DOSSIÊ TÉCNICO-CIENTÍFICO · V1.0 · 27 DE OUTUBRO DE 2026

Do 1º ao 2º turno: análise estatística da eleição presidencial brasileira de 2026

Contabilidade seção a seção, inferência ecológica, testes forenses com eleição de controle, linha do tempo da apuração e limites para desvio sistemático de votos

Preparado para: Redutto — Inteligência territorial eleitoral (redutto.com.br)

Execução: análise computacional conduzida por inteligência artificial (Claude Opus 5.5, Anthropic), sob solicitação do responsável pelo Redutto. Declaração completa na seção 9.

Dados: Tribunal Superior Eleitoral, Portal de Dados Abertos — boletins de urna por seção (2022 e 2026) e votação por município-zona (2014 e 2018). Arquivos de 2026 gerados pelo TSE em 26/10/2026.

Resumo

No 1º turno, Flávio Bolsonaro obteve 56.104.503 votos e Lula 53.879.536 (47,03% × 45,16% dos válidos). No 2º turno, Lula obteve 64.256.232 e Flávio 59.826.041 (51,79% × 48,21%). Este dossiê testa se essa virada de +6,66 mi na margem é matemática e estatisticamente compatível com os dados de 499.205 seções eleitorais pareadas entre os turnos.

Resultados. (i) A identidade contábil entre os turnos fecha em todas as seções. Em 98,7% delas, o ganho de Lula cabe nos votos livres da própria seção, sem nenhuma troca de lado. (ii) Por inferência ecológica restrita, estimada em 2.637 zonas, os finalistas retiveram 98,3% e 97,2% dos próprios eleitores. Os eleitores dos eliminados se dividiram (38,2% para Lula e 37,8% para Flávio). A vantagem veio de quem votou branco ou nulo (54,1% para Lula) ou faltou no 1º turno. (iii) O 2º turno de cada seção é previsível a partir do 1º, com R^2 fora da amostra de 0,980, igual ao da eleição de controle de 2022. (iv) Sete testes forenses não distinguem 2026 de 2022. (v) A pausa de 17h07 às 19h02 na recepção dos boletins, decorrente da prorrogação da votação pelo STF, não está associada a diferença na direção de alteração. (vi) Supondo desvio de votos de Flávio para Lula feito por software, as duas formas mais simples ficam limitadas a +0,97 mi e +2,83 mi na margem; para inverter o resultado seriam necessários +4,43 mi.

O que foi atípico. Com as taxas de transferência medidas em 2014, 2018 ou 2022, Flávio teria vencido. É a primeira virada sobre o líder do 1º turno numa eleição presidencial desde 1989. O comparecimento subiu 2.960.241 votos, o maior aumento entre turnos das quatro últimas eleições.

Palavras-chave: inferência ecológica; forense eleitoral; regressão de Goodman; modelo RxC Bayesiano; Klimek; lei de Benford; Isolation Forest; decomposição de Shapley.

Sumário

1. Introdução e pergunta de pesquisa
2. Dados
3. Métodos
4. Resultados
5. A pausa na apuração (17h07–19h02)
6. A hipótese de software adulterado na urna
7. Possíveis distorções encontradas
8. Limitações
9. Uso de inteligência artificial e reprodutibilidade
10. Conclusão do analista
11. Referências
12. Apêndices: tabelas por UF, seções mais atípicas, testes e p-valores

1. Introdução e pergunta de pesquisa

Entre o 1º turno (4/10/2026) e o 2º turno (25/10/2026), Lula ganhou 10.376.696 votos e Flávio 3.721.538. Os candidatos eliminados somavam 9.321.993 votos: Cury 3.448.569, Renan 2.675.887, Caiado 2.605.148 e demais 592.389. A pergunta que motivou o estudo é se é possível que essa virada tenha ocorrido, já que parece exigir que todo o centro e mais eleitores ausentes tenham ido para Lula.

A pergunta é decomposta em cinco hipóteses verificáveis:

- **H1 (viabilidade):** os votos observados no 2º turno de cada seção podem ser obtidos a partir do 1º turno da mesma seção.
- **H2 (transferência):** existe uma matriz de transferência plausível, com taxas entre 0 e 1, que reproduz o 2º turno.
- **H3 (previsibilidade):** o 2º turno de cada seção é previsível pelo 1º, com a mesma precisão de uma eleição de controle.
- **H4 (forense):** as assinaturas estatísticas conhecidas de adulteração estão ausentes, ou presentes no mesmo grau da eleição de controle.
- **H5 (apuração e software):** nem a pausa na recepção dos boletins nem um desvio sistemático por programa explicam a virada.

