

Pesquisas eleitorais no Brasil: tendências e desempenho

FERNANDO MEIRELES^I

GUILHERME RUSSO^{II}

Introdução

NA VÉSPERA do primeiro turno das eleições presidenciais de 2018, dois dos principais institutos de pesquisa do país, Ibope e Datafolha, apontavam Jair Bolsonaro, do PSL, com cerca de 41% dos votos válidos contra 25% de Fernando Haddad, do PT. Apuradas as urnas no dia seguinte, entretanto, descobriu-se que Bolsonaro tinha obtido nada menos do que 46,03% dos votos válidos, 5 pontos percentuais a mais, portanto, do que o indicado pelas pesquisas do dia anterior. Em alguns estados, diferenças entre pesquisas e resultados foram ainda mais elásticas: o até então desconhecido juiz Wilson Witzel (PSC), que aparecia em terceiro lugar com meros 17% de votos válidos na pesquisa Datafolha de véspera, encerrou o domingo de votação com 41,28% dos votos, isto é, cerca de 25 pontos percentuais a mais do que sugerem as pesquisas.

Resultados destoantes de pesquisas eleitorais,¹ que apontam cenários muito distintos daqueles confirmados posteriormente pelas urnas, não são exclusividade brasileira. Nos Estados Unidos, o erro da *The Literary Digest*, que coletou mais de 2 milhões de respostas e mesmo assim errou quem sairia vencedor das eleições presidenciais americanas de 1936, até hoje é mencionado como um rematado exemplo de erro de pesquisas eleitorais – nesse caso, por sobrerrepresentar em sua amostra a classe média americana, principal público da revista (Squire, 1988). A democrata Hillary Clinton, apesar de estar confortavelmente na frente de Donald Trump no voto popular de acordo com as pesquisas para as eleições presidenciais americanas de 2016, foi derrotada no colégio eleitoral; pior, esse viés pró-democratas perdurou nas eleições de 2020 vencidas por Joe Biden, quando pesquisas novamente registravam um desempenho melhor para o seu partido do que o efetivamente obtido (Gelman, 2021). O Brexit, que pesquisas disseram que perderia no Reino Unido, venceu.

A literatura sobre pesquisas eleitorais sugere que há fatores que ajudam a explicar diferenças entre pesquisas e resultados oficiais. Analisando mais de 4 mil levantamentos estaduais para a Presidência, Senado e governos estaduais nos Estados Unidos, por exemplo, Shirani-Mehr et al. (2018) mostram que margens de erro reportadas lá, em média, subestimam o erro mensurado a partir dos resultados revelados pelas urnas, o que, em parte, é decorrência de problemas logísticos e operacionais comuns que pesquisas enfrentam, isto é, erros não-

-aleatórios. Além disso, Shirani-Mehr et al. (2018) notam que margens reportadas geralmente não levam em conta detalhes como o uso de pós-estratificação, que tendem a inflá-las – algo conhecido como *design effect*. No conjunto, desse modo, pesquisas sofreriam com mais fatores de erros do que o predito apenas por variação aleatória, o que não seria adequadamente capturado por margens de erro tradicionais.

Esse diagnóstico resume bem os diferentes tipos de erros que pesquisas de *survey* enfrentam. Por um lado, levantamentos eleitorais precisam usar desenhos amostrais que abarquem toda a população votante, o que nem sempre é fácil levando-se em conta a desatualização de bases de dados eleitorais e demográficas em vários países ao redor do mundo (Lupu; Michelitch, 2018). Por outro lado, pesquisas também precisam superar problemas de abordagem que incluem obter respostas de diferentes grupos demográficos, conseguir que respondentes entendam e respondam adequadamente às questões previstas e que não omitam informações por vergonha ou medo (algo conhecido como *social desirability bias*) (Carlson, 2018; Brownback; Novotny, 2018). Essas não são tarefas simples, e há trabalhos comparativos que documentam como tais problemas afetam taxas de erro de pesquisas.

Um exemplo relevante na literatura envolve o desempenho de pesquisas pré-eleitorais quando taxas de resposta são pequenas. Especificamente, pesquisas telefônicas com o uso de discagem de números aleatórios (RDD), método probabilístico de recrutamento de respondentes canônico nos Estados Unidos e Reino Unido, passaram a sofrer com a quantia de pessoas que se recusam a dar entrevistas, o que é problemático quando os não respondentes são diferentes do restante da população em características relacionadas com as respostas analisadas. Enquanto em 1963 a taxa de resposta do *British Election Study* beirava os 80%, não passou de 55% em 2015; algo similar também aconteceu com as pesquisas telefônicas nos Estados Unidos (Curtin et al., 2005; Prosser; Mellon, 2018). Simultaneamente, mudanças de atitude na véspera de pleitos, especialmente quando há altas taxas de eleitores indecisos, também explicariam erros, preocupação que se soma à hipótese dos *shy voters*, eleitores que não revelariam suas preferências em pesquisas – o que alegadamente seria o caso de eleitores trumpistas no Estados Unidos, ainda que a evidência existente sobre esse caso seja mista (Coppock, 2017; Brownback; Novotny, 2018).

Não obstante a existência de fatores relacionados aos erros em *surveys* eleitorais, a evidência sistemática de diferentes países mostra que pesquisas nas semanas que antecedem um pleito fornecem, em média, estimativas razoavelmente precisas do que a apuração dos votos revelará – erros têm, ao contrário, apenas diminuído ao longo dos anos (Jennings; Wlezien, 2018). Em certo sentido, dadas as dificuldades envolvidas de se desenhar planos amostrais com informações nem sempre atualizadas e em meio a diferentes fontes não aleatórias de erro, é surpreendente o desempenho das pesquisas eleitorais ao antecipar

tendências e resultados, ainda que esse não seja o objetivo ou a função das pesquisas (Gelman, 2021).

