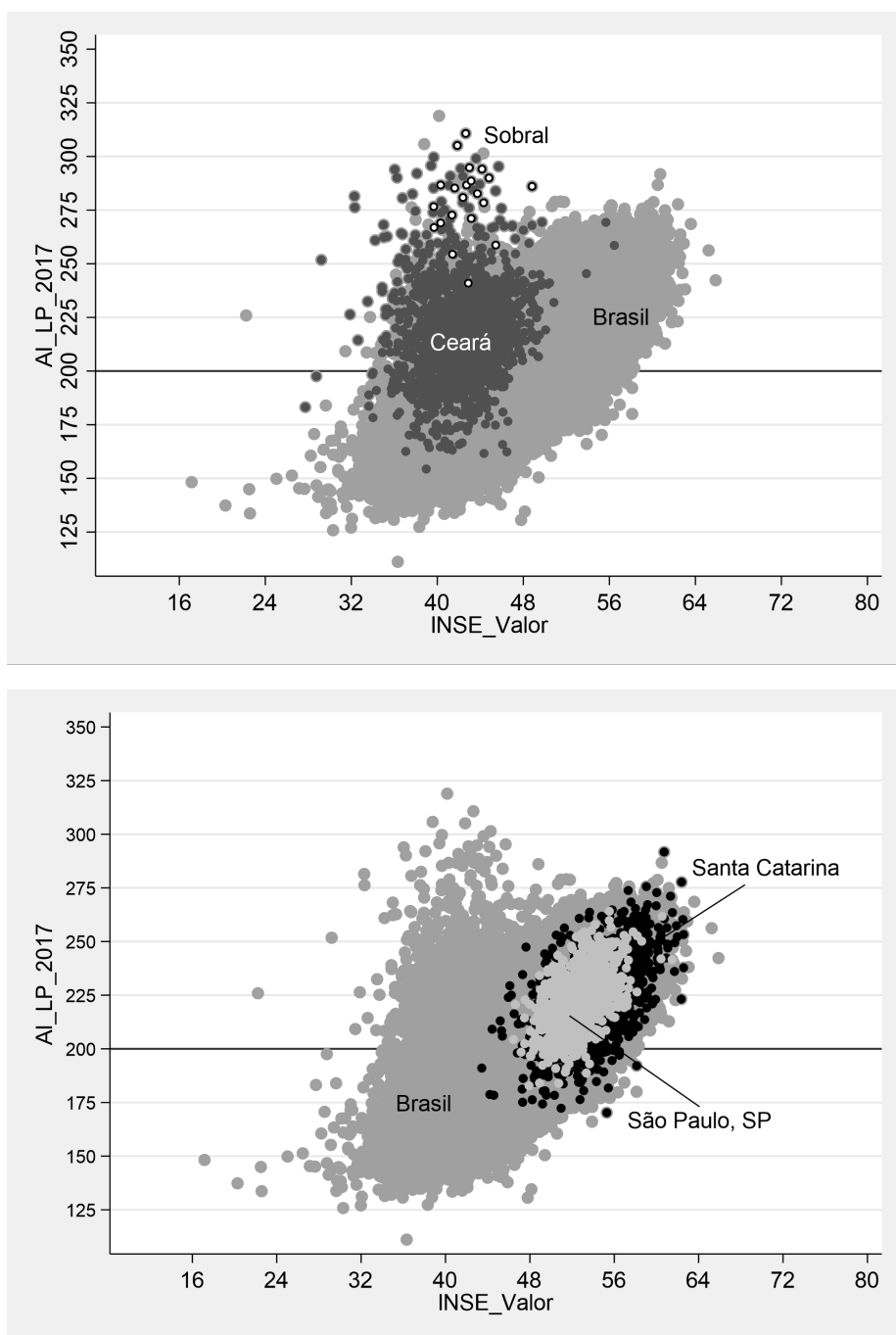


Fonte: Inep (2018). Microdados da Aneb e da Anresc 2017. Tabulação autora. Disponível em: <http://portal.inep.gov.br/basi-ca-levantamentos-acessar>. Acesso em: 31 out. 2018.

A comparação entre os dois gráficos apresentados mostra uma quase sobreposição entre as distribuições do Inse e da proficiência por UF. Alguns estados se destacam positivamente, como o Ceará, que conta com a maior parte dos alunos nos níveis mais baixos do INSE, mas que apresentam notas mais altas de proficiência, quando comparado, por exemplo, com o Rio de Janeiro, que apresenta uma distribuição parecida de proficiência, mas com uma população escolar com Inse bem mais alto e homogêneo. O Estado do Mato Grosso também aparece como um exemplo negativo com a mesma combinação de Inse alto e proficiência baixa, com metade das escolas apresentando escore abaixo da nota mínima esperada para a etapa, que são 200 pontos.

Para ilustrar melhor como se dá a relação entre nível socioeconômico e proficiência, seguem o Gráfico 3 com os resultados médios por escola para todo o Brasil e dois destaques diferentes: Ceará e Sobral (baixo NSE com médio e alto desempenho) e Santa Catarina e a cidade de São Paulo (alto NSE com médio e alto desempenho).

Gráfico 3 – Relação entre resultados médios por escola e níveis socioeconômicos



Fonte: Inep, 2018. Microdados da Aneb e da Anresc 2017. Tabulação autora. Disponível em: <http://portal.inep.gov.br/basi-ca-levantamentos-acessar>. Acesso em: 31 out. 2018.

Proficiência e nível de complexidade do ambiente educacional

Um outro aspecto que deve ser levado em consideração quando se comparam os desempenhos dos alunos entre turmas, escolas, municípios e estados é a complexidade da gestão das atividades de ensino: tamanho da clientela, nível de descentralização, distribuição da população pelo território, localização em áreas remotas, por exemplo, além do item tratado na seção anterior, obviamente. Assim, estados e municípios que atendem a contingentes muito grandes de alunos, com uma parte substancial deles dispersa em unidades escolares frequentemente de acesso difícil – como as escolas nas quais o acesso em relação à sede é feito por barco na Região Norte –, encontram maiores dificuldades para prover professores e materiais, transportar alunos, garantir a frequência docente e discente, fazer obras e garantir a manutenção dos prédios escolares que aqueles com características opostas a essas.

Para mostrar como essas condições locais, para além do Inse, podem afetar os rankings, apresenta-se uma outra comparação. Na coluna da esquerda, uma tabela com o ranking da média de proficiência em Língua Portuguesa no 5º ano do ensino fundamental (Saeb 2017) das escolas municipais (urbanas e rurais). Na coluna da direita, apresenta-se a mesma nota, só que para as escolas estaduais. Repare que o ranking é diferente para cada tipo de rede: por exemplo, Santa Catarina fica em 1º lugar com suas escolas municipais, mas se fossem consideradas as escolas estaduais, em 2017 esse lugar seria de Minas Gerais. Da mesma forma, Maranhão e Alagoas se alternariam em último.

Tabela 2 - Ranking de proficiência em Língua Portuguesa no SAEB de 2017, por dependência administrativa

Unidade da Federação	Proficiência média das municipais (2017)	Proficiência média das escolas estaduais em LP (2017)
Santa Catarina	227,41	219,54
São Paulo	225,74	225,77
Paraná	225,68	196,38
Minas Gerais	221,96	224,88
Ceará	217,91	216,95
Goiás	215,09	226,67
Espírito Santo	213,53	216,52
Rio Grande do Sul	211,19	216,19
Mato Grosso do Sul	210,75	215,19
Rio de Janeiro	209,39	217,34
Acre	207,34	220,60
Mato Grosso	205,79	205,28
Rondônia	204,68	221,52
Roraima	198,97	NA
Tocantins	197,39	210,79
Piauí	193,35	216,70
Amazonas	193,04	211,74
Pernambuco	192,50	210,42
Bahia	190,47	195,65
Alagoas	189,18	190,97
Paraíba	188,88	189,93
Rio Grande do Norte	183,68	188,89
Pará	183,60	189,6
Amapá	180,91	181,01
Sergipe	178,45	192,43
Maranhão	177,66	194,35
Distrito Federal	NA	220,24

Fonte: <http://portal.inep.gov.br/web/guest/educacao-basica/saeb/resultados>. Tabulação autora.

A Tabela a seguir mostra a distribuição de matrículas estaduais e municipais das escolas localizadas em perímetro urbano ou rural para a etapa I do ensino fundamental em 2017 e o tamanho do território de cada UF. São Paulo é o estado que conta com o maior número de matrículas administradas por um órgão central, mas apenas 2% das matrículas estão em áreas rurais. No Maranhão, essa proporção é de 40% e o ensino é praticamente todo municipalizado. Veja como um ranking por UF pode embutir diferenças substanciais das condições locais de ensino. Amazônia e Pará contam ambos com territórios muito extensos, mas com distribuição bem diferentes entre escolas municipais e estaduais, além da grande diferença entre a dimensão das clientelas.

Tabela 3 - Distribuição das matrículas de fundamental I por tipo de rede e localização em cada Unidade da Federação (2017)

UF	Matrículas em escolas municipais	Matrículas em escolas estaduais	% de municipalização	Matrículas em escolas rurais	Matrículas em escolas urbanas	% de rurais	Área em Km2
D. Federal	-	150,379	0,00%	10,292	196,527	5,00%	5,780
Sergipe	114,484	28,354	80,10%	52,535	135,984	27,90%	21,918
Alagoas	221,288	9,845	95,70%	82,435	201,293	29,10%	27,848
Rio de Janeiro	753,673	1,931	99,70%	72,039	1,055,386	6,40%	43,782
Espírito Santo	208,711	38,123	84,60%	33,164	250,242	11,70%	46,087
R. G. do Norte	165,021	37,445	81,50%	55,791	207,778	21,20%	52,811
Paraíba	209,733	30,109	87,40%	69,938	244,119	22,30%	56,468
Santa Catarina	282,891	112,833	71,50%	42,425	416,651	9,20%	95,738
Pernambuco	536,689	9,912	98,20%	171,029	575,537	22,90%	98,076
Amapá	48,266	25,664	65,30%	18,409	63,208	22,60%	142,829
Ceará	511,895	3,308	99,40%	149,788	523,336	22,30%	148,888
Acre	50,043	40,240	55,40%	33,684	61,004	35,60%	164,124
Paraná	663,420	3,454	99,50%	51,852	740,099	6,50%	199,308
Roraima	41,179	7,786	84,10%	16,205	37,928	29,90%	224,301
Rondônia	113,568	30,937	78,60%	35,115	122,853	22,20%	237,765
São Paulo	1,772,200	618,082	74,10%	61,301	2,960,669	2,00%	248,220
Piauí	235,536	2,747	98,80%	84,718	193,520	30,40%	251,612
Tocantins	109,237	15,439	87,60%	20,957	118,036	15,10%	277,720
R. G. do Sul	374,193	237,964	61,10%	74,403	638,085	10,40%	281,738
Maranhão	583,798	10,577	98,20%	271,272	404,542	40,10%	331,937
Goiás	362,367	21,920	94,30%	22,812	458,036	4,70%	340,106
M. G. do Sul	151,583	48,536	75,70%	26,923	203,517	11,70%	357,146
Bahia	924,028	4,981	99,50%	346,668	801,204	30,20%	564,732
Minas Gerais	824,395	382,472	68,30%	127,540	1,267,361	9,10%	586,521
Mato Grosso	157,789	66,685	70,30%	37,015	221,200	14,30%	903,202
Pará	716,556	59,384	92,30%	314,481	537,562	36,90%	1,247,955
Amazonas	290,381	85,939	77,20%	107,104	302,968	26,10%	1,559,147

Fonte: Inep. Sinopse Estatística da Educação Básica 2017. Disponível em: <http://portal.inep.gov.br/web/guest/sinopses-estatisticas-da-educacao-basica> e IBGE. Organização do território 2017. Disponível em: <https://www.ibge.gov.br/geociencias/organizacao-do-territorio/estrutura-territorial/15761-areas-dos-municipios.html?=&t=downloads>. Tabulação autora.

A Matemática do Ideb

Todos sabemos que não é fácil ser avaliado. Pior ainda quando essa avaliação é exposta publicamente e não é exatamente lisonjeira. Nas avaliações em larga escala para a educação, enquanto os alunos e suas turmas (professores) têm suas identidades pessoais protegidas por sigilo, o nome de cada escola, de cada município e de cada UF, associado ao desempenho de cada etapa escolar, é público e pode expor (positiva ou negativamente) a performance profissional de diretores de escola, prefeitos e governadores. Basta ver como esses dados são usados em vários tipos de disputa (não apenas em eleições, mas na hora da alocação de recursos públicos e até na disponibilidade de vagas em escolas bem avaliadas), para perceber a carga de interesses que essas informações carregam.

Não é difícil imaginar que gestores e afins podem tentar soluções heterodoxas para ficarem menos feios nas fotos, uma vez que está claro que ainda não se sabe direito como fazer com que alunos aprendam de verdade, de forma a realmente melhorarem a proficiência medida pelas provas. Vamos lembrar que, em qualquer etapa que se meça a proficiência, a maior parte dos alunos fica muito abaixo do que seria esperado que eles aprendessem em relação aos sempre medíocres padrões brasileiros. Quanto mais avança a etapa escolar, pior fica a situação dos alunos.

Se a proficiência vai ser exposta, é só fazer com que os piores alunos não façam as provas, ou que fiquem retidos em anos anteriores, ou que, “melhor ainda”, saiam da escola, caso em que não se pode errar na mão, porque tem a graninha do Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação (Fundeb), que é calculada com base nas matrículas. O fato é que, para quem quer fazer malandragem com as notas das provas, que são aplicadas e controladas por um órgão federal, precisa aprender um pouco de Matemática. Cientes dessa tendência a “reorganizar” números para deixar o cenário menos caótico, os gestores do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) criaram um índice, que combina a proficiência dos alunos com a taxa média de aprovação na etapa de ensino correspondente e explicam:

Um sistema educacional que reprova sistematicamente seus estudantes, fazendo que grande parte deles abandone a escola antes de completar a educação básica, não é desejável, mesmo que aqueles que concluem essa etapa atinjam elevadas pontuações nos exames padronizados. Por seu lado, um sistema em que os alunos concluem

o ensino médio no período correto não é de interesse caso eles aprendam muito pouco. Em suma, um sistema ideal seria aquele no qual todas as crianças e adolescentes tivessem acesso à escola, não desperdiçassem tempo com repetências, não abandonassem os estudos precocemente e, ao final de tudo, aprendessem (FERNANDES, 2007, p. 7. Destaque da autora).

Para que se possa ter uma ideia de como a Matemática pode funcionar para além do que se aprende nas escolas, veja o Gráfico 4, que mostra, por UF, a dimensão das taxas de promoção, repetência, evasão, migração para a Educação de Jovens e Adultos (EJA) – modalidade que não é avaliada pelo Saeb – e a taxa de não presença na prova de final da etapa correspondente. No caso do gráfico, o 9º ano do ensino fundamental do Saeb de 2015. As barras empilhadas mostram as diferentes componentes da taxa de fluxo (2014/15), conforme definição abaixo, e estão ordenadas pela taxa de promoção. As bolinhas mostram a proporção de alunos que não compareceu àquela determinada prova.

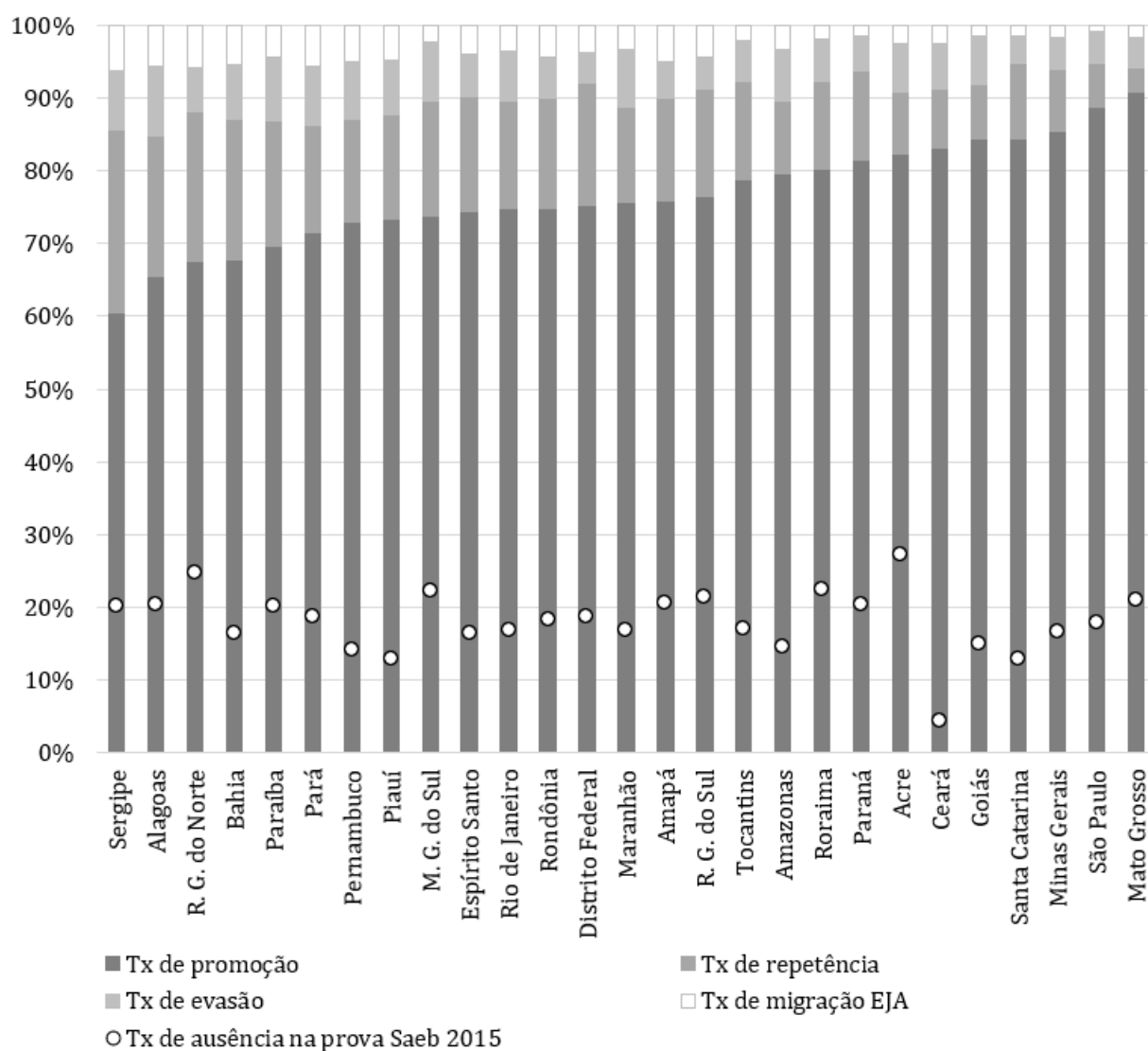
Taxa de promoção: proporção de alunos que em 2014 estavam matriculados na série k (etapa de ensino seriada do ensino fundamental ou médio) e em 2015 se matriculam em série superior a k.; [...]

Taxa de repetência: proporção de alunos que em 2014 estavam matriculados na série k (etapa de ensino seriada do ensino fundamental ou médio) e em 2015 se matriculam em série igual ou inferior a k.

Taxa de evasão: proporção de alunos que em 2014 estavam matriculados na série k (etapa de ensino seriada do ensino fundamental ou médio) e em 2015 não estavam matriculados.

Migração para EJA: proporção de alunos que em 2014 estavam matriculados na série k (etapa de ensino seriada do ensino fundamental ou médio) e em 2015 estavam matriculados na EJA. (Inep, Fluxo Escolar 2014/2015; <http://portal.inep.gov.br/indicadores-educacionais>).

Gráfico 4 - Proporção das taxas de fluxo de 2014/15, combinada com a taxa de ausência na prova de Língua Portuguesa do 9º ano no SAEB de 2015, por Unidade de Federação

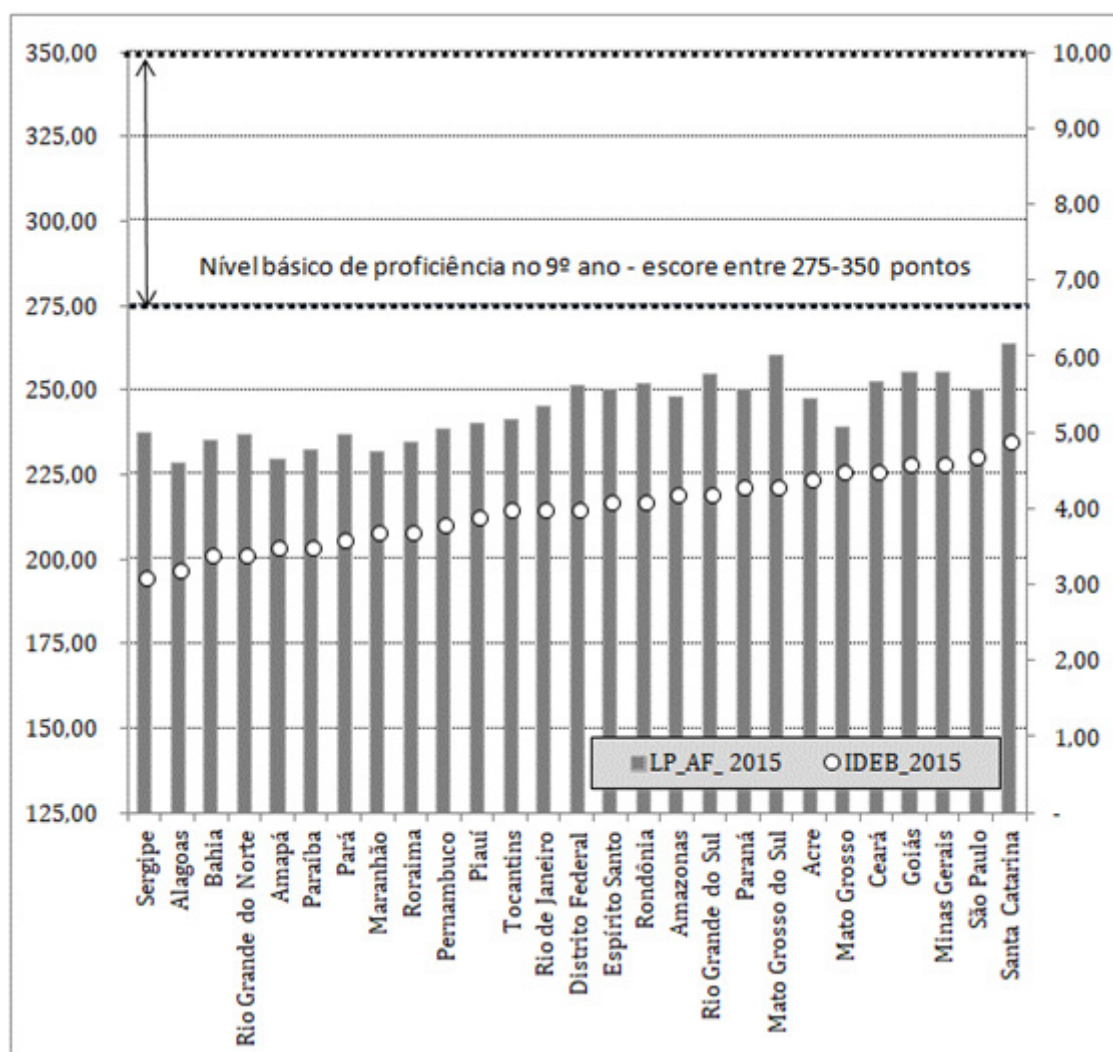


Fonte: Taxas de transição dos indicadores educacionais do Inep – 2014/15; Disponível em: <http://portal.inep.gov.br/indicadores-educacionais> e Microdados Saeb 2015; Microdados Saeb 2015; Disponível em: <http://portal.inep.gov.br/microdados>.

Agora compare o Gráfico 4, referente a fluxo, no qual se pode ver o quanto ainda é possível melhorar cada um de seus componentes, além da taxa de ausência na prova – que pode ser aleatória (em média 1/5 dos alunos não aparece na escola) ou provocada (em média as escolas dispensam 1/5 deles no dia da prova) –, com o Gráfico 5 que mostra o desempenho médio por UF em Língua Portuguesa no 9º ano. Segundo o Inep, os níveis considerados básicos para essa etapa são os de 4 a 6, o que corresponde a uma faixa de escore esperada de 275-350 pontos (Inep, 2018 – ppt de divulgação). Para o ano de 2015, percebe-se, em primeiro lugar, que nenhuma UF alcançou, em média, o nível básico para a etapa (Língua Portuguesa, 9º ano); em segundo, que o

ranking do Ideb não é exatamente o ranking de proficiência; e, em terceiro, que uma UF que esteja em uma posição alta nesse ranking pode estar com o fluxo e a presença comprometidos (como o Acre) ou o contrário (o Ceará, que teve, em média, 95% dos alunos fazendo as provas, apresenta um fluxo razoável (83% de taxa de promoção) – mas não brilhante –, além de escore de proficiência entre os mais altos (252 pontos), entretanto não está no topo do ranking).

Gráfico 5 - Proficiência média na prova de Língua Portuguesa no 9º ano no SAEB de 2015 e IDEB, por Unidade da Federação



Fonte: Inep (2017) – Planilhas para download. <http://portal.inep.gov.br/web/guest/educacao-basica/ideb/resultados>

Ideb 6,0 não é o mesmo que nível 3 no Pisa!

É corriqueiro encontrar matérias na imprensa que apresentam escolas brasileiras (e até municípios) que teriam padrão de educação “europeu”, em decorrência de terem alcançado alta pontuação de proficiência ou até mesmo apenas um alto Ideb – o que, como já foi apresentado, não é a mesma coisa. Reforçando a ideia, um Ideb alto pode carregar matemática suficiente para esconder mazelas crônicas, as quais, à partida, já distanciam, em muito, o nível de qualidade das escolas brasileiras daquele de países da Europa. Para fins da comparação educacional feita nesta seção, vamos considerar os países europeus de cultura ocidental que fizeram parte da primeira “versão” da Comunidade Europeia em 1992 (BECSKEHÁZY, 2018⁵).

Essa falácia accidental se propagou na imprensa porque o próprio Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), que é uma autarquia federal vinculada ao Ministério da Educação (MEC) – gozando, portanto, de credibilidade geral –, baseou as metas progressivas do Ideb na incorreta suposição de que as proficiências nas provas do Saeb e do Pisa poderiam ser equivalentes. Justamente o órgão que processa e informa os dados educacionais do Brasil apresenta essa correspondência impossível, mesmo ciente de que seria um exercício de “esticar a corda” dos números oriundos de avaliações em larga escala muito diferentes entre si. Houve a suposição de que seria possível fazer uma correspondência entre elas considerando apenas a distribuição das notas, pelo fato de que os alunos podem estar em faixas escolares e etárias sobrepostas. Acontece que os alunos no Brasil aprendem muito “menos” coisas que seus colegas da Europa e EUA (sem falar nos da Ásia), pois os currículos (formais ou de fato) são **muito** diferentes entre si, com clara desvantagem para os alunos brasileiros:

Existem, entretanto algumas limitações para executar esse processo. Primeiramente, as duas avaliações não estão na mesma escala de proficiência e não possuem itens comuns que possam ser utilizados para que isso seja feito. Além disso, participam do PISA somente

5 “A seleção de países participantes do Pisa 2015 para o gráfico anterior foi feita com o intuito de estabelecer uma comparação com o contexto brasileiro, situando-o entre seus pares regionais, ainda em desenvolvimento, e alguns, culturalmente mais próximos, mas já desenvolvidos. Desse modo, foram selecionados países industrializados, democráticos e de alta renda na Europa (União Europeia na configuração de Maastrich de 1992, destacando os países ibéricos) e na América do Norte, além dos países da América Latina que participaram do exame em 2015. Dos países participantes selecionados para essa comparação, apenas Colômbia, Brasil, Peru, Uruguai e Argentina não fazem parte da OCDE, entidade organizadora do Pisa.” (BECSKEHÁZY, 2018)

alunos com 15 anos, independentemente da série em que estudam. O SAEB, por sua vez, é destinado a alunos de 4ª e 8ª séries do Ensino Fundamental (EF), e de 3º ano do Ensino Médio (EM). Dessa maneira, optou-se por considerar o desempenho dos alunos brasileiros no PISA como uma aproximação do desempenho dos alunos de 8ª série no SAEB. [...]

O primeiro passo para compatibilização foi identificar para qual nota do PISA, chamada aqui de nota de referência, deseja-se encontrar correspondência aproximada no SAEB. Por exemplo, as notas correspondentes à mediana ou a média obtida pelos estudantes nos países participantes do PISA de 2003 (exemplos no quadro 1), ou a média dos países da OCDE.

(INEP, 2007. Nota metodológica sobre a compatibilização de desempenhos do PISA com a escala do SAEB. Disponível em: http://download.inep.gov.br/educacao_basica/portal_ideb/metodologias/Nota_Tecnica_n3_compatibilizacao_PISA_SAEB.pdf; p.1; destaque da autora)

Como o próprio Inep reconhece no documento, as duas avaliações não contam com itens (questões de prova) comuns. Uma análise das matrizes que embasam a construção de cada uma das provas mostra que a concepção delas é bem diferente entre si, tendo como consequência níveis muito distintos de complexidade de mobilização cognitiva necessários para se responder correta e completamente a maior parte das questões: as provas brasileiras exigem muito menos dos alunos que as provas do Pisa. Conhecendo-se os sistemas e as políticas educacionais de várias regiões do mundo – principalmente suas normativas curriculares –, as quais, de **certa forma**⁶, se refletem no Pisa, sabe-se que o patamar de ambição acadêmica de cada um deles pode ser muito diferente entre as regiões sociogeográficas, embora relativamente homogêneo dentro de cada uma.

A média dos países da OCDE em leitura está em torno de 500 pontos, o que equivale a localizar o desempenho dos alunos no nível 3 do Pisa. Normalmente, por mais “difíceis” que sejam as provas brasileiras de aferição em nível nacional (o que exclui os vestibulares isolados, por exemplo), as questões de leitura, assim como a maior

⁶ O Pisa é uma prova concebida para se fazer uma estimativa comparada do potencial de formação de capital humano futuro de cada país ou território. Não é necessariamente baseada nos currículos dos países participantes.

parte das atividades propostas nos livros didáticos oficiais, não passam do nível 2 de dificuldade, bem menos complexo que os que estão além dele. Essa é uma das razões pelas quais a nota dos brasileiros no Pisa não melhora estatisticamente (ver OCDE, 2019⁷): raramente eles são mentalmente tão desafiados quanto seus colegas de países desenvolvidos.

O Gráfico 5, que detalha a proficiência no Pisa de 2018 por nível socioeconômico, mostra como isso se distribui nas diferentes faixas de renda. O resultado dos alunos brasileiros, que aparece destacado, encontra-se de certa forma alinhado com os dos vizinhos da América Latina, mas em franca desvantagem em relação à Europa ou Ásia, sendo que mesmo os alunos de NSE mais alto apresentam proficiência não apenas mais baixa que seus pares de renda, mas dos que renda média a até baixa, em relação aos alunos da Ásia. Além disso, mostra o % de cobertura da prova – participação da população elegível⁸ – e os % de alunos participantes nos decis superior e inferior de renda. Entre os países desenvolvidos, essas diferenças vêm sendo diminuídas, exatamente (não apenas) como um desdobramento do próprio Pisa – uma vez que os países estão sempre reavaliando suas vantagens competitivas em relação à formação de capital humano –, mas ainda refletem a história e a cultura acadêmica de cada país ou território, dentro de uma determinada região. Portanto, é preciso compreender que uma nota alta em uma prova escolar feita no Brasil dificilmente encontra correspondência, quanto ao nível de exigência, com uma para a mesma faixa etária feita na Europa Ocidental ou nos Tigres Asiáticos.

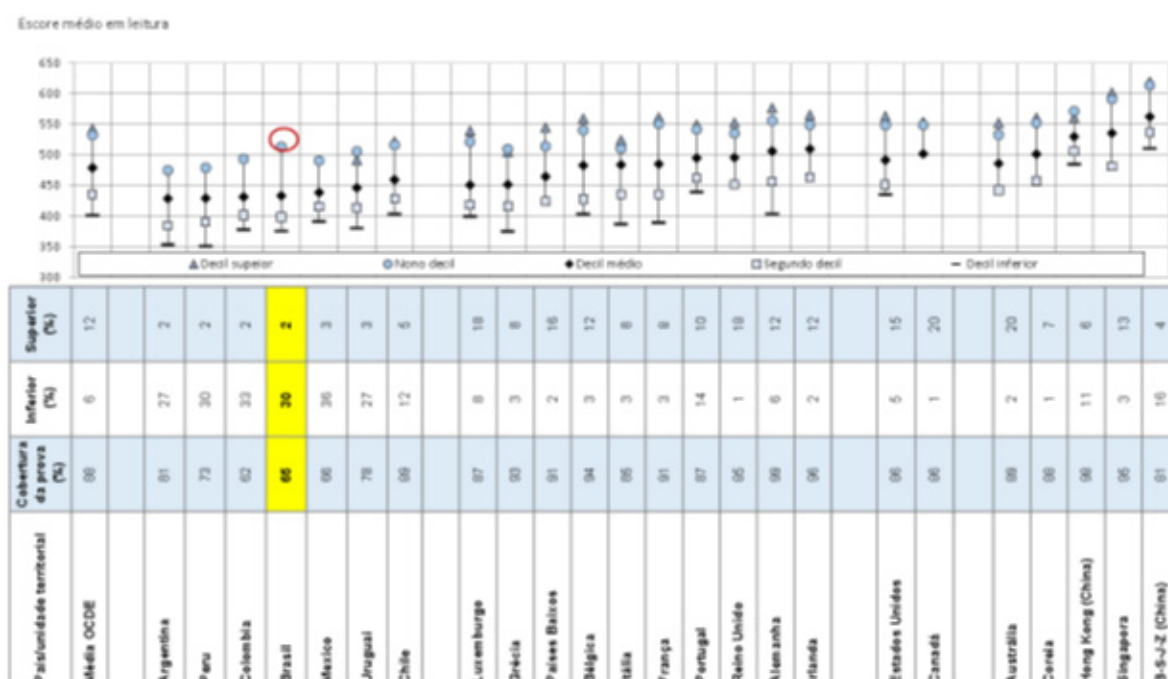
Pela mesma razão, quando reconhecemos a vergonha de não conseguir “alfabetizar” metade dos alunos de 3º ano do ensino fundamental ao analisar os dados da Avaliação Nacional de Alfabetização (ANA⁹), não levamos em conta que as provas brasileiras pressupõem uma mobilização cognitiva muito mais simplória do que se exige nas escolas de países desenvolvidos. Assim, uma realidade muito mais assustadora será “revelada” quando saírem os resultados do PIRLS, que é uma espécie de Pisa para leitura no 4º ano.

7 Ver PISA 2018 Results (Volume I) - OECD 2019 Annex B1 Results for countries and economies Version 3 - Last updated: 02-Dec-2019.

8 Alunos de faixa etária de 15 anos, matriculados em instituições formais de ensino e a partir do 7º ano. Portanto pessoas dessa faixa etária que não estão na escola ou apresentam atraso escolar muito alto, não participam dos exames do PISA. Uma baixa taxa de cobertura aponta para uma fragilidade ainda maior do sistema educacional, para além do escore médio e da distribuição das notas.

9 Prova para alunos matriculados no 3º ano do ensino fundamental, que, a partir de 2019, será simplesmente a prova de 2º ano do Saeb, o qual já avalia o 5º e o 9º ano do ensino fundamental e o último do ensino médio.

Gráfico 5 - Escore Médio em leitura por países (ou Unidades Territoriais) selecionados no PISA de 2018, agrupados por região, distribuído por decil de renda



Fonte: OECD (2019)

Habilidades socioemocionais não passam à frente do conhecimento relacionado às disciplinas escolares

O Gráfico 4 mostra o tamanho das dificuldades que os alunos de 15 anos do Brasil (os 80% da população que efetivamente fizeram a prova, pois 20% nem estava apto, segundo as regras de seleção de amostra do Pisa, por estar fora da escola ou muito atrasado) apresentam em relação à capacidade de interpretar textos de complexidade relativa como, por exemplo, textos informativos com dados e recursos gráficos variados.

Mesmo os brasileiros de renda superior (10% com NSE mais alto) mal chegam ao nível 3 do Pisa (entre 407-480 pontos), o qual representa o desempenho médio dos alunos dos países que compõem a OCDE, organizadora do exame, sem falar nos Tigres Asiáticos, que deixam o mundo educacional ocidental de cabelos em pé. A raiz desse problema está, em grande parte, na especificação curricular, que é a normativa que regula o que é obrigatório que se ensine em cada escola de cada país, a cada ano escolar. Sistemas educacionais eficazes costumam contar com currículos claros que especificam o que deve ser aprendido pelos alunos. Nesses documentos também é comum que se estabeleçam competências a serem desenvolvidas logo no início da escolarização e que auxiliam muito no desenvolvimento escolar e acadêmico dos alunos, e normalmente fazem parte dos eixos de interação discursiva ou de oralidade. Como vamos ver nos exemplos apresentados na tabela a seguir, muitas delas são chamadas no Brasil de competências socioemocionais e abriram um novo mercado para o “coaching” educacional, quando simplesmente deveriam fazer parte do cenário pedagógico de todas as escolas.

Quadro 1¹⁰ - Exemplos comparados de habilidades de comunicação presentes em currículos de etapas do início da escolarização (Irlanda; Ontário, Canadá e Singapura)

10 Fonte: Government of Ireland, 1999. Primary School Curriculum - English Language curriculum; Disponível em: https://www.curriculumonline.ie/getmedia/5b514700-e65c-46a7-a7d0-c8e-05e115bf9/PSEC01a_English_Curriculum.pdf | Ontario Ministry of Education and Training, 2016. The Kindergarten Program 2016. Disponível em: <http://www.edu.gov.on.ca/eng/curriculum/elementary/kinderprogram.html> | Singapore, & Curriculum Planning & Development Division, 2008. English language syllabus 2010: Primary & secondary (express/normal [academic]). Disponível em: <https://www.moe.gov.sg/docs/default-source/document/education/syllabuses/english-language-and-literature/files/english-primary-secondary-express-normal-academic.pdf>. Tradução autora

Irlanda – educação infantil (Infant classes¹¹)

A criança deve aprender a:

- vivenciar, reconhecer e observar comandos simples, olhar, ouvir, assistir
- ouvir uma história ou descrição e responder a ela
- ouvir, repetir e elaborar palavras, frases e frases e sentenças modeladas pelo professor
- usar e interpretar o tom de voz que expressa várias emoções
- aprender a adotar um comportamento verbal e não verbal apropriado para garantir e manter a atenção de um par, estabelecendo contato visual, usando movimentos adequados da cabeça, gestos e expressões faciais, garantindo audibilidade e clareza
- imitar e interpretar gestos, movimentos e atitudes transmitindo várias emoções.

11 The child should be enabled to:

The child should be enabled to: experience, recognise and observe simple commands look, listen, watch; listen to a story or description and respond to it; hear, repeat and elaborate words, phrases and sentences modelled by the teacher; use and interpret tone of voice expressing various emotion; learn to adopt appropriate verbal and nonverbal behaviour to secure and maintain the attention of a partner establishing eye contact using appropriate head movements, gestures and facial expressions ensuring audibility and clarity; mime and interpret gesture, movement and attitude conveying various emotions.

Canadá (Ontário) – educação infantil (kindergarten program¹²)

À medida que as crianças progridem no programa do Jardim de Infância, elas:

1.2 ouvem e respondem a outras pessoas, verbal e não verbalmente (por exemplo, usando artes, sinais, gestos e linguagem corporal), para uma variedade de propósitos (por exemplo, trocar idéias, expressar sentimentos, oferecer opiniões) e em uma variedade de contextos (por exemplo, depois de ler em voz alta e compartilhar experiências de leitura ou escrita; enquanto resolvem um problema de matemática em classe; em brincadeiras imaginárias ou exploratórias; nos espaços de aprendizagem; enquanto se envolvem em brincadeiras e brincadeiras ao ar livre; enquanto fazem observações científicas de plantas e animais ao ar livre)

3.1 agem e conversam com colegas e adultos expressando e aceitando mensagens positivas (por exemplo, com um tom de voz e gestos apropriados; fazem elogios; fazem e aceitam críticas construtivas)

5.1 demonstram respeito e consideração pelas diferenças individuais e pontos de vista alternativos (por exemplo, ajudam um amigo que fala outro idioma; adaptam o comportamento para acomodar as idéias de um colega de classe)

12 As children progress through the Kindergarten program, they: 1.2 listen and respond to others, both verbally and non-verbally (e.g., using the arts, using signs, using gestures and body language), for a variety of purposes (e.g., to exchange ideas, express feelings, offer opinions) and in a variety of contexts (e.g., after read-alouds and shared reading or writing experiences; while solving a class math problem; in imaginary or exploratory play; in the learning areas; while engaged in games and outdoor play; while making scientific observations of plants and animals outdoors). 3.1 act and talk with peers and adults by expressing and accepting positive messages (e.g., use an appropriate tone of voice and gestures; give compliments; give and accept constructive criticism). 5.1 demonstrate respect and consideration for individual differences and alternative points of view (e.g., help a friend who speaks another language; adapt behaviour to accommodate a classmate's ideas)

Singapura – Primário (Primary¹³)

OUVIR E VISUALIZAR, ATITUDES E COMPORTAMENTO

• Ouve e visualiza atentamente e por um período prolongado, por exemplo:

° olha para a pessoa que fala e mantém contato visual

° mantém postura e expressão facial adequadas

° ouve e olha durante todo o tempo de leitura de um texto (por exemplo, ouvir instruções, recontagem ou apresentação de um colega de classe, um curta-metragem, uma palestra, um discurso)

° aguarda para fazer julgamento ou comentário até que seja apropriado

• Ouve e olha com empatia e respeito (por exemplo, dando a devida atenção ao orador e sendo sensível ao que é dito)

• Antecipa a resposta adequadamente enquanto ouve e olha, por exemplo:

° concordando ou indicando entendimento

° fornecendo retorno para confirmar a compreensão e incentivar o orador (por exemplo, “Mmm”, “Sim”, “Entendo”)

° buscando esclarecimentos e elaboração

13 LISTENING AND VIEWING ATTITUDES AND BEHAVIOUR

• Listen and view attentively and for a sustained period, e.g., ° look at the person speaking and maintain eye contact; ° maintain appropriate posture and facial expression; ° listen and view for the entire duration of a text (e.g., listening to instructions, a classmate’s recount or presentation, a short film, a talk, a speech); ° withhold judgement or comment until appropriate

• Listen and view with empathy and respect (e.g., giving due attention to the speaker and being sensitive to what is said)

• Indicate response appropriately while listening and viewing, e.g., ° nod in agreement or to indicate understanding; ° provide back-channelling to confirm comprehension and encourage speaker (e.g., “Mmm”, “Yes”, “I see”); ° seek clarification and elaboration.