Estatística não prova a ausência de fraude, nem a sua presença: ela mede a compatibilidade dos dados com hipóteses definidas. O dossiê declara, a cada passo, o que cada teste é capaz de detectar e o que não é.

2. Dados

Todos os dados são arquivos abertos do Tribunal Superior Eleitoral (TSE), obtidos do CDN do Portal de Dados Abertos em 27/10/2026. A unidade de análise é a *seção eleitoral*, a menor unidade com resultado publicado e correspondente a um boletim de urna (BU).

Tabela 1. Arquivos utilizados.

Eleição	Arquivo (TSE)	Unidade	Uso
2026	votacao_secao_2026_BR (gerado em 26/10/2026, 12h07)	seção × candidato	votos por candidato, 1º e 2º turno
2026	detalhe_votacao_secao_2026 (gerado em 26/10/2026, 09h49)	seção	aptos, comparecimento, abstenção, brancos, nulos, modelo de urna, horários de recepção e totalização do BU
2022	votacao_secao_2022_BR; detalhe_votacao_secao_2022	seção	eleição de controle
2014, 2018	votacao_partido_munzona; detalhe_votacao_munzona	município-zona	comparação histórica; taxa de base de governador
2022	votacao_partido_munzona_2022 (por UF)	município-zona	taxa de base de governador

Preparação. Os arquivos foram lidos com DuckDB, filtrando o cargo 1 (presidente) e agregando os votos por seção e turno nos grupos Lula (13), Flávio (22), Cury (70), Renan (14), Caiado (55), demais candidatos e branco + nulo (95, 96). As seções foram pareadas entre turnos pela chave UF, município, zona e seção, com o exterior incluído.

Tabela 2. Checagens de integridade.

Checagem	Resultado
Seções presentes nos dois turnos (2026)	499.205 de 499.248
Seções presentes nos dois turnos (2022)	472.028 de 472.075
Soma de votos (candidatos + brancos + nulos) igual ao comparecimento	todas as seções, nos dois turnos e nos dois anos
Totais das seções pareadas	Lula 53.879.536 → 64.256.232; Flávio 56.104.503 → 59.826.041; iguais ao total publicado, salvo 2 votos de Lula no 1º turno em seções sem par no 2º
Seções com eleitorado apto diferente entre os turnos	55.722 (mediana de 1 eleitor); cada turno é medido sobre o próprio eleitorado
Disputas de governador descartadas por inconsistência	2014 AM e 2014 TO: vencedor abaixo de 50% ou fora dos dois primeiros

3. Métodos

3.1 Identidade contábil e teste de viabilidade

Em cada seção *s*, sejam L e F os votos de Lula e Flávio, BN os brancos e nulos, E os votos dos eliminados no 1º turno e C o comparecimento, com o índice 1 ou 2 indicando o turno. Como os votos só mudam de lugar dentro da seção, vale a identidade:

$$(L_{2,s} - L_{1,s}) + (F_{2,s} - F_{1,s}) + (BN_{2,s} - BN_{1,s}) = E_{1,s} + (C_{2,s} - C_{1,s}) \tag{1}$$

O estoque livre da seção é o máximo de votos que Lula pode ganhar sem que nenhum eleitor de Flávio no 1º turno troque de lado e sem rotatividade de comparecimento. A seção é viável sob essa hipótese restritiva quando o ganho de Lula não passa do estoque:

$$S_s = E_{1,s} + BN_{1,s} + \max\left(0,\ C_{2,s} - C_{1,s}\right), \qquad \Delta L_s \le S_s \tag{2}$$

3.2 Inferência ecológica: regressão de Goodman restrita

Sejam $x_{s,o}$ a fração do eleitorado da seção em cada uma das oito origens do 1º turno (Lula, Flávio, Cury, Renan, Caiado, outros, branco/nulo, abstenção) e $y_{s,d}$ a fração em cada um dos quatro destinos do 2º turno (Lula, Flávio, branco/nulo, abstenção). O modelo de Goodman (1953) supõe taxas de transferência β comuns às seções de um mesmo grupo:

$$y_{s,d} = \sum_{o=1}^{8} x_{s,o}\,\beta_{o,d} + \varepsilon_{s,d}, \qquad x_{s,o} = \frac{V_{s,o}^{(1)}}{A_s^{(1)}}, \quad y_{s,d} = \frac{V_{s,d}^{(2)}}{A_s^{(2)}} \tag{3}$$