No Brasil, evidência sobre levantamentos eleitorais é escassa, o que não impede o debate público de formar uma avaliação, em geral negativa, dos seus resultados. Como sabido, erros ao longo dos últimos anos já levaram até mesmo ao Supremo Tribunal Federal (STF) a declarar inconstitucionais leis aprovadas pelo Congresso para proibir a divulgação de pesquisas eleitorais (ver, por exemplo, Nunes et al., 1993). A despeito disso, El-Dash (2010), em trabalho que analisa pesquisas eleitorais no Brasil até 2010, mostra que pesquisas não probabilísticas por sistema de quotas – o tipo de pesquisa eleitoral mais comum no Brasil, conforme descrevemos adiante – obtiveram resultados tão bons quanto pesquisas inteiramente probabilísticas (ver também Gramacho, 2013). Salvo poucas exceções como essas, contudo, falta evidência, especialmente para o período mais recente, para avaliar se as nossas pesquisas fazem jus às críticas recorrentes que sofrem.

Neste artigo, procuramos oferecer essa avaliação. Em particular, analisamos o desempenho das pesquisas eleitorais no Brasil entre 2012 e 2020, data a partir da qual o Tribunal Superior Eleitoral (TSE) passou a publicar informações sobre todas as pesquisas eleitorais registradas em anos eleitorais. A esse catálogo que contempla informações amostrais de mais 34 mil pesquisas eleitorais registradas no país, adicionamos estimativas de mais de 2 mil dessas pesquisas coletadas a partir de duas fontes diferentes: o repositório do jornal *Poder360*, do qual obtivemos dados brutos que organizamos; e do jornal *Pindograma*, que mantém uma base própria de pesquisas eleitorais.

Nossos dados revelam padrões importantes sobre as pesquisas eleitorais no país. Em primeiro lugar, documentamos uma crescente no uso de abordagens telefônicas no período, bem como ligeiro aumento no tamanho das amostras utilizadas ao longo do tempo em eleições nacionais e estaduais. Quanto aos seus resultados, as estimativas eleitorais indicam que o erro médio das nossas pesquisas de véspera é cerca de 1.8 ponto percentual (pp), o que pode ser traduzido em uma margem de erro de 6.8 pp, maior do que o convencionalmente reportado por institutos. Isso dito, tal taxa é comparável à de outros países e, mais, há grande variação entre pesquisas: as de presidentes e de governo de estado no segundo turno cometem, em média, erros menores, assim como pesquisas com maiores amostras e em pleitos com menores taxas de eleitores indecisos.

Conforme argumentamos na conclusão, todos esses resultados sugerem que, apesar de as pesquisas brasileiras cometerem mais erros do que o previsto pelas margens de erro tipicamente registradas, têm nível de acerto similar as de outros países – e mesmo melhores, caso de eleições majoritárias nacionais, especialmente em segundo turno.

Pesquisas eleitorais no Brasil

Desenhos amostrais

No Brasil, a maioria das pesquisas eleitorais é baseada em desenhos amostrais não probabilísticos com seleção de respondentes por meio de sistemas de cotas (El-Dash, 2010). Enquanto em desenhos amostrais probabilísticos a probabilidade de sortearmos uma dada amostra dentre todas as possíveis é conhecida de antemão – e, conseqüentemente, a probabilidade de inclusão de cada pessoa também é –, a característica principal dos desenhos não probabilísticos é que não é conhecida a probabilidade de acabarmos com determinada amostra. Na prática, isso ocorre porque a seleção de pessoas a serem entrevistadas é intencional, ou seja, na ausência de listagens atualizadas e exaustivas das pessoas que integram o eleitorado, entrevistadores e entrevistadoras decidem quem será incluído na amostra. Para tanto, no Brasil assim como em outros países, conveniona-se usar quotas que determinam a proporção de pessoas de dado estrato populacional – sexo, faixa etária, renda – para orientar essa escolha (Lupu; Michelitch, 2018). Ao fim, quando tudo sai como planejado, amostras não probabilísticas feitas por meio de quotas acabam com a mesma proporção de pessoas de determinados grupos sociodemográficos que a população de interesse.

Quanto ao modo de coleta de dados, entrevistas presenciais, face a face (F2F), são o formato principal adotado pelos institutos de pesquisa brasileiros, assim como em outras democracias recentes (Lupu; Michelitch, 2018). Variações à parte, os desenhos amostrais mais usados no país em geral combinam esse formato de coleta com etapas probabilísticas de sorteio de locais de aplicação, *clusters* ou conglomerados, com uma etapa final de recrutamento intencional de respondentes por quotas. O desenho usado em várias eleições pelo Ibope e por muitos outros institutos segue, em geral, as seguintes etapas: sorteio de um determinado número de municípios a serem visitados usando-se Probabilidade Proporcional ao Tamanho (PPT), que dá maior chances a municípios com maior número de habitantes ou de eleitores registrados de serem incluídos na amostra; sorteio de um determinado número de bairros ou de setores censitários, dentro dos municípios sorteados na etapa anterior, também por PPT; e, finalmente, dentro dos locais sorteados na etapa anterior, recrutamento de respondentes que preencham as cotas demográficas definidas pela pesquisa, o que geralmente inclui um número fixo de pessoas a serem entrevistadas de diferentes sexos, faixas etárias, grau de instrução educacional e renda pessoal ou domiciliar.