Considerações finais

A análise de cinco aspectos diferentes, mas complementares, relacionados à forma de apresentar dados obtidos de avaliações nacionais e internacionais das quais participam alunos das escolas brasileiras mostra que, pelo menos para o público em geral – quando não até para profissionais do ramo e acadêmicos desavisados –, efetivamente há armadilhas cognitivas a serem superadas.

Em primeiro lugar, não se pode analisar desempenho educacional na educação básica (principalmente) sem uma medida de qualidade que traga à luz a aferição de NSE. São fatores intimamente associados que facilmente ajudam a formar uma percepção equivocada de que uma política educacional é melhor que outra, quando o que na realidade se mede é o capital cultural ou educacional (BOURDIEU; PASSERON, 1975) de determinadas famílias. Em segundo lugar, há o desafio da complexidade da gestão local de redes de ensino, que pode penalizar gestores de redes com características que tornem a operação das escolas mais desafiadora. Não se pode esquecer que a atividade educacional não é apenas de recorrência diária, mão de obra intensiva e rica em insumos a serem administrados com lupa, mas que também é cumulativa para obter resultado: um erro em qualquer ponto do processo pode trazer consequências praticamente eternas aos resultados medidos a cada final de etapa.

Em terceiro, é preciso achar formas de evidenciar a matemática matreira do Ideb, não deixando que gestores espertalhões misturem fluxo com proficiência, assim como, em quarto, não se pode aceitar a correspondência falaciosa de que Ideb 6,0 equivale a resultado de ensino europeu.

Por fim, nunca se pode perder de vista que há um conhecimento poderoso (YOUNG, 2007, 2016), o qual depende da frequência à escola e da aquisição proficiente, pelo menos, de habilidades de leitura e escrita e da capacidade de operacionalizar conceitos e raciocínio matemático. Valorizar habilidades do séc. 21 quando não se chega nem perto dessas duas conquistas elementares pode ser apenas mais uma evasiva da responsabilidade moral de compartilhar poder por meio da disseminação equitativa do conhecimento formal escolar.

Referências bibliográficas

- BECSKEHÁZY, Ilona. Institucionalização do Direito à Educação de Qualidade: o caso de Sobral, CE. Tese (Doutorado em Educação), Universidade de São Paulo, Faculdade de Educação, São Paulo, SP, 2018.
- BOURDIEU, Pierre; PASSERON, Jean-Claude. A Reprodução. Trad. Reynaldo Bairão. Rio de Janeiro, RJ: Livraria Francisco Alves Editora S. A., 1975. (Educação em questão).
- FERNANDES, Reynaldo. Índice de Desenvolvimento da Educação Básica (Ideb). Textos para Discussão, v. 26, 2007. (Série Documental).
- GOVERNMENT OF Ireland. Primary School Curriculum - English Language curriculum. Dublin, Ireland: Government of Ireland, 1999.
- INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (Inep). Indicadores. Brasília, DF: [s.n.], 2018. Disponível em: <http://portal.inep.gov.br/indicadores-e-educacionais>.
- INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (Inep). Microdados da Aneb e da Anresc 2015. Brasília, DF: [s.n.], 2016. Disponível em: <http://portal.inep.gov.br/microdados>.
- INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (Inep). Microdados da Aneb e da Anresc 2017. Brasília, DF: [s.n.], 2018. Disponível em: <http://portal.inep.gov.br/microdados>.
- INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (Inep). Nota Técnica - Indicador de Nível Socioeconômico das Escolas de Educação Básica (INSE). Brasília, DF: Inep, 2015. Disponível em: http://download.inep.gov.br/informacoes_estatisticas/indicadores_educacionais/2015/nota_tecnica/nota_tecnica_inep_inse_2015.pdf.
- INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (Inep). Nota Técnica - Indicador de Nível Socioeconômico (Inse) das Escolas. Brasília, DF: Inep, 2013. Disponível em: http://download.inep.gov.br/informacoes_estatisticas/indicadores_educacionais/2011_2013/nivel_socioeconomico/nota_tecnica_indicador_nivel_socioeconomico.pdf.
- INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (Inep) .

Nota Técnica no 8/2017/CGCQTI/DEED - Estimativas de fluxo escolar a partir do acompanhamento longitudinal dos registros de aluno do Censo Escolar do período 2007-2016. Brasília, DF: INEP, 2017. Disponível em: http://download.inep.gov.br/informacoes_estatisticas/indicadores_educacionais/2007_2016/nota_tecnica_taxas_transicao_2007_2016.pdf.

- OECD (2019), PISA 2018 Results (Volume II): Where All Students Can Succeed, PISA, OECD Publishing, Paris, <https://doi.org/10.1787/b5fd1b8f-en>.
- ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD); PISA, Programme for International Student Assessment; SCHLEICHER, Andreas et al. (Orgs.). PISA 2015 Results (Volume I): Excellence and Equity in Education. Paris: OECD Publishing, 2016. (PISA). Disponível em: <http://dx.doi.org/10.1787/9789264266490-en>.
- ONTARIO MINISTRY OF EDUCATION AND TRAINING. The Kindergarten Program 2016. Toronto: Ontario Ministry of Education : Queen's Printer for Ontario, 2016. Disponível em: <http://www.edu.gov.on.ca/eng/curriculum/elementary/kinderprogram.html>.
- RIBEIRO, Sérgio Costa. A pedagogia da repetência. Estudos Avançados, v. 5, n. 12, p. 07-21, 1991.
- SINGAPORE; Curriculum Planning & Development Division. English language syllabus 2010: primary & secondary (express/normal [academic]). Singapore: Curriculum Planning & Development Division, Ministry of Education, 2008. Disponível em: <https://www.moe.gov.sg/docs/default-source/document/education/syllabuses/english-language-and-literature/files/english-primary-secondary-express-normal-academic.pdf>.
- YOUNG, Michael. Para que Servem as Escolas? Educação e Sociedade, v. 28, n. 101, p. 1287-1302, 2007.
- YOUNG, Michael. Por que o conhecimento é importante para as escolas do século XXI? Cadernos de Pesquisa, v. 46, n. 159, p. 18-37, 2016.

GAP

CRIME E VIOLÊNCIA

Joana Monteiro¹

Bárbara Caballero²

05

1 Joana Monteiro é doutora em economia pela PUC-Rio e professora da FGV/EBAPE. Foi diretora-presidente do Instituto de Segurança Pública entre 2015 e 2018 e atualmente coordena o centro de pesquisa do Ministério Público do Rio de Janeiro.

2 Bárbara Caballero é mestre em economia pela PUC-Rio. Atualmente é coordenadora de estatística do Instituto de Segurança Pública do Rio de Janeiro (ISP-RJ) e gestora de estatística do estado do Rio de Janeiro junto à Senasp/MJSP.

Introdução

A Segurança Pública é uma das áreas centrais do setor público e vem ganhando espaço na discussão pública e na agenda política. Mas qual o tamanho do problema, quais os crimes e tipos de violência mais frequentes, quais os grupos populacionais mais afetados e como esses números têm evoluído ao longo dos anos? Para responder a essas perguntas, é preciso de dados que nos permitam mensurar e acompanhar o problema.

O sistema de informação e estatísticas da área de Segurança Pública no Brasil ainda sofre de muitos problemas e limitações, sobretudo quando comparado a sistemas de informações de outras áreas centrais de provisão de serviços públicos, como saúde e educação. As três principais fontes de informações utilizadas pela área de Segurança Pública são as estatísticas baseadas em registros administrativos das polícias, os dados de mortes violentas contabilizados pelo Ministério da Saúde e as pesquisas de vitimização. Este capítulo apresenta as fontes de dados hoje existentes, sua amplitude e deficiências. As principais fontes de dados primários são as polícias civis de cada um dos estados. No caso das mortes violentas, o Sistema de Informação de Mortalidade do Ministério da Saúde oferece uma fonte de dados consistente e comparável no nível municipal. Serão também discutidas fontes de dados secundárias, que surgiram para oferecer dados comparáveis entre estados e com atualização mais frequente.

Antes de iniciar a discussão sobre as fontes de dados, cabe frisar a distinção entre diferentes dimensões na área de segurança que podem ser medidas. Tipicamente, avaliamos e acompanhamos indicadores baseados em registros de ocorrência que refletem crimes, sejam eles contra a pessoa ou contra o patrimônio. Entretanto, medir a incidência de crimes é diferente de medir a violência. Há crimes, como estelionato e lavagem de dinheiro, que não envolvem violência. Assim como há exposições à violência, como agressões verbais ou a exposição a um tiroteio, que nem sempre se constituem crimes. Por fim, há ainda a dimensão do medo. Muitas pessoas tendem a tratar a criminalidade e o medo do crime como um único fenômeno. Embora possa existir algum grau de correlação entre ambos, dependendo do grupo social e do lugar, eles têm dinâmicas próprias e causas e consequências distintas (TRINDADE; DURANTE, 2018). Medo, especificamente, é medido através de pesquisas de vitimização, que perguntam se a pessoa se sente segura em diferentes ambientes.

Para entender como reduzir os índices de criminalidade e melhorar a Segurança Pública, é fundamental entender o funcionamento do Sistema de Justiça Criminal, o

que requer ter números sobre volumes de inquéritos, denúncias e condenações, assim como identificar entradas e saídas do Sistema Prisional e as taxas de reincidência. Essa parte é a mais carente de dados e será também discutida.

Dados de Segurança Pública no Brasil

Fontes primárias de dados

Pesquisas de vitimização

A forma mais fidedigna para obter uma boa estimativa sobre a quantidade de delitos e conflitos ocorridos em um período é por meio de pesquisas de vitimização. Nesses casos, aplicam-se questionários em uma amostra representativa da população que se deseja investigar, buscando levantar informações sobre a quantidade de delitos e conflitos ocorridos no período de referência, registrados ou não pela polícia. Essas pesquisas permitem entender as características dos crimes sofridos (se houve violência ou não, qual o bem usurpado, o local do fato etc.); as características dos agressores, quando identificados, incluindo a relação existente entre vítima e agressor (familiar, conhecido, desconhecido etc.); e as medidas preventivas utilizadas pelas vítimas reais ou potenciais. Além disso, as pesquisas de vitimização oferecem fontes únicas para determinar o percentual de subnotificação, ou seja, o volume de crime que não chega ao conhecimento da polícia (as chamadas “cifras ocultas”). Elas ajudam também a conhecer as percepções de insegurança e ter métricas de confiança nas instituições do sistema de justiça criminal. Outra vantagem associada a esse tipo de pesquisa é a possibilidade de se conhecerem melhor diferentes aspectos de crimes de natureza sexual, ou outras violências praticadas em ambientes privados ou familiares (ZILLI, 2018).

Infelizmente, pesquisas de vitimização ainda são raras no Brasil, principalmente devido ao seu alto custo financeiro e à ausência de prioridade nacional em entender o fenômeno de segurança pública em maior profundidade. Há apenas três pesquisas feitas em nível nacional no Brasil. Em 1988 e 2009, o IBGE aplicou um questionário suplementar sobre Justiça junto com a Pesquisa Nacional de Amostra de Domicílios (PNAD). Além dessas pesquisas, em 2010, o CRISP/UFMG³, em parceria com o Instituto de Pesquisas Datafolha, e financiado pelo Ministério da Justiça, realizou uma pesquisa com 78 mil pessoas em 346 municípios. Essa pesquisa revelou, por exemplo, que 30% dos furtos de veículos e 58% dos roubos de objetos pessoais não são registrados na polícia. Além disso, 18% da população disse confiar muito na polícia, com grandes variações entre os estados. Há, ainda, uma série de pesquisas com amostras menores, com abrangência mais restrita e periodicidade mais estendida já realizadas no Brasil,

3 Centro de Estudos de Criminalidade e Segurança Pública da Universidade Federal de Minas Gerais.

mas sua ausência de continuidade limita seu uso⁴. Infelizmente, nenhuma das três pesquisas são totalmente comparáveis, o que seria bastante importante para termos referências sobre a evolução temporal dos indicadores levantados. Cabe frisar, também, a iniciativa da Secretaria de Segurança Pública do Distrito Federal, que aplica anualmente, desde 2015, uma Pesquisa de Vitimização Distrital, com quase 20 mil pessoas e com representatividade no nível das 31 regiões administrativas do Distrito Federal. O questionário coleta dados relativos à vitimização das pessoas nos últimos 12 meses, nos seguintes crimes: roubo, furto, fraude, ameaça, agressão, ofensa sexual e discriminação, e tem a enorme vantagem de permitir comparabilidade ao longo do tempo, sendo um raro instrumento de acompanhamento da sensação de segurança e do nível de confiança na polícia.

Registros de ocorrência das polícias

A polícia é a porta de entrada do Sistema de Justiça Criminal. Tipicamente, uma vítima, quando sofre uma violência ou perde um bem e deseja reportá-lo a uma autoridade, busca a polícia. Como no Brasil há duas polícias, esse primeiro contato pode se dar com a Polícia Militar, que faz o patrulhamento ostensivo e emergencial, ou com a Polícia Civil, responsável pelo registro ou boletim de ocorrência⁵.

A maioria das estatísticas disponíveis sobre crimes e violências no Brasil é oriunda de registros de ocorrência junto à Polícia Civil. Cada estado brasileiro tem regras distintas de cômputo e divulgação de indicadores criminais. A Tabela no Anexo 1 indica o órgão responsável pela divulgação de estatísticas em cada estado e mostra que, geralmente, essa é uma função das Secretarias de Segurança Pública. A principal exceção é o Rio de Janeiro, onde o Instituto de Segurança Pública⁶ exerce essa função.

Sistema de Informações de Mortalidade (SIM)

Outra fonte de bastante relevância para a mensuração da violência são os dados advindos do Sistema de Saúde, que tem uma tradição muito maior em coleta e sistematização de informação. A base de dados sobre mortalidade da Secretaria de Vigilância em Saúde (SVS), do Ministério da Saúde (MS), é formada pela informação

4 Zilli (2018) apresenta uma lista de 46 pesquisas de vitimização já realizadas no Brasil.

5 Em alguns Estados, como Santa Catarina e Minas Gerais, os policiais militares lavram o Termo Circunstanciado de Ocorrência (TCO) no local do flagrante, o que dispensa as vítimas, nessas situações de emergência, de irem até uma delegacia da polícia civil para registrar a ocorrência.

6 Autarquia do governo estadual, atualmente vinculada à Secretaria de Estado de Planejamento e Gestão.

contida nas Declarações de Óbitos (DO) e contabiliza o número total de mortes no Brasil. Esses números estão disponíveis através do Sistema de Informações de Mortalidades (SIM), pelo portal Datasus. O SIM fornece as séries de dados mais longas disponíveis no Brasil e permite comparação ao nível municipal.

Sistema de Vigilância de Violências e Acidentes (Viva)

A área da saúde também fornece informações importantes sobre violência através do Sistema de Vigilância de Violências e Acidentes (Viva), que foi instituído em 2006. A partir de 2011, a notificação de violências passou a integrar a lista de notificação compulsória, o que universalizou a notificação para todos os serviços de saúde. Entretanto, o número de municípios e de serviços especializados aptos a fazerem o registro aumentou nos últimos anos, o que ainda dificulta comparações intertemporais. O Viva permite captar informações sobre os atendimentos em unidades de saúde, por meio de um questionário que deve ser preenchido pelos profissionais de saúde em unidades públicas ou privadas quando identificam um caso suspeito ou confirmado de violência, caracterizando o perfil das pessoas que sofreram violências, o local, o perfil do provável autor da agressão, dentre outras informações. O sistema classifica diferentes tipos de violência: violência física; violência psicológica; tráfico humano; trabalho infantil; tortura; violência sexual; violência financeira; negligência ou abandono; intervenção legal; e outras.

Dados do Sistema de Justiça

O registro de ocorrência de um crime é o primeiro passo na entrada do Sistema de Justiça Criminal. A partir daí, a Polícia Civil deve decidir se abre o inquérito. Esse inquérito deve ser enviado ao Ministério Público (MP) quando está relatado (ou seja, quando a polícia identificou elementos que indicam a autoria do crime) ou para pedir prazo. O Ministério Público avalia se há justa causa para denunciar os acusados perante a Justiça. Não havendo elementos suficientes, o MP pode optar por devolver o inquérito à Polícia Civil para solicitar novas diligências ou arquivar o caso. Na prática, um inquérito pode ir e voltar entre Polícia Civil e Ministério Público inúmeras vezes, seja para solicitação de prazo ou para solicitação de medidas que ajudem a elucidar o caso. Na fase judicial, o juiz responsável deve decidir se aceita a denúncia, que, uma vez aceita, será devidamente analisada e julgada.

Dados sistemáticos sobre esse processamento não são produzidos atualmente por nenhum órgão brasileiro. Os dados disponíveis sobre o funcionamento da Justiça Criminal são baseados em estudos como Nunes *et al.* (2016), Ribeiro *et al.* (2014) e Mon-

teiro et al. (2020), que analisaram a duração de fases do processamento de homicídios em diferentes capitais brasileiras. Ribeiro *et al.* (2014) revelam que o Sistema de Justiça brasileiro é muito moroso, com médias de tempo total (da descoberta do crime à sentença do júri) entre 2.058 e 3.403 dias (entre 5,6 e 9,3 anos).

Dados do Sistema Penitenciário

O Departamento Penitenciário Nacional (Depen), do Ministério da Justiça e Segurança Pública (MJSP), emite todo ano relatórios com os dados básicos do sistema prisional, como número de presos por estado, por tipo de regime (provisório, fechado, aberto, semi-aberto), por tipo penal, perfil dos detentos, entre outras informações. Os dados são coletados por meio do Levantamento Nacional de Informações Penitenciárias (Infopen), a partir de formulários preenchidos pelos estados da federação. Contudo, não há nenhum cálculo disponível sobre taxas de reincidência ao sistema prisional no Brasil.

Fontes secundárias de dados

Sinesp

As informações estaduais são centralizadas no Sistema Nacional de Informações de Segurança Pública, Prisionais, de Rastreabilidade de Armas e Munições, de Material Genético, de Digitais e de Drogas (Sinesp), gerenciado pela Secretaria Nacional de Segurança Pública do Ministério da Justiça e Segurança Pública (Senasp/MJSP). Esse sistema foi implantado em 2004, mas a lei que institui o Sinesp, no entanto, é de 2012 (Lei nº 12.681, de 4 de julho de 2012), quando então foram iniciados projetos para a coleta automatizada de dados para integração dos sistemas estaduais de registros de ocorrência policial⁷. Os dados do Sinesp só se tornaram de fácil acesso para o público em 2019, quando foi lançada uma plataforma online com dados mensais desde 2015 para os indicadores de homicídio doloso, lesão corporal seguida de morte, roubo seguido de morte (latrocínio), estupro, furto de veículo, roubo a instituição financeira, roubo de carga, roubo de veículo e tentativa de homicídio.

Anuário Brasileiro de Segurança Pública

O Fórum Brasileiro de Segurança Pública, responsável pela publicação do Anuário, é uma organização sem fins lucrativos que tem entre suas atividades a divulgação dos números de segurança pública do país de maneira organizada e padronizada. Os dados utilizados, obtidos via Lei de Acesso à Informação – LAI (Lei nº 12.527, de 18 de novembro de 2011), provêm das instituições governamentais responsáveis pelas estatísticas oficiais de crime em cada estado. O Anuário Brasileiro de Segurança Pública (ABSP) divulga números anuais por unidade da federação de 24 indicadores⁸, além de informações do sistema prisional e gastos com segurança pública. A publicação do Anuário ocorre normalmente entre agosto e setembro do ano seguinte.

7 Em 2015, foi criado o Sinesp Integração, outra solução de consolidação de dados em uma única base nacional. A adesão ao Sinesp Integração ainda não é completa por parte dos estados, devido às diferentes formas de coleta e tratamento de dados pelas polícias estaduais.

8 Os indicadores divulgados são: homicídio doloso, latrocínio, lesão corporal seguida de morte, morte por intervenção de agente do Estado, policiais civis e militares mortos em situação de confronto, mortes a esclarecer, estupro, tentativa de estupro, pessoas desaparecidas e pessoas localizadas, roubo de veículo, furto de veículo, roubo a instituição financeira, roubo a estabelecimento comercial, roubo a residência, roubo a transeunte, roubo de carga, total de roubos, suicídio, tentativa de homicídio, lesão corporal dolosa, tráfico de entorpecentes, posse e uso de entorpecentes.

Atlas da Violência

O Atlas da Violência é uma publicação do Ipea em parceria com o Fórum de Segurança que apresenta números anuais de diversos indicadores, com ênfase na violência letal, com recorte por área (região, estado e município) e perfil da vítima (sexo, idade e cor), entre outras análises. A fonte são os dados oficiais do Sistema de Informações sobre Mortalidade, do Ministério da Saúde (SIM/MS).

Monitor da Violência

O Monitor da Violência é um levantamento dos números de homicídio, lesão corporal seguida de morte e latrocínio feito pelo portal de notícias G1 em parceria com o Núcleo de Estudos da Violência da Universidade de São Paulo e com o Fórum Brasileiro de Segurança Pública, a partir de informações oficiais obtidas via LAI. Com dados mensais para cada unidade da federação, é o portal com menor defasagem de tempo disponível hoje.

Ainda que os dados sejam obtidos nas mesmas fontes primárias, os números apresentados no Monitor da Violência, no Anuário e no Sinesp divergem entre si, como mostra a Tabela 1. Isso ocorre, sobretudo, devido ao momento em que esses números são obtidos e atualizados, porque muitos estados fazem revisões periódicas das suas estatísticas, e os portais podem incorporar essas atualizações em momentos distintos. Outro ponto de divergência é a contabilização de mortes por intervenção de agentes do Estado, que não é contabilizada pelo Sinesp, conforme discutido a seguir.

Tabela 1 – Número de mortes violentas intencionais no Brasil por fonte de dados

Ano	Sistema de Informações de Mortalidade - Ministério da Saúde (SIM/MS) ¹				Anuário Brasileiro de Segurança Pública - Fórum Brasileiro de Segurança Pública ²				Sinesp - Ministério da Justiça e Segurança Pública ³		Monitor da Violência - Site G1	
	Agressões	Intervenção Legal	Mortes violentas intencionais (MVI)	Taxa de MVI	Crimes violentos letais intencionais (CVLI)	Mortes decorrentes de intervenção policial	Mortes violentas intencionais (MVI)	Taxa de MVI	Crimes violentos letais intencionais (CVLI)	Taxa CVLI	Crimes violentos letais intencionais (CVLI)	Taxa CVLI
2012	56.337	708	57.045	28,6	53.054	2.332	54.694	28,2	nd	nd	53.054	26,6
2013	56.804	592	57.396	28,6	54.163	2.212	55.847	27,8	nd	nd	54.163	26,9
2014	59.681	793	60.474	29,8	57.091	3.146	59.730	29,5	nd	nd	57.091	28,2
2015	58.138	942	59.080	28,9	55.492	3.330	58.459	28,6	52.306	25,6	55.492	27,1
2016	61.143	1.374	62.517	30,3	57.842	4.240	61.597	29,9	55.085	26,7	57.842	28,1
2017	63.748	1.854	65.602	31,6	59.523	5.179	64.078	30,9	60.478	29,0	59.128	28,5
2018	nd	nd	nd	nd	51.765	6.220	57.358	27,5	52.057	25,0	51.558	24,7
2019	nd	nd	nd	nd	nd	nd	nd	nd	42.191	20,2	41.635	20,0

Fontes: Ministério da Saúde, Fórum Brasileiro de Segurança Pública, Ministério da Justiça e Segurança Pública e Monitor da Violência (G1). Dados acessados em abril de 2020.

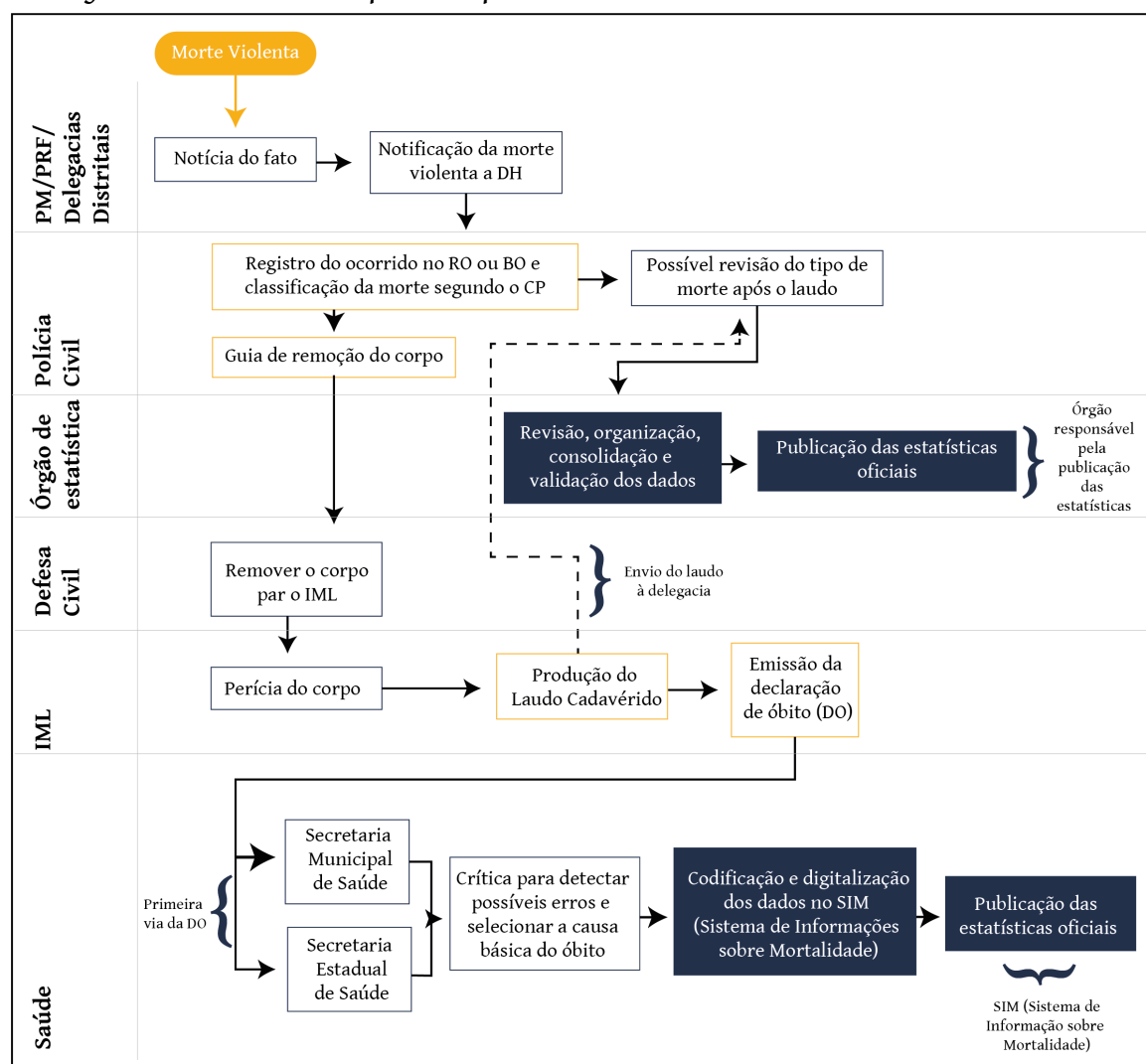
Diferenças entre dados das polícias e dados do sistema de saúde

Uma questão central no cômputo de mortes violentas é a comparação entre dados oriundos da polícia e do sistema de saúde. A classificação da morte como homicídio no SIM vem através da CID-10⁹, uma classificação internacional de doenças, passível de comparação inclusive entre países, que foge das definições penais e inclui, por exemplo, as mortes por intervenção de agente do Estado. Como explica Dirk (2017), que apresenta uma comparação entre dados da saúde e do sistema de justiça criminal no Estado do Rio de Janeiro, o indicador oriundo da polícia mais comparável ao da área de saúde é o de mortes violentas intencionais (MVI). A Tabela 1 mostra que a fonte de dados que informa as mortes violentas de forma mais próxima ao cômputo do Ministério da Saúde é o indicador de mortes violentas intencionais contabilizadas pelo Anuário Brasileiro de Segurança Pública.

Entretanto, o número de MVI contabilizadas pelo sistema de justiça e pela área de saúde tende a diferir em maior ou menor grau porque seus sistemas de informação servem a propósitos distintos. Para o sistema de segurança pública e justiça criminal importa saber se houve ou não um crime e tipificá-lo de acordo com a categoria penal correta, ao passo que para a saúde importam as informações de cunho epidemiológico relacionadas ao perfil da vítima e em que contexto houve morte. Informações sobre as circunstâncias da morte constam na Declaração de Óbito (DO), que deve ser preenchida pelo médico-legista a partir dos laudos cadavéricos produzidos nos Institutos Médico-Legais (IML) em todos os casos de mortes por causas externas, visto que há uma possibilidade de cometimento de crime. A Figura 1, na página seguinte, apresenta o fluxo de processo e notificações de uma morte violenta e sua passagem pelas diferentes instituições.

9 A Classificação Internacional de Doenças (CID-10) divide as mortes em dois grandes grupos: (i) mortes por causas naturais: são as doenças infecciosas, neoplasias, doenças do aparelho circulatório ou respiratório, entre outras; (ii) mortes por causas externas: são consideradas mortes por causas não naturais, tais como as agressões, as lesões auto-infligidas, os acidentes de todos os tipos, entre outras.

Figura 1 - Fluxo de dados para cômputo de mortes violentas



Fonte: Elaboração própria

O problema começa porque nem sempre o médico-legista tem um conjunto suficiente de informações que permita definir a intencionalidade da ação na Declaração de Óbito, resultando em uma morte cuja intenção é indeterminada. Com isso, o percentual de mortes indeterminadas tende a ser alto em estados onde os sistemas de informação são precários e onde há pouca integração de dados entre Polícia Civil e Secretaria de Saúde. O maior número de mortes indeterminadas implica diretamente em um número artificialmente mais baixo de agressões.

Outra distinção entre os sistemas é a temporalidade. Os dados oriundos da polícia tendem a ser publicados mensal ou trimestralmente, enquanto os dados da saúde têm publicação anual. Isso gera uma distinção entre os números de mortes tentadas e consumadas. Como o sistema de saúde tem um processamento de dados mais longo, ele acaba contabilizando melhor os casos de lesões e homicídios tentados que resultaram

em morte em um momento subsequente ao evento criminal¹⁰. Na prática, os dados de saúde têm uma defasagem de quase dois anos, o que faz com que os dados da polícia sejam mais utilizados para fins de monitoramento de ações na área de segurança.

Há, ainda, uma diferença relevante para quem deseja comparar dados em níveis geográficos menores. Os dados da saúde são computados segundo o local de ocorrência do óbito ou local de residência da vítima, enquanto os da polícia são computados segundo o local de ocorrência do evento.

Todos esses fatores implicam que os dados de ambas as fontes nunca serão iguais, mas precisam ser congruentes no sentido de permitirem identificar os locais com mais e menos violência e haver tendências similares ao longo do tempo.

Armadilhas no cômputo dos dados

Indicadores comparáveis

Existem diversos desafios para se medir crime e violência. Na seção anterior, discutimos as fontes de dados. O fato que grande parte das estatísticas advém de sistemas de informações estaduais requer muito cuidado na comparação de números de diferentes estados. Esse contexto gera uma série de armadilhas nos números e nas informações disponíveis hoje no Brasil sobre o tema. É preciso entender o que cada número está nos contando para sabermos compará-los da maneira correta.

O Sinesp é alimentado pelos órgãos oficiais responsáveis por esses números em cada unidade da federação e cuja lógica segue capitulações do código penal, que estabelece penas distintas de acordo com a gravidade do crime. Assim, um dos principais desafios é produzir estatísticas criminais comparáveis, garantindo que todos os estados computem os mesmos tipos de crime em cada indicador. Vamos tomar o exemplo do cômputo de homicídios dolosos, cuja definição mais simples é matar alguém com intenção. Dependendo do estado, esse indicador pode ou não incluir mortes que são decorrentes de roubos (latrocínios), mortes “por razões da condição de sexo feminino” (feminicídios), mortes decorrentes de estupro e mortes por intervenção de agente

¹⁰ Por exemplo, uma tentativa de homicídio pode se tornar um homicídio doloso consumado, e assim contar como morte no sistema da saúde, mas tal fato foi contado somente como tentativa de homicídio na publicação dos dados de polícia. Por esse motivo, é crucial haver revisões periódicas nos sistemas de dados com base em informações da polícia de forma que erratas possam compatibilizar melhor as bases de dados das duas fontes.

do Estado, entre outras. Esses crimes, para exemplificar alguns, têm capitulação específica no Código Penal e por isso muitas vezes são contabilizados de forma separada. Um levantamento de 2016 do Fórum Brasileiro de Segurança Pública em parceria com o então Ministério da Justiça mostra que das 32 unidades (dentre secretarias estaduais de segurança, polícias civil e militar) que responderam ao questionamento sobre as categorias contabilizadas em homicídio doloso, 88% afirmaram incluir feminicídio; 50% informaram contabilizar civis mortos em confronto com a polícia; 34% incluíam lesão corporal seguida de morte; 31% contabilizavam roubo seguido de morte; e 19% somavam também homicídios culposos¹¹.

Para dar conta do tipo de morte contida na categoria homicídio em um sentido mais amplo, iniciativas nacionais de contabilização utilizam conceitos como Crimes Violentos Letais Intencionais (CVLI) e Mortes Violentas Intencionais (MVI). A categoria CVLI inclui homicídio doloso, latrocínio e lesão corporal seguida de morte e passou a ser usada no Anuário Brasileiro de Segurança Pública (ABSP) e por Estados como Pernambuco, Ceará e Espírito Santo no desenho dos seus Sistemas de Gestão por Resultados. Em 2015, o ABSP também criou a categoria de MVI, agregando homicídio doloso, latrocínio, lesão corporal seguida de morte e mortes por intervenção de agente do Estado, cujo principal objetivo é incluir o número de pessoas mortas pela polícia. Este é um ponto polêmico, que será discutido mais adiante.

Neste contexto, é fundamental o estabelecimento de protocolos para construção de indicadores, de forma a criar comparabilidade entre eles. Há duas iniciativas que merecem destaque. O Protocolo de Bogotá estabelece critérios internacionais para o cômputo de estatísticas de homicídios. O protocolo é resultado de um empenho de um grupo de organizações da sociedade civil e acadêmicos que se reuniu em 2015, na cidade de Bogotá, com o intuito de tornar as estatísticas de homicídio latino-americanas mais comparáveis¹². No âmbito nacional, no fim de 2018, o então Ministério da Segurança Pública publicou portaria padronizando os nomes e as classificações incluídas nos principais indicadores criminais (Portaria nº 229, de 10 de dezembro de 2018). Esforços nesse sentido são fundamentais para garantir, por exemplo, que os estados computem quase todas as mortes violentas dentro do indicador “homicídio doloso” e que haja um indicador para contabilizar de forma separada as mortes cometidas por ação de agentes do Estado. Mas a publicação tardia da normativa e as notas dos ges-

11 Fórum Brasileiro de Segurança Pública (2016).

12 Disponível em http://www.forumseguranca.org.br/storage/publicacoes/FBSP_Estudo_classificacoes_estatisticas_tipificacoes_penais_2016.pdf. Acessado em agosto de 2019.

tores estaduais disponíveis no portal do Sinesp demonstram como é necessário ainda evoluir na padronização das estatísticas entre estados no Brasil.

A Portaria define, para fins estatísticos, os homicídios como “(a) Morte de alguém em que há indício de crime ou sinal de agressão externa, exceto ‘Feminicídio’, ‘Lesão Corporal Seguida de Morte’, ‘Roubo Seguido de Morte (Latrocínio)’ e crimes culposos; (b) Morte violenta provocada por acidente de trânsito, desde que haja dolo; e (c) Morte com indício de crime ou sinal de agressão externa qualificada como ‘encontro de ossada’, ‘encontro de cadáver’, ‘morte a esclarecer’, ‘morte suspeita’, ‘morte por causa desconhecida’ e congêneres”. Porém, tal definição ainda não é utilizada pela maioria das polícias estaduais.

Os feminicídios e as mortes de autoria de policiais são duas categorias que demandam bastante atenção pública e enfrentam desafios especiais no seu cômputo. Os feminicídios, que são mortes “por razões da condição de sexo feminino”, são uma titulação criminal incluída no artigo 121 do Código Penal Brasileiro, que trata do crime de homicídio, pela Lei nº 13.104, de 9 de março de 2015. Portanto, em termos jurídicos, esta classificação está necessariamente contida no indicador de homicídio doloso. Por isso, Estados como Bahia, Ceará, Mato Grosso e Rio de Janeiro contabilizam os casos de feminicídio como homicídio doloso, ainda que seja permitido obter o número em separado também.

Mas a nova titulação fez surgir um novo indicador em muitos estados e comparações errôneas em muitos meios de comunicação, indicando crescimentos percentuais de mais de 50% no número de feminicídios, decorrentes do simples fato de que no período base de comparação esses casos não eram bem identificados. De fato, o cômputo do fenômeno de feminicídios apresenta desafios muito maiores do que a aplicação da Lei nº 13.104. Um registro de uma morte com essa titulação pressupõe que a autoridade policial responsável pelo caso identifique, no momento do registro, a motivação daquele homicídio, entendendo que o crime contra a mulher ocorreu “por razões da condição de sexo feminino”. Por mais que, muitas vezes, esses casos sejam rapidamente identificados, nem sempre isso é confirmado no momento da lavratura do registro de ocorrência. Dessa forma, são necessárias iniciativas como a desenvolvida pela Polícia Civil da Paraíba, que realiza um trabalho contínuo de análise criminal, que permite classificar as motivações dos crimes e, com isso, identificar o número de mortes “por razões da condição de sexo feminino”, assim como o número de mortes por LGBTfobia.

Como denominar e contar mortes provocadas pela polícia

Contabilizar mortes envolvendo policiais e outros agentes do Estado tem se mostrado bastante desafiador no Brasil. A primeira dificuldade se dá na padronização do nome. No passado recente, essa categoria já foi (e, na prática, ainda é) chamada de “morte por intervenção policial”, “morte decorrente de intervenção policial”, “homicídio decorrente de intervenção policial”, “homicídio decorrente de oposição à intervenção policial”, “morte decorrente de oposição à violência policial”, “resistência seguida de morte” e “auto de resistência” – este último até hoje utilizado pela mídia e sociedade em geral, independentemente das mudanças de denominações. A discussão em torno do nome envolve alguns pontos. Primeiro, há uma questão conceitual, visto que mortes envolvendo agentes do Estado não seriam crimes intencionais, mas sim resultado de sua função legal e, por isso, não poderiam ser chamadas de homicídios. Segundo, a questão da resistência, cuja menção visa contabilizar apenas as mortes onde a vítima apresentou resistência, e não todas as mortes cuja autoria é de um policial. Por fim, há a menção a polícia ou agente do Estado, cuja nomenclatura implica se mortes por agentes penitenciários ou membros das Forças Armadas devem ser computadas ou não.

A Portaria nº 229, de 10 de dezembro de 2018, padroniza o nome para “morte por intervenção de agente do Estado”, cuja definição é “Morte por intervenção de agente de segurança pública, do sistema prisional ou de outros órgãos públicos no exercício da função policial, em serviço ou em razão dele, desde que a ação tenha sido praticada sob quaisquer das hipóteses de exclusão de ilicitude”.

Na prática, há uma enorme variabilidade em como os estados contabilizam essas mortes. Estados como Acre, Mato Grosso, Mato Grosso do Sul, Minas Gerais, Paraíba, Pernambuco, Piauí, Rio Grande do Sul e Rondônia incluem as mortes por intervenção de agente do Estado como homicídio doloso¹³, enquanto outros contabilizam de forma separada. Essa é uma questão muito relevante, visto que em Estados como Rio de Janeiro, Pará, São Paulo e Sergipe, mortes envolvendo agentes do estado chegam a representar mais de 20% no total de mortes violentas intencionais. Para piorar, muitas vezes a mídia escolhe as terminologias com vistas a facilitar o entendimento público, e acaba cometendo erros. É o caso de matéria publicada pelo jornal O Globo em janeiro de 2019 (ver Figura 2), onde a confusão entre os termos já fica explícita no título “Autos de resistência representam 43,8% dos homicídios registrados na região

de Niterói em 2018”, sugerindo, equivocadamente, que as mortes por intervenção de agente do Estado fossem um subconjunto dos casos de homicídio.

Figura 2 - Matéria do jornal O Globo em 27 de janeiro de 2019



Fonte: Jornal O Globo de 27 de janeiro de 2019. Disponível em: <https://oglobo.globo.com/rio/bairros/autos-de-resistencia-representam-438-dos-homicidios-registrados-na-regiao-de-niteroi-em-2018-23403503>. Acessado em agosto de 2019.

A armadilha principal aqui é entender que esse número não significa o número de “pessoas mortas pela polícia”, como costuma sair na imprensa, mas sim o número de pessoas mortas em confronto com a polícia ou que resistiram à ação policial. Mortes ocorridas em situações de confronto, porém sem envolvimento da vítima, isto é, os casos de “bala perdida”, mesmo que vindas da polícia, costumam ser registrados como homicídio doloso, uma vez que a vítima não ofereceu resistência.

A diferença entre a data do crime e a data do registro

Um critério simples, porém de extrema relevância para análise de dados criminais, é a definição do período de apuração dos números: os crimes podem ser contabilizados pela data em que ocorreram ou pela data em que foram registrados nos sistemas das polícias.

Para efeitos de planejamento operacional e estratégico das polícias, para análise criminal e para os setores de inteligência, é bastante natural que os crimes devam ser medidos pelo momento em que ocorreram¹⁴. Por outro lado, as estatísticas oficiais costumam se referir ao número de crimes registrados, independentemente de

¹⁴ Uma exceção é o Estado do Mato Grosso do Sul, que divulga o número de delitos pela data do fato. Além disso, divulgam também as ocorrências que foram registradas no estado, independentemente de onde ocorreram. Já o Distrito Federal divulga os números tanto pela data do fato quando pela data do registro. Fonte: Sinesp (agosto de 2019).

quando ocorreram. Isso faz sentido também, pois, caso contrário, nunca existiria um número final – os números estariam constantemente sujeitos a mudança, sempre que novos registros fossem feitos.