Em cada zona eleitoral *z*, as taxas são estimadas por mínimos quadrados ponderados pelo eleitorado, com restrições de probabilidade:

$$\hat{B}_z = \arg\min_B \sum_{s \in z} w_s \left\| y_s - B^{\top} x_s \right\|^2 \quad \text{s.a.} \quad B \ge 0,\ B1 = 1, \qquad w_s = \frac{A_s}{\sum_{s \in z} A_s} \tag{4}$$

O problema é um programa quadrático convexo, resolvido com CVXPY (Diamond e Boyd, 2016) e o solver de ponto interior Clarabel (Goulart e Chen, 2024). Foram ajustadas 2.637 zonas com 30 ou mais seções, todas com status ótimo; as zonas menores e o exterior formam um grupo à parte. O fluxo nacional soma os fluxos das zonas:

$$F_{o,d} = \sum_{z} V_{z,o}^{(1)}\,\hat{\beta}_{z,o,d}, \qquad \hat{\beta}_{o,d} = \frac{F_{o,d}}{\sum_{d'} F_{o,d'}} \tag{5}$$

O intervalo de confiança de 95% vem de 2.000 reamostragens bootstrap das zonas, pelo método dos percentis (Efron e Tibshirani, 1993). A mesma estimação foi repetida com unidades município-zona agrupadas por UF, o único nível comum às quatro eleições (2014–2026).

3.3 Inferência ecológica Bayesiana RxC

Para relaxar a hipótese de taxa constante, usou-se o modelo multinomial-Dirichlet de Rosen, Jiang, King e Tanner (2001), no pacote PyEI (Knudson et al., 2021) sobre PyMC (Abril-Pla et al., 2023). Nele, cada seção tem a sua própria matriz de transferência, puxada de uma distribuição Dirichlet comum. A amostragem usa NUTS (Hoffman e Gelman, 2014), via NumPyro (Phan et al., 2019), sobre amostras aleatórias de seções com eleitorado igual nos dois turnos. A convergência foi avaliada pelo R-hat dividido (Gelman e Rubin, 1992; Vehtari et al., 2021).

$$V^{(2)}_{s,\cdot} \sim \text{Multinomial}\left(A_s,\ \sum\nolimits_o x_{s,o}\,\beta_{s,o,\cdot}\right), \qquad \beta_{s,o,\cdot} \sim \text{Dirichlet}(\alpha_o) \tag{6}$$

3.4 Validação fora da amostra

Em cada zona, as seções foram sorteadas em duas metades (semente 2026). As taxas estimadas numa metade preveem a outra, e vice-versa, de modo que nenhuma seção é prevista por um modelo que a viu. A medida é o R^2 ponderado:

$$\hat{y}_s = \hat{B}^{\top}_{z(-s)} x_s, \qquad R^2 = 1 - \frac{\sum_s w_s\,(y_s - \hat{y}_s)^2}{\sum_s w_s\,(y_s - \bar{y})^2} \tag{7}$$

3.5 Regressões com efeitos fixos

Os testes de modelo de urna, de horário de chegada do boletim e de composição do comparecimento extra usam regressão linear com efeitos fixos de alta dimensão, no PyFixest. As regressões são ponderadas pelo eleitorado e têm erro-padrão agrupado por município (Cameron e Miller, 2015):

$$\Delta L_s = \alpha_{g(s)} + \gamma\, T_s + \sum_o \theta_o\, x_{s,o} + u_s \tag{8}$$

O termo α_g é o efeito fixo de município ou zona, T_s é a variável testada e os $x_{s,o}$ controlam a composição do 1º turno.