Esse tipo de desenho amostral, que combina diferentes etapas probabilísticas ou não probabilísticas, tem algumas vantagens no contexto brasileiro (El-Dash, 2010). Do ponto de vista logístico, é menos custoso restringir geograficamente a aplicação de questionários em um número pequeno de municípios e de bairros ou setores do que tentar cobrir uma grande área geográfica – o que, em um país com as dimensões do Brasil, é inviável para pesquisas que têm campos com duração de apenas 3 ou 4 dias. Adicionalmente, com uma população

marcada por uma elevada desigualdade de renda e falta de acesso à internet e à telefonia móvel, por muito tempo o melhor meio de acesso à maior parte da população foi o contato pessoal, especialmente domiciliar. Por outro lado, esse tipo de desenho amostral tem negativos, como a dificuldade de acesso a determinadas localidades – condomínios fechados, zonas de milícia, áreas rurais distantes. De toda a forma, conforme sustenta El-Dash (2010), não é possível afirmar que esse formato de amostra seja pior que outra totalmente probabilística do ponto de vista empírico.

Vale notar que as pesquisas eleitorais no Brasil têm alguns de seus aspectos regulados pela legislação eleitoral. Entre outros, institutos de pesquisa que queiram divulgar suas estimativas precisam fazer o registro do plano amostral a ser implementado com antecedência, reportando as datas em que a coleta dos dados será feita, a data de divulgação dos resultados, o tamanho da amostra e detalhes de metodologia como uso de quotas e discriminação dos locais onde entrevistas são aplicadas. Em razão disso, por exemplo, margens de erro divulgadas no momento do registro das pesquisas não podem contemplar o efeito do uso de pós-estratificação para corrigir desvios nas amostras, uma vez que o efeito de tais procedimentos nas margens de erro só podem ser quantificados após o encerramento da coleta de dados – o que contribui para subestimar possíveis erros aleatórios e não aleatórios, isto é, erros totais (Gramacho, 2013).

Tendências nas pesquisas eleitorais registradas

Para mapear tendências no uso de pesquisas eleitorais no Brasil, recorremos à base de pesquisas registradas mantidas pelo Tribunal Superior Eleitoral (TSE). Por força da legislação eleitoral, como dito anteriormente, institutos registram seus desenhos amostrais em relação à população alvo das pesquisas, ao método de coleta, plano amostral, tamanho da amostra e discriminação de locais de aplicação e cotas sociodemográficas. A partir das eleições de 2012, esses dados foram disponibilizados pelo TSE em seu repositório.

Dos arquivos das eleições de 2012, 2014, 2016, 2018 e 2020, coletamos 36.608 registros de pesquisas. Esses registros, entretanto, não necessariamente correspondem a pesquisas únicas: por conta de problemas de preenchimento ou de registro, algumas pesquisas acabam recebendo mais de um número de registro; além disso, levantamentos que divulgam estimativas para eleições presidenciais e de governos estaduais, ainda que baseadas no mesmo plano amostral, podem receber dois registros diferentes. Por essa razão, agregamos os registros por data de início e fim do período de coleta de dados, local de aplicação dos questionários e texto de descrição de metodologia e plano amostral, o que resultou em número menor de pesquisas, 34.373. É com base nesse banco agregado de registros que reportamos estatísticas descritivas sobre os padrões de pesquisas eleitorais no Brasil.

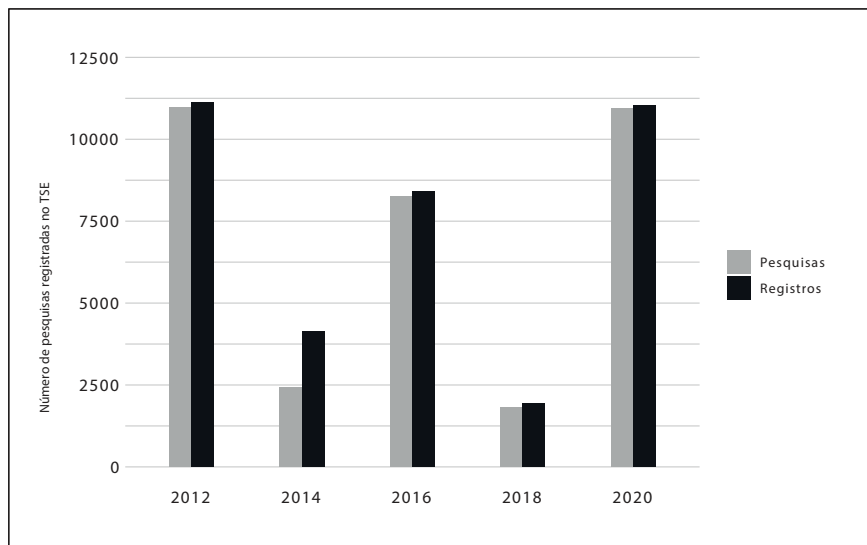
No período analisado, 4.220 pesquisas, 12% do total, foram realizadas em eleições nacionais e estaduais, contra 30.153 realizadas em eleições municipais,

o que dá uma média de cerca de 10 mil pesquisas por eleição municipal contra 2.110 em eleições nacionais e estaduais. O Gráfico 1 exibe a distribuição dessas pesquisas ao longo dos pleitos analisados, separando entre pesquisas agregadas e registros individuais – o que deixa evidente que, em 2014, houve muitos mais registros do que pesquisas, isto é, algumas pesquisas foram registradas mais de uma vez.