O lapso de tempo entre a data do fato e a data do registro varia de acordo com o tipo de crime. Análise de dados de registro de ocorrência no Rio de Janeiro mostra que 90% dos casos de roubo a transeunte são registrados no mesmo mês do fato. Esse percentual é bem mais baixo, entretanto, para casos de estupro, onde cerca de 60% dos casos foram registrados no mesmo mês do fato. Essa questão já não é problemática quando olhamos para os registros de homicídio, crimes de trânsito e roubo de veículo, por exemplo. Os dois primeiros por razões óbvias, já que envolvem vítimas fatais, com a obrigatoriedade do atendimento policial ao evento. E, no caso de roubo de veículo, pelo interesse da própria vítima em registrar para poder acionar a seguradora o quanto antes.

Contudo, mesmo nos casos em que a diferença é grande, se esse percentual é estável ao longo do tempo, não há muito com o que se preocupar em termos estatísticos, e as análises temporais não serão prejudicadas. As análises comparativas podem ser afetadas quando queremos acompanhar diariamente quantos crimes estão ocorrendo.

Vejamos o exemplo dos casos de roubo a transeunte. É muito comum que as vítimas deixem para registrar o ocorrido em um final de semana, ou em um feriado, apenas no próximo dia útil. As segundas-feiras costumam apresentar o maior número de registros de roubo da semana.

É compreensível a procura da imprensa por um furo de reportagem, para dar uma notícia em primeira mão. Assim como é importante para o gestor de segurança pública o monitoramento das ocorrências criminais diariamente. Entretanto, para fins estatísticos, os dados devem passar por um processo de consolidação, onde são tratadas inconsistências, duplicidades e mudanças de titulação que possam ocorrer entre o registro inicial e o decorrer do inquérito. Essa divergência entre o dado parcial e o consolidado pode ser bem relevante devido a dois motivos. Primeiro, em grandes feriados e grandes eventos que são comumente de interesse da mídia, é comum as pessoas demorarem a registrar os crimes ocorridos. A Tabela 2 mostra a diferença entre dados parciais consultados na sexta-feira de manhã após o Carnaval de 2018 no Estado do Rio de Janeiro e os dados consolidados para o mesmo período. Como referência, são exibidos também os dados relativos ao período de Carnaval de 2015, 2016 e 2017, este último artificialmente baixo porque a Polícia Civil se encontrava em greve

no mês em questão. Note que, entre a consulta parcial e os números consolidados, houve um aumento de 70% nos roubos de rua e de 62% nos furtos de rua, o que mostra a necessidade de aguardar a consolidação de números para ter uma real dimensão do fenômeno.

Tabela 2 - Roubos e furtos no Carnaval - Estado do Rio de Janeiro - 2015 a 2018

Delito	13 a 18 fev 2015	05 a 10 fev 2016	24 fev a 01 mar 2017	09 a 14 fev 2018 (parcial)	09 a 14 fev 2018 (consolidado)	Diferença entre o consolidado e o parcial (2018)
Roubo de Rua ¹	1.811	2.185	1.820	1.527	2.595	70%
Furto de Rua ²	2.983	2.495	1.168	1.030	1.670	62%
Letalidade Violenta ³	76	114	108	92	103	12%

1 - O indicador Roubo de Rua representa a soma dos casos de roubo a transeunte, roubo de aparelho de celular e roubo em coletivo.

2 - O indicador Furto de Rua representa a soma dos casos de furto a transeunte, furto de aparelho de celular e furto em coletivo.

3 - O indicador de Letalidade Violenta representa a soma das vítimas de homicídio doloso, morte por intervenção de agente do Estado, latrocínio e lesão corporal seguida de morte.

Fonte: ISP

Contando vítimas ou casos

Outro critério metodológico relevante é a contabilização dos crimes pelo número de vítimas ou pelo número de casos. Convém que os crimes contra o patrimônio (roubos, furtos, estelionatos etc.) sejam contabilizados pelo número de ocorrências, independente do número de vítimas, evitando, assim, a dupla (ou múltipla) contagem do mesmo fato. Já os crimes contra a pessoa (homicídio, lesão corporal, ameaça, injúria etc.), mesmo que não sejam letais, devem ser contabilizados por vítima.

De acordo com o Anuário Brasileiro de Segurança Pública 2018, a contagem de homicídio doloso no Estado do Mato Grosso do Sul era por número de casos, e não por número de vítimas; o mesmo para o Estado de São Paulo, nos casos de lesão corporal seguida de morte; já em Minas Gerais, não havia distinção entre homicídio doloso e homicídio culposo, mas apenas a contagem de homicídio consumado; o Amapá, por sua vez, não excluía duplicidades de ocorrências provenientes de aditamentos, geran-

do um número de casos maior que o número de vítimas¹⁵.

Em pesquisa do Fórum Brasileiro de Segurança Pública de 2016, já citada anteriormente, foi identificado que, das 36 unidades respondentes, dez afirmaram contabilizar os homicídios por número de casos, e não de vítimas¹⁶. Essa distinção pode prejudicar a comparação entre estados, visto que a contabilidade de homicídios por registro, e não por vítima, pode gerar números cerca de 8% menores¹⁷.

Essa diferença se dá devido aos casos de homicídios múltiplos, quando mais de uma morte é registrada em um único evento criminal. Inclusive, uma das formas encontradas pela literatura para identificar chacinas é contabilizar o número de vítimas por registro e classificar como chacina quando há três ou mais vítimas na mesma ocorrência. Entretanto, tal cálculo só é possível de ser feito quando se tem acesso aos microdados e, mesmo assim, não permite identificar a totalidade de casos de homicídios múltiplos, visto que há casos em que mortes que ocorreram em um mesmo evento podem não ser registradas no mesmo boletim de ocorrência. O mais correto seria buscar o total de vítimas que morreram na mesma data, no mesmo horário e no mesmo local. Esta última informação é a que traz maior dificuldade para ser apurada, pois muitas vezes é parcialmente preenchida ou até reservada para a polícia.

Mais registros, menos crimes

Outra armadilha que ocorre ao analisarmos a evolução de um crime se dá quando o registro do mesmo é influenciado por campanhas sociais ou governamentais. Neste caso, serão afetados aqueles crimes cujo registro pode envolver questões de cunho pessoal para a vítima, seja por vergonha, medo de represália, receio de estigmatização etc. – são os crimes de estupro, de violência contra a mulher, crimes motivados por LGBTfobia e por intolerância religiosa, por exemplo.

Nas últimas duas décadas, pode-se observar no Brasil um grande avanço nas discussões sobre esses temas, aumentando o diálogo na sociedade, com manifestações públicas e até mesmo mudanças na legislação que incentivaram o debate sobre esses crimes, tendo como consequência um aumento no número de registros, seja de fatos

15 Informações apuradas no Portal Sinesp e no Anuário Brasileiro de Segurança Pública 2018, em agosto de 2019.

16 Fórum Brasileiro de Segurança Pública (2016).

17 Valor calculado através da comparação entre a soma de registros de ocorrência e o número de vítimas de homicídio em 2017 e 2018, conforme disponível no Portal Sinesp.

ocorridos no passado ou no presente.

Especificamente para o crime de estupro, a Lei nº 12.015, de 7 de agosto de 2009, fez a junção deste ao crime de atentado violento ao pudor (AVP), revogando o artigo 214 e alterando o artigo 213 do Código Penal (estupro) – e também as estatísticas desse crime. O AVP representava cerca de 60% do crime de estupro¹⁸.

Em que pese que todas essas medidas tenham o objetivo de combate à violência contra a mulher, com a consequente redução do número de crimes, e, de fato, há estudos que comprovam a eficácia dessas políticas (CERQUEIRA *et al.*, 2015), esses eventos contribuíram sobretudo para a redução da subnotificação desses crimes. Por isso, deve-se ter muito cuidado ao analisar o aumento recente desses indicadores, pois não sabemos, *a priori*, se de fato esta violência está ocorrendo mais, ou se só estão sendo mais registradas devido à maior conscientização da população e o maior acesso às delegacias.

Mais usuários, mais registros

De maneira análoga, também podemos citar a armadilha em que podemos cair ao estudar a evolução de números apurados por aplicativos de uso recente. São os casos em que, na ausência de estatísticas oficiais sobre tiroteio, visto que este evento não configura uma capitulação penal (o evento, se registrado, será contabilizado como homicídio, se houver vítima fatal, ou pelo menos uma tentativa de homicídio e/ou disparo de arma de fogo), muitos meios de comunicação buscam os aplicativos de celular como fonte de dados. No entanto, como a contabilização desse número depende da disseminação do uso da ferramenta, é natural que nos primeiros anos os números sejam sempre crescentes, não por haver cada vez mais tiroteios, mas por haver cada vez mais usuários inserindo informação.

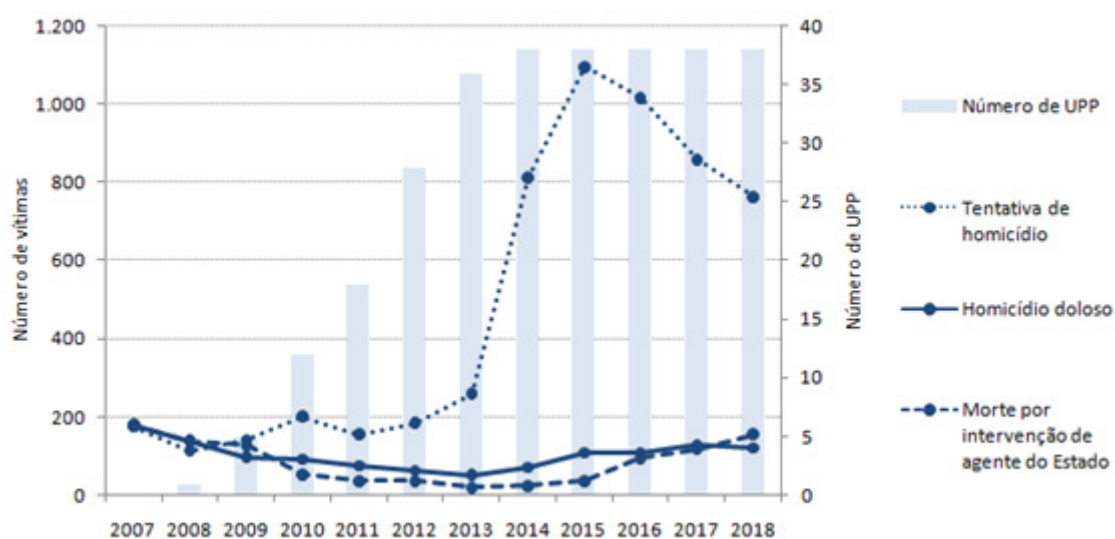
Mudanças nas regras de registros

Outra armadilha na análise de dados surge porque as estatísticas são baseadas em dados administrativos, cuja função primordial é coletar informações para abertura de inquéritos. Assim, é possível que mudanças procedimentais causem variações abruptas nos números. Por exemplo, em 2014, a Polícia Militar do Rio de Janeiro passou a instruir seus membros que atuavam em áreas de UPP que lavrassem, junto à Polícia Civil, um registro de tentativa de homicídio toda vez que um policial se

18 Fonte: ISP. Números para o Estado do Rio de Janeiro, de janeiro de 2006 a julho de 2009.

envolvesse em um tiroteio. Com isso, o número de vítimas de tentativa de homicídio em áreas de UPP aumentou sobremaneira em 2014, como demonstrado no Gráfico 1, tornando muito difícil a identificação de quanto do aumento se devia ao novo procedimento e quanto dele refletiria um aumento do fenômeno. Esse exemplo ajuda a mostrar que todas as variações abruptas em um indicador devem ser analisadas com bastante cautela, sendo necessário investigar se houve mudanças na forma de registrar um incidente.

Gráfico 1 - Número de vítimas em área de UPP e número de UPP implantadas - Estado do Rio de Janeiro - 2007 a 2018



Fonte: ISP

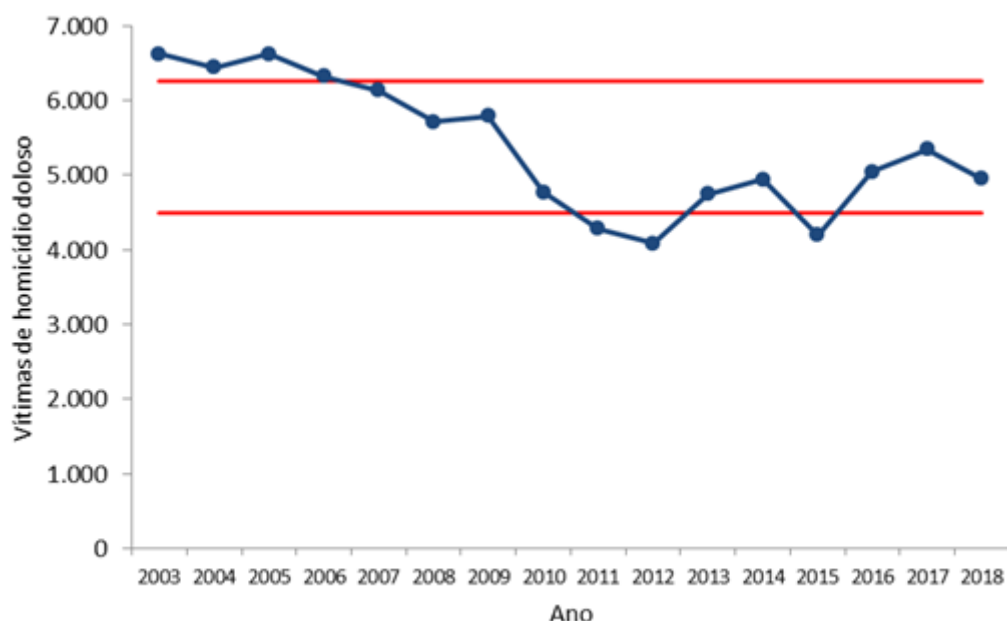
Armadilhas na interpretação dos dados

Tendo tomado o cuidado necessário para buscar os dados nas fontes oficiais e entender o que cada número está contando, é importante agora ficar atento às armadilhas inerentes à análise dos indicadores criminais.

O crime está aumentando ou diminuindo?

Quando queremos saber se a ocorrência de um crime está aumentando ou diminuindo, precisamos escolher um período de comparação. Em casos de crimes onde há grandes variações ao longo do tempo, a escolha do período base pode gerar conclusões totalmente distintas. Tome como exemplo o número de homicídios no Rio de Janeiro nos últimos 15 anos, conforme mostra o Gráfico 2. O número de vítimas em 2018 caiu em relação a 2017 e cresceu em relação a 2015. Por mais que isso pareça uma questão óbvia, é muito comum ver na mídia apenas menções a variações percentuais, que, dependendo do ponto da matéria, pode buscar uma imagem positiva ou negativa. Isso deixa o leitor sem uma real noção do patamar da criminalidade e indica a necessidade de se usar cada vez mais gráficos e menos tabelas comparativas.

Gráfico 2 - Vítimas de homicídio doloso - Estado do Rio de Janeiro - 2003 a 2018



Nota: Linhas horizontais em vermelho representam o valor da média mais um desvio-padrão (linha superior) e média menos um desvio-padrão (linha inferior). Fonte: ISP.

Um exemplo desse tipo de cobertura pode ser visto na Figura 3, cujo título da matéria apontava um aumento de 107% nos roubos de rua em comparação ao mesmo mês do ano anterior. Uma análise um pouco mais detalhada dos dados revelaria que, no primeiro trimestre de 2017, a Polícia Civil do Rio de Janeiro esteve em greve, gerando uma enorme subnotificação. Assim, o período base do ano anterior não era uma referência adequada para calcular a variação percentual.

Figura 3 - Matéria do jornal O Dia em 16 de fevereiro de 2018

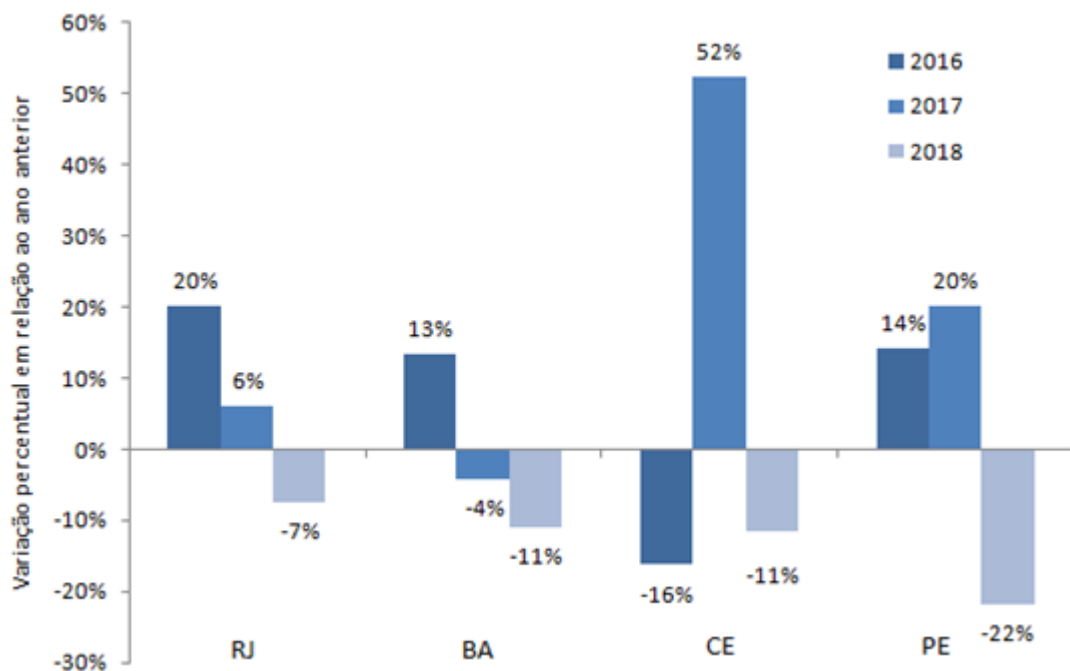


Fonte: : Jornal O Dia de 16 de fevereiro de 2018. Disponível em <https://odia.ig.com.br/rio-de-janeiro/2018/02/5514824-no-carnaval-roubos-de-rua-subiram-107-em-relacao-a-festa-de-2017.html>. Acessado em agosto de 2019.

Como identificar se uma variação é muito alta ou muito baixa?

Variações nos crimes são muito mais comuns do que a experiência de um cidadão que anda nas ruas parece revelar. Veja o exemplo do número de homicídios dolosos nos quatro estados brasileiros selecionados no Gráfico 3. Existem fortes variações entre os anos e com direção e magnitude bastante diferentes entre os estados. De 2015 para 2016, o Estado do Rio apresentou um aumento de 20% no número de vítimas, enquanto o Estado do Ceará teve redução de 16%. De 2016 para 2017, ambos os estados apresentaram crescimento – no Rio, de 6%; no Ceará, de expressivos 52%. No Brasil, o aumento nesse período foi de 9%. Já no ano seguinte, de 2017 para 2018, houve redução de 13% no número de vítimas no país, percentual próximo da queda de 11% observada na Bahia e no Ceará, mas bastante menor que os 22% do Estado de Pernambuco. Essas variações abrem espaço para um comportamento quase bipolar de todos os governos: quando os números caem, todos correm para se apropriar da queda, mas quando os números sobem, a justificativa é que violência é um fenômeno multicausal e o aumento não é responsabilidade do governo.

Gráfico 3 – Variação percentual do número de vítimas de homicídio doloso em relação ao ano anterior – estados selecionados – 2016 a 2018



Fonte: Sinesp. Acessado em abril de 2020.

Como saber se uma variação é especialmente alta ou baixa? Uma forma de responder a essa pergunta é entender qual a variação típica de uma série de dados, algo informado pelo desvio-padrão. Números que ficam acima da média histórica mais um desvio-padrão podem ser considerados altos, enquanto números abaixo da média menos um desvio-padrão sugerem que o patamar verificado está baixo para o padrão. O Gráfico 2 apresenta a série anual de homicídios no Rio de Janeiro. As linhas horizontais em vermelho apresentam os marcadores mencionados, indicando que os anos 2011, 2012 e 2015 registraram patamares bem baixos, enquanto todos os anos até 2007 tiveram valores bem elevados. Note que a média e o desvio-padrão variam de acordo com a série de dados analisada e, quanto maior o tamanho da série de dados, melhor.

Como identificar as causas das variações nos índices?

Um erro muito comum que vemos na mídia e no discurso político é associar as melhoras ou pioras de indicadores a fatores específicos, como às gestões de governo. Identificar causalidade, ou seja, identificar que a redução de um crime Y foi provocada por uma ação ou política específica X, é uma das tarefas mais difíceis de um analista de dados e exige muito mais que uma variação temporal. Isso ocorre porque inúmeros fatores afetam os indicadores criminais, desde a ostensividade policial, a capacidade investigativa da polícia, a liderança dos comandantes e chefes de unida-

des, as guerras entre facções, a introdução de novas leis, e as condições socioeconômicas. Para identificar a contribuição de um desses fatores em especial, é necessário fazer uma avaliação de impacto. A avaliação de impacto busca isolar o efeito de uma ação que se deseja analisar de todos os outros fatores que influenciam um indicador. Para isso, é preciso haver um contexto em que seja possível identificar regiões e grupos de pessoas que não são afetados pela ação ou política em análise, para utilizá-los como grupo de comparação, permitindo entender o que teria ocorrido na ausência da política (o chamado *contrafactual*). Existem muitos métodos para construir um grupo de comparação, e sua aplicação irá depender do desenho da política e da disponibilidade de dados. Esses estudos são feitos por acadêmicos e envolvem técnicas e análises mais sofisticadas do que as tipicamente feitas pela imprensa.

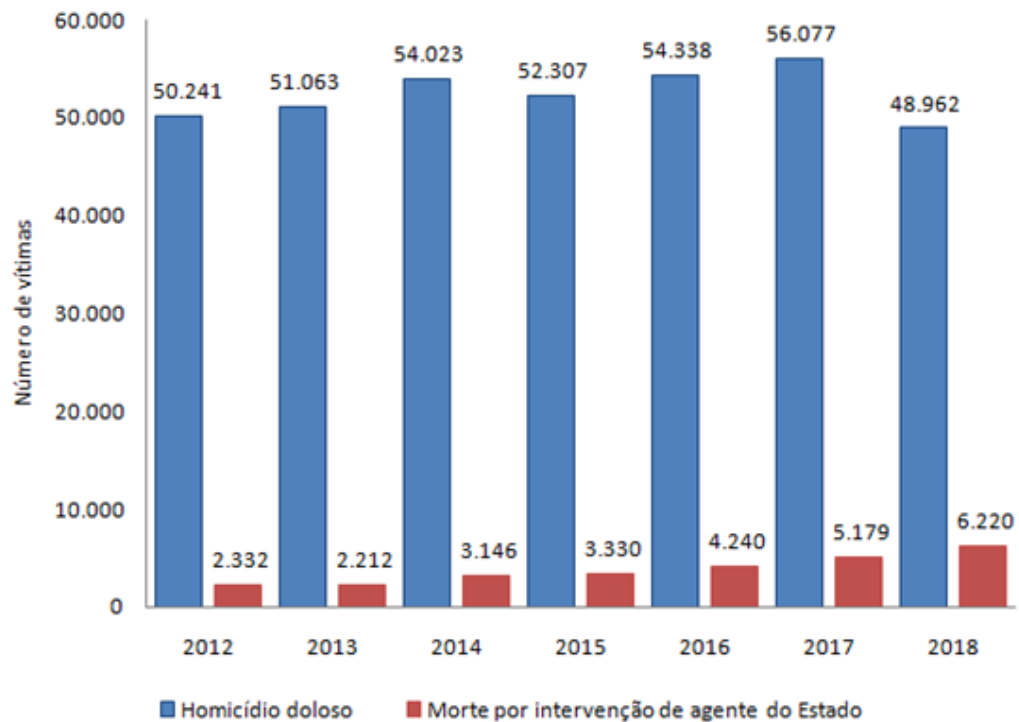
É importante frisar que, na ausência desses estudos, não é possível dizer que o aumento de um indicador acompanhado da queda de outro indica uma relação de causa-efeito. Por outro lado, há análises que podem ser feitas para levantar hipóteses sobre as possíveis causas e analisar se os dados indicam que elas são plausíveis ou se não há qualquer fundamento.

A primeira análise necessária para identificar se uma ação do governo provocou uma melhora nos indicadores é perguntar o que o governo fez. O governo precisa dizer o que está fazendo diferente ou melhor para poder argumentar que a variação no indicador se deve à sua gestão. Segundo, é preciso identificar se os números indicam uma mudança de trajetória posterior ao início da ação. Mudar trajetória significa inverter uma tendência de crescimento ou acelerar uma queda dos indicadores. Terceiro, é necessário observar uma melhora contínua entre os meses, e não picos ou vales isolados na série de dados. Quarto, os locais onde a ação em questão tem sido mais executada têm que apresentar reduções maiores que os locais onde a ação não ocorreu. No jargão técnico, a ação do governo tem que estar correlacionada com indicadores criminais no nível local. Note que essa análise busca comparar variações de crimes, e não níveis de criminalidade, visto que ela busca analisar mudanças decorrentes da introdução de uma ação específica.

Tomemos como exemplo um tema bastante debatido em 2019: o que explica a queda de homicídios no Brasil verificada em 2018. Uma das hipóteses levantadas e que há bastante destaque na mídia é o aumento do uso da força policial, que poderia ser medido pelo aumento das mortes por agente do Estado. A primeira análise a ser feita é ver se as duas séries caminham juntas. O Gráfico 4 indica que, enquanto as mortes por agente do Estado vêm crescendo continuamente ao longo dos anos no Brasil, o número de homicídios tem variado entre anos, com quedas em 2015 e, notadamente,

em 2018.

Gráfico 4 – Vítimas de homicídio doloso e de morte por intervenção de agente do Estado – Brasil – 2012 a 2018



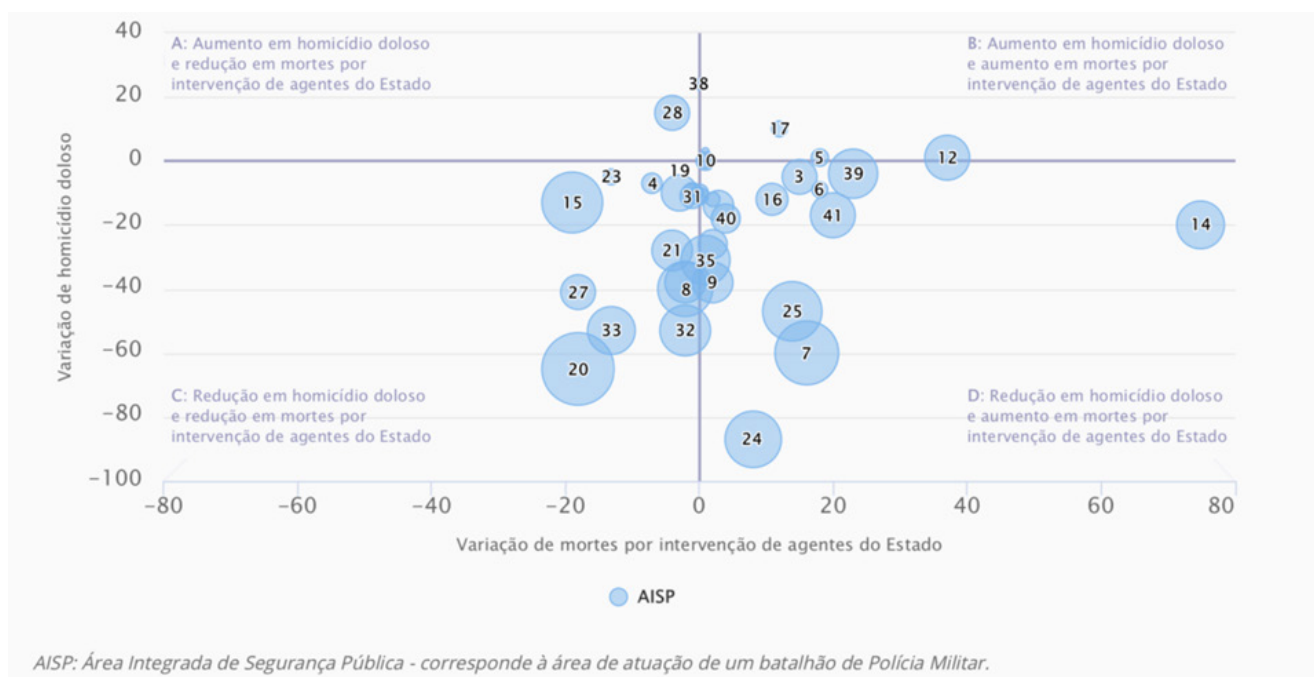
Fonte: Fórum Brasileiro de Segurança Pública.

Entretanto, essa não é a melhor forma de estudar a questão porque análises deste tipo devem ser feitas com o menor nível de agregação possível e com períodos de tempo longos. Além disso, o fato de não haver padrão para o Brasil não implica que não possa haver um padrão para algum estado.

Para exemplificar o tipo de análise relativamente simples que pode ser feita, o Gráfico 5 apresenta a relação entre a variação no número de homicídios dolosos no Rio de Janeiro entre 2018 e 2019 e a variação no número de mortes por intervenção de agente do Estado no mesmo período. Cada bola no gráfico representa uma área integrada de segurança pública (Aisp), que corresponde à área de um batalhão de Polícia Militar. O tamanho da bola indica o número de homicídios dolosos em cada área. O fato de haver muitas bolas tanto à direita quanto à esquerda do eixo vertical indica que há muitos casos em que houve queda de homicídios e aumento de mortes por intervenção de agente do Estado, mas também muitas áreas em que os dois indicadores caíram. O fato de as bolas estarem horizontalmente distribuídas sugere que ambas as variáveis não são correlacionadas. Entretanto, a ausência de correlação só pode ser

atestada com exercícios estatísticos com períodos mais longos e que isolem os efeitos sazonais e locais. Monteiro et al. (2020) fazem esse exercício e apontam que, historicamente, não há evidência de correlação entre as duas variáveis no Rio de Janeiro.

Gráfico 5 – Gráfico de dispersão da variação de homicídios dolosos e mortes por intervenção de agente do Estado por Aisp – 2019/2018 (janeiro a agosto)



Nota: O tamanho da bola indica o número absoluto de homicídios dolosos em cada Aisp em 2019. Fonte: Cenpe (2019).

Outro exercício possível para levantar hipóteses das causas associadas à queda de homicídios é levantar as causas das mortes. Existem poucos estudos informativos sobre circunstância ou motivação da morte. Uma pesquisa do Conselho Nacional do Ministério Público (CNMP), de 2012, só conseguiu obter informações sobre o número de homicídios que possuíam causa provável identificada para 15 estados, além do Distrito Federal, como pode ser visto na página seguinte, Tabela 3.

Tabela 3 – Homicídios com identificação de causa provável por unidade da Federação

UF ou região	Homicídios com identificação de causa		Homicídios sem identificação de causa		Período
	Nº absoluto	Valor percentual	Nº absoluto	Valor percentual	
AC	127	100,0%	0	0,0%	jan/2011 a dez/2012
AL - Maceió	108	50,2%	107	49,8%	jan/2012 a out/2012
AP	81	43,8%	104	56,2%	jan/2011 a out/2012
BA - Salvador	154	12,9%	1.044	87,1%	jan/2011 a out/2012
DF	297	40,3%	440	59,7%	jan-jul/2011 e jan-jul/2012
ES - Grande Vitória	1.188	98,4%	19	1,6%	jan/2011 a dez/2012
GO	n.d.	100,0%	n.d.	0,0%	jan/2011 a set/2012
MS - Campo Grande	91	91,0%	9	9,0%	jan/2011 a out/2012
MT - Cuiabá e Várzea Grande	383	73,9%	135	26,1%	jan/2011 a jul/2012
PA	1.785	83,6%	349	16,4%	jan-dez/2012
PE	4.541	64,7%	2.475	35,3%	jan/2010 a dez/2011
PR	n.d.	90,0%	n.d.	10,0%	n.d.
RJ	1.136	14,6%	6.663	85,4%	jan/2011 a set/2012
RS	779	100,0%	0	0,0%	jan-dez/2011
SC	650	48,5%	691	51,5%	jan/2011 a dez/2012
SP	n.d.	81,9%	n.d.	18,1%	jan/2011 a dez/2012

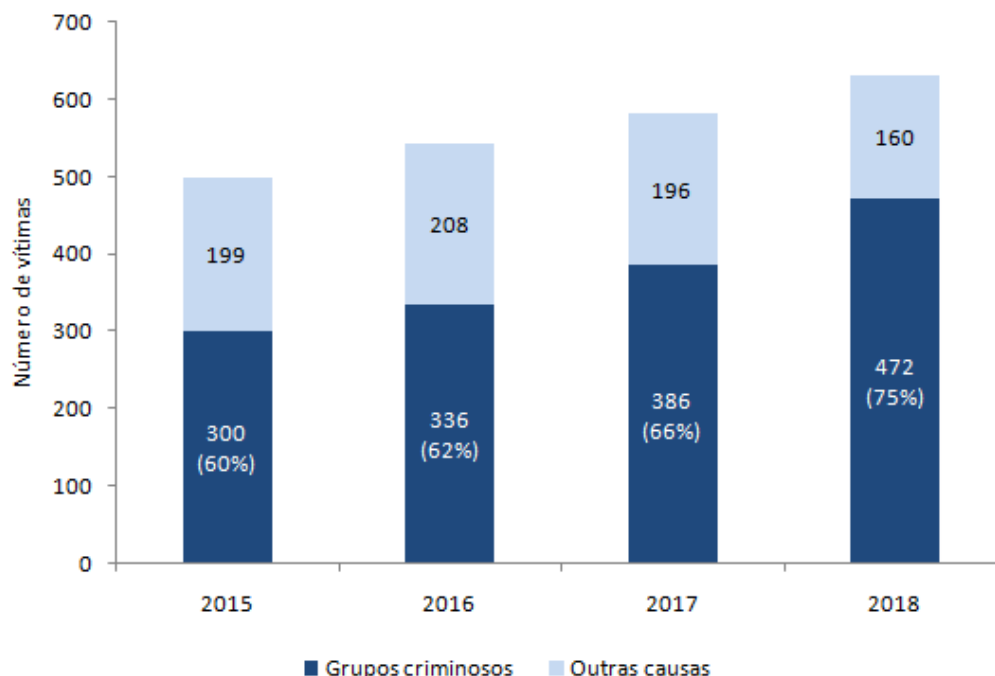
Fonte: CNMP (2012).

A dificuldade em obter tal tipo de informação ocorre porque, tipicamente, as estatísticas são baseadas no primeiro relato sobre o crime no momento do registro de ocorrência. Mesmo que durante o curso das investigações sejam identificadas as circunstâncias ou motivações das mortes, raramente essa informação será aditada nos registros de ocorrência, seja pela falta de tradição em armazenar dados, ou seja pelo fato de que muitas vezes a circunstância é identificada, mas não provada, o que gera receio em registrá-la em uma peça do inquérito. A forma de lidar com esse desafio é instituir uma rotina de análise criminal, onde os casos de mortes violentas são analisados e qualificados com o objetivo de compreender melhor o fenômeno. Esse tipo de análise é feita em Medellín, na Colômbia, pela Polícia Nacional em parceria com o governo municipal, e permite ter as informações como no Gráfico 6. O Gráfico mostra que o crescimento recente dos homicídios em Medellín é puxado pelas mortes de membros de grupos criminosos, o que sugere um aumento de disputa entre eles.

Aqui no Brasil, a Polícia Civil da Paraíba tem feito um trabalho importante nessa linha de qualificar as circunstâncias e motivações das mortes. Algo muito importante nesse tipo de análise é qualificar todos os casos de homicídio. Um erro bastante co-

mum é analistas afirmarem o percentual de homicídios atribuídos a grupos criminosos baseado em um subconjunto de casos analisados.

Gráfico 6 – Homicídios por causa – Colômbia – 2015 a 2018



Fonte: SISC (2018)

Qual o município mais violento do Brasil?

Quando falamos de políticas públicas nacionais, além de saber identificar os problemas que queremos combater e saber priorizá-los, devemos conhecer onde os mesmos se apresentam com maior gravidade. Para poder comparar localidades de tamanhos tão distintos, é necessário padronizar os indicadores para uma mesma unidade. Sabemos que, estatisticamente falando, a ocorrência de 10 homicídios no município de São Gabriel da Cachoeira (AM) é muito mais impactante do que 10 homicídios na capital de São Paulo – em que pese o óbvio argumento de que nenhuma vida vale mais ou menos por conta do seu endereço. Assim, o uso de taxas se faz necessário porque, caso contrário, São Paulo iria liderar qualquer ranking, simplesmente devido à grande concentração de pessoas naquele município, logo, a maioria dos eventos que envolvem pessoas ocorre ali.

Uma maneira simples e consensual de comparar áreas (países, estados, municípios, bairros etc.) é padronizando seus números pelo tamanho da sua respectiva população, calculando, assim, a taxa por 100 mil habitantes.

Contudo, a armadilha aqui não é fazer o ranking com números absolutos ou taxas, mas saber quais são as limitações do uso desse tipo de indicador. Populações muito pequenas podem gerar valores distorcidos – ainda que não estejam incorretos –, induzindo a conclusões equivocadas, com inversão de prioridades, pela falsa relevância dos resultados.

A Senasp restringe o uso de taxas a municípios com população igual ou maior que 100 mil habitantes, ainda que seja comum observar o uso de taxas para áreas com pelo menos 50 mil. Uma alternativa é o uso de taxas bayesianas das microrregiões estabelecidas pelo IBGE¹⁹, em vez de usar as taxas brutas por 100 mil habitantes para cada município. A simples escolha de qual a população mínima para participar de um ranking de taxas de homicídios implica em listas muito distintas sobre as 10 cidades mais violentas, que não é resultado de erros de cálculos, mas sim de diferenças de amostras.

Outra armadilha está no cálculo das taxas usando a população residente do local. Assim, quando comparamos zonas de uma cidade, é normal que o centro comercial da cidade apresente as maiores taxas, visto que essas áreas têm elevada circulação de pessoas (população flutuante), mas uma pequena população residente. Não está errado que as regiões comerciais apareçam em primeiro lugar, mas isso não significa que esta localidade seja a mais violenta. O mesmo problema se aplica a cidades e regiões turísticas que apresentam índices altos devido a uma população residente baixa. O ideal é que a taxa seja calculada pela população flutuante, isto é, a que transita em cada área, independentemente de onde reside. Infelizmente, ainda não há estatísticas oficiais de população flutuante disponíveis para o Brasil.

Por fim, cabe frisar que o objetivo de calcular taxas é oferecer uma comparabilidade entre grupos populacionais que têm diferentes exposições à vitimização ou entre locais que oferecem diferentes oportunidades para o crime. Assim, é possível calcular taxas pelo tamanho da frota (por 100 mil veículos) para fazer comparações entre roubos e furtos de veículos; pelo efetivo policial (por 100 mil policiais) para calcular vitimização policial; ou por subgrupos populacionais divididos por raça, faixa etária ou gênero.

19 As microrregiões são conjuntos de municípios limítrofes com características econômicas e sociais similares.

Impacto relativo

Analogamente ao uso de taxas, também é recomendado que usemos variações percentuais, e não variações absolutas, quando queremos comparar a evolução de um indicador em diferentes regiões. Ao usar variações percentuais, não estamos olhando para a região que mais contribuiu para a redução ou aumento da violência, mas conseguimos entender qual área tem apresentado resultados mais impactantes em relação ao histórico dela mesma.

Como toda regra, a aplicação de variação percentual para medir o agravamento ou redução de um crime em determinada região também tem suas armadilhas. Neste caso, a limitação se refere ao uso de análises percentuais quando falamos de crimes ou locais com baixa casuística. Faz sentido se queremos comparar a evolução de indicadores com centenas ou milhares de ocorrências. No entanto, se os números são baixos, os valores percentuais podem induzir a uma interpretação exagerada da evolução criminal.

Na página seguinte, Tabela 4, estão destacadas em vermelho as maiores variações de 2017 para 2018 nos números de latrocínios em cada Unidade da Federação. Os casos mais graves estariam no Estado de Roraima e no Distrito Federal, com crescimento de 100% e 51,4%, respectivamente. No entanto, estas variações significam sete e 19 vítimas a mais, na mesma ordem. Bastante menos grave que os aumentos no Paraná e no Rio Grande do Norte (mais 27 vítimas em cada estado), que tiveram aumento de 40,3% e 33,3%, respectivamente.

Tabela 4 – Vítimas de latrocínio por unidade da Federação – Brasil - 2017 e 2018

UF	2017	2018	Variação 2018 x 2017
AC	15	20	33,3%
AL	59	39	-36,2%
AM	77	56	-27,3%
AP	22	22	0,0%
BA	176	136	-22,7%
CE	88	53	-39,8%
DF	37	56	51,4%
ES	45	34	-24,4%
GO	122	96	-21,3%
MA	98	71	-27,6%
MG	105	82	-21,9%
MS	25	30	-9,1%
MT	50	37	-26,0%
PA	228	194	-14,9%
PB	38	37	0,0%
PE	250	131	-47,6%
PI	43	46	7,0%
PR	67	94	40,3%
RJ	237	174	-26,6%
RN	81	108	33,3%
RO	21	18	-14,3%
RR	7	14	100,0%
RS	156	110	-29,5%
SC	63	42	-33,3%
SE	59	32	-45,8%
SP	340	278	-18,2%
TO	12	17	41,7%

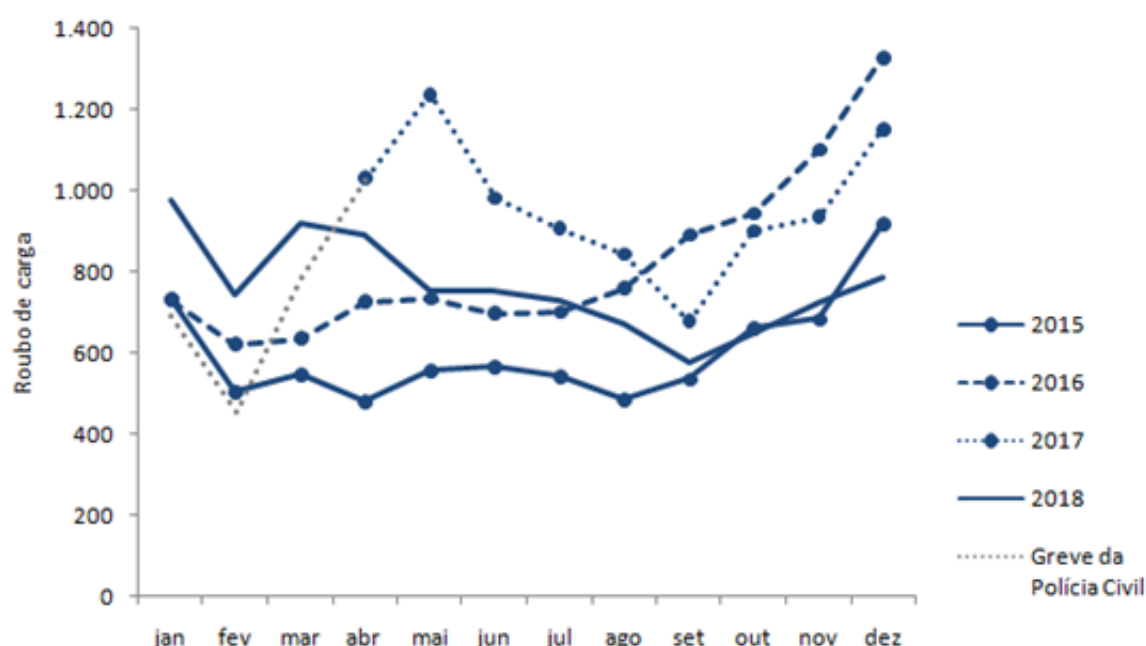
Fonte: Sinesp. Acessado em abril de 2020

As voltas que o crime dá

Sazonalidade é um conceito usado na estatística muito associado a atividades econômicas, mas com forte presença também nos indicadores criminais. Trata-se da repetição cíclica dos fatos, podendo ocorrer com diferentes periodicidades: ao longo do ano, do mês, da semana ou mesmo ao longo do dia.