O Gráfico 1 também mostra padrões relevantes ao longo do período. Em primeiro lugar, é possível notar uma diminuição no número de pesquisas entre as eleições nacionais e estaduais de 2014 e de 2018, de 2.419 para 1.801 pesquisas, respectivamente. Nesse último pleito, realizado depois da aprovação reforma eleitoral pela Lei n.13.165/2015, o período de campanha foi reduzido de 90 para 45 dias, o que pode ter contribuído para a ligeira variação – ainda que pesquisas eleitorais possam ser divulgadas, desde que registradas no TSE, ao longo de todo o ano eleitoral, mesmo antes do período oficial de campanha, portanto. De 2012 para 2016, também houve redução no número de pesquisas, de 10.964 para 8.250, movimento que só é revertido em 2020, quando 10.939 pesquisas foram registradas, aumento de 32%. O ano 2020 também foi o de maior número de pesquisas telefônicas registradas, 290, contra 170 de 2018, e apenas 9, em 2016. Apesar do crescimento no período, estas ainda compreendem apenas uma minoria das pesquisas eleitorais registradas no TSE.

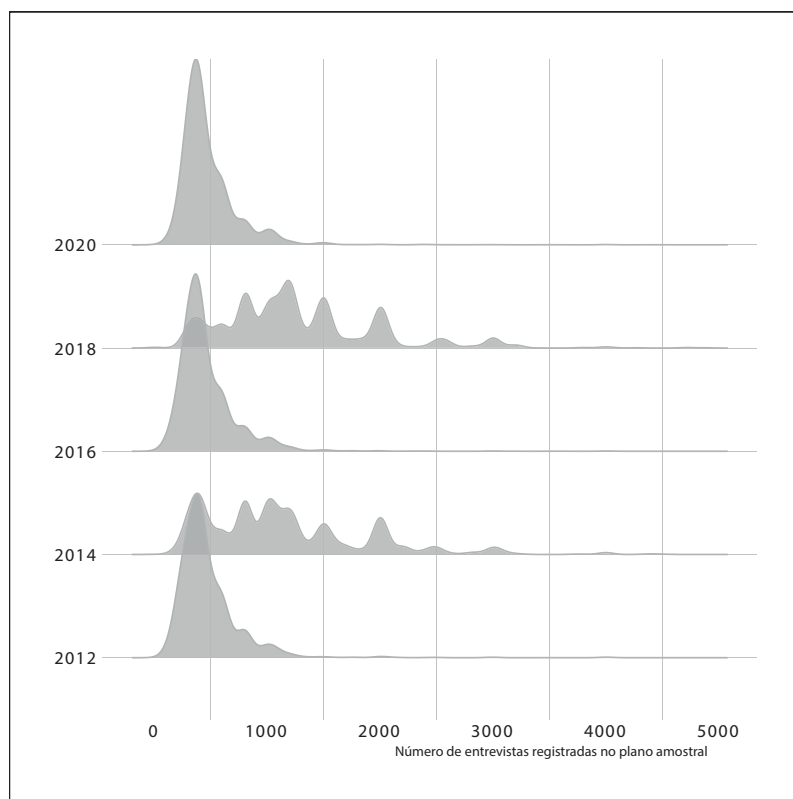
Outra informação importante nos dados no TSE, reportada no Gráfico 2, é o do tamanho das amostras registradas, um dos fatores que influencia a precisão das estimativas de uma pesquisa (mas que não influencia o viés dessas estimativas).² Em anos de eleição municipal, a maioria dos planos amostrais previa cerca de 450 entrevistas, o que, a título de ilustração, resulta em uma margem máxima de erro de 4.6% para totais amostrais.³ Exceções incluem pesquisas realizadas por alguns institutos tradicionais, no geral, em grandes capitais, como São Paulo e Rio de Janeiro. De outro modo, eleições nacionais/estaduais são marcadas por uma distribuição multimodal em pontos como 1.200, 1.500 e 2.000 entrevistas. Nos dois casos, entretanto, houve aumento na média de entrevistas no período: em pleitos municipais, de 18 entrevistas entre 2016 e 2020; e de 202 entrevistas entre 2014 e 2016.

Em resumo, esses dados mostram que, após uma queda em 2016 e 2018, o número de pesquisas registradas voltou a patamares de 2012; pesquisas telefônicas, embora ainda pouco praticadas, vêm ganhando espaço rapidamente, e houve pequeno aumento no tamanho das amostras registradas no TSE, especialmente em ano de eleições nacionais e estaduais. Essas são tendências que sinalizam alterações no mercado de pesquisas eleitorais no Brasil: para uma maior pluralidade de modos de coleta de dados e, ao mesmo tempo, de desenhos amostrais adotados pelos institutos de pesquisa.



Fonte: Elaboração própria a partir de dados do Tribunal Superior Eleitoral (TSE).

Gráfico 1 – Distribuição de pesquisas e registros eleitorais no Tribunal Superior Eleitoral, 2012–2020.



Fonte: Elaboração própria a partir de dados do Tribunal Superior Eleitoral (TSE).

Gráfico 2 – Distribuição do tamanho das amostras registradas no TSE, 2012–2020.

Comparando estimativas e resultados eleitorais

Para comparar estimativas de pesquisas eleitorais com os resultados efetivamente obtidos nas urnas nas eleições brasileiras, usamos duas fontes de dados. Primeiro, do repositório de pesquisas eleitorais do Poder360,⁴ que disponibiliza o maior conjunto público de dados de pesquisas no país, coletamos as estimativas divulgadas de 4303 pesquisas realizadas entre 2012 e 2022.⁵ Esses dados, apesar de volumosos, cobrem apenas 12% das pesquisas registradas no período e, além disso, contêm diversos erros de imputação manual. Para organizar os dados e compará-los com os resultados eleitorais do TSE, primeiro selecionamos apenas cenários estimulados das pesquisas e removemos aqueles em que a soma das intenções de voto divulgadas não continham os nomes das candidaturas que efetivamente concorreram ou que não somavam entre 98 e 102 pontos percentuais; finalmente, cruzamos as pesquisas restantes com a base de pesquisas registradas do TSE por meio do código de registro e, também, com dados de resultados eleitorais do TSE. Considerando apenas pesquisas realizadas 30 dias antes dos pleitos, isso resultou em uma base de 790 pesquisas, 2% do total registrado no período, nas quais foi possível identificar registro no TSE, tamanho da amostra utilizada, estimativas divulgadas e resultados nas urnas.