O caso mais emblemático para o estudo da sazonalidade dos crimes são as ocorrências de roubo de carga. Ao longo do ano, as tendências são bastante previsíveis. O crime aumenta quando aumenta a oportunidade, no caso, o número de caminhões transportando cargas. Assim, todo ano, é possível observar um aumento do indicador nos últimos meses, sobretudo em dezembro, devido às compras de Natal, e, ainda que em volume menor, também no mês de maio, por conta do Dia das Mães. Logo, os meses subsequentes, janeiro e junho, respectivamente, costumam apresentar queda em relação ao mês anterior. O Gráfico 7 ilustra esses movimentos para os casos do Estado do Rio de Janeiro.

Gráfico 7 – Número de roubo de carga por mês – Estado do Rio de Janeiro – 2015 a 2018



Fonte: ISP

O roubo de carga também apresenta sazonalidade ao longo da semana. Como o transporte de cargas ocorre geralmente nos dias úteis, os finais de semana e os feriados apresentam números significativamente mais baixos. Isso pode gerar interpretações errôneas, principalmente quando algum fato político leva ao acompanhamento

de um indicador em unidades de tempo pequenas, como o exemplo apresentado na Figura 4. Na segunda-feira, 31 de julho de 2017, o jornal O Globo publicou uma matéria de capa anunciando a queda de 30% nos roubos de carga no Estado do Rio de Janeiro após a operação das tropas federais – iniciada em 28 de julho, sexta-feira. Os números estavam corretos, mas faltou informar que o “resultado relâmpago” foi medido pelo número de ocorrências em apenas dois dias, sábado e domingo, quando há pouca circulação de carga. Ao longo da semana seguinte, os números voltaram ao mesmo patamar da semana anterior.

Figura 4 – Capa do jornal O Globo em 31 de julho de 2017



Fonte: Jornal O Globo, 31 de julho de 2017. Disponível em <https://oglobo.globo.com/rio/reforco-na-seguranca-reduz-roubo-de-carga-no-rio-na-sexta-sabado-21649820>. Acessado em agosto de 2019.

Por fim, ainda é possível observar uma forte sazonalidade dessas ocorrências em horários específicos do dia, mais concentradas na parte da manhã (entre 8h e 13h). Os números do Estado do Rio de Janeiro podem ser observados na Tabela a seguir. As colunas indicam os dias da semana e as linhas indicam as horas do dia em que ocorreram os casos de roubo de carga no ano de 2018. Quanto mais vermelho, maior a concentração de casos no período indicado.

Tabela 5 – Roubo de carga e roubo a transeunte por dia da semana e hora do dia – Estado do Rio de Janeiro – 2018

(a) Roubo de carga								(b) Roubo a transeunte							
	Seg	Ter	Qua	Qui	Sex	Sáb	Dom		Seg	Ter	Qua	Qui	Sex	Sáb	Dom
0h	0,09%	0,13%	0,14%	0,14%	0,13%	0,05%	0,02%	0h							
1h	0,14%	0,14%	0,15%	0,13%	0,09%	0,05%	0,02%	1h	0,34%	0,18%	0,26%	0,22%	0,32%	0,40%	0,42%
2h	0,08%	0,09%	0,13%	0,10%	0,13%	0,09%	0,00%	2h	0,28%	0,16%	0,17%	0,15%	0,23%	0,29%	0,32%
3h	0,10%	0,22%	0,21%	0,19%	0,23%	0,09%	0,01%	3h	0,27%	0,15%	0,14%	0,15%	0,23%	0,28%	0,31%
4h	0,23%	0,19%	0,37%	0,19%	0,24%	0,07%	0,00%	4h	0,69%	0,55%	0,64%	0,60%	0,69%	0,54%	0,48%
5h	0,39%	0,60%	0,88%	0,77%	0,57%	0,19%	0,00%	5h	1,48%	1,11%	1,12%	1,24%	1,32%	0,95%	0,76%
6h	0,46%	0,83%	0,92%	0,84%	0,77%	0,37%	0,04%	6h	1,21%	1,04%	0,98%	1,05%	1,16%	0,93%	0,73%
7h	0,77%	1,12%	1,23%	1,24%	1,13%	0,39%	0,05%	7h	0,79%	0,65%	0,66%	0,65%	0,60%	0,50%	0,45%
8h	0,73%	1,91%	2,05%	1,73%	1,62%	0,84%	0,03%	8h	0,57%	0,51%	0,55%	0,52%	0,48%	0,41%	0,35%
9h	1,26%	1,86%	2,40%	2,43%	2,12%	0,69%	0,07%	9h	0,45%	0,43%	0,42%	0,43%	0,37%	0,29%	0,27%
10h	1,33%	2,49%	3,26%	2,87%	2,59%	0,88%	0,03%	10h	0,53%	0,47%	0,45%	0,45%	0,38%	0,31%	0,29%
11h	1,39%	1,99%	2,47%	2,52%	2,29%	0,82%	0,03%	11h	0,43%	0,40%	0,40%	0,42%	0,40%	0,29%	0,29%
12h	1,23%	1,84%	1,99%	1,88%	2,03%	0,68%	0,03%	12h	0,42%	0,40%	0,48%	0,46%	0,41%	0,29%	0,26%
13h	1,14%	1,38%	1,21%	1,56%	1,37%	0,46%	0,03%	13h	0,43%	0,46%	0,45%	0,41%	0,45%	0,32%	0,32%
14h	0,81%	1,22%	1,27%	1,26%	0,98%	0,37%	0,07%	14h	0,46%	0,48%	0,49%	0,46%	0,39%	0,37%	0,32%
15h	0,68%	0,88%	0,80%	0,99%	0,87%	0,22%	0,05%	15h	0,49%	0,44%	0,45%	0,42%	0,46%	0,37%	0,42%
16h	0,49%	0,60%	0,62%	0,64%	0,69%	0,16%	0,07%	16h	0,48%	0,45%	0,42%	0,42%	0,42%	0,42%	0,44%
17h	0,25%	0,28%	0,37%	0,50%	0,46%	0,05%	0,03%	17h	0,59%	0,48%	0,51%	0,53%	0,52%	0,39%	0,51%
18h	0,17%	0,33%	0,32%	0,29%	0,25%	0,08%	0,07%	18h	0,83%	0,83%	0,77%	0,78%	0,75%	0,53%	0,70%
19h	0,16%	0,15%	0,22%	0,20%	0,17%	0,04%	0,08%	19h	1,18%	1,11%	1,10%	1,01%	0,96%	0,71%	0,89%
20h	0,13%	0,16%	0,13%	0,21%	0,10%	0,00%	0,08%	20h	1,32%	1,25%	1,32%	1,17%	1,13%	0,90%	1,09%
21h	0,06%	0,07%	0,10%	0,10%	0,07%	0,04%	0,10%	21h	1,20%	1,27%	1,20%	1,20%	1,06%	0,94%	1,17%
22h	0,08%	0,09%	0,10%	0,10%	0,04%	0,02%	0,08%	22h	1,12%	1,15%	1,13%	1,04%	0,99%	0,86%	1,10%
23h	0,08%	0,10%	0,14%	0,09%	0,04%	0,02%	0,00%	23h	0,78%	0,77%	0,79%	0,81%	0,86%	0,79%	0,86%

Fonte: ISP

A sazonalidade na verdade é bastante evidente nos crimes contra o patrimônio. Para além do roubo de carga, outro exemplo seriam os roubos a transeunte, que apresentam maior incidência durante a semana nas áreas centrais da cidade ou aos finais de semana nas áreas de lazer – pois são nestas datas que estas regiões são mais frequentadas. Os crimes de rua, incluindo aí também os furtos, sempre irão apresentar mais casos onde e quando houver mais gente, sendo conhecidos também como “crimes de oportunidade”. Repare, na Tabela anterior, que os horários de pico desse crime são os horários de traslado das pessoas entre casa e trabalho (na parte da manhã, entre 5h e 7h, e na parte da noite, entre 18h e 0h), já que são os horários em que há maior número de pessoas nas ruas.

Os roubos de veículo, por sua vez, não apresentam sazonalidade ao longo da semana, mas possuem uma concentração bem forte de ocorrências no horário da noite, entre 19h e 0h, conforme Tabela 6.

Tabela 6 – Roubo de veículo e furto de veículo por dia da semana e hora do dia – Estado do Rio de Janeiro – 2018

(a) Roubo de veículo								(b) Furto de veículo							
	Seg	Ter	Qua	Qui	Sex	Sáb	Dom		Seg	Ter	Qua	Qui	Sex	Sáb	Dom
0h	0,52%	0,41%	0,50%	0,55%	0,55%	0,65%	0,61%	0h	0,77%	0,89%	0,72%	0,78%	0,86%	0,78%	0,74%
1h	0,40%	0,22%	0,33%	0,35%	0,35%	0,49%	0,52%	1h	0,37%	0,33%	0,37%	0,34%	0,44%	0,39%	0,36%
2h	0,28%	0,18%	0,22%	0,20%	0,28%	0,40%	0,35%	2h	0,25%	0,25%	0,31%	0,32%	0,23%	0,33%	0,39%
3h	0,17%	0,13%	0,12%	0,21%	0,24%	0,31%	0,24%	3h	0,20%	0,30%	0,25%	0,25%	0,23%	0,27%	0,33%
4h	0,23%	0,22%	0,25%	0,25%	0,35%	0,31%	0,24%	4h	0,34%	0,30%	0,26%	0,30%	0,25%	0,27%	0,22%
5h	0,51%	0,61%	0,61%	0,63%	0,77%	0,51%	0,42%	5h	0,39%	0,48%	0,42%	0,44%	0,39%	0,35%	0,23%
6h	0,52%	0,55%	0,51%	0,60%	0,70%	0,62%	0,59%	6h	0,55%	0,77%	0,73%	0,62%	0,72%	0,33%	0,36%
7h	0,28%	0,34%	0,46%	0,48%	0,43%	0,46%	0,55%	7h	0,63%	0,98%	0,90%	0,93%	0,78%	0,46%	0,41%
8h	0,23%	0,27%	0,33%	0,26%	0,34%	0,33%	0,42%	8h	0,93%	0,92%	1,03%	0,87%	0,97%	0,63%	0,56%
9h	0,20%	0,31%	0,29%	0,26%	0,28%	0,25%	0,31%	9h	0,72%	0,84%	0,77%	0,69%	0,75%	0,65%	0,60%
10h	0,21%	0,30%	0,32%	0,33%	0,30%	0,29%	0,36%	10h	0,71%	0,81%	0,81%	0,75%	0,80%	0,73%	0,87%
11h	0,22%	0,28%	0,33%	0,33%	0,36%	0,30%	0,36%	11h	0,53%	0,51%	0,73%	0,70%	0,60%	0,65%	0,62%
12h	0,28%	0,38%	0,31%	0,33%	0,39%	0,34%	0,41%	12h	0,54%	0,65%	0,58%	0,79%	0,65%	0,72%	0,72%
13h	0,27%	0,41%	0,41%	0,47%	0,42%	0,44%	0,46%	13h	0,51%	0,58%	0,62%	0,56%	0,51%	0,45%	0,58%
14h	0,28%	0,37%	0,42%	0,49%	0,50%	0,46%	0,55%	14h	0,51%	0,60%	0,58%	0,51%	0,65%	0,52%	0,49%
15h	0,33%	0,38%	0,46%	0,49%	0,54%	0,49%	0,68%	15h	0,49%	0,46%	0,49%	0,43%	0,53%	0,45%	0,63%
16h	0,42%	0,46%	0,50%	0,46%	0,58%	0,58%	0,58%	16h	0,42%	0,44%	0,36%	0,48%	0,52%	0,51%	0,46%
17h	0,47%	0,56%	0,57%	0,61%	0,61%	0,67%	0,73%	17h	0,52%	0,51%	0,48%	0,58%	0,49%	0,39%	0,66%
18h	0,71%	0,82%	0,83%	0,84%	0,92%	0,73%	0,84%	18h	0,72%	0,82%	0,84%	0,84%	0,74%	0,53%	0,74%
19h	0,98%	1,27%	1,38%	1,38%	1,32%	1,13%	1,01%	19h	0,78%	0,77%	0,83%	0,79%	0,68%	0,71%	0,84%
20h	1,15%	1,50%	1,51%	1,56%	1,47%	1,35%	1,20%	20h	0,82%	0,87%	0,78%	0,86%	0,71%	0,80%	0,74%
21h	1,23%	1,49%	1,72%	1,69%	1,52%	1,40%	1,21%	21h	0,67%	0,51%	0,63%	0,60%	0,68%	0,63%	0,64%
22h	0,93%	1,28%	1,47%	1,38%	1,38%	1,23%	1,10%	22h	0,81%	0,93%	0,78%	0,92%	0,85%	0,74%	0,81%
23h	0,70%	0,89%	0,94%	0,98%	1,04%	1,01%	0,94%	23h	0,67%	0,48%	0,60%	0,61%	0,68%	0,67%	0,65%

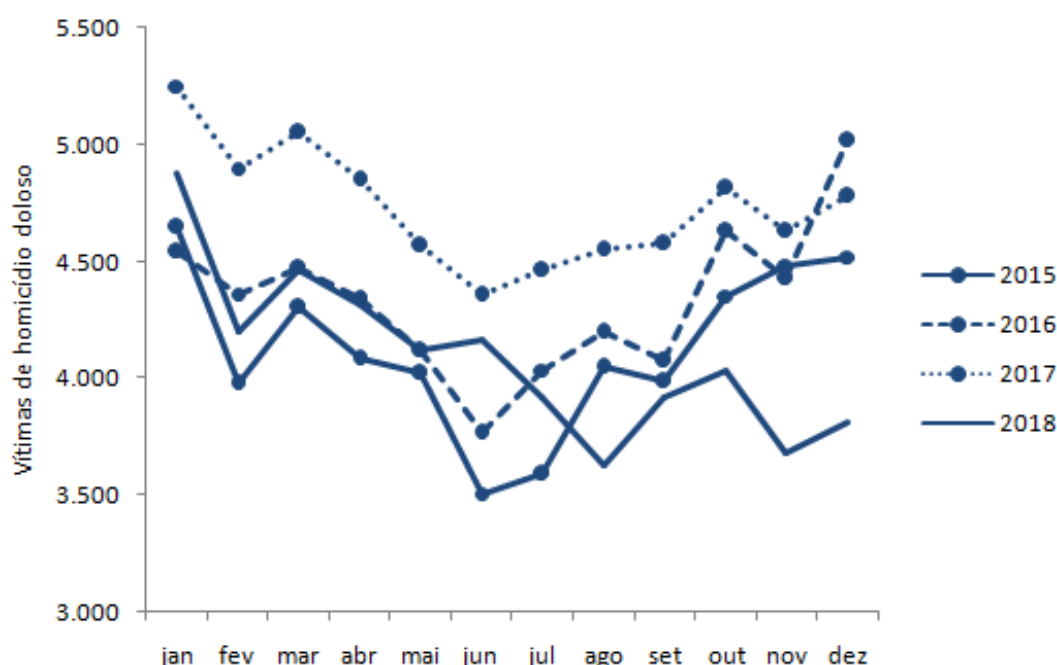
Fonte: ISP

Já os furtos de veículo, pela natureza desse tipo de crime, que ocorre normalmente na ausência da vítima, em tese não possuem um padrão de horário tão bem definido. A verdade é que não podemos dizer que os casos ocorrem com distribuição mais uniforme ao longo do dia, como pressupõe a análise do Quadro acima, porque a própria vítima muitas vezes não consegue dizer o horário exato da ocorrência. Suponha um motorista que tenha estacionado seu carro na rua às 19h, e somente voltou ao local no dia seguinte, às 08h. Como saber em que horário seu carro foi furtado?

A mesma previsibilidade de concentração por dia da semana ou horário já não é encontrada entre os crimes contra a pessoa, que, em geral, apresentam sazonalidade mais marcada ao longo dos meses do ano. Casos de homicídio e de lesão corporal ocorrem mais nos meses de verão, e são menos frequentes nos meses de junho e julho. Assim, é esperado que esses números apresentem tendência de queda no decorrer do primeiro semestre do ano – e tendência contrária no segundo semestre. Esse compor-

tamento fica evidente no Gráfico 8, com a exceção do segundo semestre de 2018, que teve comportamento atípico.

Gráfico 8 – Vítimas de homicídio doloso por mês - Brasil – 2015 a 2018



Fonte: Sinesp. Acessado em abril de 2020.

Portanto, deve-se ter cautela para não cair na armadilha de interpretar erroneamente variações nos números como se fossem resultado de alguma política ou programa de governo, enquanto na verdade são oscilações intrínsecas à ocorrência desses eventos.

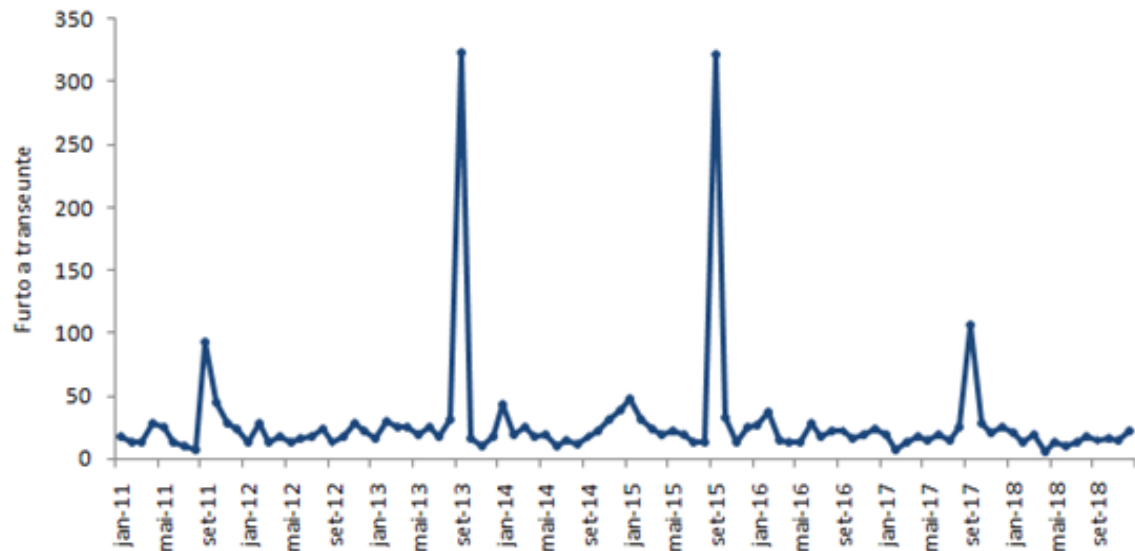
A influência de grandes eventos

Por outro lado, mudanças bruscas nos indicadores, que fogem ao padrão da sazonalidade, também devem ser interpretadas com cuidado, pois muitas vezes há circunstâncias específicas em determinados locais que justificam tais variações, não sendo prudente alardeá-los como aumento ou redução da criminalidade.

No Gráfico 9, temos a série mensal de furto a transeunte no bairro do Recreio dos Bandeirantes, no Rio de Janeiro, em que observamos picos totalmente fora do padrão em alguns meses específicos – fora do padrão, sim, inesperados não. Os meses em questão são os meses de setembro dos anos 2011, 2013, 2015 e 2017 – datas do evento

internacional Rock in Rio, que multiplica a população naquela região durante as quase duas semanas de evento.

Gráfico 9 – Furto a transeunte – Bairro Recreio dos Bandeirantes (RJ) – 2011 a 2018



Fonte: ISP

Diversos outros eventos impactam os números de crime em regiões específicas. Além dos grandes eventos, podemos citar o carnaval, réveillon e manifestações que aumentam o número de pessoas em um lugar por um período específico. Eventos pontuais também podem interferir, como partidas de futebol, greves, início das aulas escolares etc.

Considerações finais

Neste capítulo, apresentamos as fontes de dados de Segurança Pública do país, levando em conta as diversas dimensões com que o tema pode ser abordado (crime, polícia, violência, saúde, justiça etc.), e considerando, contudo, a oferta extremamente restrita e problemática de dados a respeito. Foram expostas, ainda, as principais armadilhas para obter e interpretar os dados existentes, com o intuito de contribuir para a qualificação do debate e reforçar a importância de que as análises sejam feitas dentro do contexto correto.

Os exemplos mostrados apontam três armadilhas fundamentais na análise de dados de crime e violência. Primeiro, devemos ter a preocupação de comparar regiões ao longo do tempo usando indicadores que meçam a mesma coisa, para evitarmos somar maçãs com laranjas. Segundo, há fortes sazonalidades nas ocorrências criminais devido à presença de fatores de risco ou de oportunidades para o crime. Isso implica que os dados têm que ser interpretados como altos ou baixos com referência ao padrão comum para o período em análise. Ao mesmo tempo, isso revela quão fundamental é a análise criminal para haver uma ação efetiva de combate ao crime. Um dos preceitos mais defendidos pela literatura internacional para combater o crime é muito simples: focar esforços nos locais e horários onde o crime é mais frequente. Terceiro, é necessário ter muito cuidado em atribuir variações nos crimes a mudanças de governo ou introdução de políticas. Como discutido em diversos capítulos deste livro, o simples fato de que uma política se inicia em um mês em que há redução de indicadores não permite atribuir a queda à política. No caso das análises dos indicadores de criminalidade e violência, o problema é ainda mais grave, visto que o debate público hoje é recheado de afirmações sobre relações de variáveis que não são nem correlacionadas.

O escasso compromisso de autoridades estaduais e federais para produção e disponibilização de estatísticas mostra como é limitado o interesse dos governos em ter uma abordagem científica para o tema. O maior problema de ter um sistema de informação limitado é que temos que despender grandes esforços só para obter o mínimo de informação sobre o tema, ficando prejudicados trabalhos mais dedicados para avaliar políticas e construir evidências sobre o que pode ajudar na promoção da Segurança Pública do Brasil. Não há estatísticas oficiais consolidadas, acessíveis e transparentes de quantos inquéritos policiais são relatados ao Ministério Público, quantos destes viram denúncias nos tribunais de justiça e quantos, finalmente, são julgados. Ou seja, temos um longo caminho ainda a percorrer para entendermos toda a trajetória do

crime no país e como atua cada uma das instituições envolvidas. Ter dados confiáveis e comparáveis são o primeiro passo para a identificação de prioridades, a qualificação do problema e a avaliação de políticas públicas. Sem isso, não conseguiremos aumentar a eficiência no uso dos recursos públicos e reduzir a violência no país.

Referências bibliográficas

- BRASIL. Lei nº 12.015, de 07 de agosto de 2009. Altera o Título VI da Parte Especial do Decreto-Lei no 2.848, de 7 de dezembro de 1940 - Código Penal, e o art. 1º da Lei nº 8.072, de 25 de julho de 1990, que dispõe sobre os crimes hediondos, nos termos do inciso XLIII do art. 5º da Constituição Federal e revoga a Lei nº 2.252, de 1º de julho de 1954, que trata de corrupção de menores. Diário Oficial da União. Brasília, DF, 07 de agosto de 2009.
- BRASIL. Lei nº 12.527, de 18 de novembro de 2011. Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei nº 8.112, de 11 de dezembro de 1990; revoga a Lei nº 11.111, de 5 de maio de 2005, e dispositivos da Lei nº 8.159, de 8 de janeiro de 1991; e dá outras providências. Brasília, DF, 18 de novembro de 2011.
- BRASIL. Lei nº 12.681, de 04 de julho de 2012. Institui o Sistema Nacional de Informações de Segurança Pública, Prisionais e sobre Drogas - Sinesp; altera as Leis nºs 10.201, de 14 de fevereiro de 2001, e 11.530, de 24 de outubro de 2007, a Lei Complementar nº 79, de 7 de janeiro de 1994, e o Decreto-Lei nº 3.689, de 3 de outubro de 1941 - Código de Processo Penal; e revoga dispositivo da Lei nº 10.201, de 14 de fevereiro de 2001. Brasília, DF, 04 de julho de 2012.
- CENTRO DE PESQUISA DO MINISTÉRIO PÚBLICO DO RIO DE JANEIRO. Letalidade Policial no Rio de Janeiro em 10 pontos. 2019. Disponível em: <http://www.mprj.mp.br/conheca-o-mprj/centro-de-pesquisas/letalidade-policial>. Acessado em setembro de 2019.
- CERQUEIRA, Daniel et al.. Avaliando a efetividade da Lei Maria da Penha. 2015. Disponível em http://www.ipea.gov.br/portal/images/stories/PDFs/TDs/td_2048.pdf. Acessado em agosto de 2019.
- CONSELHO NACIONAL DO MINISTÉRIO PÚBLICO (CNMP). Motivação dos crimes de homicídio. 2012. Disponível em <http://www.cnmp.gov.br/portal/images/stories/Noticias/2012/Apresentao2.pdf>. Acessado em agosto de 2019.
- DIRK, Renato. Comparação entre os Registros de Ocorrência e as Declarações de Óbitos. 2017. Disponível em http://arquivos.proderj.rj.gov.br/isp_imagens/uploads/ComparacaoSegurancaSaude2015.pdf. Acessado em agosto de 2019.
- FÓRUM BRASILEIRO DE SEGURANÇA PÚBLICA. Projeto: pesquisa e análise de dados vinculados ao campo da segurança pública e sistema penitenciário. 2016. Disponível em http://www.forumseguranca.org.br/storage/publicacoes/FBSP_Elaboracao_relatorios_semestrais_descritivos_2016.pdf. Acessado em agosto de 2019.
- MINISTÉRIO DA SEGURANÇA PÚBLICA. Portaria nº 229 de 10 de dezembro de 2018. Dispõe sobre a unificação e padronização das classificações e o envio de dados, definidos pelos entes federados, a serem implementados e fornecidos pelo Sistema Nacional de Informações de Segurança Pública, Prisionais e de Rastreabilidade de Armas e Munições, de Material Genético, de Digitais e de Drogas - Sinesp. Brasília, DF, 11 de dezembro de 2018.
- MONTEIRO, J. et al. Entre a rua e o tribunal: uma análise do processamento de casos de homicídio no Estado do Rio de Janeiro, Nota Técnica nº 1. Rio de Janeiro: Cenpe/MPRJ, 2020.
- MONTEIRO, J.; CARVALHO, E.F.; FERNANDES, J.G. Letalidade Policial e Criminalidade Violenta no Rio de Janeiro, manuscrito aceito para publicação na Revista de Administração Pública.

- NUNES, Marcelo Guedes et al.. O processamento de homicídios no Brasil e a Estratégia Nacional de Justiça e Segurança Pública em três Estados: Alagoas, Santa Catarina e São Paulo. 2016. mimeo disponível em: http://www.soudapaz.org/upload/pdf/artigo_processamento_de_homicidios_20161116.pdf. Acessado em agosto de 2019.
- RIBEIRO, Ludmila et al.. O Tempo do processo de homicídio doloso em cinco capitais. 2014. Crisp/UFGM; PNUD; Ministério da Justiça: Brasília.
- TRINDADE, Arthur; DURANTE, Marcelo. Medo do crime e vitimização no Distrito Federal: analisando as vulnerabilidades de gênero, idade, raça e renda, Dilemas: Revista de Estudos de Conflito e Controle Social, v. 12, n. 2, p. 239-265, 2018.
- ZILLI, Luis Felipe. Mensurando a violência e o crime: potencialidades, vulnerabilidades e implicações para políticas de segurança pública, Revista Brasileira de Segurança Pública, São Paulo. v. 12, n. 1, p. 30-48, fev./mar. 2018.

Fontes de dados

- FÓRUM BRASILEIRO DE SEGURANÇA PÚBLICA. Anuário Brasileiro de Segurança Pública. 2018. Disponível em: <http://www.forumseguranca.org.br/wp-content/uploads/2019/03/Anuario-Brasileiro-de-Seguran%C3%A7a-P%C3%BAblica-2018.pdf>. Acessado em agosto de 2019.
- INSTITUTO DE PESQUISA ECONÔMICA APLICADA (Ipea); Fórum Brasileiro de Segurança Pública. Atlas da Violência. 2019. Disponível em http://www.ipea.gov.br/portal/images/stories/PDFs/relatorio_institucional/190605_atlas_da_violencia_2019.pdf. Acessado em agosto de 2019.
- MINISTÉRIO DA JUSTIÇA E SEGURANÇA PÚBLICA. Infopen. Disponível em: <http://depen.gov.br/DEPEN/depen/sisdepen/infopen/relatorios-sinteticos/relatorios-sinteticos>. Acessado em agosto de 2019.
- MINISTÉRIO DA JUSTIÇA E SEGURANÇA PÚBLICA. Portal Sinesp. Disponível em <https://www.justica.gov.br/sua-seguranca/seguranca-publica/sinesp-1/bi/dados-seguranca-publica>. Acessado em agosto de 2019.
- MINISTÉRIO DA SAÚDE. Sistema de Informações de Mortalidade (SIM). Disponível em <http://datasus.saude.gov.br/sistemas-e-aplicativos/eventos-v/sim-sistema-de-informacoes-de-mortalidade>. Acessado em agosto de 2019.
- MINISTÉRIO DA SAÚDE. Vigilância de Violências e Acidentes (Viva). Disponível em <http://www.saude.gov.br/vigilancia-em-saude/vigilancia-de-violencias-e-acidentes-viva>. Acessado em agosto de 2019.
- MONITOR DA VIOLÊNCIA. Disponível em <http://especiais.g1.globo.com/monitor-da-violencia/2018/mortes-violentas-no-brasil/>. Acessado em agosto de 2019.
- SISTEMA DE INFORMACIÓN PARA LA SEGURIDAD Y LA CONVIVENCIA - SISC (2018).

Anexo 1- Fonte de divulgação de dados por Estado

Estado	Órgão que consolida informação	Periodicidade	Maior nível de desagregação	Sítio na web
AC	Secretaria de Estado da Segurança Pública (SESP)	Não tem dados	Não tem dados	http://www.seguranca.ac.gov.br
AL	Secretaria de Estado da Segurança Pública (SSP)	Mensal	Municípios	http://seguranca.al.gov.br/estatisticas/35/
AM	Secretaria da Segurança Pública (SSP)	Não tem dados	Não tem dados	http://www.ssp.am.gov.br
AP	Secretaria de Estado de Justiça e Segurança Pública (SEJUSP)	Não tem dados	Não tem dados	https://portaldaseguranca.portal.ap.gov.br/
BA	Secretaria da Segurança Pública (SSP)	Mensal	Municípios	http://www.ssp.ba.gov.br/modules/conteudo/conteudo.php?conteudo=102
CE	Secretaria da Segurança Pública e Defesa Social (SSPDS)	Mensal	Municípios/Bairros	https://www.sspds.ce.gov.br/estatisticas-2/
DF	Secretaria de Estado de Segurança Pública (SSP)	Mensal	RISP/Região Administrativa	http://www.ssp.df.gov.br/estatisticas/
ES	Secretaria de Estado da Segurança Pública e Defesa Social (SESP)	Mensal	Municípios	https://sesp.es.gov.br/estat%C3%ADsticas-criminais-2
GO	Secretaria da Segurança Pública (SSP)	Mensal	Municípios	https://www.seguranca.go.gov.br/painelOcorrencias.html
MA	Secretaria de Segurança (SSP)	Diário	Capital/Interior	https://www.ssp.ma.gov.br/estatisticas/estatisticas-da-grande-sao-luis/
MG	Secretaria de Estado de Segurança Pública (SESP)	Mensal	Municípios	http://www.seguranca.mg.gov.br/2018-08-22-13-39-06/dados-abertos
MS	Secretaria de Estado de Justiça e Segurança Pública (SEJUSP)	Mensal	Região	http://estatistica.sigo.ms.gov.br/
MT	Secretaria de Estado de Segurança Pública (SSP)	Mensal	RISP	http://www.seguranca.mt.gov.br/transparencia-estatisticas.php
PA	Secretaria de Segurança Pública e Defesa Social do Pará (SEGUP/PA)	Mensal	RISP	http://transparencia.segup.pa.gov.br/#6/-3.464/-47.336
PB	Secretaria da Segurança e da Defesa Social (SESDES)	Não tem dados	Não tem dados	https://paraiba.pb.gov.br/diretas/secretaria-da-seguranca-e-defesa-social/
PE	Secretaria de Defesa Social (SDS)	Mensal	Municípios	http://www.sds.pe.gov.br/estatisticas
PI	Secretaria de Segurança Pública (SSP)	Mensal	Capital/ Interior	http://www.ssp.pi.gov.br/estatisticas/
PR	Secretaria da Segurança Pública e Administração Penitenciária	Mensal	Municípios	http://www.seguranca.pr.gov.br/modules/conteudo/conteudo.php?conteudo=38
RJ	Instituto de Segurança Pública (ISP)	Mensal	Delegacias	http://www.ispdados.rj.gov.br

Estado	Órgão que consolida informação	Periodicidade	Maior nível de desagregação	Sítio na web
RN	Secretaria Estadual de Segurança Pública e Defesa Social (SESED)	Mensal	Municípios	http://defesasocial.rn.gov.br/Conteudo.asp?TRAN=ITEM&TARG=54041&ACT=&PAGE=&PARM=&L-BL=MAT%C9RIA
RO	Secretaria de Estado da Segurança, Defesa e Cidadania (SESDEC)	Mensal	Municípios	http://www.rondonia.ro.gov.br/sesdec/publicacoes/estatisticas/
RR	Secretaria de Estado da Segurança Pública (SESP)	Não tem dados	Não tem dados	http://www.sesp.rr.gov.br/
RS	Secretaria da Segurança Pública (SSP)	Mensal	Municípios	https://www.ssp.rs.gov.br/indicadores-criminais
SC	Secretaria de Estado de Segurança Pública (SSP)	Semanal	Municípios	http://www.ssp.sc.gov.br/index.php/component/content/article/88-servicos/184-seguranca-em-numeros-2?Itemid=437
SE	Secretaria de Estado da Segurança Pública (SSP)	Não tem dados	Não tem dados	https://www.ssp.se.gov.br/Transparencia
SP	Secretaria da Segurança Pública (SSP)	Mensal	Delegacias	http://www.ssp.sp.gov.br/Estatistica/Pesquisa.aspx
TO	Secretaria da Segurança Pública (SSP)	Trimestral	Estadual	https://www.ssp.to.gov.br/estatisticas-criminais/

Fonte: elaboração própria.

GAP

MACROECONOMIA

Guilherme Tinoco

06

Introdução

A macroeconomia é uma das áreas mais fascinantes da ciência econômica. Trata-se de um campo bastante amplo, que estuda as variáveis econômicas no agregado, destacando-se a produção, o emprego, o comércio externo, a inflação, dentre outras. Nos cursos de economia, a macroeconomia contrapõe-se à microeconomia, área da ciência que se dedica a estudar comportamentos individuais (de empresas e consumidores, por exemplo) e as particularidades dos diferentes setores da economia, embora, no mundo real, a micro e a macroeconomia sejam intimamente ligadas e dependentes uma da outra.

As principais questões na macroeconomia, desde seu surgimento, dizem respeito ao nível de produto e de emprego. Grosso modo, pode-se dizer que, no curto prazo, os objetivos seriam reduzir as flutuações econômicas e manter o pleno emprego, para os quais seriam utilizadas as políticas econômicas, principalmente a fiscal e monetária. No médio e longo prazo, a preocupação volta-se mais intensamente para o crescimento econômico, relacionado com a prosperidade material de uma sociedade¹.

Na academia, a macroeconomia possui uma infinidade de perguntas e linhas de pesquisa. No campo mais prático, contudo, pode-se afirmar que um dos pilares que melhor representa a macroeconomia é a análise de conjuntura, a qual se dedica este capítulo.

Por análise de conjuntura, entende-se a análise do ambiente econômico de um país, por meio dos dados que vão sendo divulgados a cada mês, trimestre ou ano. Além dos números, torna-se importante acompanhar as decisões de governo, do setor privado, e os acontecimentos em geral, que se difundem na sociedade por meio do noticiário do dia a dia e afetam as decisões de milhões de famílias e empresas.

Em realidade, esse conjunto amplo de informação diz respeito à realidade da economia, que afeta a vida de todos os cidadãos, mesmo que nem todos consigam perceber. Apesar disso, nos últimos anos parece haver um aumento no interesse pela macroeconomia na nossa sociedade. Os anos de baixo crescimento econômico (e até mesmo recessão), o acirramento das disputas políticas e a intensificação do debate via redes sociais são fatores que ajudam a explicar esse maior interesse.

¹ Para quem tiver interesse em aprofundar um pouco mais sobre a macroeconomia, não se dispensa um bom livro-texto, como o de Blanchard (2017). Além desta, sugerimos algumas leituras adicionais nas referências.

Nesse sentido, com o objetivo de contribuir para o aperfeiçoamento da discussão econômica, a ideia deste capítulo é apresentar os principais dados da macroeconomia, as principais armadilhas nos números e em sua interpretação, além de responder perguntas frequentes que podem interessar ao leitor. O capítulo está estruturado da seguinte forma: na próxima seção, apresentam-se os principais temas para a análise de conjuntura. Em seguida, na seção 3, expõem-se os principais dados de cada tema, incluindo a fonte e disponibilidade. Na seção 4, busca-se chamar atenção para alguns pontos que julgamos importantes, com o objetivo de evitar armadilhas e interpretação equivocada de dados, enquanto na seção 5 apresentam-se algumas perguntas frequentes. Por último, seguem as considerações finais.

Medidas usuais na conjuntura econômica

A análise de conjuntura macroeconômica engloba vários temas. No fundo, o analista quer saber como está o desempenho da economia. O país vai crescer mais? Os salários vão aumentar? Existe a probabilidade considerável de uma crise à frente? O que o governo pode ou deveria fazer para melhorar o desempenho econômico do país?

Para analisar o quadro geral e formular respostas para essas perguntas, economistas precisam olhar para diferentes áreas da conjuntura, conhecidas como as “caixinhas da conjuntura”. Normalmente, os departamentos econômicos de grandes bancos dividem seus analistas dentro desses temas, cabendo ao economista-chefe juntar as peças do quebra-cabeça para traçar cenários e projeções completas e consistentes.

Nesta seção, o objetivo é apresentar quais são essas áreas. Listamos as principais: (i) atividade econômica, (ii) inflação e taxa de juros, (iii) crédito, (iv) setor externo (balanço de pagamentos), (v) contas do governo, (vi) variáveis financeiras e (vii) economia internacional.

Antes de iniciar, vale o alerta de que a análise fica mais rica quando o analista compreende bem o funcionamento do sistema macroeconômico, pois assim é possível combinar melhor a análise que sai de cada caixinha, uma vez que existem muitas interdependências entre elas.

Em um regime de metas de inflação, por exemplo, quando a expectativa de inflação fica acima da meta, passa-se a esperar uma reação do Banco Central no sentido de aumentar a taxa de juros. O analista mais experiente, que sabe que as consequências desse aumento costumam ser negativas para a atividade, poderá então se preparar para uma desaceleração na produção e um desaquecimento no mercado de trabalho, por exemplo. Tentaremos explicar melhor essas relações à medida que apresentamos os temas, começando pela atividade econômica.

Atividade econômica

A atividade econômica acaba sendo o coração da macroeconomia e da análise de conjuntura. Suas medidas, as quais se destacam a produção e o emprego, são materializadas através de indicadores agregados, dos quais o melhor exemplo é o Produto Interno Bruto (PIB) e sua composição. O que o país está produzindo? Em quanto a produção está aumentando? Quais setores estão produzindo e crescendo

mais e quais setores estão produzindo e crescendo menos?

O nível de produção de bens e serviços é o coração da macroeconomia porque, em última análise, é o que vai determinar o nível de consumo daquela sociedade. O ponto de partida conceitual para entender o que determina o volume produzido é a função de produção, bastante conhecida entre os economistas, que relaciona os insumos principais com o resultado final.

Desses insumos, um dos mais importantes é justamente o trabalho (que diz respeito não só à quantidade, mas também à qualidade). A conclusão, de certa maneira óbvia, é que a produção da economia está intimamente ligada ao emprego.

Logo cedo, os estudantes de economia aprendem sobre a Lei de Okun, uma “lei” que estabelece uma relação (inversa) entre desemprego e produção. A lei é uma homenagem ao economista norte-americano Arthur Okun, que estudou esta relação nos anos 1960, sugerindo que, quando o desemprego caía, a produção aumentava e vice-versa. Para o aumento de produção, portanto, torna-se importante perseguir o pleno emprego e, por isso, a análise do mercado de trabalho torna-se tão relevante².

A produção de uma economia não é determinada somente pelo nível de emprego: de acordo com a função de produção, ela também depende do estoque de capital (estoque de investimento) e da produtividade (um conceito amplo que tem a ver tanto com as tecnologias utilizadas quanto com a eficiência alocativa dos fatores de produção).

Dessa maneira, para que o economista faça uma boa análise a respeito da produção futura, de longo prazo, deverá acompanhar todos esses temas. Precisaria passar pelos indicadores de investimentos das empresas e dos investimentos em infraestrutura, pelas novas tecnologias, pela qualidade das instituições, pelo desenho de mercado (responsáveis pelos incentivos ao setor produtivo), dentre outros. Para o curto prazo, além dos números da própria produção, o analista precisa entender aspectos da demanda, como, por exemplo, as vendas no varejo e os indicadores de confiança, e aspectos da oferta, como os indicadores de ociosidade (nível de utilização da capacidade e o próprio desemprego), que podem ajudar a sinalizar algumas tendências de curto prazo.

2 Vale notar que uma economia saudável é aquela que utiliza todos os recursos disponíveis, especialmente o trabalho, que se relaciona mais intimamente com a qualidade de vida da população.

Inflação e taxas de juros

A inflação é o fenômeno relacionado a mudanças no nível de preços da economia. Economistas em geral concordam que uma inflação baixa (ou estabilidade de preços) é um valor a ser preservado, dentre outras coisas, porque é relacionado com maior crescimento econômico.

De fato, a elevação constante e significativa dos preços é bastante prejudicial à atividade econômica. O Brasil já viveu tempos de hiperinflação, dos quais ninguém tem saudade. A inflação anual teve um pico de mais de 2.000% em 1993, um fato impensável para os dias de hoje. Como esse passado ficou para trás, vamos focar no presente.

Desde 1999, o Brasil adota o regime de metas de inflação, um sistema utilizado por diversos países do mundo. De maneira simples, esse regime faz com que o Banco Central persiga uma taxa de inflação pré-determinada, que hoje está em torno de 4% ao ano. Para atingir esta meta de inflação, o principal instrumento é a taxa de juros. Economistas modelam o comportamento do Banco Central através da chamada regra de Taylor³: se a inflação e as expectativas de inflação estão abaixo (acima) da meta, o BC pode reduzir (aumentar) os juros.

Desta maneira, percebe-se a importância do acompanhamento da inflação corrente, pois, no fundo, ela pode revelar o futuro da política monetária. Uma inflação pressionada, por exemplo, pode indicar aquecimento da economia, exigindo uma política monetária mais contracionista. E sabe-se que, tudo mais constante, maiores níveis de juros significam menores níveis de atividade.

Além disso, a inflação diz respeito ao poder de compra da população. Como os salários nominais são estáveis, ou rígidos (no jargão da profissão), uma inflação maior retira poder de compra das pessoas. Por isso, inflações elevadas também são associadas a queda de popularidade de governantes, o que, para casos extremos, pode ter alguma consequência política mais significativa⁴.

O acompanhamento da inflação é muito importante para o mercado financeiro e isso ocorre não só pelos efeitos que ela gera na taxa de juros e na economia real, mas também por ser insumo para a precificação de ativos, principalmente de renda fixa.

3 Em homenagem ao economista John Taylor, que criou a regra em 1993.

4 Essa relação fica clara, por exemplo, por meio da taxa de miséria, que é simplesmente a soma da taxa de desemprego com a taxa de inflação e que tem uma boa correlação com as taxas de popularidade de um governante.

Pequenas variações no IPCA podem gerar elevados ganhos ou perdas para investidores e, por essa razão, instituições do mercado financeiro normalmente possuem um acompanhamento bastante aprofundado dos dados de inflação.

Como falado anteriormente, a inflação tem bastante relação com as taxas de juros. E taxas de juros têm grande ligação com o custo do crédito, sendo este um dos principais canais de transmissão da política monetária para a atividade.

Crédito

O crédito é uma variável fundamental e intimamente ligada com o nível de atividade, pois diz respeito às condições de financiamento de famílias e empresas para que possam consumir, investir ou simplesmente conduzir os seus negócios.

Para analisar o comportamento do crédito da economia, os analistas tomam como ponto de partida as estatísticas do crédito bancário, com destaque para dados de estoque (saldo), volume de concessões, modalidades, taxas de juros, spreads, prazos e taxa de inadimplência.

O saldo se refere ao estoque que já foi emprestado, no qual interessa não só o montante total, como também o prazo, o custo, a taxa de inadimplência e, como veremos a seguir, a modalidade. As concessões indicam o fluxo de novas operações, para as quais também é importante conhecer as condições (juros, prazo, finalidade etc.).

Os grandes recortes do crédito dão-se entre pessoa física e pessoa jurídica, ou entre crédito livre e crédito direcionado. Ao analista de crédito, não interessa somente o saldo de crédito ou o volume de concessões em um dado período, mas também as suas características e modalidades. Tomemos, como exemplo, o crédito às famílias (pessoa física). Faz muita diferença se o novo crédito (ou mesmo o estoque) está sendo destinado para o financiamento habitacional, para a compra de automóveis, para pequenas aquisições ou ainda para refinanciamento de empréstimos.

No bloco de crédito também podemos abordar a análise da situação financeira de consumidores e empresas. O nível de endividamento de cada um deles pode indicar espaço ou não para novas contratações de crédito no curto prazo, o que influencia no nível de atividade. Caso os agentes estejam endividados, por exemplo, contratarão menos crédito, o que implica em menor consumo e investimento à frente.

Ao mesmo tempo, outra questão importante é a sustentabilidade do crédito. Os economistas sabem que problemas no mercado de crédito podem desencadear gran-

des crises, a exemplo da recessão de 2008. Por isso, devem sempre estar atentos quanto aos sinais de fragilidade financeira, embora, para isso, ao menos podem contar com a supervisão do Banco Central e demais autoridades do setor⁵.

Falou-se muito do crédito bancário até aqui, mas podemos ir além e encarar a caixinha do crédito como um grande bloco de financiamento de famílias e empresas. Neste caso, é importante olhar para outras formas de financiamento, tais como os instrumentos no mercado de capitais. Em realidade, o crédito nada mais é do que a transferência de recursos entre poupadores e tomadores. Logo, todas as formas que os conectam devem compor o objeto de análise deste tema de conjuntura.

Setor externo (Balanço de Pagamentos)

É por setor externo que a análise de conjuntura chama as relações de um país com o exterior. A principal variável de atenção consiste no volume de comércio (uma soma), ou balança comercial (uma diferença), que levam em conta as exportações e importações de bens e serviços.

No caso da balança comercial, diversas análises podem ser realizadas, como, por exemplo, quais os tipos de produto o país está vendendo ou comprando. Adicionalmente, pode ser verificado se o país está dependendo demasiadamente de moeda estrangeira para fechar as suas contas externas, o que poderia sinalizar algum problema de restrição externa.

De maneira mais abrangente, as contas do setor externo são reportadas no Balanço de Pagamentos (BP), um conceito apresentado aos alunos de economia no início do curso de graduação. O BP engloba não só a balança comercial, como a totalidade da conta corrente do país (que inclui a balança comercial, a balança de serviços fatores e não fatores e as transferências) e a conta de capital (fluxos de capital, como, por exemplo, investimentos em renda fixa e investimento externo direto).

As relações de um país com o exterior não se resumem às contas de fluxo, como aquelas mapeadas pelo BP. Há também as contas de estoque, como, por exemplo, o passivo externo líquido. Nessas contas, inclui-se também o saldo das reservas internacionais, um ativo de liquidez elevada, extremamente ligado às condições de solvência externa de uma economia.

⁵ É por conta das graves consequências que podem ser causadas por crises no crédito bancário é que este setor é tão regulado. Apesar disso, nem sempre as fragilidades são detectadas a tempo, como no caso da crise do subprime de 2008, nos Estados Unidos.

A ligação das contas externas com a atividade e com o resto da macroeconomia, além de óbvia, é fortíssima. As importações e exportações definem as características de muitos setores econômicos. O comportamento dos investimentos estrangeiros (diretos e indiretos) é fundamental para diversas áreas da economia. E, por fim, a evolução das contas externas acaba tendo consequências na taxa de câmbio, que, como vamos ver à frente, afeta uma série de variáveis macroeconômicas.

Contas do governo

As contas do governo são também muito importantes para compor o quadro macroeconômico do país. Governos arrecadam impostos para financiar gastos públicos. Quando os impostos não são suficientes para honrar as obrigações, governos também se endividam com o setor privado para financiá-las.

A análise das contas públicas é relevante por várias razões. A primeira delas é por questão de solvência: um governo precisa apresentar sustentabilidade intertemporal em seu orçamento. Quando isso não acontece, os agentes passam a esperar: (i) aumento de impostos, (ii) calote do governo ou (iii) aceleração da inflação, todas consequências indesejáveis e que têm potencial para desorganizar a economia.

Outra razão é que, por meio do gasto público, o governo implementa suas políticas públicas, que afetam diferentemente regiões, setores e extratos populacionais. Da mesma maneira, o desenho da estrutura tributária, que determina a arrecadação, influencia a economia real, seja na organização da produção ou na distribuição de renda, para ficar em dois exemplos.

De uma maneira geral, portanto, para os diferentes fins, acompanham-se os dados de estoque (nível de dívida) e dados de fluxo (receitas e despesas). Uma análise importante passa pela projeção da sustentabilidade (ou não) da dívida pública, bem como o cumprimento das regras fiscais em vigor (por exemplo, as metas de superávit primário e, mais recentemente, o teto de gasto).

Como este tema será abordado em capítulo à parte, deixemos os detalhes para lá e passemos à próxima caixinha.

Variáveis financeiras

O bloco financeiro é um dos mais importantes para o cenário macroeconômico. Por bloco financeiro, vamos nos concentrar na taxa de câmbio, no risco país, na bolsa de valores e na curva de juros.

Esses indicadores são fundamentais para atividade. A taxa de câmbio tem uma relação íntima com o nível de preços da economia (inflação), além de influenciar a situação patrimonial (de pessoas, empresas e governos) e afetar os fluxos de comércio externo (e investimentos) da economia, sem falar na própria demanda interna por consumo e investimento.

O risco país, por sua vez, ajuda a influenciar a taxa de câmbio e, por consequência, todas as variáveis que dependem do câmbio, além de sinalizar fluxos e apetite para o investimento externo.

Já a bolsa de valores representa um preço que não interessa apenas a investidores e especuladores. Muitas vezes, seus índices também podem estar sinalizando a saúde das empresas, as expectativas de crescimento econômico futuro e o custo de capital, dentre outras coisas. O mesmo ocorre com a curva de juros, indicador fundamental para o setor financeiro e também para o setor real da economia.

Economia internacional

Até agora, tratou-se das variáveis que compõem o cenário doméstico. Contudo, é importante ressaltar que acontecimentos no cenário internacional têm grande reflexo na economia interna, normalmente transmitidos, inicialmente, por meio das variáveis financeiras apresentadas anteriormente. Por essa razão, analistas de conjuntura precisam acompanhar de perto os acontecimentos no exterior.

Tudo é importante: desde as perspectivas para o crescimento na China, passando pela política monetária dos EUA e o ritmo de atividade na Europa, até a evolução de preços de commodities, negociações comerciais, disputas geopolíticas, fenômenos climáticos/naturais relevantes e evolução na fronteira tecnológica.

Os dados no Brasil

Na seção anterior, apresentaram-se os principais temas, ou caixinhas, da análise de conjuntura. Nesta seção, concentram-se os principais dados de cada tema. Como será visto, as fontes mais importantes são o IBGE e o Banco Central. Assim, recomenda-se ao leitor mais interessado que explore não só os respectivos websites, mas também as notas técnicas e explicativas disponibilizadas para cada dado ou estatística, principalmente quando houver dúvidas e questionamentos.

Os temas apresentados nesta seção seguem, de maneira geral, a seção anterior: (i) atividade econômica, (ii) inflação e política monetária, (iii) crédito, (iv) setor externo, (v) variáveis financeiras e (vi) economia internacional. Não serão abordados os temas que serão tratados com maior detalhe em capítulos específicos, como mercado de trabalho.

Atividade econômica

a) PIB

O principal dado de produção é o PIB, divulgado trimestralmente pelo IBGE por meio das contas nacionais trimestrais. O PIB é a soma de todos os bens e serviços produzidos no país em um determinado tempo. O volume de produção é medido em termos nominais: por exemplo, o PIB brasileiro alcançou R\$ 7.257 bilhões em 2019, ou seja, mais de R\$ 7 trilhões.

Quando estamos falando de PIB, também queremos saber o quanto a produção em determinado tempo variou em relação a outro período. No caso do PIB de 2019, por exemplo, podemos verificar que houve crescimento de 1,1% em relação ao PIB de 2018.

Quando verificamos o valor do PIB nominal em 2018, contudo, encontramos a marca de R\$ 6.889 bilhões, o que pode suscitar a primeira dúvida: o crescimento não teria sido de 5,3%, ao invés de 1,1%? Na verdade, como os preços variaram no período, não é correto compararmos as medidas de PIB nominal: para obter o crescimento real, devemos fazer a comparação levando em conta os preços da economia. Para a sorte do analista, o IBGE já calcula as variações reais, que têm destaque grande na divulgação

das contas nacionais⁶.

Os dados do PIB não se restringem ao valor da produção total. Em realidade, a divulgação das contas nacionais apresenta dados desagregados, isto é, por componente do produto. Em geral, as desagregações mais importantes são definidas pelo lado da oferta e pelo lado da demanda.

Pelo lado da oferta, são três categorias (agropecuária, indústria e serviços) com suas subdivisões. Já o lado da demanda é desagregado em consumo das famílias, investimento (formação bruta de capital fixo), consumo do governo e setor externo (diferença entre exportações e importações de bens e serviços). A Tabela 1 apresenta os componentes de oferta e de demanda.

Tabela 1 – Contas Nacionais (ótica de oferta e demanda)

PIB oferta	PIB demanda
Agropecuária	Despesa de consumo das famílias
Indústria	Despesa de consumo da administração pública
Extrativa mineral	Formação bruta de capital fixo
Transformação	Exportação de bens e serviços
Produção e distrib. eletricidade, gás, água, esgoto e limpeza urbana	Importação de bens e serviços (-)
Construção civil	
Serviços	
Comércio	
Transporte, armazenagem e correio	
Serviços de informação	
Interm. financeira, seguros, previdência comp. e relacionados	
Atividades imobiliárias e aluguéis	
Outros serviços	
Administração, saúde e educação públicas	

(*) Componentes de oferta são mensurados por valor adicionado e acrescidos de impostos líquidos para a obtenção do PIB total. Fonte: IBGE

Por fim, vale um alerta: o PIB é a soma de toda a produção de um país, mas seu valor total depende fortemente do tamanho da população do país. Para comparações internacionais, portanto, é importante ter em mente o conceito de PIB per capita, que divide o PIB de um país por sua população. Este indicador vai ser mais correlacionado

⁶ Os dados de contas nacionais são divulgados em volume físico (índice real) e em valores nominais. Comparando à variação entre os dois, podemos obter a inflação do PIB, também chamada de deflator do PIB. Para o exemplo acima, o deflator foi de 4,2% em 2019 em relação à 2018, calculado da seguinte forma: $((1 + \text{variação nominal}) / (1 + \text{variação real}) - 1)$. Essa conta fica, portanto, $1,053 / 1,011 - 1 = 1,042 - 1 = 4,2\%$.

com a riqueza e nível de vida de um país do que simplesmente o PIB total.

b) Produção

O PIB é, sem dúvida, o indicador mais relevante e completo sobre a produção de uma economia. Contudo, é divulgado com uma defasagem temporal grande: em geral, no Brasil, a divulgação ocorre somente dois meses após o encerramento do trimestre. Por isso, algumas pesquisas divulgadas com antecedência menor, e em bases mensais, são bastante importantes para indicar o ritmo de produção, ou de atividade, em um período mais curto, ainda que sejam destinadas a algum setor específico.

Vamos destacar duas delas, que nos parecem mais abrangentes: indústria e serviços, todas divulgadas mensalmente pelo IBGE⁷.

Indústria: a Pesquisa Industrial Mensal (PIM) é uma pesquisa realizada mensalmente pelo IBGE e que cobre a indústria brasileira, com destaque para a produção física por setor. São duas aberturas principais: (i) seções e atividades (com 26 “setores”, sendo um referente à indústria extrativa e 25 referentes à indústria de transformação) e (ii) categorias de uso (com 4 categorias: bens de capital, bens intermediários, bens de consumo duráveis e bens de consumo semi e não-duráveis, seguidas de suas subdivisões).

Serviços: a Pesquisa Mensal dos Serviços (PMS) é uma pesquisa realizada mensalmente pelo IBGE e que cobre setor de serviços, com destaque para o volume e receita nominal por setor. Os resultados são divulgados para cinco categorias principais, seguidas de suas divisões.

Antes de terminar essa subseção, é importante falar do IBC-Br, informalmente conhecido como PIB mensal. Enquanto analistas esperam a divulgação do PIB, com defasagem significativa, vimos que existem algumas pesquisas para ajudar a medir o comportamento de setores específicos. Contudo, também existem indicadores que tentam medir a atividade como um todo, praticamente funcionando como uma prévia do PIB. A estes damos o nome de indicadores agregados de atividade, dos quais o mais importante é o IBC-Br, calculado mensalmente pelo Banco Central⁸.

7 Um ponto importante a ser destacado é que, embora essas pesquisas sejam insumos para o cálculo do PIB, elas possuem algumas diferenças. No caso dos serviços, por exemplo, a PMS abrange apenas cerca de 20% dos serviços cobertos pelo PIB. Assim, é importante ter cuidado ao interpretar essas pesquisas como uma prévia do PIB do respectivo setor.

8 Existem outros indicadores agregados de atividade, como o monitor do PIB, calculado pelo

O IBC-Br é um indicador de volume, divulgado com e sem o ajuste sazonal. Seu cálculo leva em conta várias pesquisas setoriais, incluindo as que citamos acima. Funciona, portanto, como uma boa aproximação do PIB, embora tenha mais aderência com o número do IBGE em janelas mais longas, como veremos na seção 4.

Por fim, vale mencionar que existe uma série de indicadores setoriais, que muitas vezes são divulgados com defasagem mínima e ajudam a interpretar o quadro macroeconômico. Um bom exemplo está no setor de veículos. Mensalmente, são divulgados dados não apenas de produção, como também de vendas, exportações e importação de auto veículos, ônibus e caminhões pela Anfavea, a associação dos fabricantes do setor. Outro exemplo, diz respeito aos chamados indicadores antecedentes, que ajudam a prever o ritmo de atividade, como a produção de papel ondulado, o fluxo de caminhões nas estradas e o consumo de energia elétrica. Como o objetivo do capítulo é tratar do agregado, contudo, não vamos nos aprofundar nesses dados.

c) Demais

Procurou-se destacar acima os principais indicadores de atividade, focando na produção. Quando falamos em nível de atividade, outro bloco importantíssimo é o mercado de trabalho, onde o analista se debruça nos dados de emprego, tipo de ocupação, níveis salariais etc., assunto que será explorado em maior detalhe em capítulo posterior.

No bloco de atividade, também é importante falar dos indicadores de vendas (varejo), mais relacionados com a demanda. Destacamos a Pesquisa Mensal do Comércio (PMC), realizada mensalmente pelo IBGE e que cobre o comércio brasileiro, com destaque para o volume e receita nominal por setor. São divulgados resultados para o varejo restrito (com 8 grandes setores) e para o varejo ampliado (inclui os setores do varejo restritos e mais dois: automóveis e materiais de construção).

Em realidade, dentro de atividade, existe uma infinidade de indicadores importantes, além da produção, do emprego e das vendas. São medidas que auxiliam o analista a investigar melhor o nível e as tendências para a economia no curto prazo. Destacamos duas, que julgamos bem relevantes.

A primeira é a ociosidade. Saber se as empresas estão operando a pleno vapor ou com bastante folga pode ajudar a pensar na evolução do investimento à frente, bem

como emitir sinais para ajudar a prever a inflação futura. A FGV divulga o Nuci (Nível de Utilização da Capacidade Instalada) para a indústria e para os serviços, em bases mensais.

A segunda são os níveis de confiança. Diferentes institutos possuem pesquisas para captar os níveis de confiança para setores variados da economia. A FGV, por exemplo, divulga mensalmente para cinco categorias: consumidores, empresários da indústria, dos serviços, do comércio e da construção. São dados que acabam se relacionando de alguma forma, mesmo que com defasagens, com o nível de atividade.

Inflação, taxas de juros, política monetária

a) Inflação

Existem diversos índices de inflação no Brasil, sendo que os principais são o IPCA (Índice Nacional de Preços ao Consumidor Amplo), do IBGE, e os IGPs (Índice Geral de Preços), da FGV.

Os IGPs são índices gerais de preços, compostos pelo IPC (Índice de Preços ao Consumidor), IPA (Índice de Preços ao Produtor Amplo) e INCC (Índice Nacional de Custo da Construção), com pesos respectivos de 30%, 60% e 10%. São denotados por um sufixo (M, DI ou 10), que indica qual o período de coleta dos preços.

O mais importante, contudo, é o IPCA, divulgado mensalmente pelo IBGE, e utilizado como referência para o regime das metas de inflação. A população-objetivo do índice é formada por famílias com rendimentos de 1 a 40 salários mínimos, residentes em áreas urbanas compostas por 13 regiões metropolitanas ou municípios brasileiros. A cesta de consumo, que tem por base a POF (Pesquisa de Orçamentos Familiares), possui 377 itens, cada um com um peso diferente. A média ponderada da inflação dos itens desta cesta corresponde à inflação cheia do período⁹.

A cesta de consumo abrange itens bem diferentes, que sofrem as mais distintas influências. Portanto, para entender melhor a dinâmica dos preços, é importante que o analista tenha um olhar desagregado.

A principal desagregação é entre preços livres e preços administrados. Por pre-

9 É importante mencionar que o IBGE também divulga o IPCA-15, com metodologia praticamente igual ao do IPCA, mas com o período base variando do dia 15 de um mês ao dia 15 do mês seguinte. O IPCA tem como base o mês calendário (dia 01 até o final do mês).

ços administrados, entendem-se preços que são regulados, normalmente associados à prestação de serviços, tais como energia elétrica ou produtos como medicamentos e combustíveis.

O IBGE divulga o IPCA com uma desagregação que engloba 9 categorias (alimentação e bebidas; habitação; artigos de residência; vestuário, transportes; saúde e cuidados pessoais; despesas pessoais; educação e comunicação). Contudo, uma divisão proposta pelo Banco Central também é bastante útil para análises, contando com quatro divisões (alimentos no domicílio, preços administrados, serviços e bens industriais) e utilizado no modelo de pequeno porte daquela instituição¹⁰.

Para analisar o comportamento da inflação, os economistas também se preocupam com os núcleos. Os núcleos são medidas de inflação que expurgam alguns itens mais voláteis das métricas de inflação, dando uma noção mais fidedigna da evolução dos preços com base nas condições da economia. O Banco Central calcula e divulga algumas medidas de núcleos de inflação para o IPCA.

b) Taxa de juros e política monetária

O Copom (Comitê de Política Monetária), do BCB, reúne-se 8 vezes por ano. A cada reunião é definida a taxa de juros que vigorará na economia até a reunião seguinte. A série 432, do sistema gerenciador de séries temporais, apresenta o histórico da taxa Selic definida pelo Copom¹¹. Além dos dados em si, é de grande interesse o comunicado e a ata das reuniões, em que o Copom informa e fundamenta suas decisões ao mercado, trazendo inclusive suas próprias projeções de inflação.

Além disso, outro material bastante importante para analistas que acompanham a política monetária é o relatório trimestral de inflação, divulgado trimestralmente pelo Banco Central. Nele, o BC explicita o seu cenário, faz análises e muitas vezes aproveita para apresentar inovações em seus modelos de previsão, em um esforço de transparência inerente ao regime de metas de inflação.

¹⁰ O modelo de pequeno porte do Banco Central é um modelo macroeconômico de curto prazo, que contém as equações de equilíbrio do modelo novo-keynesiano, como a curva IS e a curva de Phillips. Em uma versão desagregada, apresenta curvas de Phillips para diferentes categorias dos preços livres, cada uma com as suas especificidades. Para mais detalhes, ver Lima et al. (2011).

¹¹ O Sistema Gerador de Séries Temporais (SGS) do Banco Central é um banco de dados que disponibiliza as principais séries para a economia brasileira (não só as produzidas pelo próprio BC).

Crédito

As principais estatísticas de crédito bancário são divulgadas mensalmente pelo Banco Central, através das notas econômico-financeiras para a imprensa (estatísticas monetárias e de crédito).

Como já explicitamos na seção anterior, os dados mais relevantes são aqueles de estoque (saldo) de crédito, volume de concessões, modalidades, taxas de juros, spreads, prazos e taxa de inadimplência, recortados entre pessoa física e pessoa jurídica ou crédito livre e direcionado.

Como falamos, pode-se agregar outras informações no bloco de crédito, como, por exemplo, a situação patrimonial de famílias e empresas. Nesse caso, existem bases privadas, como o Economatica ou o Valor Data, que compilam dados corporativos das maiores empresas brasileiras.

Setor externo

As principais estatísticas do setor externo são divulgadas mensalmente pelo Banco Central, através das notas econômico-financeiras para a imprensa (estatísticas do setor externo).

O balanço de pagamentos é o coração da análise do setor externo e os valores são divulgados em dólares. Quando acumulados em 12 meses, os números normalmente são apresentados em percentual do PIB, principalmente para facilitar as comparações internacionais.

Variáveis financeiras

Como apresentado na seção anterior, as variáveis mais relevantes nessa caixa são a taxa de câmbio, o risco país, a bolsa de valores e a curva de juros. Vários sites disponibilizam esses dados, como o portal Infomoney. Contudo, para um banco de dados mais completo (séries históricas longas, tempo real etc.) normalmente as melhores fontes são pagas, como o famoso Terminal da Bloomberg. Para a maioria das análises de conjuntura, contudo, esses serviços podem não ser necessários, sendo suficiente a oferta de dados gratuita, disponível na internet.

Economia internacional

As estatísticas de economia internacional estão fora do escopo do capítulo, mas podemos indicar algumas referências. Uma delas é o relatório World Economic Outlook, divulgado trimestralmente pelo Fundo Monetário Internacional (FMI). Essa publicação ajuda a entender em linhas gerais os grandes assuntos relevantes para a economia global, além de trazer projeções macroeconômicas para diversos países. Acompanha também uma base de dados importantes, com estatísticas macroeconômicas (produção, contas públicas, comércio, inflação, etc.) para diversos países ao longo de vários anos.

Outros

Até agora, focou-se nas séries históricas das variáveis, isto é, números realizados. Uma parte importante da análise de conjuntura, contudo, diz respeito à previsão e à elaboração de cenários. Quanto a economia vai crescer no ano que vem? Qual vai ser a taxa de juros no final deste ano?

Os economistas partem do passado para projetar cenários futuros, combinando as variáveis descritas aqui com premissas exógenas e algum tipo de modelagem econômica. Uma famosa frase diz que economistas são como meteorológicos: pródigos em errar suas previsões. Verdade ou não, o fato é que o trabalho de fazer previsão é de fato muito difícil.

Não é objetivo deste capítulo falar sobre previsão, mas é importante expor que, fora a projeção de cada instituição, existe a projeção de mercado. Trata-se, portanto, de uma projeção “média” que acaba dando o norte do cenário para muitos analistas de conjuntura.

Nesse sentido, uma publicação particularmente importante é o Boletim Focus, divulgado semanalmente pelo Banco Central e que traz as projeções médias (e medianas) do mercado para uma série de variáveis, destacando-se inflação (IPCA), PIB, juros e câmbio. Existe o ranking das instituições que mais acertam, que gera as projeções Top 5, que podem ser encaradas como projeções de maior qualidade. As previsões de inflação são especialmente importantes, uma vez que acabam sendo um insumo para as decisões de política monetária por parte do Banco Central.

Armadilhas dos dados e na interpretação dos dados

Sazonalidade

Como já explicado em capítulos anteriores, um dos principais pontos de atenção nos dados macroeconômicos diz respeito à sazonalidade. Por exemplo, o ritmo de produção (PIB) não é constante em todos os trimestres do ano, da mesma forma que a inflação tem também um componente sazonal muito forte. O mercado de trabalho, como será visto adiante, é outra caixinha da conjuntura que possui uma sazonalidade bem particular.

Muitos dados econômicos são divulgados também com ajuste sazonal (exemplo: PIB, PIM, PMC, PMS, IBC-Br), mas nem todos são assim. Além disso, há casos específicos: inflação do IPCA é divulgada com e sem o ajuste sazonal, mas a grande maioria dos analistas e jornalistas preferem trabalhar com o número sem ajuste. Assim, quando alguém diz que a inflação se acelerou em relação ao mês anterior, temos que tomar certo cuidado, porque é possível que a inflação não tenha efetivamente se acelerado quando descontado o fator sazonal.

As séries de produção citadas acima são apresentadas em índice de volume. Essas séries são disponibilizadas com e sem ajuste sazonal. Para realizar comparações na margem (isto é, de um período em relação ao período imediatamente anterior), é um pouco óbvio que devemos utilizar as séries com ajuste sazonal. Em comparações acumuladas, contudo, o analista pode ficar em dúvida sobre qual série utilizar. Veremos esse ponto a seguir.

Comparações acumuladas (trimestre, 12 meses, acumulado no ano): qual série utilizar?

Quando se deparar com dados macroeconômicos, principalmente aqueles ligados à produção, o analista vai querer estabelecer as comparações e, em geral, o maior interesse está na variação do produto entre dois períodos.

A comparação mais demandada é aquela que chamamos de variação na margem, isto é, entre dois períodos adjacentes. Contudo, existe outro tipo de comparação bastante utilizado, que diz respeito à variação entre períodos iguais de anos diferentes. Qual série, então, usar para cada tipo de comparação? A Tabela 2 responde à pergunta.

Tabela 2 – Comparações e tipos de série (*)

Comparação	Exemplo	Série a ser utilizada
Mês (trimestre) contra mês (trimestre) anterior	maio de 2020 contra abril de 2020	com ajuste sazonal
Mês (trimestre) contra mesmo mês (trimestre) do ano anterior	maio de 2020 contra maio de 2019	sem ajuste sazonal
Acumulado no ano	jan a mai/2020 contra jan a mai/2019	sem ajuste sazonal
Acumulado em 12 meses	jun-19 a mai-20 contra jun-18 a mai-19	sem ajuste sazonal

(*) Regra para séries relacionadas a fluxo de produção, vendas e outras similares. Fonte: elaboração própria.

A regra, portanto, é clara: se estamos comparando meses iguais entre diferentes anos, não faz sentido usar a série com ajuste sazonal. Nesse sentido, comparar maio de um ano com maio do ano anterior, usando a série com ajuste sazonal, é um erro. Fique atento!

No entanto, vale uma ressalva: em determinados casos, o analista pode querer corrigir o dado para a diferença de dias úteis. Por exemplo, para comparar a produção de maio de 2020 com a de maio de 2019, deve ser utilizada a série sem ajuste sazonal. No entanto, se maio de 2020 teve dois dias úteis a mais que maio de 2019, o valor do primeiro, tudo mais constante, tende a ser maior do que o segundo. Uma opção, portanto, seria comparar o valor total dividido pelo número de dias úteis. Vale notar, contudo, que essa correção, dependendo do objetivo da análise, pode ser contraproducente. Em jornais de grande circulação, por exemplo, pode acabar confundindo o leitor. Melhor deixá-la, portanto, para análises bastante específicas.

Variáveis reais X variáveis nominais

A palavra real tem um significado especial na macroeconomia (análise de conjuntura). As variações reais dizem respeito a variações que descontam a variação de preços.

Um exemplo deixa isso muito claro. Se no ano 1 forem produzidas 20 laranjas ao preço de cinco reais, a produção nominal terá sido de 100 reais. Se no ano 2 forem produzidas 22 laranjas ao preço de seis reais, a produção nominal terá sido de 132 reais. O aumento nominal da produção terá sido 32%, mas o aumento real terá sido de apenas 10%. Ou seja, uma grande diferença que, se negligenciada na análise, pode levar a conclusões equivocadas.

Para a maioria das séries de produção, que são divulgadas em índices de volume ou quantidade física, isso não é um problema. Para outros tipos de divulgação, no entanto, é importante fazer a correção para evitar comparações equivocadas.

Percentual do PIB

Muitas das variáveis econômicas são medidas em termos de percentual em relação ao PIB. Os exemplos abundam: variáveis de crédito (como saldo), variáveis do setor externo (como déficit em transações correntes) e variáveis fiscais (como dívida, receita e despesa).

Normalmente, essas variáveis já são divulgadas em percentual do PIB pela própria pesquisa que lhes dá origem, mas em alguns casos essa tarefa caberá ao próprio analista (principalmente quando está trabalhando em projeções para o futuro). Nesse caso, a regra é clara. A fração deve ser calculada a partir de valores comparáveis: variável nominal sobre variável nominal.

“No Brasil, até o passado é incerto”

A frase acima, atribuída ao grande economista Pedro Malan, ex-ministro da Fazenda, é bastante conhecida entre os macroeconomistas brasileiros e diz respeito ao fato de que, vez por outra, as séries macroeconômicas são revisadas, afetando números do passado. Uma parte da explicação é inerente ao próprio processo de mensuração das variáveis econômicas. No caso do PIB, por exemplo, assim como em outros países, a própria metodologia da pesquisa prevê a revisão do número preliminar depois de um certo período de tempo, antes da divulgação definitiva. A outra parte da explicação, contudo, tem a ver com o fato de que, algumas vezes, essas mudanças ocorrem de maneira não esperada, sem muita fundamentação técnica e, portanto, acompanhadas de críticas do grande público.

Se as críticas são justas ou não, trata-se de outra história, que não precisamos explorar aqui. A recomendação é, portanto, sempre utilizar as séries mais atualizadas de cada pesquisa, prevenindo-se contra mudanças marginais que podem continuar ocorrendo em dados do passado, mesmo que sejam no futuro. Em português claro: nunca incorpore apenas o último número de uma série às suas planilhas, sempre copie a série inteira a cada nova divulgação.

Por último, vale um alerta: embora as séries de produção sem ajuste sazonal tenham uma estabilidade muito maior, as séries de produção com ajuste sazonal normalmente mudam a cada novo dado na série histórica. Isso é normal e inerente ao próprio processo de ajuste sazonal. Quanto a esses casos, a recomendação vira uma ordem: a cada nova divulgação, sempre trate a série como se fosse nova, baixando os dados desde o início.

Inflação baixa não significa queda de preços e inflação não é uniforme

Na seção 2 foi possível observar que a inflação impacta o custo de vida, o que é rapidamente percebido pela população. Apesar disso, é importante ter em mente que a inflação não é uniforme a todos os bens da economia.

A inflação é calculada para uma dada cesta de consumo. A inflação de cestas diferentes (consumidores diferentes), portanto, pode evoluir de maneira distinta. Por exemplo, se o preço de bens essenciais sobe mais (menos) do que a média, é possível que a inflação percebida pela população mais pobre fique maior (menor) do que para a população mais rica.

Outro erro comum diz respeito à interpretação da inflação. Como inflação consiste em aumento de preços, uma inflação baixa não significa uma redução de preços, mas sim um pequeno aumento dos preços. Esse exemplo não foi escolhido por acaso: em períodos recentes, de baixa inflação, foi comum ouvir consumidores na mídia afirmando que não percebiam os efeitos benignos da baixa inflação no dia a dia, “uma vez que os preços continuavam subindo”.

Valores absolutos e per capita

Já falamos sobre isso na seção sobre produção, mas é importante reforçar, por se tratar de um erro comum no debate econômico brasileiro. O PIB é a soma de toda a produção de um país, mas depende do tamanho da população de cada um. Para comparações internacionais, portanto, é importante ter em mente o conceito de PIB per capita, que divide o PIB do país pela população. Esse indicador vai ser muito mais relacionado com a riqueza e com o nível de vida de um país do que simplesmente o PIB total.

A Tabela 3, na página seguinte, apresenta os 10 países com maiores PIB e PIB per capita do mundo, com dados do FMI, referentes a 2018. Podemos ver que a ordem dos países é bem diferente. O Brasil, que ocupava a nona posição dentre as maiores economias do mundo, ficava na apenas 78ª posição na lista dos maiores PIB per capita. Dessa maneira, fica claro que não conseguiremos melhorar o bem-estar da população somente distribuindo melhor a renda: precisamos também de crescer (e muito!).

Tabela 3 – Lista dos 10 maiores PIB e PIB per capita do mundo

Posição	PIB	PIB per capita
1	Estados Unidos	Luxemburgo
2	China	Suíça
3	Japão	Macau
4	Alemanha	Noruega
5	Reino Unido	Irlanda
6	França	Islândia
7	Índia	Catar
8	Itália	Singapura
9	Brasil	Estados Unidos
10	Canadá	Dinamarca

Fonte: FMI

Dados mensais tem muito ruído: parcimônia na interpretação!

O analista de conjuntura econômica deve ter em mente que o quadro macroeconômico tem uma grande inércia. As tendências, na maioria das vezes, mudam lentamente. E quando começam a mudar, isto acaba se refletindo em uma série de indicadores, o que nos permite identificar melhor a mudança de tendência.

Manchetes de jornais, muitas vezes, adotam um tom sensacionalista para estampar a divulgação de um dado. Contudo, às vezes, um indicador sobe mais justamente porque caiu mais anteriormente. Alguns economistas brincam que a análise de conjuntura é uma análise de elevador, um sobe e desce sem fim. Brincadeiras à parte, há um fundo de verdade e, nesse caso, é bom ter parcimônia, principalmente em relação a dados mensais.

Fora isso, algumas perguntas ajudam a nortear o que é mais importante. Por exemplo, o que deve ser mais comemorado: maior volume de produção ou maior taxa de crescimento? Ou, o que é melhor? Um valor ruim, mas acima do esperado, ou um valor bom, mas abaixo do esperado? Não existe resposta certa e cabe ao analista julgar com base nos objetivos de sua análise, mas a visão crítica sempre deve sobressair.

Cuidado com as regressões!

Como vimos, na análise de conjuntura macroeconômica são muitas as relações entre as variáveis: impacto das variações do câmbio na inflação; impactos da mudança da taxa de juros no consumo, impactos da atividade no mercado de trabalho e por aí vai.

A tentação para estabelecer correlações e, em seguida, causalidades, é grande. Contudo, muitas vezes as variáveis macroeconômicas são interdependentes (ou endógenas, no jargão da profissão). Além disso, muitas relações podem mudar ao longo do tempo, reagindo a mudanças de políticas econômicas, aparatos regulatórios, dentre outros. Nesse caso, convém adotar bastante cautela na hora de estimar regressões e estabelecer relações entre variáveis.

Análise de equilíbrio parcial ou equilíbrio geral

Lembram-se do efeito-borboleta, aquele que diz que a atividade das borboletas na Ásia pode desencadear efeitos aqui no Brasil?

Pois é, em economia também temos algo do tipo: mudanças em algumas variáveis podem influenciar todas as outras. Para analisar o sistema como um todo, portanto, economistas devem pensar em abordagens de equilíbrio geral. O problema é que nem sempre isso acontece e muitas vezes caímos na análise de equilíbrio parcial, o que pode ocasionar erros importantes.

Tecnicamente, análises de equilíbrio parcial ou geral são ferramentas da microeconomia, que envolvem a otimização de funções com o objetivo de determinar quantidades e preços de equilíbrio. Para o objetivo deste capítulo, contudo, podemos pensar de forma mais prática e intuitiva.

Vamos pensar, por exemplo, em uma política de subsídios. Suponha que o governo conceda subsídios para uma fábrica de automóveis se instalar em uma cidade específica. Aos olhos de muita gente, pode parecer uma decisão ótima (e sem custos): uma fábrica será construída, empregos serão gerados e carros serão produzidos.

O que não estamos vendo, contudo, são os demais efeitos que essa medida acarreta: para o pagamento dos subsídios, o dinheiro precisar vir de algum lugar, o que pode prejudicar outros setores e atividades, que precisarão pagar maior volume de impostos.

Outro efeito estaria relacionado à mão de obra: de repente, a atração da mão de obra pela nova fábrica pode prejudicar a oferta de mão de obra para outros setores da região, impedindo-os de florescer. Pode também aumentar o preço dos insumos da região, incluindo salário, afetando novamente outros setores.

Com esse exemplo, não queremos dizer que o subsídio não deva ser concedido nem que a fábrica não deva ser construída. Queremos apenas dizer que análises desse tipo devem ser realizadas preferencialmente com análises de equilíbrio geral, ao invés de análises de equilíbrio parcial, de maneira a abordar efeitos em todo o ambiente macroeconômico.

Perguntas frequentes

O Brasil é um país rico? Como crescer mais?

Como vimos acima, enquanto o PIB brasileiro está entre os 10 maiores do mundo, seu PIB per capita se encontra em posições bem menos nobres. Assim, não se pode afirmar que o país é rico: infelizmente, somos, na melhor das hipóteses, um país de renda média, que possui imensos problemas sociais, englobando áreas básicas como o saneamento, onde temos uma precária cobertura de água e esgoto. Para resolver muitos de nossos problemas, não há como fugir do óbvio: precisamos elevar o crescimento econômico.

Assim, chega-se a uma nova pergunta: como aumentar o crescimento? Esta é uma questão que mobiliza os economistas há bastante tempo. Uma resposta completa poderia ocupar um único capítulo (ou mesmo um livro!), mas, para não decepcionar o leitor, recorreremos à já mencionada função de produção.

Por essa abordagem, o produto depende fundamentalmente de seus inputs, dos quais se destacam a força de trabalho (quantidade e qualidade), o estoque de capital e a produtividade total dos fatores. Aumentar o crescimento passa, portanto, por melhorar a educação, aumentar e qualificar o investimento e por todas as medidas em prol da produtividade, que englobam desde promover maior abertura comercial, até melhorar o regime tributário, promover a competitividade das empresas, melhorar a regulação, dentre outras.

O assunto é complexo e, para mais detalhes, sugerimos os livros de Marcos Mendes, “Por que o Brasil cresce pouco?”, de 2014, e de Fabio Giambiagi e Mansueto Almeida, “A retomada do crescimento”, de 2017.

O que é produto potencial?

O produto potencial é um conceito bastante utilizado por economistas e se refere à capacidade de produção efetiva de uma economia compatível com o equilíbrio macroeconômico. Trata-se de uma variável não observável, que leva em conta o estoque de capital existente, a quantidade e qualificação dos trabalhadores e a produtividade da economia (novamente, a função de produção!).

Apesar das dificuldades de mensuração, o produto potencial é um conceito muito

importante para a análise de conjuntura. É, por exemplo, uma ferramenta utilizada por bancos centrais para a condução da política monetária, pois se sabe que a inflação depende do nível de ociosidade da economia no curto prazo, que é dada justamente pelo produto potencial (ou, em outro jargão dos economistas, pelo hiato do produto, que é a diferença entre o produto efetivo e o produto potencial).

Assim, uma economia próxima ou acima do produto potencial pode gerar pressões inflacionárias, enquanto uma economia abaixo do potencial pode ser estimulada via política monetária sem muitos riscos para a inflação.