Apesar de consistente, essa base de pesquisas, após exclusões e cruzamentos com dados oficiais, cobre uma parcela diminuta das pesquisas registradas no país. Por essa razão, para termos uma alternativa com cobertura maior, analisamos a base de pesquisas eleitorais criada e mantida pelo Pindograma, plataforma jornalística que coleta e organiza pesquisas eleitorais no Brasil realizadas desde 2012. Em particular, essa base inclui estimativas que foram extraídas tanto de relatórios publicados pelos institutos de pesquisa quanto de notícias buscadas no *Bing* a partir dos códigos de registro no TSE. Segundo o Pindograma, 97% dos dados coletados independentemente e que são comuns à base do Poder360 são idênticos. Ao todo, essa base inclui 6.936 pesquisas eleitorais registradas, 20% do total registrado no TSE no período.

Para avaliar o desempenho de cada uma das pesquisas nas nossas bases com os resultados eleitorais finais, adotamos o mesmo indicador de Shirani-Mehr et al. (2018) que consiste em calcular a diferença de cada estimativa e o resultado oficial (e.g., estimativa de intenção de votos da candidatura menos sua votação efetivamente obtida) e, destas diferenças, calcular a raiz do erro quadrado médio (ou RMSE, de *root mean squared error*):

$$RMSE = \sqrt{\frac{1}{N} \sum_{i,p=1}^N (y_{ip} - v_{ip})^2}$$

onde i indica as candidaturas e p cada uma das eleições cobertas pelas pesquisas nas nossas bases de dados. Diferentemente de outras medidas, como erro

absoluto médio (El-Dash, 2010; Mosteller et al., 1949), essa penaliza desvios maiores (em função do termo quadrático) e, além disso, serve como uma referência empírica de variação que pode ser usada para se aproximar margens de erro reais.

A Tabela 1 exibe os resultados do RMSE médio para cada eleição nas nossas bases. Em pesquisas de véspera, o RMSE fica, em média, em cerca de 0.035 nas duas bases, o que corresponde a uma margem de erro de 6.8 pontos percentuais. Em perspectiva comparada, essa margem *vis-à-vis* resultados é comparável a dos Estados Unidos e, importante ressaltar, é compatível com o erro médio de pesquisas encontradas em outras democracias (Jennings; Wlezien, 2018). Dizendo de outra forma, mesmo ultrapassando margens de erro tradicionais, calculadas para amostras probabilísticas, a evidência de duas bases de pesquisas eleitorais de véspera no Brasil indica que elas têm desempenhos similares ao de pesquisas conduzidas em outros países.

Tabela 1 – Raiz do erro quadrado médio (RMSE) das pesquisas eleitorais no Brasil, 2012–2020

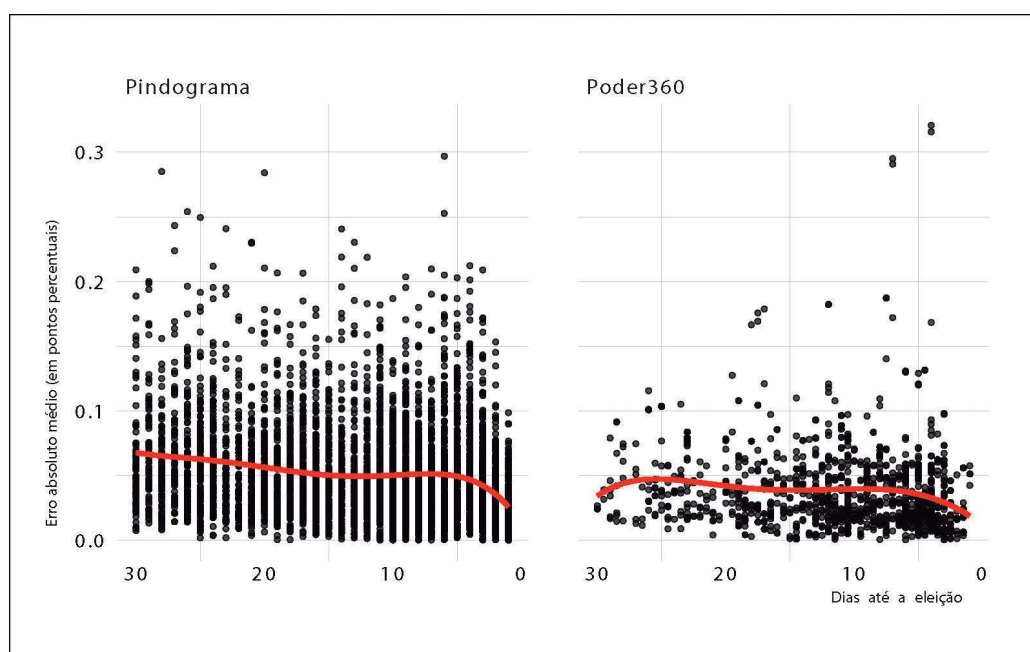
Eleição	Poder360		Pindograma	
	Véspera	7 dias antes	Véspera	7 dias antes
2012	0.027	0.04	0.036	0.039
2014	0.008	0.019	0.035	0.038
2016	0.025	0.033	0.039	0.053
2018	0.031	0.04	0.037	0.041
2020	0.035	0.057	0.047	0.057

Fonte: elaboração própria a partir de dados do Tribunal Superior Eleitoral (TSE), Poder360 e Pindograma.