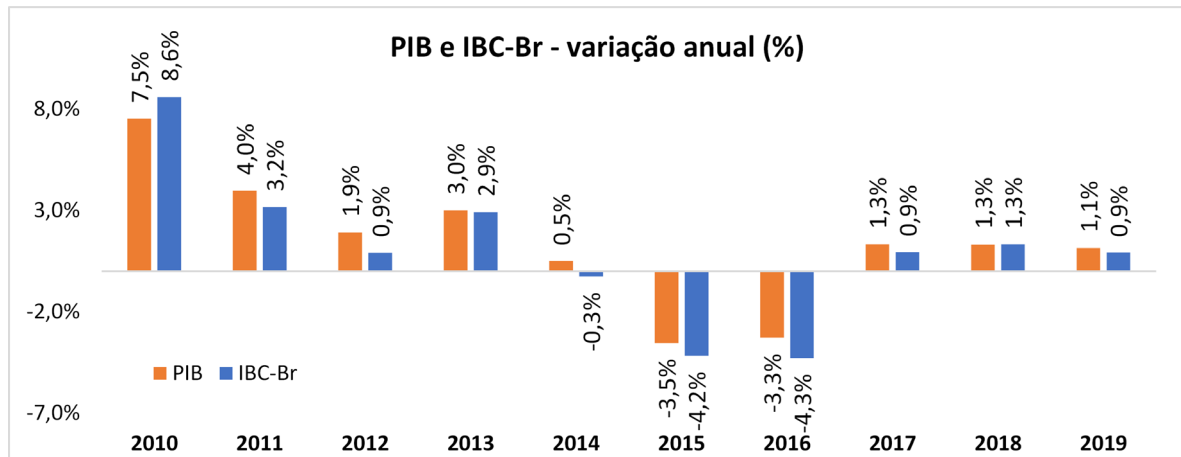
Como os cálculos são complicados e dependem de uma série de premissas, não existe uma fonte única e nem uma fonte oficial que disponibilize essa variável. Contudo, algumas casas disponibilizam estimativas, destacando-se, entre elas, o Ipea (Instituto de Pesquisa Econômica Aplicada), a IFI (Instituição Fiscal Independente), o Ibre-FGV (Instituto Brasileiro de Economia, da Fundação Getúlio Vargas) e a LCA Consultoria.

O IBC- Br é uma boa medida para o PIB?

A medida de atividade mais importante e abrangente, como visto anteriormente, é o PIB, mas ele é divulgado com uma grande defasagem. Viu-se também que o Banco Central divulga o IBC-Br, um indicador mensal de atividade que teria o papel de antecipar o PIB. Uma pergunta frequente, portanto, é: o IBC-Br é um bom previsor, ou uma boa medida para o PIB?

Pode-se responder essa pergunta simplesmente olhando para as séries históricas de ambas as estatísticas. Começamos pela variação anual dos indicadores, exposta no Gráfico 1. Nos últimos 10 anos, isto é, de 2010 a 2019, elas tiveram o mesmo sinal em 9 anos (somente em 2014 eles divergiram: o PIB cresceu 0,5%, enquanto o IBC-Br recuou 0,3%). No resto da série, observamos anos com valores bem próximos (como 2013, 2018 e 2019), ao mesmo tempo em que outros exibem uma diferença maior (como 2010, 2012 e 2016).

Gráfico 1 – PIB e IBC-Br – 2010 a 2019



Fonte: IBGE e BCB

Para dados trimestrais, o comportamento segue mais ou menos similar. Nos últimos 28 trimestres, do primeiro de 2013 ao último de 2019, tivemos sete deles onde o sinal da taxa de crescimento divergiu entre as duas pesquisas. Para os outros, observamos trimestres onde os números foram bem similares, enquanto em outros a divergência foi mais notável.

A conclusão é, portanto, que, embora na maioria das vezes os números do PIB e do IBC-Br caminhem juntos, há alguma discrepância entre seus valores. O IBC-Br seria um indicador importante, mas não infalível, sobre o comportamento do PIB. Usem, portanto, com moderação!

Para onde vai o câmbio?

Um economista é frequentemente abordado por amigos e familiares com a pergunta de um milhão de dólares: será que está na hora de comprar (ou vender) dólares? Qual vai ser a cotação da moeda estrangeira no final do ano?

Infelizmente, é uma pergunta muito difícil, para a qual não temos boas respostas. Economistas tentam criar modelos para estimar o câmbio há muitas décadas. Temos, por exemplo, modelos teóricos, como o da paridade de juros, e modelos econométricos, que estabelecem o câmbio como função de variáveis como diferencial de juros, preços de commodities e risco-país.

O fato, contudo, é que a profissão não teve muito sucesso com isso, tanto é que, muitas vezes, a melhor previsão para o câmbio no futuro é justamente o valor do

câmbio no presente.

Dessa maneira, você, economista, não se sinta pressionado com a pergunta do seu tio sobre palpites para a taxa de câmbio. Poderá responder de dois jeitos: ou explicando que, assim como o preço de ações, a dificuldade de previsão é imensa ou simplesmente dizendo que, caso você fosse capaz de prever, já estaria rico há muito tempo. Com qualquer uma das duas, pelo menos, você se livra do problema (e também do risco de uma previsão errada!).

Considerações finais

Este capítulo se dedicou à macroeconomia, definida aqui como análise de conjuntura econômica. Tratamos, portanto, de um tema que afeta o cotidiano de todos os cidadãos do país, mesmo que nem todos consigam perceber.

Nos últimos anos, tem-se observado um aumento no interesse pela macroeconomia na nossa sociedade, explicado, talvez, pelos anos de baixo crescimento econômico, pelo acirramento das disputas políticas e pela intensificação do debate via redes sociais.

Nesse sentido, o objetivo do capítulo foi o de fornecer ferramentas e informações para o aperfeiçoamento desta discussão econômica, que será complementada por muitos capítulos deste livro. Espero que tenhamos alcançado os objetivos propostos. Esperamos, também, que o país volte a crescer mais rapidamente nos próximos anos: analisar conjuntura em períodos de crescimento é bem mais fácil e agradável.

Referências bibliográficas e leituras adicionais

- BLANCHARD, Olivier. Macroeconomia. Ed. Pearson, 7ª edição, 2017.
- GIAMBIAGI, Fabio; ALMEIDA, Mansueto. A retomada do crescimento: diagnóstico e propostas. Elsevier, 2010.
- FRANCO, Gustavo. Cartas a um jovem economista. Elsevier, 2010.
- MENDES, Marcos. Por que o Brasil cresce pouco? Elsevier, 2014.
- SCHMIDT, Cristiane; GIAMBIAGI, Fabio. Macroeconomia para executivos. Campus, 2014.
- LIMA, Eduardo; ARAÚJO, Fabio; SILVA, José. Previsão e modelos macroeconômicos no Banco Central do Brasil. In: Dez anos de metas de inflação no Brasil, 1999-2009. Banco Central, Brasília, 2011.

GAP

MERCADO DE TRABALHO¹

Daniel Duque

07

¹ Agradeço encarecidamente a ajuda do meu ex estagiário e atual colega de trabalho Bernardo Esteves, sem o qual não teria sido em absoluto possível terminar este trabalho (pelo menos a tempo).

Introdução

O mercado de trabalho é uma das áreas na economia das mais influentes sobre o bem-estar dos indivíduos no seu dia a dia. Afinal, poucas coisas são mais importantes para as pessoas do que seus empregos (ou a falta destes), com seus benefícios – salários e previdência – e custos – tempo e intensidade no trabalho – associados.

Isso não impede, no entanto, que as pessoas tenham diversas dúvidas e cometam muitos enganos sobre esse tema. Conceitos misturados, relações duvidosas e interpretações de tendências equivocadas são comuns no meio de quase tudo que é dito e escrito sobre o tema.

Principais indicadores

a) Taxa de desemprego

A taxa de desemprego é, possivelmente, o indicador mais importante sobre o mercado de trabalho. É, pelo menos, o mais utilizado, junto com a renda média. A variável é muito importante para medir tanto ociosidade da economia, quanto bem-estar da população. Sua definição é expressa pela fórmula abaixo:

$$\text{Pessoas procurando ocupação em um período}$$

$$(\text{Pessoas procurando ocupação em um período} + \text{Pessoas ocupadas no mesmo período})$$

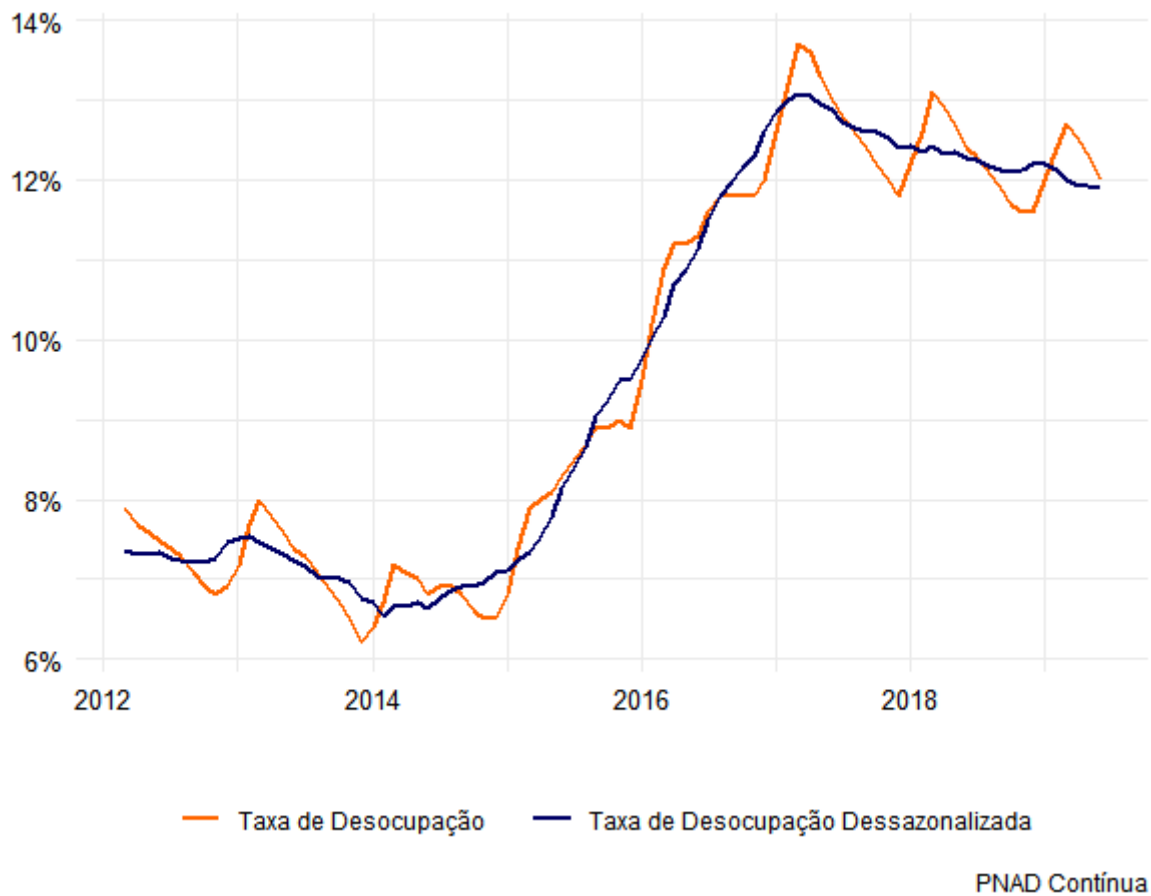
O denominador da equação (a soma entre pessoas procurando ocupação e pessoas ocupadas) é também chamado de força de trabalho (conforme o IBGE) ou população economicamente ativa. Ou seja, o indicador mede quantos dos que estão querendo trabalhar e fazendo algo sobre isso não obtiveram sucesso.

É importante notar que, apesar de comumente chamada de “taxa de desemprego”, há uma imprecisão no termo. Em inglês, unemployment rate seria mais corretamente traduzido como taxa de desocupação, e é dessa forma que hoje o IBGE reporta esse indicador – ainda que popularmente ainda se use o primeiro. Isso porque há diferenças entre os conceitos de ocupação e “emprego”.

Quando uma pessoa está ocupada, significa que ela está utilizando o seu trabalho em uma atividade produtiva por pelo menos algumas horas. Note que essa definição de ocupação contempla muitas atividades, mesmo aquelas em que o trabalhador não tem chefe: ele é o chamado ocupado por conta própria, uma categoria que abarca hoje cerca de ¼ de todo mercado de trabalho brasileiro. Já quem está na categoria empregado precisa, necessariamente, ter um empregador, normalmente uma empresa.

A taxa de desocupação, como quase todas as variáveis econômicas no Brasil, tem um alto grau de sazonalidade. Isso porque, no último trimestre de todos os anos, aumenta o número de contratações para as festas de final de ano, de modo que sempre se encontra uma taxa menor de desocupados nesses meses. Passadas estas, a taxa de desocupação aumenta rapidamente no primeiro trimestre do ano seguinte. O Gráfico 1 mostra essa relação.

Gráfico 1: Taxa de Desocupação



A taxa de desocupação é importante porque, além de ser um componente relevante para o bem-estar da população, também faz parte das estimativas do grau de ociosidade da economia, ou seja, o quanto não se está usando fatores de produção já disponíveis. A intuição é simples: se a taxa estiver muito alta, significa que muitas pessoas que estão disponíveis para trabalhar não o estão fazendo. Se, ao contrário, a economia estiver em pleno emprego, caso o nível de atividade continuar acelerando, espera-se também que os preços comecem a subir mais aceleradamente.

Uma economia no chamado pleno emprego, no entanto, não significa que a taxa de desocupação estará em 0%, ou algo muito próximo disso. Na verdade, costuma-se considerar uma economia em pleno emprego quando a taxa está entre 3 e 5%, a depender das características do mercado de trabalho do país. Isso ocorre porque, quando você mede o desemprego em determinado período, você está captando tanto pessoas que estão, de fato, sendo mal-sucedidas em encontrar uma ocupação, quanto aquelas que estão apenas em um período curto de busca, entre uma ocupação e outra. Esse é o chamado desemprego friccional.

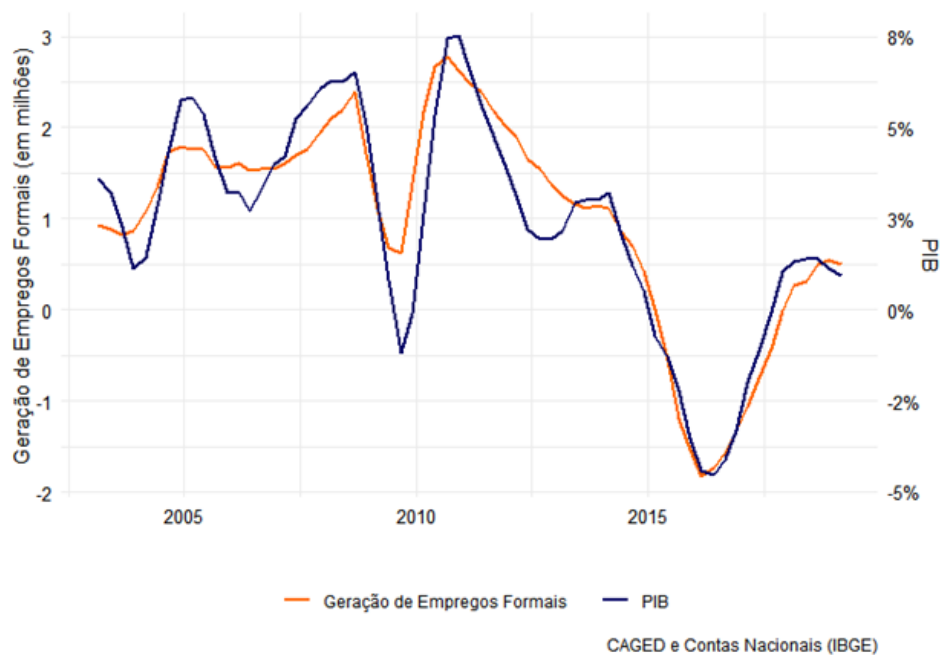
Há ainda um conceito chamado taxa natural de desemprego. No entanto, essa parte será coberta no capítulo de Macroeconomia do livro.

b) Geração de empregos formais

A geração de empregos formais é, diferentemente da taxa de desemprego, uma variável de fluxo: ela denota quantos novos empregos formais foram criados, considerando o saldo das admissões e demissões. Ela pode ser medida tanto pela PNAD (Pesquisa Nacional por Amostra de Domicílios) Contínua, quanto pelo Cadastro Geral de Empregados e Desempregados (Caged). Nesse último caso, os novos empregos são registrados pelas empresas legais, tendo assim, portanto, caráter censitário. Observe que, quando falamos de taxa de desemprego, estamos falando do estoque, o total de trabalhadores, não da geração de vagas, e perceba também que ela abrange não só os formais, mas também os informais

A geração de empregos formais tem grande correlação com a variação do PIB, como mostra o Gráfico 2. Isso significa que o setor formal é um forte respondente da atividade econômica e muitas vezes é usado como um preditor dessa medida.

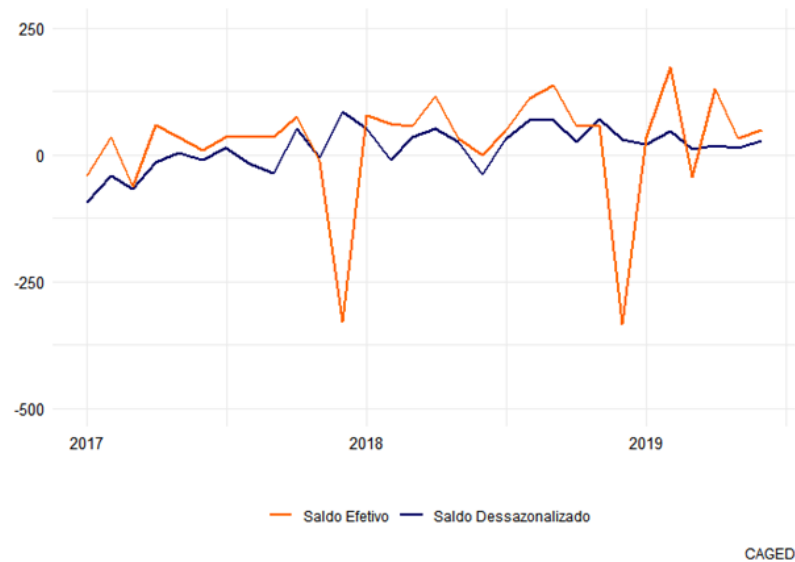
Gráfico 2: Geração de Empregos Formais e PIB: Acumulado em 4 Trimestres



Tal como a taxa de desemprego, a geração de empregos formais tem bastante sazonalidade, mas isso também depende da base de dados utilizada. O Gráfico 3 mostra o

comportamento médio dessa variável ao longo de cada mês, medida pelo Caged. Como se vê, há sempre uma grande queda em dezembro. Isso ocorre porque as empresas registram muitas de suas demissões apenas no final do ano, por razões contábeis-legais.

Gráfico 3: CAGED: Salário Menal (em milhares)



c) Taxa de formalidade

A taxa de formalidade mede quanto do mercado de trabalho de um país é dominado pelo setor formal. Seu cálculo é feito pela seguinte relação:

$$\frac{\text{Trabalhadores do setor formal}}{(\text{Trabalhadores do setor formal} + \text{Trabalhadores do setor informal})}$$

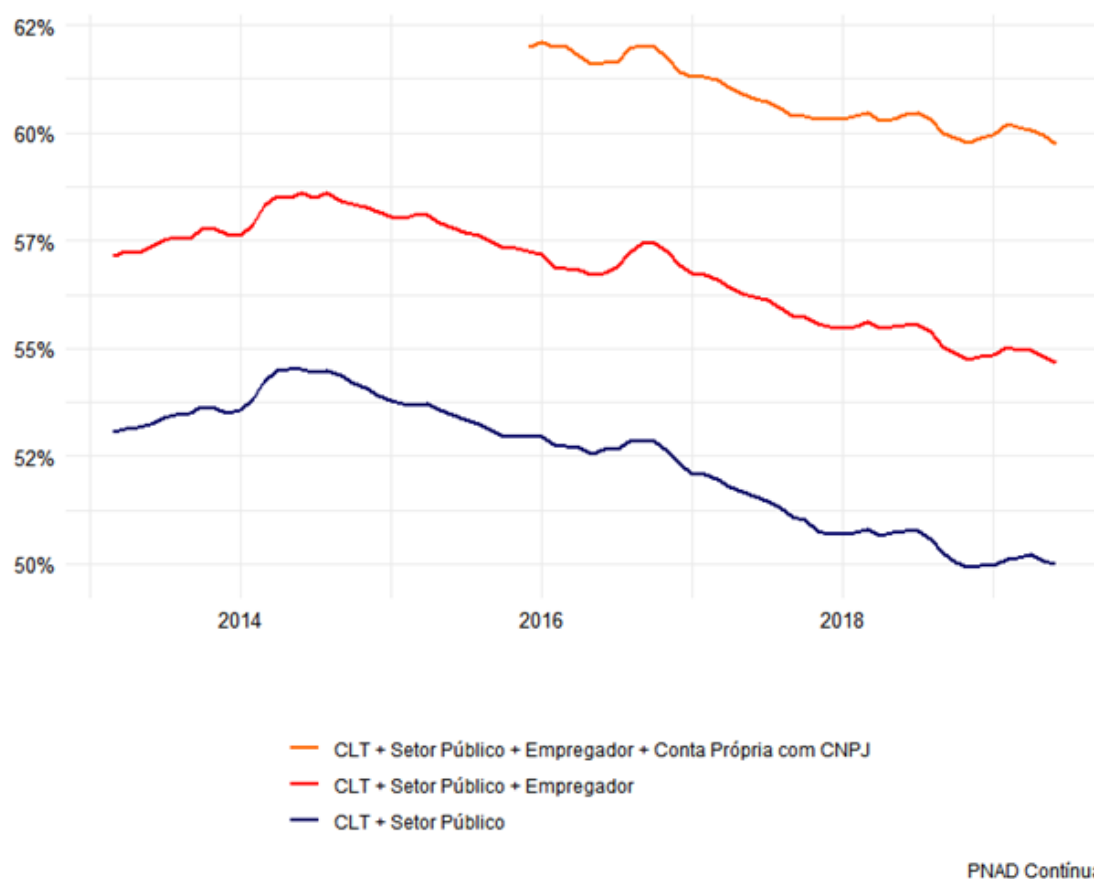
O problema, no entanto, está muitas vezes em definir o que é formalidade, e não há consenso ainda sobre esse conceito, principalmente nos tempos atuais. Temos por certo, é claro, que empregados com carteira assinada e funcionários públicos deveriam se enquadrar nessa categoria, mas será que deveria parar por aí? Empregadores supostamente deveriam entrar, mas e se seus empregados não tiverem qualquer relação legal de trabalho? E trabalhadores que não têm carteira assinada, mas se registram como microempreendedores individuais (MEIs), tão presentes hoje no mercado de trabalho?

Tal definição, portanto, de fato não é clara, e, a depender de qual se utiliza, te-

mos diferentes trajetórias da taxa de formalidade no Brasil. Na verdade, até mesmo diferentes possibilidades de análise: isso porque o IBGE acrescentou apenas no final de 2015 à sua pesquisa de mercado de trabalho, a PNAD Contínua, a pergunta de se a pessoa teria CNPJ.

O Gráfico 4 mostra a taxa de formalidade desde 2012 no país, seguindo diferentes definições. Como se vê, apesar de todos apresentarem tendências parecidas, seus níveis mudam significativamente a depender de quem é considerado formal. A definição mais ampla, que inclui ocupados por conta própria com CNPJ, também diverge recentemente na tendência, mostrando alguma estabilização, devido ao grande aumento dessa população ocupada.

Gráfico 4: Taxa de Formalidade por Diferentes Definições



d) Taxa de participação

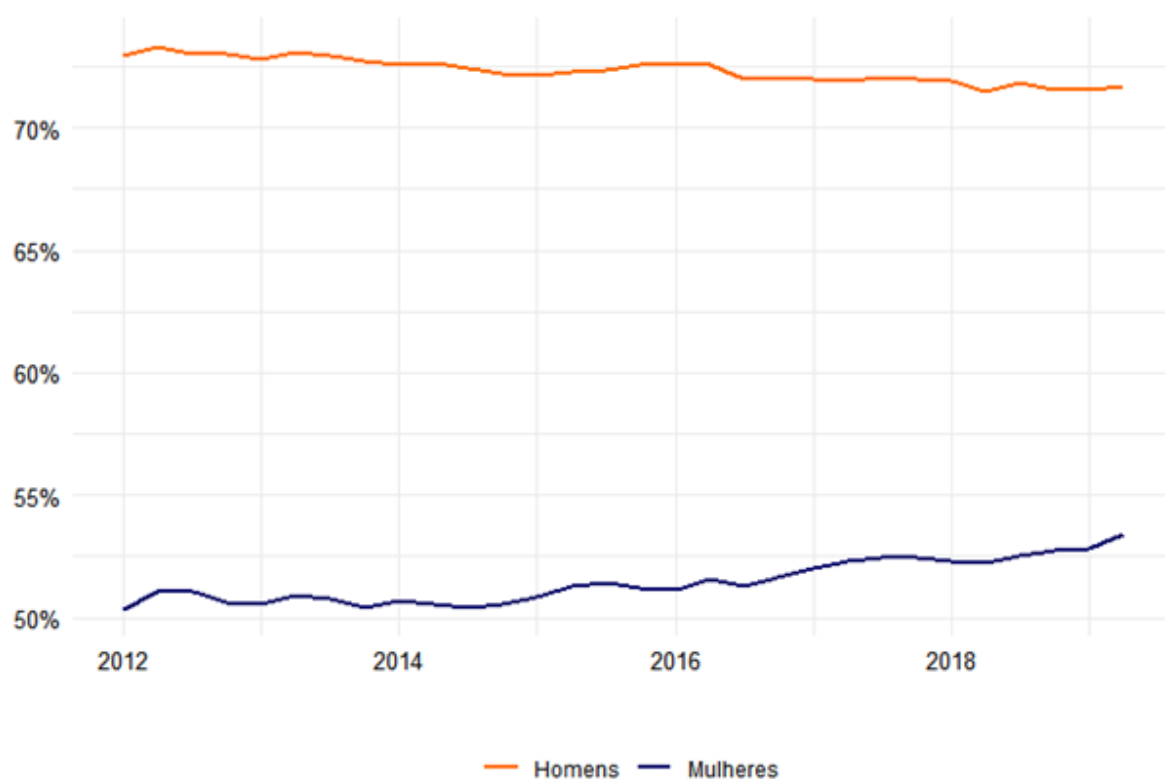
A taxa de participação significa quantos daqueles em idade para trabalhar estão, efetivamente, trabalhando ou procurando uma ocupação. Seu cálculo segue a fórmula abaixo:

(Pessoas procurando ocupação + Pessoas ocupadas)
(Pessoas em idade para trabalhar)

O entendimento do que seria idade para trabalhar varia em cada sociedade, mas, normalmente, se assume 15 a 64 anos como a faixa para tal – ainda que, no Brasil, a PNAD Contínua considere entre estes todos aqueles com pelo menos 14 anos de idade. Normalmente, se interpreta esse indicador como “oferta de trabalho”, ou algo como a propensão média da população adulta a trabalhar.

Em muitos trabalhos econométricos, busca-se entender o que leva as pessoas a ofertar ou não sua mão de obra. Sabe-se, por exemplo, como mostra o Gráfico 5, que mulheres ofertam menos trabalho do que os homens, e que há um ciclo de vida da oferta de trabalho. Portanto, pode-se esperar, em um país com muitos jovens, uma baixa taxa de participação que aumentará conforme aumenta a proporção de adultos – gerando, assim, o que se chama de bônus demográfico. Com seu envelhecimento, no entanto, a taxa de participação se reduz novamente.

Gráfico 5: Taxa de Participação por Sexo



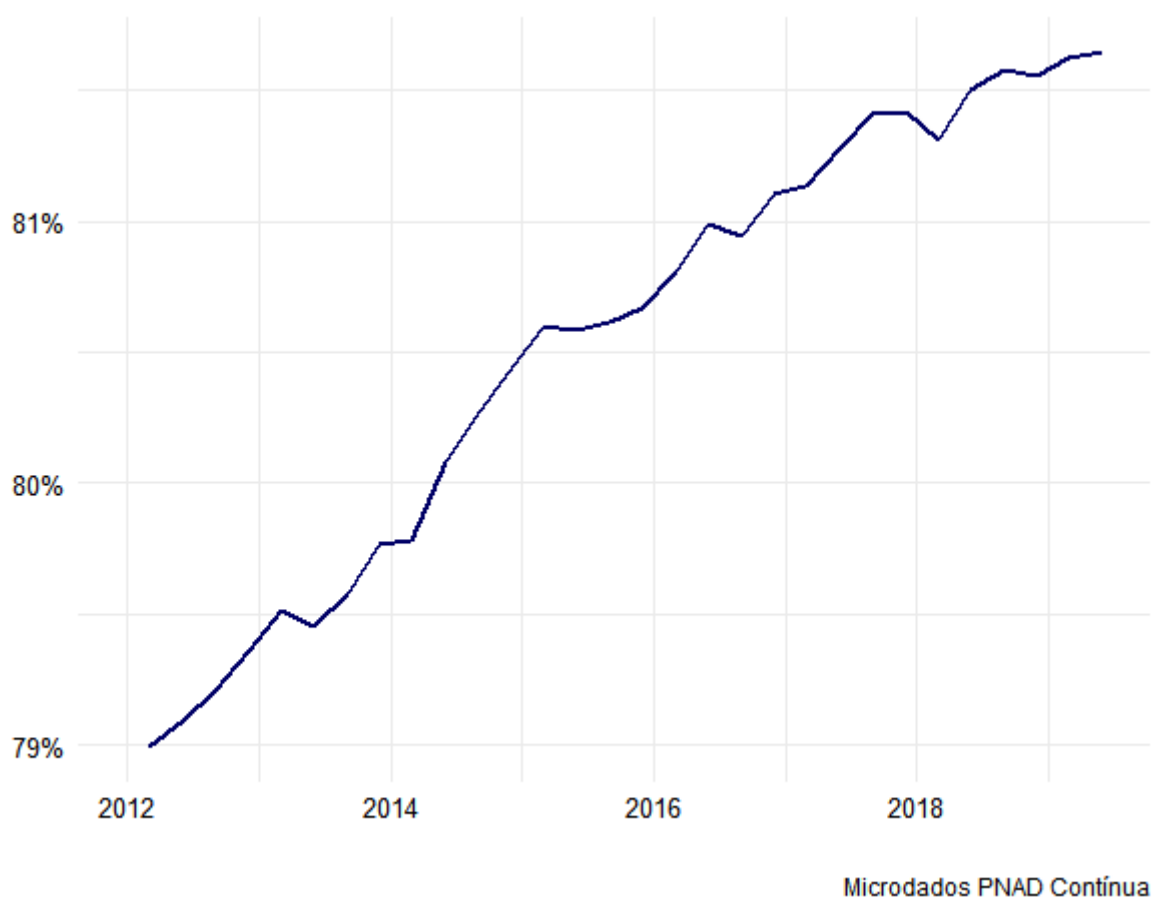
e) *Taxa de atividade*

A taxa de atividade é uma variável mais diretamente ligada à situação demográfica de um país: ela informa a proporção da população que está em idade para trabalhar. Seu cálculo segue a fórmula abaixo:

$$\frac{\text{Pessoas em idade para trabalhar}}{\text{População total}}$$

Tem-se o bônus demográfico quando esse indicador cresce: significa que a população em idade para trabalhar está crescendo mais do que a população total. Isso significa uma oportunidade de aumentar a renda per capita via demografia, que foi o que ocorreu no Brasil em anos recentes, como mostra o Gráfico 6.

Gráfico 6: PIA em relação à População Total



f) *Desalento*

O desalento é um tipo de categoria daqueles adultos (ou, no caso da PNAD Contínua, acima de 14 anos) que estão fora da força de trabalho. Ou seja, não estavam ocupados ou procurando ocupação quando da coleta do dado. Um desalentado é uma pessoa que gostaria de estar trabalhando, mas, por alguma razão, não foi em busca de se ocupar.

As razões para não se procurar trabalho, mesmo desejando, podem ser diversas: desde estar fisicamente incapacitado a estar desencorajado a procurar emprego – o caso dos desalentados, que, na Organização Internacional do Trabalho, chamam-se *discouraged workers*. Um adulto pode crer que não encontrará trabalho, ou mesmo que não deve, por razões culturais ou religiosas, procurar trabalho – como as mulheres que ficam em casa para cuidar da família, apesar do desejo de trabalhar que porventura tivessem.

Tal como no desemprego, é preciso ter um período de referência. No caso do IBGE, os desalentados devem informar terem tido desejo de trabalhar nos últimos 30 dias, sem que tenham feito nada para tal.

g) *Subutilização da força de trabalho*

A subutilização da força de trabalho é uma medida alternativa à taxa de desemprego para medir o grau de ociosidade da economia e o bem-estar de uma população. Sua definição é menos direta, pois inclui diferentes categorias, mas sua interpretação é mais simples: ela busca mensurar quanto da população poderia estar trabalhando mais (incluindo aqueles que não trabalham), caso houvesse demanda. Seu cálculo segue a fórmula abaixo:

$$\frac{(\text{Pessoas procurando ocupação em um período} + \text{Pessoas trabalhando menos de 40 horas por semana que gostariam e estariam disponíveis para trabalhar mais} + \text{Pessoas desalentadas})}{(\text{Pessoas procurando ocupação em um período} + \text{Pessoas ocupadas no mesmo período} + \text{Pessoas desalentadas no mesmo período})}$$

A fórmula acima auxilia a interpretar o significado desse indicador: ele inclui pes-

soas desocupadas, as que querem trabalhar mais horas (e não chegam a trabalhar 40 horas semanais) e as que querem trabalhar, mas não procuraram trabalho. Esse é o total de pessoas que podem usar mais da sua própria força de trabalho do que de fato estão fazendo no momento.

A razão da existência de tal indicador é simples: um choque de demanda pode fazer com que as empresas tenham que, de repente, produzir muito mais rapidamente. Elas têm duas opções: contratar mais trabalhadores, ou aumentar a carga horária daqueles que já estão empregados. Ela só poderá aderir à segunda opção se lá houver quem trabalhe menos do que 40 horas e queira trabalhar mais. Caso isso ocorra em nível nacional, é possível, portanto, que haja um aumento da demanda das empresas por trabalho, mas que a taxa de desocupação não caia.

h) Renda média e mediana

A renda média é uma das variáveis mais utilizadas para avaliar o bem-estar de uma população. Normalmente, o rendimento do trabalho médio, também chamado de salário médio (ainda que salário seja mais diretamente associado a emprego, o que, como vimos, é diferente de ocupação), mostra quanto de poder de compra médio os ocupados conseguem obter no mercado de trabalho.

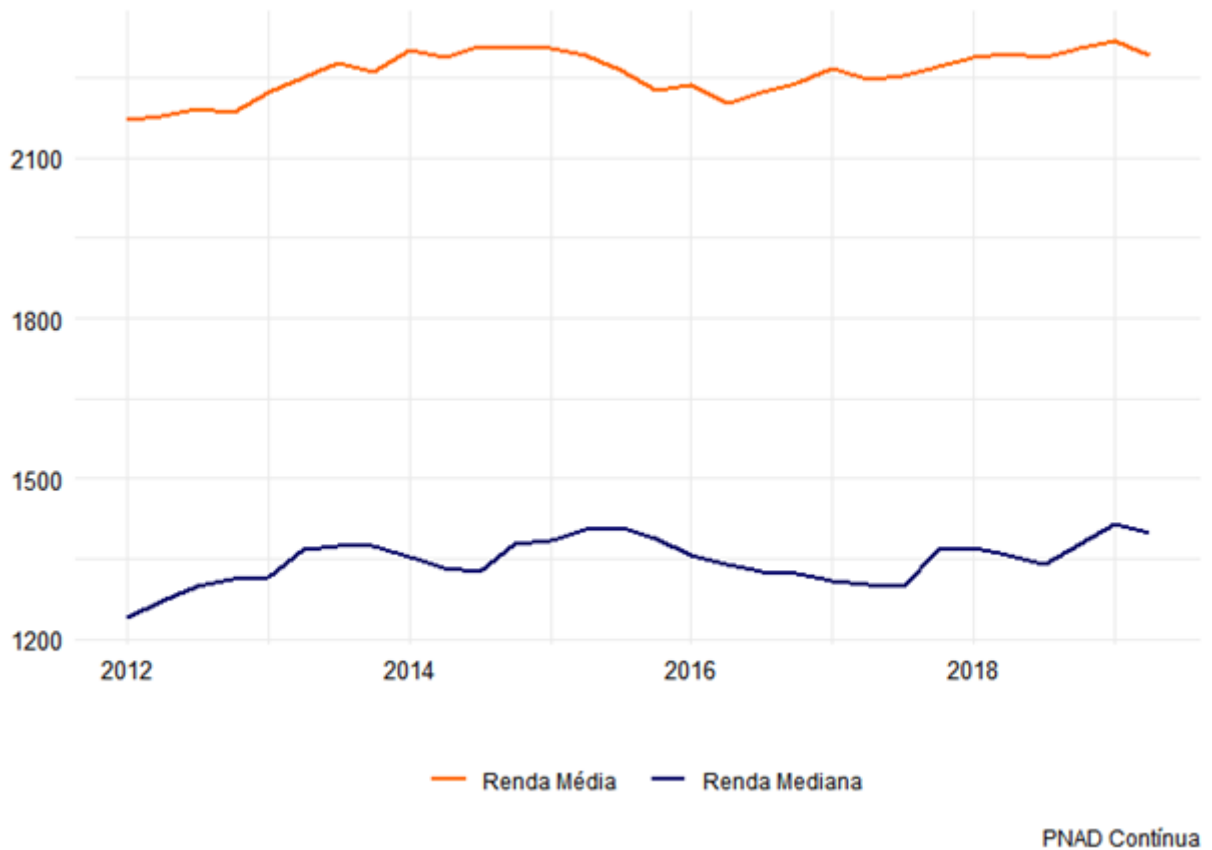
A fórmula para o cálculo da renda média do trabalho é a mesma de qualquer média: a soma de todos os rendimentos do trabalho em um período, dividida por todos ocupados daquele período. A renda média, portanto, multiplicada pelo número de ocupados mostrará o que se chama de massa de rendimentos do trabalho – isso, se forem excluídos aqueles que estão ocupados, mas não declaram renda.

A renda mediana, por outro lado, mostra qual era o rendimento do trabalhador exatamente no meio da distribuição, que está entre a metade mais pobre e a metade mais rica do país. Por que se pensar em mediana, ao invés de média? Porque essa última pode ser viesada para cima por valores muito altos no extremo superior. Caso apenas os muito ricos tenham sua renda em crescimento, a média apresentará uma sensível elevação (uma vez que eles correspondem a grande parte do total dos rendimentos), mesmo que a imensa maioria da população não tenha ganho algum.

A renda mediana, por outro lado, pode ser uma alternativa para medir com mais fidelidade o “bem-estar geral” da população, por estar próxima de percentuais elevados da população (a não ser que a distribuição se separe entre um grupo de baixíssima

e outro de altíssima renda). O Gráfico 7 mostra, a partir da PNAD Contínua trimestral, a evolução da Média e da Mediana.

Gráfico 7: Renda Média e Renda Mediana (em Reais)



Como se vê no Gráfico 7, a renda mediana fica sistematicamente abaixo da média, que é muito puxada para cima pelos altíssimos rendimentos de poucos. É possível também ver que, enquanto a renda média voltou a crescer ainda em meados de 2016, a renda mediana só iria fazê-lo a partir da segunda metade de 2017.

Dados no Brasil

a) PNAD Contínua Mensal

A PNAD Contínua Mensal entrou em cena a partir de 2012, para ser uma substituta intermediária entre a PNAD (anual, com abrangência nacional) e a Pesquisa Mensal de Emprego PME (mensal, com abrangência de 6 regiões metropolitanas). A nova pesquisa tem apenas dados relativos ao mercado de trabalho – tal como a PME –, mas abrange todo o Brasil – tal como a PNAD.

A PNAD Contínua Mensal não tem microdados disponíveis, apenas agregados e médias nacionais tabulados pelo próprio IBGE (não há informações para estados ou Regiões). E, apesar de sair mês a mês, seu período de referência não é mensal, mas trimestral. Isso significa que a PNAD Contínua de abril será relativa aos meses de fevereiro, março e abril. Já a de agosto será relativa a junho, julho e agosto.

Ou seja, a PNAD Contínua Mensal apresenta dados por trimestres móveis, o que suaviza suas séries. Portanto, caso se busque uma média anual para algum indicador, não se deve usar todos os meses da pesquisa, mas seus quatro trimestres fechados, dos meses de março, junho, setembro e dezembro. Assim, evita-se dupla contagem de vários meses, além da inclusão de meses que estão fora do ano de referência (por exemplo a PNAD Contínua de janeiro de um ano seria relativa também a novembro e dezembro do ano anterior).

b) PNAD Contínua Trimestral

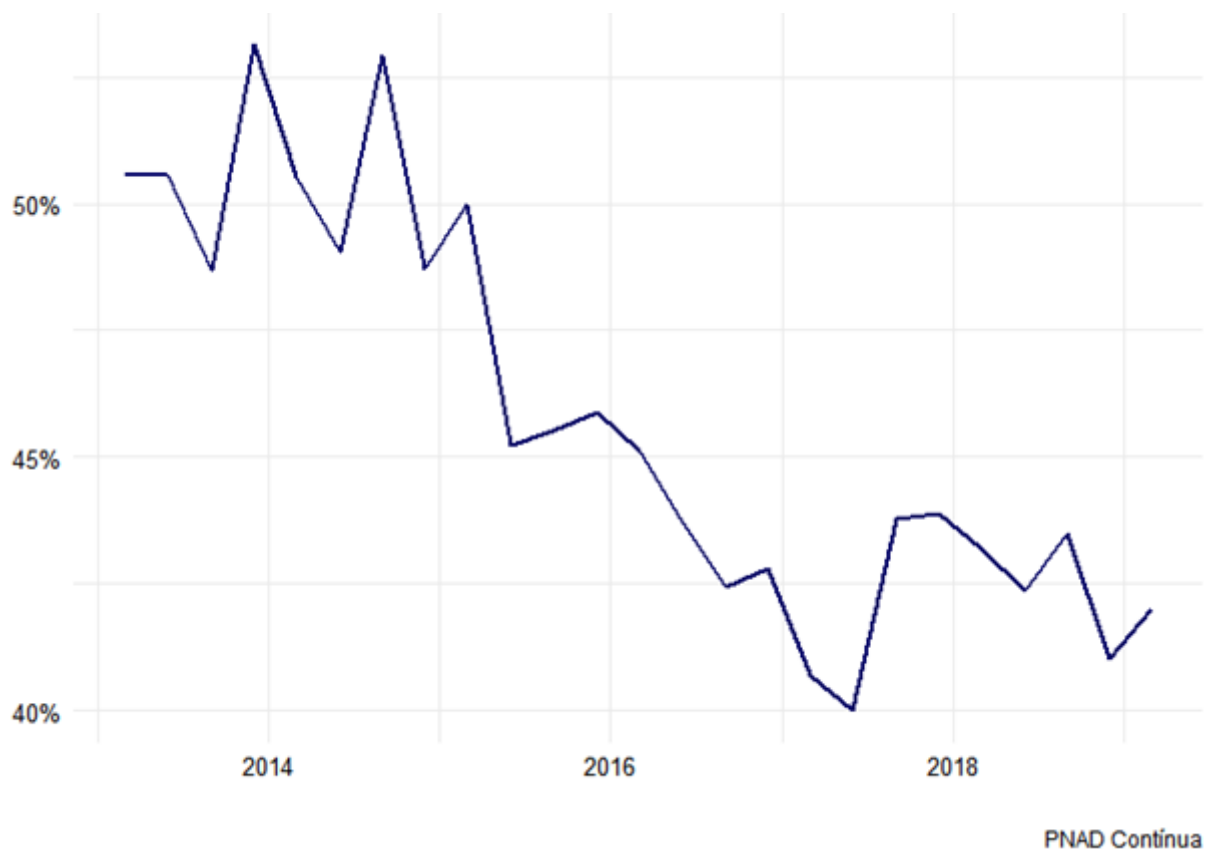
A PNAD Contínua Trimestral é divulgada sempre alguns dias após as pesquisas mensais de março, junho, setembro e dezembro. Nela, tem-se os microdados referentes a cada trimestre, com abrangência nacional, contendo informações de estados, capitais e regiões metropolitanas (além de Regiões Integradas de Desenvolvimento Econômico).