Esse desempenho médio das pesquisas, no comparativo com resultados nas urnas, varia em razão do cargo em disputa e do turno analisado. Considerando o RMSE calculado para pesquisas de véspera em alguns destes recortes usando dados da base do Pindograma, que possui maior número pesquisas, é possível constatar, por exemplo, que pesquisas de segundo turno apresentam erros menores do que as de primeiro turno, de 0.029 contra 0.038, respectivamente (o que se traduziria em margens de erro de 5.8 pp e de 7.9 pp, na mesma ordem). Refazendo essa análise por cargo, a diferença é similar: em pesquisas para a Presidência, o RMSE de primeiro turno (0.039) é maior do que o de segundo (0.030); para governos estaduais, idem (0.036 contra 0.026). O número de pesquisas que se enquadram nesses critérios, contudo, é muito menor.

Entre outros fatores não aleatórios que afetam as pesquisas no Brasil, mudanças de atitudes e opiniões ao longo do tempo importam. Para se ter uma dimensão desse efeito, calculamos o erro absoluto médio de cada pesquisa nas nossas amostras e estimamos o efeito predito dos dias até eleição sobre essa quantia, conforme no Gráfico 3. Na imagem, a curva laranja indica um modelo

linear com polinômio 5, o que capta melhor tendências locais, como fica nítido na faixa final do eixo X que é quando faltam poucos dias para as eleições. Particularmente, pesquisas de véspera têm erros absolutos médios preditos de 1.8 pontos percentuais (pp), o que fica dentro de margens de erro convencionais para amostras com 2.000 entrevistas, por exemplo, com um intervalo de confiança superior (95%) de 3pp. Quanto mais distante a data de realização de uma pesquisa da data do pleito, maior é o erro predito: 3 dias antes do pleito, é 2.9 pp; 15 dias antes, de 3.8 pp. Isso reforça a importância de se tomar pesquisas eleitorais como retratos, *snapshots*, de momentos de uma corrida eleitoral, e não como uma ferramenta de previsão de resultados eleitorais, especialmente muitos dias antes das datas de votação.



Fonte: elaboração própria a partir de dados do Tribunal Superior Eleitoral (TSE), Poder360 e Pindograma.

Gráfico 3 – Erro absoluto médio das estimativas de pesquisas eleitorais por data de realização, 2012–2020.

O que explicam esses erros das pesquisas ligeiramente acima do que o previsto por margens de erro convencionais? Em parte, como já abordamos, tanto por razões de legislação eleitoral, que obrigam institutos a registrar margens de erro antes de ter dados totalmente coletados – com frequência, antes mesmo do término de campos –, quanto pelo uso de desenhos amostrais não probabilísticos, parte da explicação tem a ver com a régua utilizada para avaliar pesquisas: se formos usar margens de erro que não são totalmente adequadas, a comparação desfavorece as pesquisas brasileiras, ainda que elas, como vimos, apresentem erros similares ao de outras conduzidas em outros contextos (Gramacho, 2013).

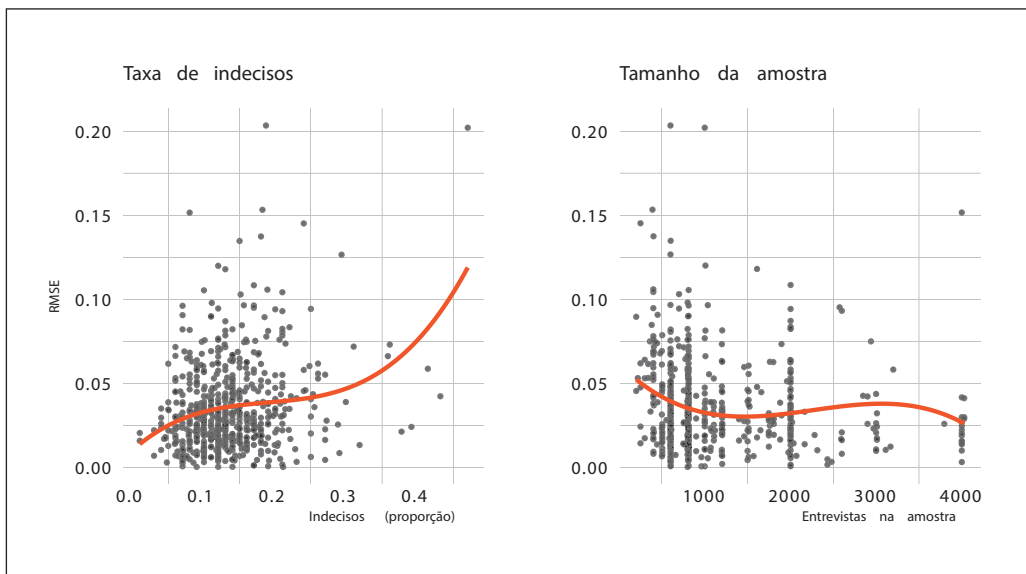
Sob essa perspectiva, o problema reside mais nas margens de erro, e menos nas estimativas divulgadas pelos nossos institutos.

Para além das considerações sobre o cálculo de margens de erro, há outros fatores relacionados com os erros de pesquisas que conseguimos investigar com os dados da base de pesquisas do Pindograma. O primeiro deles é o tamanho das amostras utilizadas – que impactam diretamente a variação esperada das estimativas obtidas. Já o segundo fator tem a ver com atitudes, especificamente de eleitores(as) indecisos que, com a proximidade dos pleitos, podem pender para um lado ou outro no intervalo entre pesquisas de véspera e dias de votação em função da aquisição de novas informações no decorrer da campanha ou por movimentos do tipo bumerangue, quando eleitores decidem seus votos em função do conhecimento de quem está na frente nas pesquisas (Gelman; King, 1993; Skalaban, 1988; Rothschild; Malhotra, 2014).

O Gráfico 4 oferece evidência visual para a relação entre estes fatores e erros mensurados por RMSE em pesquisas eleitorais de véspera. Conforme as curvas do modelo linear com polinômio indicam, taxa de indecisos elevada está positivamente associado com erros maiores, o que é consistente com a ideia de que essa parcela do eleitorado, ao optar por alguma candidatura no dia de votação, pode destoar de tendências detectadas nas vésperas (Shirani-Mehr et al., 2018). Quanto ao tamanho das amostras, embora a relação não seja saliente, efeitos preditos são sugestivos: um modelo polinomial prediz que amostras com 500 entrevistas, como na média das pesquisas em eleições municipais no país, têm RMSE de 0.044, enquanto esse valor cai para 0.033 e 0.032 para amostras de 1.000 e 2.000 entrevistas, respectivamente. Assim como os trabalhos de El-Dash (2010) e de Gramacho (2013) para períodos anteriores ao considerado aqui, portanto, encontramos evidência que reforçam a relação entre opções amostrais e erros nas pesquisas eleitorais brasileiras.

Considerações finais

Em ano de eleição, pesquisas eleitorais voltam para o centro do debate público, em parte para serem alvos de críticas. O que procuramos mostrar neste artigo é que, salvo casos excepcionais, as pesquisas eleitorais brasileiras são, em média, tão boas quanto as realizadas em outros países, o que é algo considerável dadas as barreiras logísticas, geográficas e de qualidade e atualização de informações sociodemográficas sobre a população brasileira. Especificamente, com duas bases de dados diferentes, examinamos as diferenças entre estimativas de pesquisas e resultados nas urnas apurados pelo TSE e encontramos que, em média, erros absolutos das pesquisas de véspera chegam a 1.8 pp, taxa que fica abaixo de margens de erro convencionalmente divulgadas pelos institutos. Esse resultado, no entanto, esconde uma variação maior do que a prevista por tais margens, mas é compatível, para dizer novamente, com os erros encontrados em pesquisas nos Estados Unidos e países europeus.



Fonte: Elaboração própria a partir de dados do Tribunal Superior Eleitoral (TSE).

Gráfico 4 – Relação entre taxa de indecisos, tamanho das amostras e erros em pesquisas eleitorais, 2012–2020.

Também documentamos outros padrões relevantes. Pesquisas com amostras algo maiores tornaram-se mais comuns nas últimas eleições, bem como pesquisas telefônicas, apesar de essas ainda terem pouco espaço relativamente. Entre fatores relacionados com erros, corroboramos achados de outros trabalhos e mostramos que, também em anos recentes e em uma base mais extensa de pesquisas, pesquisas oferecem estimativas mais próximas de resultados oficiais quanto mais próximo do dia da eleição, em eleições para a Presidência, em corridas de segundo turno e quando há menos eleitores indecisos. Parte do erro das pesquisas que quantificamos, portanto, têm relação com fatores não aleatórios, que não são previstos por margens de erro tradicionais. Com essa conclusão, não queremos sugerir que as pesquisas brasileiras não errem e que casos excepcionais não ocorram. Acreditamos, ao contrário, que colocar esses casos em perspectiva é útil para termos uma avaliação mais sistemática do desempenho de nossas pesquisas.

Agradecimento – O primeiro autor agradece o financiamento da Fundação de Amparo à Pesquisa de São Paulo (Fapesp), por meio de bolsa de Pós-Doutorado no Processo n.2021/01393-2, que viabilizou a realização desta pesquisa.

Notas

- 1 A rigor, pesquisas eleitorais geralmente se referem a pesquisas que estudam decisão de voto após eleições, como as pesquisas de boca de urna ou eleitorais conduzidas ao redor do mundo. Para simplificar o texto, o que chamamos de pesquisas eleitorais são, mais precisamente, pesquisas pré-eleitorais, ou *trial heat polls*, na expressão em inglês. Outra consideração importante é que, ao falar de erro de pesquisas eleitorais, estamos nos referindo a diferenças entre estimativas de pesquisas *vis-à-vis* resultados nas urnas, mas essa é uma definição particular, que depende do pressuposto de que não há variação entre as datas de realização de pesquisas e das eleições, o que dificilmente é o caso. O que fazemos, de todo o modo, é explorar fatores que nos ajudam a entender as limitações de se extrapolar estimativas de *survey* para se fazer predição.
- 2 Possivelmente por conta de erros de digitação, algumas pesquisas registraram no TSE número negativos de entrevistas ou excessivamente grande, como de mais de 400 mil entrevistas. Para extrairmos estatísticas descritivas dessa informação, excluímos amostras com menos de 0 entrevista ou com mais de 50 mil entrevistas (algumas pesquisas de véspera ou de boca de urna chegam a passar de 30 mil casos, por isso o teto mais alto de corte).
- 3 Há extenso debate sobre o cálculo de margens de erro em pesquisas eleitorais não-probabilísticas como a maioria das realizadas no Brasil. Por não ser possível derivar as probabilidades de obtenção de uma amostra não probabilística, cálculos tradicionais de margem não deveriam se aplicar totalmente, embora institutos os usem como aproximações de erros que poderiam ser obtidos. Pesquisas presenciais, além disso, sofrem efeito da conglomeração dos dados, isto é, da correlação entre respostas de pessoas entrevistadas nos mesmos setores censitários ou bairros, o que em geral aumenta a variação de estimativas.
- 4 Disponível em: <<https://www.poder360.com.br/pesquisas-de-opinio/>>. Acesso em: 4 sete. 2022.
- 5 O Poder360 disponibiliza pesquisas conduzidas em anos anteriores, mas, por conta de o TSE divulgar os registros das pesquisas feitas apenas a partir de 2012, adotamos esse recorte para podermos cruzar estimativas de pesquisas com informações efetivamente registradas no Tribunal.