Há pouco mais de uma centena de variáveis nos microdados da PNAD Contínua Trimestral, com informações de família, educação e mercado de trabalho. Um dos fatores mais interessantes da base de dados, no entanto, é o acompanhamento fixo de 20% dos domicílios por cinco trimestres. Assim, temos informações longitudinais de um grande número de indivíduos.

Com essa estrutura, a PNAD Contínua Trimestral permite fazer estudos de mobilidade de renda, ocupações, classe, condições no mercado de trabalho etc., ainda que em um período curto de tempo. Identificando as pessoas por sua data de nascimento (disponível na base de dados), por exemplo, pode-se saber quantos daqueles que estavam desocupados em um trimestre tinham conseguido emprego um ano depois (e vice-versa). Uma das maiores limitações, porém, dessa abordagem é que se perdem aqueles indivíduos que se mudaram do domicílio no período – algo comum especialmente para os mais pobres.

O Gráfico 8 mostra o percentual de indivíduos que, desocupados em um trimestre, migraram para o grupo de ocupados 12 meses depois. Ou seja, é a probabilidade de uma pessoa desocupada acabar ocupada um ano depois. Como se vê, esse indicador estava em torno de 50% até 2015, quando cai rapidamente para cerca de 40% até 2017, recuperando dois a três pontos percentuais nos trimestres seguintes, ao longo de 2018 e 2019.

Gráfico 8: Transição da Condição de Desocupado para Ocupado



c) PNAD Contínua Anual

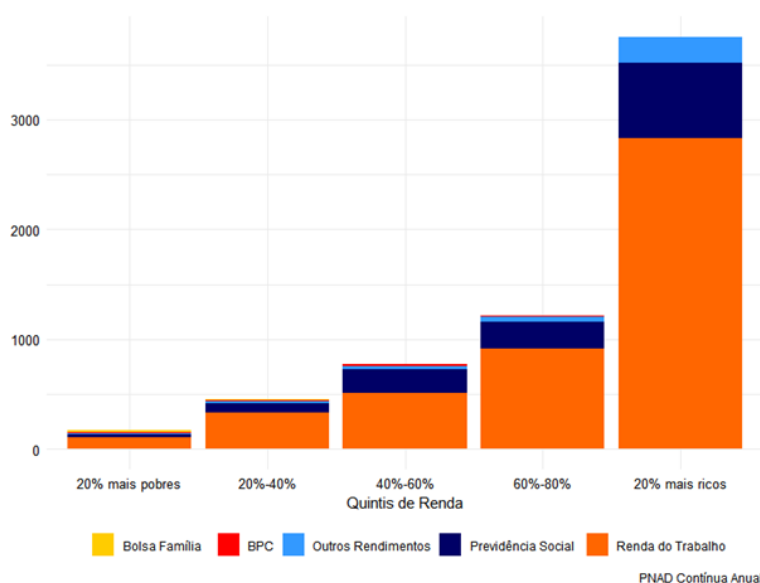
A PNAD Contínua Anual conta com estrutura parecida com a Trimestral, mas com algumas diferenças e adições. A primeira é divulgada em duas edições todo ano: a da primeira entrevista no domicílio, e a da quinta. Ou seja, em vez de se dividir em trimestres – há os quatro presentes na amostra –, a divisão se faz por número da entrevista, que pode ser a primeira ou a última do domicílio.

Assim, em 2016, temos as 1ª e 5ª entrevistas, tal como em 2017, 2018 e assim por diante. Caso juntemos a 1ª entrevista de 2016 e a 5ª de 2017, por exemplo, temos novamente dados longitudinais, mas com duas observações em cada período do tempo.

A grande vantagem dessa base em relação à Trimestral, no entanto, está nas suas informações de renda para além do trabalho. Nela, temos dados de renda do Benefício de Prestação Continuada (BPC), Bolsa Família (PBF), aposentadorias ou pensões e outros rendimentos, como demais programas sociais, aluguéis etc.

O Gráfico 9 mostra a renda domiciliar per capita do país em 2017 por fonte. Como se vê, a renda do trabalho é predominante para todos os quintis de rendimento. Ainda assim, os proventos previdenciários também correspondem a parte considerável dos ganhos da população. Programas sociais como o BPC e Bolsa Família adicionam pouco às famílias em termos absolutos (o PBF adiciona cerca de 10 reais per capita para os 20% mais pobres, enquanto o BPC provê aproximadamente 15 reais per capita para os que estão no segundo quintil), mas fariam muita falta para uma parte da população que tem muita volatilidade na sua renda do trabalho.

Gráfico 9: Renda Domiciliar per capita Média (em Reais) por Categorias de Renda por Quintis

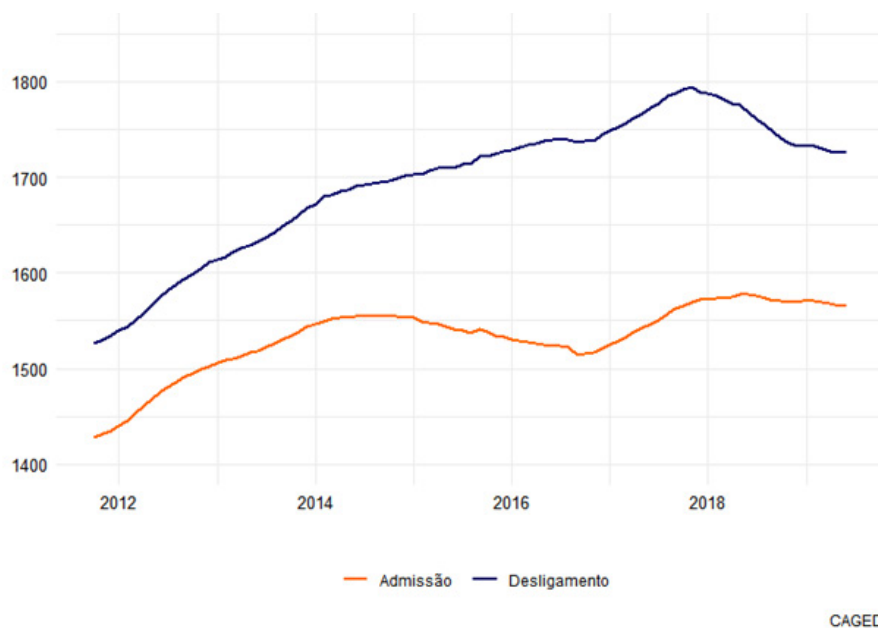


d) Caged

O Cadastro Geral de Empregados e Desempregados (Caged), nas palavras da Secretaria do Trabalho, “foi criado como registro permanente de admissões e dispensa de empregados, sob o regime da Consolidação das Leis do Trabalho (CLT)”. Desse modo, é um dado administrativo de fluxo de empregados e desempregados em relação ao setor formal (excluindo servidores públicos estatutários e militares).

Além de informar sobre o número de admitidos e demitidos todo mês (por estado, sexo, faixa de idade, entre outras informações), o Caged também apresenta os salários de admissão e de demissão. O Gráfico 10 mostra como, sistematicamente, o salário de demissão é superior ao de admissão, o que se explica em grande parte pelos diferenciais de experiência.

Gráfico 10: CAGED: Média Móvel Anual de Salários de Admissão e Desligamento (em Reais)

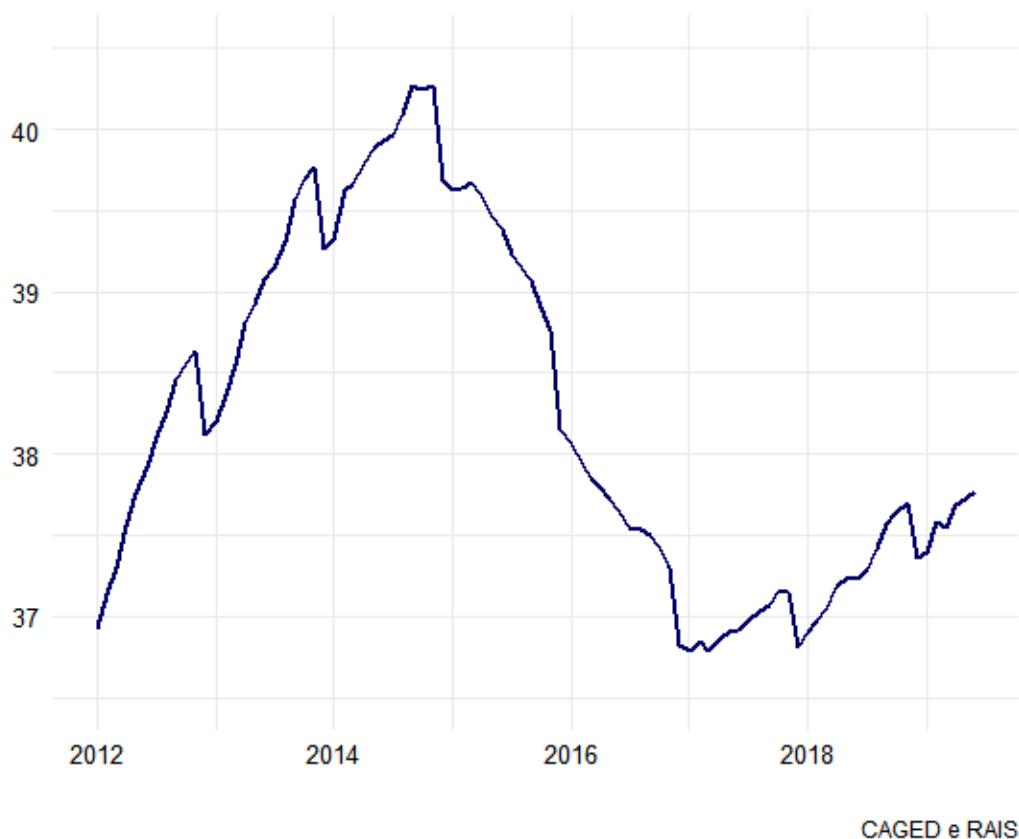


e) Rais

A Relação Anual de Informações Sociais (Rais) é o estoque do qual o Caged é o fluxo. Ela provê diversas informações dos trabalhadores formais no Brasil – nesse caso, inclusive os servidores públicos estatutários e militares. É uma base de dados muito rica, com diversas informações, que, quando disponibilizada especialmente para pesquisas, identifica (anonimamente) os empregados e as firmas nas quais eles estão trabalhando. No entanto, ela é apenas uma base anual, com referência no mês de dezembro.

Fazendo as devidas exclusões, a Rais é compatível com o Caged, de modo que é possível assim criar uma base de dados mensal de estoque de trabalho via CLT, com precisão censitária. O Gráfico 11 mostra a evolução desta base, com grande aumento entre 2012 e 2014, e subsequente queda até 2017, quando havia voltado ao nível de cinco anos antes, começado a se recuperar desde então.

Gráfico 11: CAGED e RAIS: Estoque de Empregos Formais (em milhões)



f) Censo Demográfico

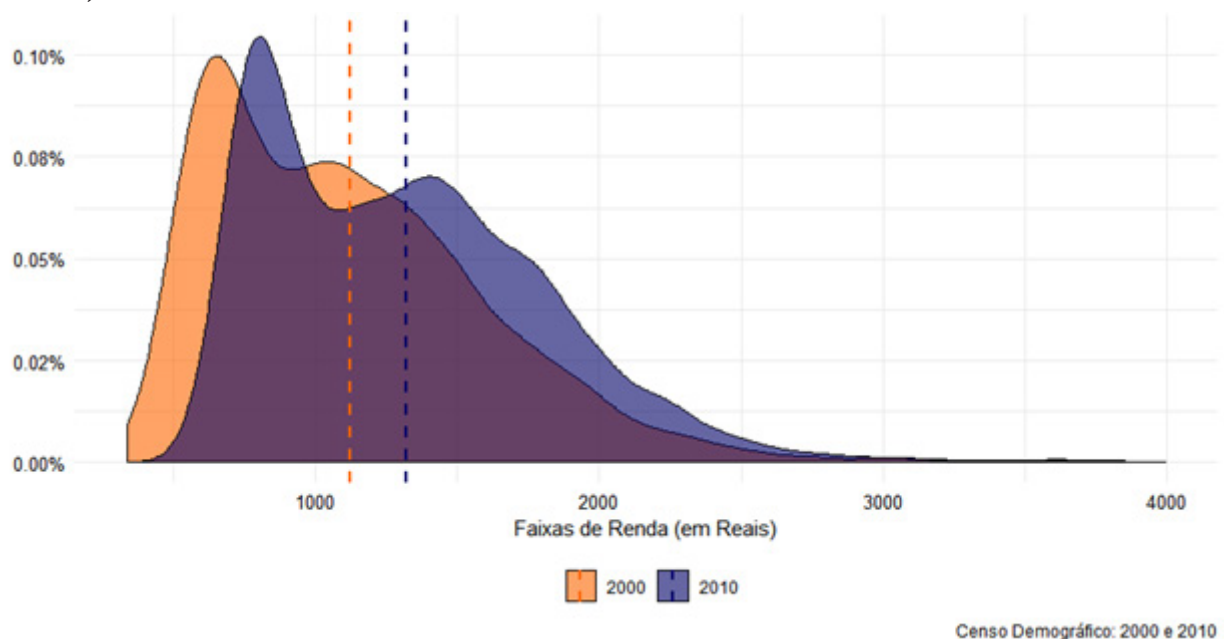
O Censo Demográfico é a maior pesquisa domiciliar do país, a única com caráter censitário. Como são grandes os custos de sua implementação, ele é realizado geralmente apenas a cada 10 anos. No entanto, apenas poucas perguntas de seu questionário são feitas para toda população. A maioria das perguntas é feita em uma amostra menor (ainda que com uma quantidade de entrevistados muitas vezes maior do que qualquer outra pesquisa).

Uma das principais vantagens do Censo é sua representatividade espacial: enquanto, na maioria das vezes, as pesquisas amostrais só chegam a ser representativas ao

nível do estado e de suas capitais, o Censo não apenas chega aos municípios, mas também às chamadas Áreas de Ponderação destes, que são algo relativamente próximo aos bairros.

O Gráfico 12, por exemplo, é um gráfico de densidade dos municípios por renda média em 2000 e 2010. Apenas com pesquisas como a do Censo Demográfico é possível realizar esse tipo de análise.

Gráfico 12: Distribuição de Renda Média nos Municípios por Faixas de Renda (a preços de 2018)



O último Censo foi realizado em 2010: o Censo de 2020 foi adiado por conta da pandemia do novo coronavírus, para 2021.

c) PNAD e PME

A PNAD e PME são duas pesquisas domiciliares já extintas – substituídas pela PNAD Contínua. A última edição de ambas ocorreu no ano de 2015.

A PNAD era uma pesquisa anual, que ia a campo em setembro, com abrangência nacional desde 2004. Antes disso, desde 1992, essa pesquisa só não abrangia os municípios rurais da Região Norte (com exceção do Tocantins). Há ainda edições anteriores até 1976, ainda que com menor área de cobertura.

A PNAD contava com diversas perguntas sobre renda (do trabalho e do não traba-

lho), família, fecundidade, migração etc., com diversos suplementos. Por exemplo, em 2003 e 2008, houve um suplemento de saúde e fumo. Em 2004, 2009 e 2013, um suplemento de segurança alimentar. Em 1996 e 2014, um suplemento de mobilidade intergeracional. Ainda é possível encontrar suplementos de atividades esportivas, litígios, associação a sindicatos etc.

Já a PME era uma pesquisa mensal, realizada em 6 regiões metropolitanas, com informações sobre mercado de trabalho (emprego, salários, participação etc.). Tal como a PNAD Contínua, ela também tinha uma estrutura longitudinal, com uma pequena diferença: os domicílios eram contemplados na amostra por 4 meses e, depois de passado o mesmo período, eles retornavam nos seguintes até completar um ano.

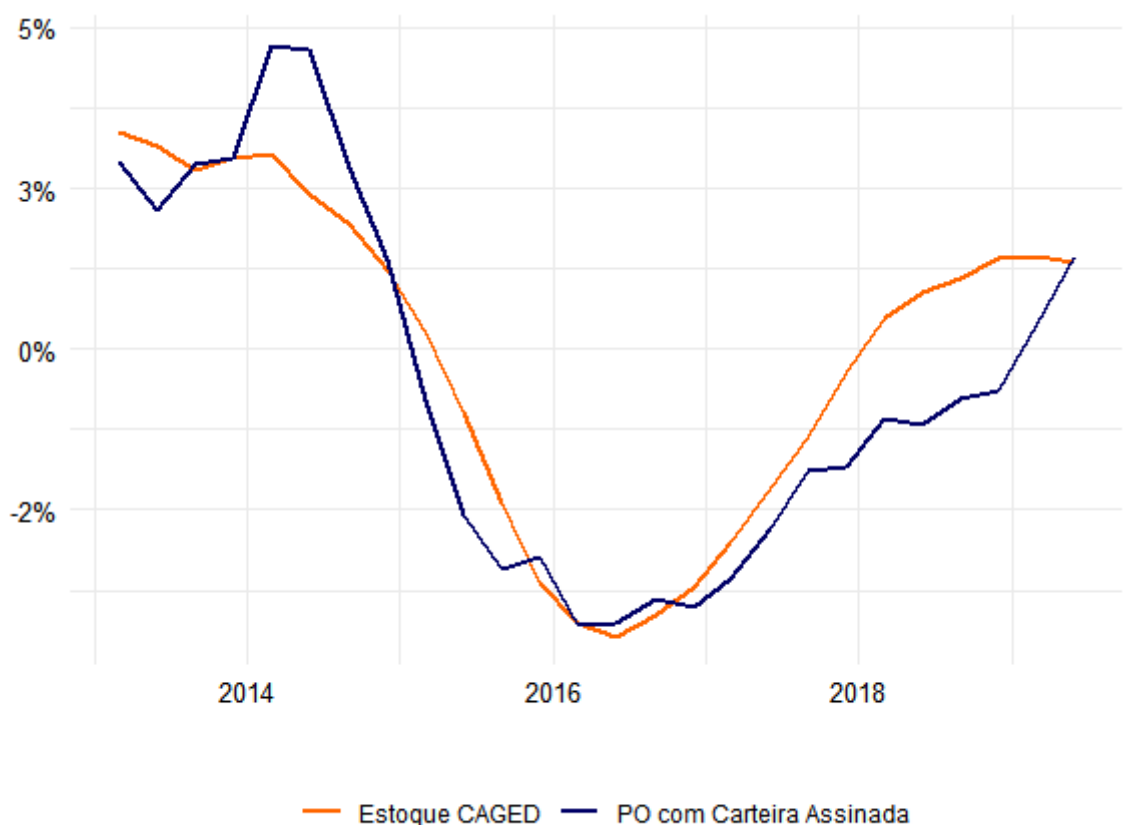
Armadilhas dos dados: interpretação, uso e questões recentes

a) Caged vs PNADC

Como dito anteriormente, o Caged e a PNAD Contínua contêm ambas um mesmo tipo de informação: o número de empregos celetistas criados em um dado período de tempo. Portanto, seus números deveriam ser próximos – mas não é isso que ocorre sempre.

O Gráfico 13 mostra a evolução da média móvel anual dessa variável em ambas as pesquisas. Como se vê, há períodos com significativas diferenças.

Gráfico 13: CAGED x PNAD Contínua: Variação em 12 Meses



CAGED e PNAD Contínua

É bom se ter em mente que, de fato, as duas pesquisas têm metodologias bastante distintas. Enquanto a PNADC é baseada em entrevistas, uma pesquisa por amostra de domicílios, escolhida de forma a ser representativa para a economia brasileira como um todo, o Caged, por outro lado, é uma pesquisa cujos dados são obtidos via registros

administrativos, preenchidos pelos empregadores. As empresas são obrigadas, por lei, a informar à Secretaria de Trabalho, todos os meses, quantos trabalhadores foram demitidos e quantos foram contratados.

Entre um e outro, normalmente o Caged seria o correto, afinal, trata-se de dados administrativos oficiais (a não ser que haja algum problema com o próprio cadastro). Os economistas Fernando de Holanda Barbosa Filho e Tiago Carreira, adicionalmente também mostram que comparar o Caged com a PNADC do mês seguinte aproxima as duas séries, uma vez que o cadastro preenchido pelas empresas acaba captando antes demissões que só serão relatadas pelos trabalhadores à PNADC nos três meses posteriores. Por causa dessa diferença temporal, o Caged tem pico de demissões em dezembro, enquanto na Pnad, o momento mais grave para desligamentos acontece em março.

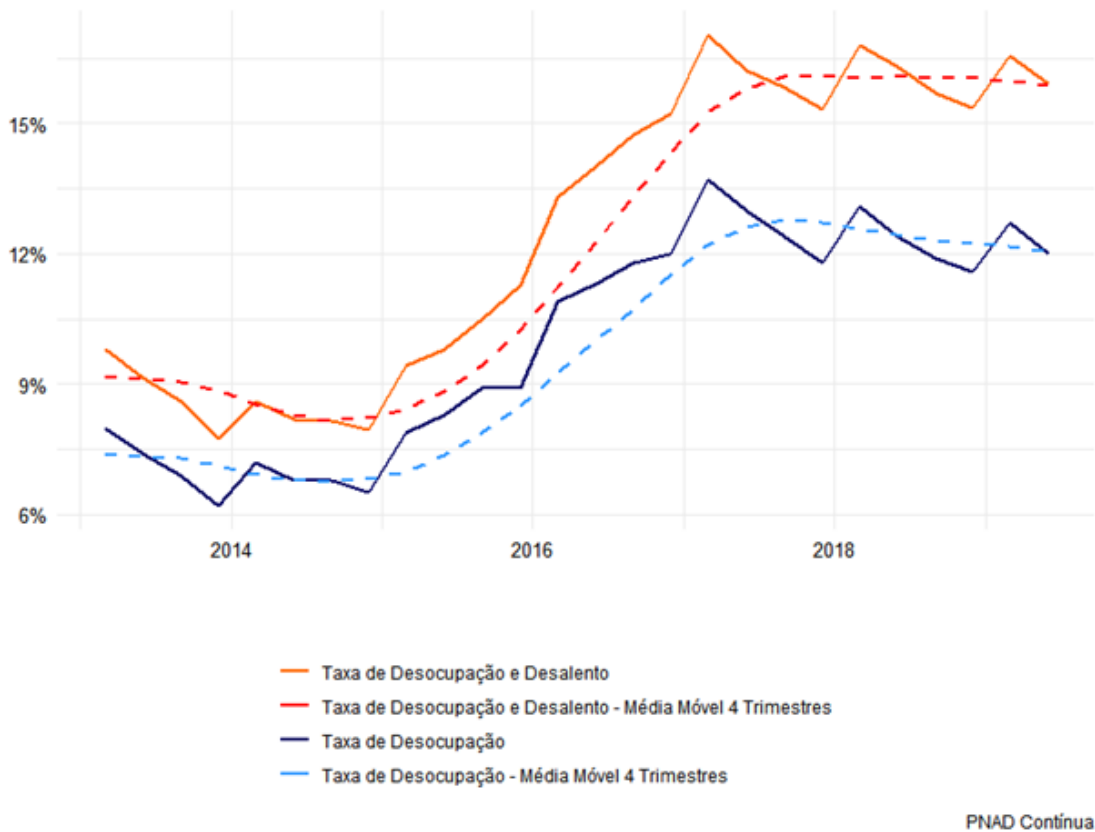
No entanto, apesar dos desacordos ocasionais, Caged e a PNADC acabam convergindo, como mostra o próprio Gráfico 13.

b) Desalento vs Desemprego

Ainda que haja algum debate acerca do tema, a Organização Internacional do Trabalho (OIT) não considera os desalentados como desocupados (ORGANIZAÇÃO INTERNACIONAL DO TRABALHO, 2013), uma vez que a razão pela qual os indivíduos não procuram trabalho é subjetiva e pode estar associada ainda a barreiras culturais. Desse modo, sua inclusão agregaria ruído à capacidade do indicador de medir aquecimento do mercado de trabalho. No entanto, a própria OIT e o IBGE consideram os desalentados como força de trabalho potencial, de modo a possibilitar análises mais amplas.

A redução da taxa de desemprego observada no Brasil nos últimos dois anos tem sido acompanhada, surpreendentemente, de uma contínua e rápida alta também do número de desalentados – brasileiros que, apesar de não terem procurado trabalho nos últimos 30 dias de referência, têm o desejo de estar trabalhando. Como se vê no Gráfico 14, caso se considerasse o desalento como uma forma de desemprego, a taxa de desocupação estaria muito mais alta e estável nos últimos anos.

Gráfico 14: Desocupação e Desalento



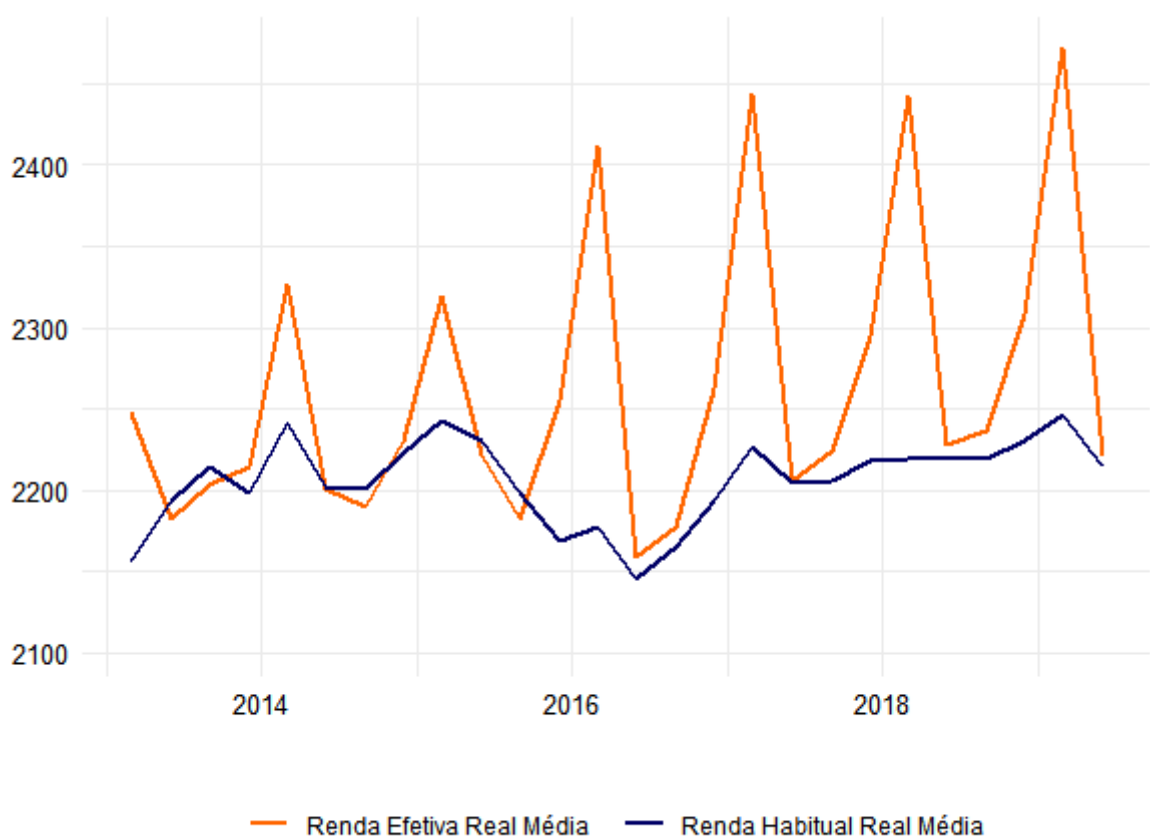
c) Renda efetiva vs Renda habitual

Na PNAD Contínua, há duas formas de se medir a renda: pelo seu valor habitual e pelo seu valor efetivo. Enquanto, na primeira, o IBGE pergunta o rendimento que se ganha habitualmente, no último, a pergunta é, especificamente, qual foi o valor recebido no último mês. Essa diferenciação tem importantes implicações.

O rendimento efetivo reportado, como pode se concluir no parágrafo anterior, será sempre referente a um mês antes da pesquisa. Ou seja, se a PNAD Contínua de julho de 2019 reporta uma renda efetiva média de R\$ 2,295, isso significa que, na realidade, essa foi a de junho (na verdade, do trimestre terminado em junho).

A outra questão é em relação à sazonalidade. O rendimento habitual geralmente é reportado como se fosse uma média mensal dos ganhos anuais, de modo que há pouca variação ao longo dos meses a não ser por um erro de medida dos respondentes em meses que eles ganham mais ou menos. Já o efetivo tem uma sazonalidade clara, com picos no primeiro e último trimestres do ano, como mostra o Gráfico 15.

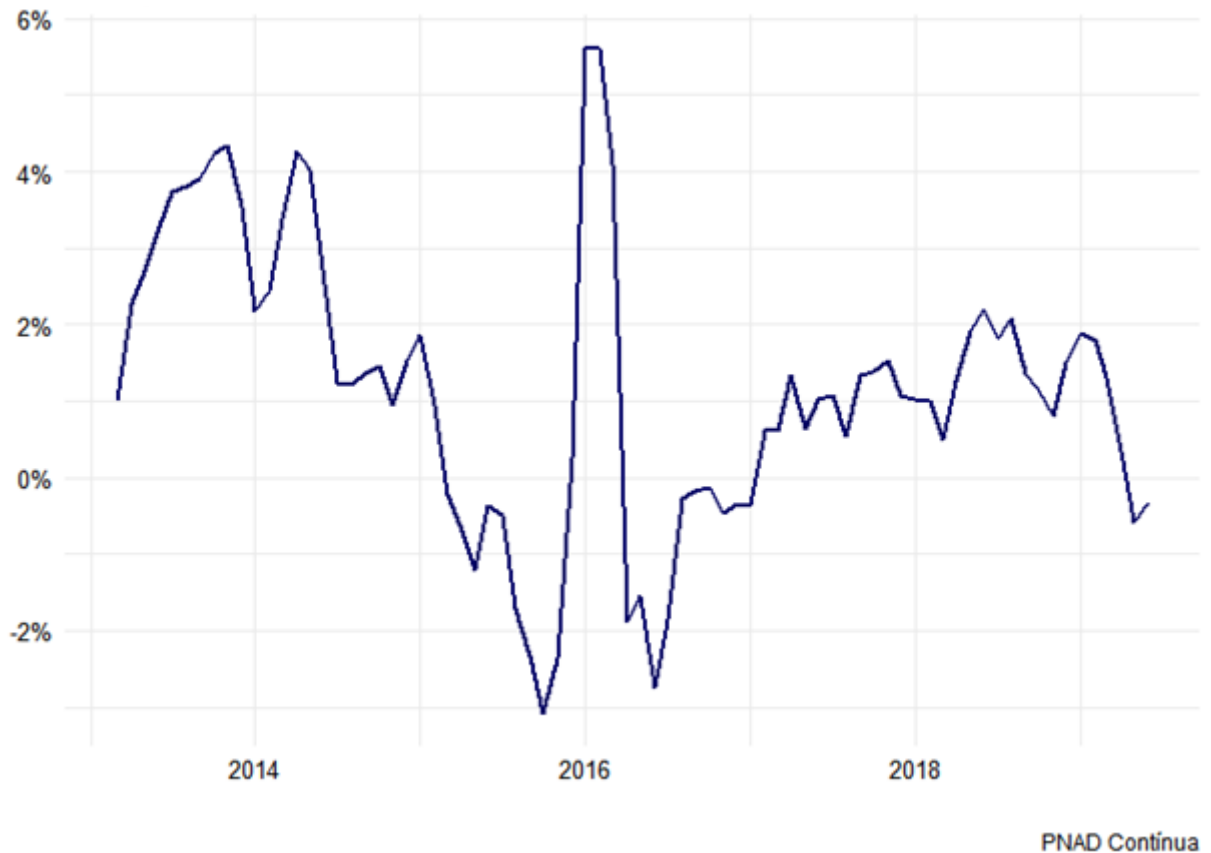
Gráfico 15: Renda Habitual e Renda Efetiva



PNAD Contínua

Há ainda um outro ponto na renda efetiva, devido a uma mudança metodológica do IBGE ocorrida em 2015. Na realidade, foi uma mudança da própria forma de captar a variável. Percebendo que grande parte dos entrevistados repetiam o mesmo valor do rendimento habitual e efetivo, houve uma reformulação do questionário de modo que, antes de fazer a pergunta relacionada ao segundo tipo de rendimento, era também perguntando o que tinha sido ganho de férias, 13º salário, abono salarial etc. A mudança ocorrida pode facilmente ser percebida no Gráfico 16, na página seguinte, em que se mostra a variação interanual da renda efetiva. Durante 2015, em um momento de crise econômica, o último trimestre registrou um grande e improvável aumento, de quase 6%!

Gráfico 16: Variação Interanual da Renda Efetiva Média



d) Outlier e renda média

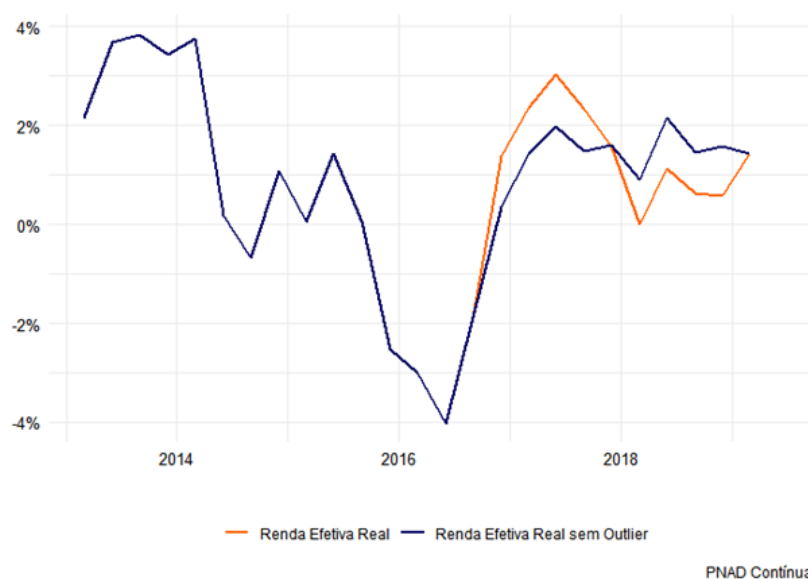
O que é uma média, no sentido aritmético? Nada mais do que a soma de todos os elementos, dividida pelo número (n) de elementos. O que ocorre, portanto, se um dos elementos destoa dos demais? De fato, basta que haja um ou poucos pontos muito acima dos outros para que a média suba sensivelmente. É que o se chama de efeito outlier.

Agora, por que essa introdução sobre estatística? Porque esse foi um dos fatores que ajudaram a explicar o aumento da renda do final de 2016 até o fim de 2017, e também uma das razões pelas quais o crescimento da média dos rendimentos em 2018 foi tão baixa.

No quarto trimestre de 2016, em um dos domicílios de uma zona de baixa renda de São Paulo, um empregador do setor de transportes declarou ao entrevistador do IBGE seu rendimento habitual e efetivo: nada menos do que 1 milhão mensal.

O problema é que ninguém nunca registrou qualquer valor mesmo próximo desse em todo histórico da pesquisa. A segunda pessoa mais rica do país, naquele trimestre, declarou receber R\$ 300.000, muito abaixo do primeiro. E, para piorar a situação, como ele morava em um lugar de baixa renda, seu peso na amostra era significativo, de modo a alterar bastante o comportamento interanual da renda média do país por nove trimestres – os cinco em que ele ficou na amostra (puxando para cima), e os quatro seguintes (puxando para baixo).

Gráfico 17: Renda Habitual Média: Variação Interanual com e sem Outlier

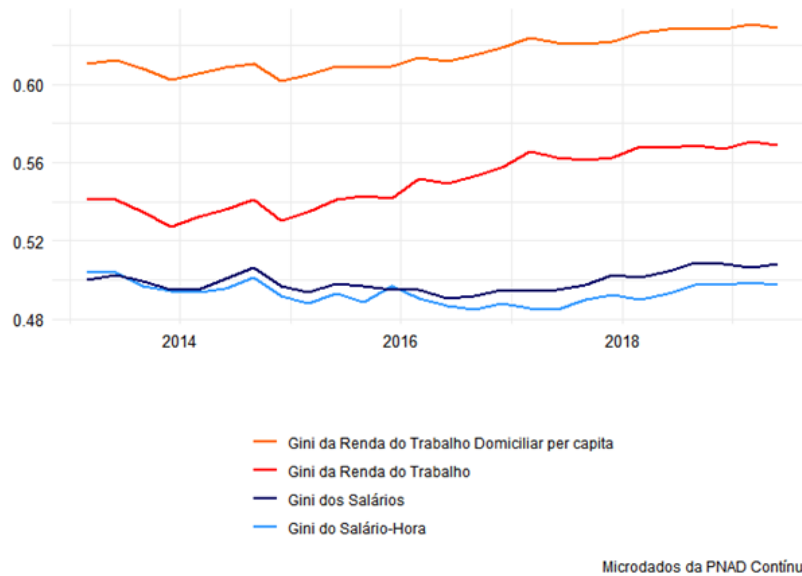


Felizmente, nas últimas atualizações, o IBGE retirou o outlier de sua amostra mesmo nos arquivos originais. Atualmente, qualquer um que baixar os dados da PNAD Contínua verá apenas a linha azul na evolução da renda média.

e) Desigualdade do mercado de trabalho.

Com que indicador se deve calcular a desigualdade no mercado de trabalho? Na verdade, isso depende de qual universo o pesquisador busca abranger. O Gráfico 18 mostra a evolução do Índice de Gini dos diferentes conceitos de renda do trabalho no Brasil de 2012 a 2019. Como se vê, todas apontam ainda crescimento, com exceção da renda do trabalho domiciliar per capita, que está a ponto de se estabilizar.

Gráfico 18: Gini por Diferentes Conceitos de Renda



Essas diferentes dinâmicas se explicam devido aos novos trabalhadores que estão sendo incorporados no mercado de trabalho, resultando em maior distribuição de renda. Quando se considera a soma de todos os rendimentos do domicílio, essa dinâmica puxa a desigualdade para baixo. Entre os ocupados, no entanto, como aumentam as diferenças de qualificação, tal fator aumenta a desigualdade de salários.

Já para a renda do trabalho domiciliar per capita, reduções da taxa de desemprego em geral fazem a desigualdade cair, por reduzir o número de zeros de renda nos domicílios. Essa variável pode ser decomposta pelos seguintes componentes:

Renda do trabalho domiciliar per capita = Salário por hora médio x Jornada de trabalho média x Taxa de ocupação x Taxa de participação x Taxa de atividade

Portanto, tal indicador na verdade incorpora praticamente todos os demais do mercado de trabalho, permitindo uma análise mais rica das dinâmicas relativas desse segmento da economia. Por exemplo, mesmo que a desigualdade do salário por hora esteja estável ou caindo, se a jornada tiver aumento das diferenças e o desemprego estiver crescendo, a desigualdade vai crescer.

Referências bibliográficas e leituras adicionais

- Salário Mínimo e Desemprego: O que diz a teoria e os dados?
- MINCER, Jacob. Unemployment effects of minimum wages. *Journal of Political Economy*, v. 84, n. 4, p. S87-S104, 1976. Disponível em: www.jstor.org/stable/1831104.
- CARD, David; KRUEGER, Alan B. Minimum wages and employment: a case study of the fast food industry in New Jersey and Pennsylvania. *American Economic Review*, v. 84, n. 4, p. 772-793, set. 1994. Disponível em: <http://ideas.repec.org/a/aea/aecrev/v84y1994i4p772-93.html>
- CORSEUIL, Carlos Henrique; CARNEIRO, Francisco Galvão. Os impactos do salário mínimo sobre emprego e salários no Brasil: evidências a partir de dados longitudinais e séries temporais. Brasília: IPEA, 2001. http://repositorio.ipea.gov.br/bitstream/11058/2054/1/TD_849.pdf
- FOGUEL, Miguel; ULYSSEA, Gabriel; CORSEUIL, Carlos Henrique. Salário mínimo e mercado de trabalho no Brasil. In: Monasterio, Leonardo; Neri, Marcelo C.; Soares, Sergei. *Brasil em desenvolvimento 2014 : estado, planejamento e políticas públicas*. Brasília: Ipea, 2014. Disponível em: <http://repositorio.ipea.gov.br/handle/11058/3590>.

GAP

OPINIÃO PÚBLICA¹

Pedro Masson

Álvaro J. Pereira Filho

08

¹ Agradecemos ao Doutor Alessandro Freire pela revisão de parte conteúdo do capítulo.

Introdução

Opinião pública é um conceito chave para o entendimento das democracias modernas. Na maioria dessas democracias, os governos se atentam às opiniões dos cidadãos, ajustando-as com as expectativas dos seus cidadãos. Por sua vez, a opinião pública também pode reagir aos atos governamentais e pronunciamentos dos seus representantes. Torna-se difícil, portanto, compreender o funcionamento das democracias sem considerar o peso da opinião pública.