Referências

- BROWNBACK, A.; NOVOTNY, A. Social desirability bias and polling errors in the 2016 presidential election. *Journal of Behavioral and Experimental Economics*, v.74, p.38-56, 2018.
- CARLSON, E. The perils of pre-election polling: Election cycles and the exacerbation of measurement error in illiberal regimes. *Research & Politics*, v.5, n.2, 2018.
- COPPOCK, A. Did shy Trump supporters bias the 2016 polls? Evidence from a nationally-representative list experiment. *Statistics, Politics and Policy*, v.8, n.1, p.29-40, 2017.
- CURTIN, R.; PRESSER, S.; SINGER, E. Changes in telephone survey nonresponse over the past quarter century. *Public opinion quarterly*, v.69, n.1, p.87-98, 2005.
- EL-DASH, N. A. *Avaliação metodológica das pesquisas eleitorais brasileiras*. São Paulo, 2010. Tese (Doutorado em Estatística) – Instituto de Matemática e Estatística, Universidade de São Paulo.

- GELMAN, A. Failure and success in political polling and election forecasting. *Statistics and Public Policy*, v.8, n.1, p.67-72, 2021.
- GELMAN, A.; KING, G. Why are American presidential election campaign polls so variable when votes are so predictable? *British Journal of Political Science*, v.23, n.4, p.409-451, 1993.
- GRAMACHO, W. G. À margem das margens? a precisão das pesquisas pré-eleitorais brasileiras em 2010. *Opinião Pública*, v.19, p.65-80, 2013.
- JENNINGS, W.; WLEZIEN, C. Election polling errors across time and space. *Nature Human Behaviour*, v.2, n.4, p.276-83, 2018.
- LUPU, N.; MICHELITCH, K. Advances in survey methods for the developing world. *Annual Review of Political Science*, v.21, n.1, p.195-214, 2018.
- MOSTELLER, F. The pre-election polls of 1948; report to the Committee on Analysis of Pre-election Polls and Forecasts. *Public Opinion Quarterly*, v.12, n. 4, p. 599-622, 1949.
- NUNES, M. C.; OLSÉN, Ö.; STRAUBHAAR, J. O uso de pesquisas eleitorais em decisões de voto. As eleições brasileiras de 1989. *Opinião Pública*, v.1, n.2, p.98-119, 1993.
- PROSSER, C.; MELLON, J. The twilight of the polls? A review of trends in polling accuracy and the causes of polling misses. *Government and Opposition*, v.53, n.4, p.757-90, 2018.
- ROTHSCHILD, D.; MALHOTRA, N. Are public opinion polls self-fulfilling prophecies? *Research & Politics*, v.1, n.2, 2014.
- SHIRANI-MEHR, H. et al. Disentangling bias and variance in election polls. *Journal of the American Statistical Association*, v.113 n.522, p.607-14, 2018.
- SKALABAN, A. Do the polls affect elections? Some 1980 evidence. *Political Behavior*, v.10, n.2, p.136-50, 1988.
- SQUIRE, P. Why the 1936 Literary Digest poll failed. *Public Opinion Quarterly*, v.52, n.1, p.125-33, 1988.

RESUMO – Este artigo analisa as estimativas de mais de 2 mil pesquisas eleitorais com os resultados de cinco eleições municipais e nacionais no Brasil entre 2012 e 2020. Em particular, examinamos como fatores previstos nos planos amostrais, como tamanho da amostra e modo de aplicação de entrevistas, e outros como a distância temporal da data de realização das pesquisas até o dia do pleito, predizem diferenças entre estimativas e resultados oficiais. Entre outros, mostramos que pesquisas de véspera com amostras maiores tiveram resultados mais próximos dos apurados nas urnas no período, assim como pesquisas conduzidas perto do dia dos pleitos. Também documentamos que estimativas de pesquisas em eleições nacionais, especialmente para a Presidência e Governos estaduais em segundo turno, tendem a ser mais próximas dos resultados do que em outras disputas. No geral, portanto, os achados indicam que as pesquisas eleitorais no Brasil têm desempenho similar às realizadas em outros contextos.

PALAVRAS-CHAVE: Pesquisas eleitorais, Survey, Eleições.

ABSTRACT – This paper compares the results of more than two thousand pre-electoral polls between 2012 and 2020 with the official electoral results in five Brazilian elections. We examine the effects of random and non-random factors, such as sample size and the time before the elections, on differences between the polls' estimates and the ballot results. Among other findings, we show that later polls, conducted just before the election, with large samples, are better at minimizing errors. We also document that polls for presidential and gubernatorial elections and second-round elections exhibit lower error rates. Taken together, our results suggest that Brazilian pre-electoral polls perform similarly to polls conducted in other countries.

KEYWORDS: Pre-electoral polls, Survey, Elections.

Fernando Meireles é pesquisador de pós-doutorado no Centro Brasileiro de Análise e Planejamento (Cebrap). @ – meirelesff@hotmail.com / <https://orcid.org/0000-0002-7027-2058>.

Guilherme Russo é lecturer na Escola de Economia de São Paulo da Fundação Getúlio Vargas (EESP/FGV). @ – guilhermeazzirusso@gmail.com / <https://orcid.org/0000-0001-8763-5698>.

Recebido em 5.9.2022 e aceito em 16.9.2022.

^I Centro Brasileiro de Análise e Planejamento, São Paulo, Brasil.

^{II} Fundação Getúlio Vargas, Escola de Economia, São Paulo, Brasil.