As pesquisas e os estudos acadêmicos consideram opinião pública como a captação de respostas aos estímulos feitos por meio de um instrumento de coleta, o *survey* (ou questionário). Seu objetivo é capturar opiniões por meio de perguntas. Dessa maneira, o *survey* é como um termômetro capaz de medir o posicionamento de uma população sobre uma série de assuntos naquele momento.

O presente capítulo trará alguns passos necessários para realização de uma pesquisa de *survey*. Na medida em que os *surveys* tornaram-se a própria opinião pública, essa opinião pode variar de forma substancial a depender de como esse instrumento é tratado. Por isso, alguns cuidados devem ser tomados pelos pesquisadores, ou curiosos sobre o assunto, no momento de se realizar uma pesquisa dessa natureza.

Podemos dividir uma pesquisa de opinião pública em 3 (três) fases principais: elaboração, aplicação do *survey*, e análise dos dados. Primeiro, na elaboração, iremos debruçar sobre questões fundamentais como a elaboração de perguntas e suas escalas, a disposição das perguntas ao longo do *survey* e potenciais riscos de vieses.

Segundo, iremos nos aprofundar sobre a fase da aplicação que é intimamente ligada à capacidade de generalização dos resultados de uma pesquisa. Definiremos o que é amostragem e o que isso implica na consideração dos resultados. Da mesma maneira, apresentaremos uma pesquisa de opinião e avaliaremos se a utilização da internet como meio de coleta de dados é realmente o futuro em pesquisas de opinião.

Por fim, descreveremos as melhores práticas nas análises dos resultados de um *survey*. Focaremos nas possíveis falácias e erros cometidos por pesquisadores e leitores de pesquisas de opinião. Ainda, discutiremos aqui sobre o impacto de erros cometidos nas fases anteriores que podem repercutir e dificultar a análise dos seus resultados.

Este capítulo não pretende ser um guia exaustivo das possibilidades do estudo de

atitudes e de comportamento político, mas espera-se que ao seu fim o leitor tenha um panorama dos principais riscos e cuidados que devem ser tomados para se trabalhar com uma pesquisa de opinião pública e dessa forma evitar erros simples, porém determinantes para uma boa pesquisa.

Elaboração de survey

A elaboração do *survey* é uma etapa que irá impactar todas as demais etapas seguintes. De fato, a elaboração cuidadosa de um questionário é determinante para a garantia de qualidade do instrumento de pesquisa. Isso porque um *survey* elaborado de forma descuidada pode comprometer seriamente as análises e fazer com que se desperdice todo um esforço empreendido para sua elaboração e coleta. Análises estatísticas incorretas com dados de alta qualidade podem ser corrigidas posteriormente, porém, quando os dados coletados são de baixa qualidade, mesmo análises estatisticamente corretas podem levar a resultados distantes do ideal. Por isso, vários cuidados devem ser tomados nessa etapa, de forma que as demais etapas não sejam prejudicadas.

A primeira questão que o pesquisador deve ter consciência ao desenhar um *survey* é que este não é um bom instrumento para processos de pesquisas indutivos. Isto é, o *survey* não é uma ferramenta de exploração da realidade para que a partir dos dados coletados sejam construídas hipóteses. Na verdade, o *survey* é um excelente instrumento quando usado em modelos de pesquisas hipotético-dedutivas, onde o pesquisador parte de uma teoria que fundamenta suas hipóteses, e, a partir dessas hipóteses, elabora um *survey* capaz de captar do seu objeto as informações necessárias para testá-las.

Em um processo de pesquisa indutivo, caso o pesquisador queira realizar um estudo sobre um ator que tomou atitudes que fugiam às atribuições do seu cargo para mudar o resultado de uma política pública, o *survey* não é o instrumento mais indicado. O pesquisador não sabe previamente como determinado ator agiu, podendo até mesmo não saber corretamente quem é esse ator. A partir de um método indutivo de análise, indo a campo, o pesquisador conseguirá identificar esse ator e realizar entrevistas em profundidade, para só depois formular uma teoria sobre a atuação desse indivíduo dentro de determinada instituição.

No caso de um processo de pesquisa **hipotético-dedutivo**, o pesquisador elabora a teoria antes do estudo empírico, baseando-se em estudos ou teorias pré-existentes. Por exemplo, o pesquisador pode querer investigar o efeito da raça/cor dos indivíduos sobre a opinião pública. Para realizar essa pesquisa, o pesquisador deve previamente fazer uma revisão de literatura sobre a temática e formular uma hipótese que faz uma relação de causa e efeito entre a raça/cor e a opinião sobre determinado assunto. Após levantar suas hipóteses, ele deve utilizar um instrumento de pesquisa

que seja capaz de testá-las. Nesse processo, o survey se apresenta como uma ferramenta adequada para realização da pesquisa.

O *survey* é um instrumento que capta dos entrevistados a reação aos estímulos do pesquisador. O *survey* operacionaliza conceitos abstratos (por exemplo, apoio ao governo) que devem ser medidos da forma mais objetiva possível a fim de gerar dados comparáveis que sejam válidos e consistentes. Nas ciências sociais, em geral, falamos de conceitos “úteis” em vez de conceitos “corretos”. Isto é, podemos operacionalizar esses conceitos abstratos de diversas formas, porém algumas operacionalizações são mais úteis que outras.

Dessa forma, conscientes de que o *survey* é um instrumento adequado para pesquisas com processo hipotético-dedutivo, e que operacionaliza, muitas vezes, conceitos abstratos, o restante desta seção trará diferentes formas de se captar as dimensões do conceito estudado.

Elaboração da pergunta

As perguntas que compõem um *survey* podem apresentar erros tanto de sintática quanto de semântica. Quando se fala em sintática, pensa-se em gramática, ou seja, na construção correta de uma pergunta de acordo com a norma culta da língua. Quando se fala em semântica, pensa-se em significado, isto é, como a pergunta formulada será interpretada pelo entrevistado.

Os problemas que podem ocorrer em relação à sintática podem ser resolvidos de forma mais simples com uma revisão da pergunta realizada, e com sua reformulação, de forma a deixá-la mais sintética e objetiva. Por exemplo, a pergunta “Que horas você acorda para sair de casa?” pode causar dúvida no entrevistado (se é a hora que ele acorda ou a hora que ele sai de casa). Se a intenção é perguntar sobre a segunda opção, então a pergunta pode ser reformulada para “Que horas você sai de casa?”, ou, caso seja a primeira, para “Que horas você acorda?”.

Os problemas relacionados à semântica podem ser mais complexos de serem resolvidos, pois podem ser: problemas com pressuposto; falta de familiaridade; ou vagueza em relação à pergunta. Em relação ao pressuposto, o formulador do questionário pode fazer uma pergunta que parte de um pressuposto que é falso para o respondente. Por exemplo, quando perguntamos “Que horas você vai trabalhar?”, o pressuposto é de que o respondente trabalha. Para evitar problemas de pressuposto pode-se fazer

questões de filtro, como, por exemplo, “Você possui um trabalho?”. Também podem ser incluídas respostas como “não sei” ou “não tenho opinião” para evitar que o respondente assuma um pressuposto apenas para responder à pergunta.

São comuns, também, perguntas duplas (double-barreled) que podem provocar confusão no seu entendimento pelo entrevistado. Por exemplo, caso perguntemos: “você apoia o aumento de gastos com educação e aposentadorias pelo governo?”. Aqueles que são favoráveis a apenas uma dessas medidas não teriam como expressar corretamente sua opinião. O ideal é realizar as perguntas de forma separada.

Os problemas de falta de familiaridade podem ocorrer quando o tema perguntado é muito complexo para grande parte dos cidadãos. Por exemplo: “Qual sua opinião sobre o sistema distrital de ocupação de cadeiras legislativas?”. Muitas pessoas podem não ter familiaridade com o assunto, sendo que essa pergunta pode levar a uma grande taxa de não-resposta ou de respostas erradas (quando o respondente assinala uma opção sem saber realmente do que se trata).

Por último, o problema de vagueza acontece quando não há objetividade naquilo que se pergunta, ou quando há a utilização de conceitos que são por natureza vagos, como, por exemplo, “Você é a favor da democracia?”. Esse problema também pode ser solucionado com uma melhor delimitação da pergunta ou com uma delimitação do conceito. No exemplo anterior, o melhor seria detalhar alguma característica do que se acredita como democracia – como eleições livres ou alternância de poder – e fazer uma pergunta diretamente a isso. Perguntas complexas demandam soluções mais sofisticadas. No caso de apoio à democracia, seria interessante juntar uma série de posicionamentos e daí construir um índice de apoio ou não.

Além da preocupação com uma pergunta individualmente, há também as preocupações com a sequência de pergunta. Quando analisamos o questionário como um todo, estamos preocupados com sua consistência e validade, que permitirá uma melhor análise dos dados coletados e uma posterior generalização. A seguir, iremos discutir sobre a estrutura do survey e a disposição das suas perguntas.

A estrutura do survey

Levando em consideração os cuidados que devem existir na elaboração de uma pergunta, o pesquisador deve analisar qual a melhor estratégia para operacionalizar o seu conceito de interesse, de forma que seu instrumento seja confiável e, ao

mesmo tempo, válido.

Quando falamos em confiabilidade, referimo-nos à capacidade da reprodutibilidade dos resultados do questionário, sua homogeneidade e o grau em que ele está livre de erros aleatórios. Por outro lado, quando falamos em validade, estamos falando do quanto o instrumento está medindo o que ele se propõe a medir. Naturalmente, uma boa medida deve ser tão confiável quanto válida. Isso porque medidas pouco confiáveis produzem resultados inconclusivos e medidas inválidas descrevem fenômenos que não aqueles de interesse.

Nesse sentido, o pesquisador deve refletir sobre: **o formato do survey; a quantidade de perguntas; a ordem e o tipo de pergunta**. Em relação ao **formato do survey**, os tipos mais comuns são os questionários fechados com opções de respostas variadas, ou o questionário aberto onde a resposta do entrevistado é anotada. Ainda há a possibilidade de uma versão mista de fechado com espaço para pergunta aberta, porém esse tipo misto apenas junta as duas categorias e não possui nenhuma característica própria. Vamos focar nas duas principais categorias: fechados e abertos.

Os questionários fechados foram e são usados em larga escala dada a facilidade de se fazer análises estatísticas a partir dos dados coletados, como análises de regressões e análises fatoriais. Os questionários abertos trazem dificuldades para se analisar as respostas, pois codificá-las e analisá-las estatisticamente implicam mais esforços. Contudo, com o desenvolvimento de softwares capazes de fazer análises de sentimentos das respostas e de técnicas de processamento de linguagem natural, a tarefa de categorizar respostas abertas se tornou possível, ainda que ainda pouco usada na maioria dos estudos de opinião pública por survey.

Nos questionários fechados, o pesquisador deve se atentar à **quantidade de perguntas**, podendo utilizar uma pergunta simples ou de uma bateria de perguntas para operacionalizar seu conceito de interesse. A vantagem de uma pergunta simples é sua praticidade, o tempo de coleta e o custo. Porém, esse tipo de estratégia pode trazer problemas de validade ao questionário, pois conceitos complexos, geralmente, possuem mais de uma dimensão a ser analisada. Por exemplo, operacionalizar conceitos como raça (no caso brasileiro especialmente) e apoio à democracia por meio de uma única pergunta simples pode revelar apenas uma de suas dimensões tornando a medida com pouca validade. Há também problemas relacionados à confiabilidade do instrumento, a qual pode ser reduzida em virtude das grandes variações que podem existir entre um contexto e outro, quando se usa apenas uma pergunta simples.

Dessa forma, uma bateria de perguntas é preferível quando há alta complexidade no conceito operacionalizado. Escalas usando múltiplos itens são, em geral, mais estáveis e confiáveis, pois viabilizam a abordagem do conceito de diversos “ângulos” ou perspectivas. Além disso, o uso de múltiplos itens permite que erros de mensuração aleatórios se cancelem, resultando em um aumento de confiabilidade e precisão.

O pesquisador pode fazer um questionário com uma pergunta simples para avaliar o governo. Por exemplo: “Qual a sua avaliação sobre o desempenho do atual governo?”, e oferecer opções de respostas ao respondente. No caso de uma bateria de perguntas, o pesquisador pode fazer uma série de perguntas sobre o mesmo tema, para captar suas diversas dimensões. Por exemplo: “Qual sua avaliação sobre a atuação na área de economia do atual governo?” e, em seguida, “Qual sua avaliação sobre a atuação junto ao Congresso do atual governo?”, depois “Qual sua avaliação sobre as políticas sociais do atual governo?”. Fazendo uma bateria de perguntas, o pesquisador consegue examinar diferentes dimensões sobre um determinado tema, no caso, a avaliação dos cidadãos sobre o governo, chegando a uma medida mais confiável e precisa.

Entretanto, cabe notar que perguntas adicionais para formar um índice tendem a apresentar ganhos pequenos de confiabilidade. Ou seja, a partir de certo ponto, o acréscimo de itens pode não aumentar a confiabilidade de uma medida. Além disso, muitas perguntas podem induzir o viés de satisficing (KROSnick; PRESSER, 2010), que consiste em reduzir a motivação e o esforço cognitivo do entrevistado ao responder ao survey.

Em relação à **ordem** das perguntas, o método mais utilizado é o chamado aproximação de funil, isto é, os blocos de perguntas começam com uma questão ampla e se reduzem progressivamente a questões específicas; esse processo envolve necessariamente o uso de perguntas de filtro para que os respondentes aos quais as perguntas específicas não se aplicam sejam direcionados para a próxima pergunta que se aplica a eles.

As principais regras levantadas por Bowling (2005) para a ordem das perguntas são: fazer perguntas simples e básicas (não sensíveis ou ameaçadoras) primeiro; como as respostas podem ser influenciadas por respostas anteriores, fazer as perguntas mais importantes primeiro onde nenhuma outra regra se aplica; perguntas sobre o comportamento (ações do indivíduo) devem ser feitas antes de perguntas sobre atitudes (opiniões do indivíduo).

Diversos experimentos de survey demonstraram que as respostas a uma pergunta podem ser afetadas por perguntas anteriores, o que a literatura define como “efeitos de contexto” (TOUREANGEAU; RASINSKI, 1988). Esses efeitos podem ser minimizados colocando-se questões gerais antes de perguntas específicas.

O **tipo de pergunta** varia de acordo com o conceito operacionalizado. O primeiro tipo são conceitos categóricos ou nominais (estado civil, gênero, local de residência etc.); para esse tipo de conceito as perguntas são feitas tendo como respostas categorias distintas que poderão ser posteriormente codificadas numericamente, mas não hierarquizadas em um continuum.

Um conceito também pode ser operacionalizado em intervalos (como renda, idade etc.) assim as respostas serão faixas onde o respondente deve se localizar (no caso da renda, por exemplo, de 0 a 1 salário mínimo, de 1 a 3 salários mínimos, e aí por diante). Há também os chamados dados de proporção, que têm as características de dados de intervalos, porém com um zero absoluto (por exemplo, peso, altura, etc.).

Por último, os tipos mais utilizados de perguntas são os que se referem a conceitos de variáveis ordinais. A escala ordinal mais utilizada no campo do comportamento político é a escala Likert.

A escala Likert geralmente varia de 5 a 7 itens de resposta e representa um contínuo sobre a pergunta feita. Esse contínuo pode variar de um grau menor (maior) de concordância até um grau maior (menor), como de “discordo muito” a “concordo muito”. Esse tipo de escala é muito utilizado, pois é fácil de entender e analisar e diminui os erros em relação às escalas dicotômicas (concordo/discordo), ao dar mais opções de respostas ao entrevistado.

No caso da escala Likert, há uma escolha entre colocar um ponto médio ou não nas opções de respostas. Escolher não colocar o ponto médio pode facilitar na hora de se somarem os itens, porém pode induzir o entrevistado a dar uma resposta errada. No caso da inserção de um ponto intermediário (“nem concordo, nem discordo”), deve-se valorá-lo com um valor mediano, por exemplo: discordo totalmente = 5; discordo = 4; nem concordo, nem discordo = 3, concordo = 2, concordo totalmente = 1.

Há um debate na literatura acerca da utilização do ponto mediano. Há evidências empíricas de que a inclusão de uma categoria intermediária reduz o viés de aquiescência, ou seja, a tendência de se concordar com afirmações feitas em um questionário. Porém, também há evidência que os respondentes podem escolher o ponto mediano

como uma forma de escapar da tarefa de tomar posição sobre o assunto (satisficing).

Tomadas essas decisões e escolhas sobre as perguntas que serão realizadas, o pesquisador pode partir para a aplicação do survey.

Aplicação de surveys

População e amostra

A população, no âmbito deste capítulo, é um conjunto de pessoas com uma característica, comportamento ou algo em comum. De fato, é possível definirmos diversas populações em uma população maior. Por exemplo, dentro da população brasileira há a população de eleitores, ou o eleitorado brasileiro. O eleitorado brasileiro compartilha duas características em comum: é cidadão brasileiro e tem idade mínima de 16 anos, como previsto no Código Eleitoral. Quanto maior o número de características atribuídas a um grupo de pessoas, menor será o total de pessoas dentro dessa população. Portanto, a definição da população é o primeiro passo de um estudo com survey.

A definição da população é necessária, porém não trabalhamos diretamente com ela. Dependendo do tamanho dessa população, o estudo se torna complexo e inviável economicamente. Por conta disso, pesquisas diretamente com populações são extremamente raras, sendo conduzidas, na maioria das vezes, por governos de 10 em 10 anos (censos demográficos).

As pesquisas de opinião, todavia, são realizadas, preferencialmente, em ciclos curtos, o que envolve, além de tudo, um menor gasto de recursos. Pesquisadores selecionam alguns membros da sua população de interesse para a partir desse pequeno grupo (amostra) tirar conclusões sobre o grupo todo (população). Esse exercício chama-se inferência, e é a mesma técnica que utilizamos, por exemplo, quando fazemos um exame de sangue: tira-se uma pequena quantidade para inferir como está todo o sangue do seu corpo. A inferência somente é possível por conta de uma propriedade estatística conhecida como teorema do limite central.

De forma bem simplificada, o teorema do limite central diz que se uma variável tem os valores possíveis conhecidos, então sua distribuição se tornará mais próxima de uma curva normal quanto maior o número de amostras dessa população. Por exemplo, eu jogo um dado de 6 lados 10 vezes e tiro a média desses valores; fazendo

isso 1000 vezes, a distribuição dessas médias tende a formar um histograma com o formato de um sino, com a maioria das médias em torno de um ponto central, no caso, 3,5, que é a média da soma das facetas do dado.

O teorema do limite central e as suas aplicações são muito importantes para pesquisas utilizando estatística inferencial, mas por aqui ficaremos somente com essa noção mais básica. O mais importante é que raramente fazemos várias amostras de uma mesma população. Na verdade, fazemos uma amostra e inferimos os seus valores e estimamos a diferença entre os valores da amostra e da população: também conhecido como erro amostral. Sem entrar em detalhes, aquela informação sobre a margem de erro para mais e para menos, é o intervalo que provavelmente a média da população estará contida em 95% das amostras feitas para uma população (por isso a maioria das pesquisas trazem o erro amostral e o intervalo de confiança, no caso 95%).

Uma outra propriedade interessante dessa margem de erro é que ela varia com o tamanho da amostra. Imagine que temos uma amostra com o mesmo tamanho da população. Isso significa que eu tenho uma margem de erro muito pequena, quase mínima, porque não preciso inferir qual seria a média populacional se temos a população como amostra. Ao contrário, portanto, se temos uma amostra de uma pessoa, a margem de erro será a maior possível. Afinal, por que fazemos pesquisas com amostras pequenas de uma população muito maior e tentamos explicar com o número reduzido de pessoas o que a maioria pensa?

Primeiro, o tamanho amostral, além de estar relacionado à margem de erro através do cálculo da dispersão, relaciona-se com esforço e custo de pesquisa. Assim, como grandes amostras significam muito dinheiro e tempo expendido, as pesquisas buscam o tamanho amostral mais prático e menos oneroso. Segundo, e não menos importante, a relação entre o tamanho da amostra e a precisão da amostra não é linear. Isto é, em determinado ponto, os ganhos de precisão das estimações com o aumento da amostra diminuem ao ponto que são praticamente constantes, enquanto o custo sobe cada vez mais. Assim sendo, não é preciso suspeitar dos tamanhos de amostras que você provavelmente lê em jornais e vê no noticiário, nem se você não foi uma das pessoas pesquisadas. Os tamanhos amostrais buscam otimizar tanto em termos de tempo e dinheiro quanto em termos de retorno da eficiência das estimações dos efeitos. É possível, portanto, que uma amostra de 2 mil pessoas faça inferências sobre o conjunto de 200 mil, por exemplo.

A relação entre o tamanho de amostra e a capacidade de fazer conclusões para a população são diretamente conectadas. Quanto maior o tamanho amostral, mais pró-

ximo está da verdadeira média populacional. Entretanto, a composição da amostra e o quanto essa realmente representa a população de interesse implicam em uma série de considerações. Se imaginarmos uma pesquisa como uma fotografia, o tamanho da amostra seria a qualidade da foto enquanto a representatividade da amostra significa o seu enquadramento. Para entender melhor o conceito de representatividade da amostra, vamos falar sobre tipos de amostragem.

Tipos de amostragem

A amostragem tem dois tipos aplicados no campo da opinião pública: as amostragens probabilísticas e as amostragens não probabilísticas. Esses tipos de amostragem se diferenciam na capacidade de montar amostras representativas, isto é, que se aproximam ao máximo das características gerais da população de interesse e permitem estatísticas inferenciais mais confiáveis.

Primeiramente, nas amostras probabilísticas, cada unidade da população tem chance igual de ser selecionada. As amostras que são representativas, portanto, permitem maior grau de generalizações dos resultados, por conta da distribuição do erro e da capacidade de evitar vieses. Por fim, podemos estimar melhor a ‘distância’ entre a média da amostra e a populacional, o que chamamos de erro amostral.

O mais básico das técnicas de amostragem probabilística é a amostragem aleatória simples. Essa técnica lista todas as possíveis unidades, enumera-as consecutivamente e então usa uma tabela numérica para selecionar uma porção ou porcentagem do total das unidades. Uma vez selecionada, as unidades devem ser contatadas uma a uma.

Há basicamente dois problemas na amostragem aleatória simples. Em geral, pesquisadores não têm acesso a todas unidades de uma população, sendo que a probabilidade de se omitir parcelas de baixa renda é ainda maior, pois se trata de uma população ora de difícil acesso, ora com baixa propensão a participar de pesquisas. No caso de uma unidade selecionada não responder, temos um outro problema, pois essa unidade não pode ser substituída e exige, assim, quantidade maior de selecionados antes da amostragem.

As pesquisas realizadas no Brasil, em geral, são do tipo amostragem aleatória estratificada. Pesquisas feitas em institutos de pesquisas de opinião populares utilizam essa técnica na grande maioria das vezes, por exemplo as pesquisas de intenção de voto realizadas pelo Datafolha e Ibope. Essa técnica é, em suma, uma série de amos-

tragens aleatórias realizadas em uma mesma população. Garante-se assim a representatividade de subgrupos da população, praticidade com a técnica e confidencialidade das unidades. A priori, essa técnica exige a informação dos grupos que dividem a população, que pode ser retirada de censos e outras pesquisas prévias.

As amostragens não probabilísticas, por outro lado, não produzem amostras representativas e, provavelmente, conterão algum tipo de viés por não ter um ‘sorteio’ justo das unidades. Ainda, em algumas dessas pesquisas, os participantes são quem se selecionam para o estudo, o que pode privilegiar um tipo de característica no universo populacional. Portanto, o uso dessa amostragem implica dificuldade de generalização dos resultados.

Em sua grande maioria, essas amostragens são utilizadas por pesquisadores que não têm como objetivo principal amostras representativas e, por consequência, não tem como foco de pesquisa a generalização dos resultados. Esse procedimento pode ser escolhido também por falta de recursos para uma amostragem probabilística, que em geral é mais cara e realizada por institutos especializados.

Em experimentos, é bastante comum encontrar amostras não probabilísticas, como a amostra de conveniência. Em suma, essas amostras têm como alvo apenas indivíduos que possuem características que os fazem mais acessíveis ao pesquisador. Por isso, são amplamente utilizadas em experimentos laboratoriais por conta da dificuldade do recrutamento para esse tipo de estudo, por conta da sua duração e a atividade que se demanda nesses estudos.

As amostras de conveniência também são úteis em estudos exploratórios como, por exemplo, para se ajustar alguma pergunta em que não se sabe como montar a escala de respostas ou como será interpretada pelos participantes. Essa amostra também pode ser utilizada para ajustar fatores técnicos, como plataforma da aplicação e os entrevistadores.

Outra amostra não probabilística bastante conhecida é a amostragem do tipo snowball – ou bola de neve. Basicamente, essa amostragem funciona da seguinte maneira: o pesquisador seleciona alguns informantes para começarem as pesquisas. Em seguida, esses informantes compartilham o questionário com outros participantes das suas próprias redes e assim por diante.

O snowball é ideal para acessar populações que não são facilmente identificadas, ou que possuem grande resistência em participar de pesquisas, como grupos muito

religiosos ou mais fechados por difícil contato. Entretanto, ainda é uma amostra que apresenta grande viés, pois as redes de contatos compartilham também uma série de características comuns e a depender das características dos informantes, a característica geral da amostra será bastante contaminada. Na seção 3 serão apresentados exemplos que ilustrarão os riscos de realizar inferências a partir de uma amostra snowball.

Meios de aplicação

A amostragem diz respeito à seleção dos participantes. Como você os seleciona implica na característica da amostra e nas possibilidades de conclusão do trabalho como um todo. Já o meio de aplicação dos surveys diz respeito à abordagem dos selecionados e implica basicamente na taxa de resposta e na postura do indivíduo em responder tal pesquisa. Abordaremos aqui os principais meios de aplicação dos surveys e a implicação tanto na taxa de resposta quanto nas respostas em si.

A pesquisa face a face é um meio de pesquisa amplamente utilizado no Brasil. Basicamente, os entrevistadores aplicam o survey presencialmente, na maioria das vezes, lendo e preenchendo as respostas no questionário. O questionário, por sua vez, deve carregar uma série de instruções, como pular uma pergunta em caso de uma resposta específica até mesmo incluir o julgamento do próprio entrevistador.

A opção face a face demanda uma decisão sobre onde os entrevistadores serão distribuídos, o que pode ter impacto na pesquisa. Por exemplo, se os entrevistadores são enviados para locais onde o pesquisador julga de maior circulação, a depender da cidade, isso pode implicar em amostras com certo perfil demográfico – excluir pessoas que se locomovem por carro, ou que não frequentam tal região. Uma opção é utilizar entrevistas em domicílio, o que pode implicar em problema na taxa de resposta.

A taxa de resposta é o número de pessoas que respondem ao survey, respostas válidas e completas, sobre o número total de pessoas contatadas pela pesquisa. Geralmente, a taxa de resposta é um bom indicador de representatividade da população que o survey deseja atingir. No caso de face a face, a vantagem desse método é que existe baixo índice de abandono – quando os indivíduos desistem durante a aplicação da pesquisa – e a taxa de resposta da pesquisa tende a aumentar, quando se considera

o total de respostas válidas como respostas completas².

Por outro lado, pesquisas face a face podem ter um custo muito alto, com a contratação de entrevistadores e o envio deles para diferentes regiões, treinamento e implicações nas respostas do survey a depender do tema da pesquisa. Por exemplo, temas polêmicos como uso de drogas, e até a intenção de voto, podem ter suas respostas influenciadas. Número de fumantes, em perguntas face a face, tende a ser subestimado, assim como apoio a democracia e as cotas raciais, por exemplo, tende a ser sobrestimado. As características individuais dos entrevistadores tendem a influenciar na postura e nas respostas dos entrevistados. Por exemplo, perguntas sobre cotas raciais podem sofrer efeito da raça do entrevistador durante a aplicação do survey.

Pesquisas por telefone resolveram, em parte, os problemas de custo das pesquisas presenciais, ou face a face. Não seria preciso mais deslocar entrevistadores para os domicílios ou ponto de fluxo nas cidades para aplicação da pesquisa. Assim, com diversas linhas telefônicas, alguns números e um galpão, os pesquisadores encontrariam uma forma mais barata de aplicar pesquisas. Entretanto, de um tempo para cá, o que parecia uma solução sustentável para aplicação de surveys para grandes amostras foi vendo sua taxa de resposta diminuir.

Uma solução das pesquisas utilizando telefones foi incluir linhas de celulares. Existem diferenças importantes entre as linhas fixas e de celulares. Em linhas fixas, por exemplo, a maioria dos usuários que permanecem nesse modo são pessoas mais velhas, por ser uma tecnologia mais antiga. Outra questão é que a linha fixa é única para o domicílio, enquanto o celular é individual. Por isso, institutos de pesquisa mundo afora utilizam uma porção de linhas fixas e outra de celulares, para mesclar o acesso à amostra e aumentar a taxa de resposta.

Mais recentemente a internet entrou em cena como um meio para pesquisa. Existe um debate caloroso sobre representatividade e eficiência desse meio para pesquisas. Há uma fartura de pesquisas retiradas de redes sociais, prometendo registrar finalmente o comportamento de uma população e com baixo custo. Por outro lado, uma parcela importante não tem acesso ou não utiliza certas redes sociais, o que coloca em xeque a capacidade de formar amostras representativas – sem contar a quantidade de robôs (bots), ou perfis falsos. Por fim, a utilização da internet deve ser vista com

2 Existem diferentes maneiras de calcular a taxa de resposta. Alguns utilizam respostas completas, porém outros somente respostas. A boa prática é sempre ter ambas as informações, pois número de pessoas que começam, mas não terminam, pode indicar algum problema no meio do survey ou sua própria extensão pode desestimular o preenchimento até o final.

cautela pelas possíveis violações éticas, como o risco à anonimidade, como pesquisas com snowball em redes sociais podem ocasionar.

A próxima seção traz as principais armadilhas na análise de dados sobre opinião pública que levam a erros de interpretação e de inferências. Tendo em vista o conteúdo apresentado até então (elaboração de survey e aplicação de survey), serão analisados casos práticos onde houve erro do analista, que não levou em conta diferenças tanto do questionário aplicado quanto do método de coleta dos dados.

Análise dos dados

Na área da opinião pública, muito raramente os microdados³ das pesquisas são disponibilizados para o público, de modo que a maioria dos erros cometidos pelos analistas e comentaristas políticos são na hora de comparar pesquisas de opinião já consolidadas. Os institutos de pesquisas também são passíveis de erros, porém esses erros são oriundos mais de inadequação de instrumentos/metodologia do que de erros de análise propriamente dito. Ao longo desta seção vamos destacar alguns casos e tentar explicar os erros cometidos para que eles sejam evitados em novas análises.

As pesquisas mais populares sobre opinião pública são as realizadas para verificar a intenção de voto em eleições. Diversos institutos de pesquisa, em épocas próximas às eleições, realizam esses levantamentos sendo que muitos utilizam metodologias distintas: alguns realizam pesquisas com entrevistador presencial; outros realizam pesquisas por telefone; há ainda os que utilizam a internet para fazer levantamento. O tamanho e tipo da amostra também variam de acordo com o instituto. Além disso, diferentes perguntas são feitas. Todas essas variáveis devem ser levadas em consideração quando realizada a comparação dos resultados das pesquisas.

Um erro cometido por apoiadores do Partidos dos Trabalhadores durante as eleições de 2018 foi relativo à intenção de voto no candidato Fernando Haddad. Em pesquisa realizada em julho de 2018, pelo instituto Ipespe, contratada pela XP Investimentos, o candidato Fernando Haddad aparecia, quando informado que era apoiado

3 São chamados de microdados aqueles que discriminam por indivíduo, ou por unidade de análise, a informação coletada. Ou seja, em uma pesquisa de opinião, os microdados trariam informações sobre aquele respondente (sem necessariamente identificá-lo por nome ou CPF), como região, raça, idade, gênero, e à opinião sobre o tema.

por Lula, com 12% das intenções de voto. A pesquisa foi realizada via telefone, utilizando amostragem estratificada, onde primeiro realizou-se um sorteio dos municípios e depois um sorteio dos moradores desse município. Foram entrevistadas 1.000 pessoas⁴, e o intervalo de confiança estimado foi de 95,45%, com uma margem de erro estimada em 3.2 pontos percentuais.

Em outra pesquisa realizada em setembro de 2018 pelo instituto Vox Populi, contratado pela Central Única dos Trabalhadores (CUT), o candidato Fernando Haddad aparece, quando informado que era apoiado por Lula, com 22% das intenções de voto. A pesquisa foi realizada presencialmente por entrevistadores, sendo que foram entrevistadas 2.000 pessoas, utilizando amostragem estratificada, com aleatoriedade de municípios e eleitores dentro dos municípios. O intervalo de confiança estimado foi de 95% e a margem de erro estimada em 2.2 pontos percentuais⁵.

Em relação às duas pesquisas citadas acima, a imagem abaixo foi amplamente divulgada nas redes sociais.



Fonte: <https://www.ocafezinho.com/2018/09/13/cut-vox-poe-haddad-a-frente-de-bolsonaro/>

O erro foi comparar duas pesquisas que utilizaram metodologias de coleta de dados diferentes. A pesquisa de julho/2018 foi realizada por telefone, e a pesquisa de setembro/2018 foi realizada presencialmente. Há uma grande discussão na literatura,

4 As informações constam no site do TSE e a pesquisa foi registrada com o seguinte identificador: BR-07756/2018.

5 As informações constam no site do TSE e a pesquisa foi registrada com o seguinte identificador: BR-01669/2018.

na área de opinião pública, sobre a diferença nos resultados de pesquisas realizadas presencialmente ou por telefone. O que os estudos indicam é que as pesquisas realizadas por telefone reduzem o viés de desejabilidade social, isso é, a atitude, até mesmo inconsciente, do entrevistado de moldar sua opinião para o que ele supõe ser mais aceitável socialmente (GRIMM, 2010). Assim, a comparação de pesquisas com diferentes metodologias de coleta de dados, indicando um “crescimento” da intenção de votos em Fernando Haddad, é falha. No caso são duas fotos, de momentos diferentes, usando diferentes lentes na câmera, impossibilitando inferir que a mudança não tenha ocorrido por causa da diferença de metodologia, e não por outros fatores como o tempo de campanha (e no caso mais específico a indicação oficial de Haddad como candidato do PT).

Outro erro cometido pelos apoiadores do PT e pelo diretor do instituto Vox Populi foi considerar o resultado da pesquisa realizada em setembro de 2018 como um indicador da intenção real de voto em Fernando Haddad naquele momento. Em entrevista para Carta Capital⁶ o diretor diz: “Esconder o fato de que o ex-prefeito foi indicado e tem o apoio do ex-presidente tornaria irreal o resultado de qualquer levantamento. É uma referência relevante para uma parcela significativa dos cidadãos. Chega perto de 40% a porção do eleitorado que afirma votar ou poder votar em um nome apoiado por Lula”. Assim, o que o instituto e a pesquisa mostram, na verdade, é a capacidade máxima de transferência de votos de Lula para Haddad naquele momento.

Essa diferença de análise ocorre pelo simples fato de que a pesquisa informou ao eleitor que Haddad era apoiado por Lula, ou seja, dentre aqueles entrevistados selecionados todos tinham essa informação, o que difere da realidade. Apesar de campanhas eleitorais, supor que todo cidadão brasileiro possuía a informação de que Lula apoiava Haddad era uma suposição errônea. Ao realizar uma pesquisa com a pergunta indicando o apoio de Lula à Haddad, o entrevistador (ou pesquisador) estava dando ao respondente uma informação que ele poderia não ter. Não seria possível, então, concluir com uma pesquisa realizada dessa forma que os resultados representavam a intenção real de votos em Haddad naquele momento. Essa diferença na pergunta pode explicar a diferença de resultado entre essa pesquisa e as demais pesquisas que não colocaram na pergunta o fato de Haddad ser apoiado por Lula.

Outro caso que é recorrente na mídia é a divulgação de pesquisas usando amostras não-probabilísticas e sua comparação com amostras probabilísticas. Numa re-

6 <https://www.cartacapital.com.br/politica/associado-diretamente-a-lula-haddad-soma-22-e-ultrapassa-bolsonaro/>

portagem da revista Metr  poles⁷ do dia 11 de abril de 2019, divulgaram-se resultados de uma pesquisa sobre o apoio ao Presidente Jair Bolsonaro. Na referida pesquisa o Presidente aparecia com aprova  o de 79%, diferente de outras pesquisas realizadas sobre aprova  o presidencial. Ao longo da reportagem, o jornalista compara esses resultados com os resultados de outras pesquisas realizadas de forma tradicionais (e.g. instituto Ibope), ou seja, com uma amostragem probabil  stica. Ainda, o jornalista defende que a pesquisa apresentada na reportagem seria mais vantajosa que as pesquisas tradicionais no tocante ao seu objetivo, investigar a aprova  o presidencial nas redes sociais.

Na parte em que descreve a metodologia, que    o que nos interessa aqui, alguns pontos podem ser abordados com o que j   foi visto nas se  es anteriores. No primeiro, pode-se concordar com a reportagem: pesquisas tradicionais que utilizam amostras probabil  sticas possuem objetivo de serem representativas, desde que cada indiv  duo de uma popula  o tenha probabilidade igual de responder ao survey. Assim, utilizando amostras probabil  sticas e representativas temos um retrato mais fiel das varia  es na popula  o.

O problema da reportagem, por  m,    quando se justifica a amostragem snowball. J   dissemos aqui, o snowball    utilizado para o acesso de pessoas que dificilmente s  o acessadas ou quando o financiamento da pesquisa n  o permite a realiza  o de uma amostragem probabil  stica. N  o inclui, portanto, o objetivo de incluir o m  ximo de pessoas poss  veis. Podemos ter como objetivo essa inclus  o em amostras probabil  sticas, por exemplo, por um pre  o muito alto. Assim, se o objetivo era relacionar o snowball com o n  mero de pesquisados, seria melhor reformular e relacionar a t  cnica com o objetivo de atingir o maior n  mero de pessoas com um baixo custo.

O problema mais grave, por  m, aparece na senten  a: “Assim, a cada novo participante que convida seus conhecidos, a abrang  ncia dos resultados aumenta”. Se o significado de abrang  ncia for de generaliza  o, a reportagem est   completamente errada. Amostras n  o-probabil  sticas, como o snowball, n  o permitem generaliza  o, pois as amostras cont  m vieses pela falta de representatividade. A precis  o das medidas tem rela  o com o tamanho da amostra, como vimos, por  m    preciso ter variabilidade, o que amostras como snowball t  m grandes chances de n  o proporcionar. Portanto, a afirma  o    imprecisa e n  o tem embasamento nos estudos cient  ficos no campo de opini  o p  blica.

7 Para acessar    reportagem: <https://www.metropoles.com/brasil/na-contram  o-de-institutos-pesquisa-mostra-bolsonaro-aprovado-por-72>

Outra afirmação que consta na entrevista e que é imprecisa é que o acúmulo de dados se torna útil para a pesquisa. Em termos de margem de erro, novamente, essa afirmação é errônea. O acúmulo de participantes é sempre bom e o resultado da amostra, no caso, é de 7 mil, como exalta a reportagem. Entretanto, esse acúmulo nem sempre é útil, já que a relação entre a eficiência da medida (erro-padrão) e o tamanho amostral não é linear. Em determinado ponto, próximo aos 2.000 participantes, os retornos em nível de precisão diminuem ao ponto que o esforço em acrescentar mais unidades não gera mais eficiência na medida.

Como ressaltado anteriormente, há um debate sobre o uso das redes sociais como uma fonte de dados. Novamente, a quantidade não significa qualidade, e uma amostra não-probabilística não tem a capacidade de representar uma população de pessoas, mesmo que essa população seja os usuários da rede social. Existe uma probabilidade muito grande de que pessoas com características diferentes desta pequena rede acessada pela pesquisa da reportagem tenha opiniões próprias e destoantes sobre o Presidente da República. A utilização do snowball, mesmo na internet, pode concentrar uma característica apenas para uma amostra e superestimar alguns dos valores de comportamento.

Considerações Finais

A mensagem central deste capítulo é que o analista de pesquisa de opinião deve ler com cuidado os questionários aplicados e a metodologia. Esse cuidado deve anteceder qualquer análise. Já para os interessados em aplicar surveys, não se fez aqui um guia exaustivo, mas sim um roteiro simplificado sobre preocupações básicas ao usar o instrumento de pesquisa. Seguindo a ordem apresentada ao longo deste texto, destacamos os principais pontos levantados: perguntas no questionário, amostragem e a aplicação.

Perguntas e respostas: ao analisar os resultados de uma pesquisa de opinião, o leitor deve observar qual foi a pergunta realizada pelo instituto de pesquisa/pesquisador e quais as respostas possíveis. Será que a pergunta foi escrita da forma correta? Será que aquilo que o pesquisador perguntou é aquilo que ele queria medir? O conceito medido é controverso ou complexo? As opções de resposta eram plausíveis? Alguma opção continha informação que outra não continha?

Amostragem e forma de aplicação: o leitor deve, também, ao analisar os dados de uma pesquisa, verificar qual tipo de amostra foi utilizada e qual foi seu meio de aplicação. A amostra utilizada é representativa da população (probabilística x não probabilística)? Qual a implicação da minha amostragem nos resultados da pesquisa? Os resultados podem ser interpretados como a população ao todo? O tamanho da amostra é razoável? A aplicação da pesquisa foi por telefone, face a face, ou pela internet? Qual é a implicação disso em como as pessoas se portariam na pesquisa?

Ao se atentar para essas perguntas, o interessado pode fazer uma melhor análise dos dados fornecidos pelos institutos de pesquisa. Como aprendemos no ensino básico, temos que comparar bananas com bananas, e maçãs com maçãs. Por mais que seja difícil realizar inferências, até mesmo com métodos estatísticos avançados, quando comparamos pesquisas de opinião já consolidadas, devemos tomar cuidado para comparar pesquisas que sejam minimamente comparáveis, afinal, dizer que uma banana difere de uma maçã, pois choveu menos naquele período, é uma conclusão muito pouco confiável.

Referências bibliográficas

- BOWLING, A. Techniques of questionnaire design. Handbook of health research methods: investigation, measurement and analysis, p. 394-427, 2005.
- GRIMM, P. Social desirability bias. Wiley international encyclopedia of marketing. 2010.
- KROSNICK, J. A.; PRESSER, S. Handbook of Survey Research, p. 1432-1033. 2010.
- TOURANGEAU, R.; RASINSKI, K. A. Cognitive processes underlying context effects in attitude measurement. Psychological bulletin, v. 103, n. 3, p. 299, 1988.
- TOURANGEAU, R.; RIPS, L. J.; RASINSKI, K. The psychology of survey response. Cambridge University Press. 2000.



Brasília
-2021-