3.6 Testes forenses

- **Impressão digital de Klimek** (Klimek et al., 2012): distribuição conjunta de comparecimento e votação, e a inclinação votação ~ comparecimento dentro do município.
- **Dígitos:** último dígito, esperado quase uniforme (Beber e Scacco, 2012), e segundo dígito contra a lei de Benford, no teste 2BL (Mebane, 2006). Com centenas de milhares de seções, o qui-quadrado rejeita qualquer desvio mínimo. Por isso, a distribuição de referência é a do mesmo teste em 100 contagens binomiais simuladas, com os tamanhos e proporções observados.
- **Isolation Forest** (Liu, Ting e Zhou, 2008), com 300 árvores e contaminação de 0,1%, sobre o erro de previsão de cada finalista, a variação de comparecimento e a de branco e nulo.

$$P(d) = \sum_{k=1}^{9} \log_{10}\left(1 + \frac{1}{10k + d}\right), \quad d = 0,\ \ldots, 9 \qquad \chi^2 = \sum_d \frac{(O_d - E_d)^2}{E_d} \tag{9}$$

3.7 Contrafactual e decomposição de Shapley

As taxas nacionais de cada eleição foram alinhadas por campo político: candidato do PT, adversário, eliminados, branco/nulo e abstenção. Em 2018 o candidato do PT foi o perdedor. Essas taxas foram aplicadas ao 1º turno de 2026. A diferença entre a margem contrafactual e a real foi repartida entre quatro blocos de comportamento pelo valor de Shapley (1953), a média das contribuições marginais sobre as 24 ordens:

$$\varphi_k = \frac{1}{|K|!}\sum_{\pi}\left[v\left(P_k^{\pi}\cup\{k\}\right) - v\left(P_k^{\pi}\right)\right] \tag{10}$$

3.8 Correção para testes múltiplos

Os p-valores dos testes de irregularidade foram corrigidos pelo método de Holm (1979), que controla a taxa de erro por família.

4. Resultados

4.1 Contabilidade e viabilidade

A identidade (1) vale em 499.205 de 499.205 seções. No total: 10.376.696 + 3.721.538 − 1.816.000 = 9.321.993 + 2.960.241 = 12.282.234.

Pelo critério (2), 492.666 seções (98,7%) são viáveis sem nenhuma troca de lado. Nas 6.539 restantes, o excesso soma 18.032 votos, 0,17% do ganho de Lula. Esse excesso é compatível com troca de lado ou com rotatividade de comparecimento. Para Flávio, 201 seções excedem o estoque, somando 557 votos. Lula perdeu votos em 9.590 seções (1,9%); Flávio, em 82.180 (16,5%).

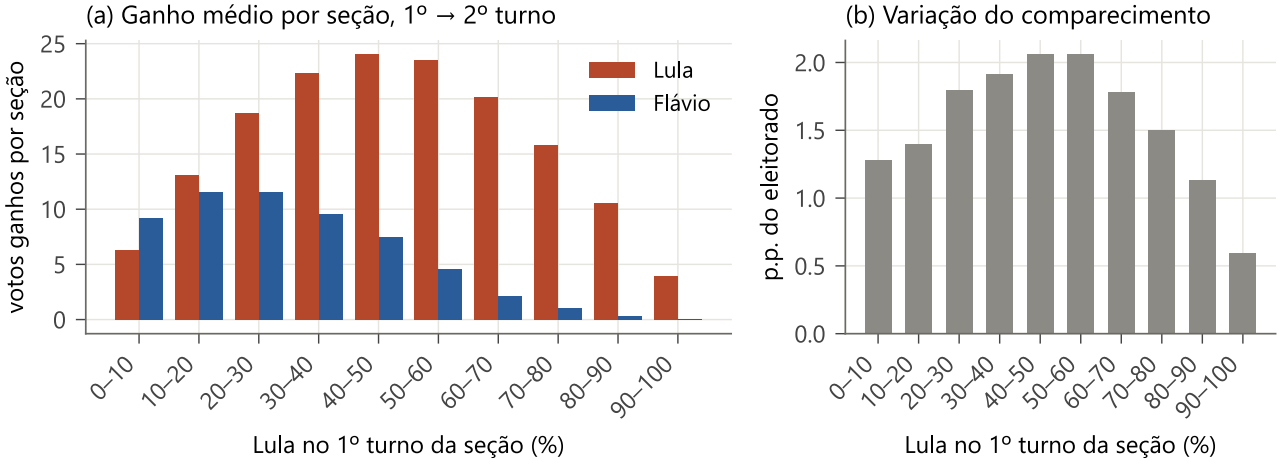


Figura 1. Ganho médio por seção e variação de comparecimento, pela votação de Lula no 1º turno da própria seção. O maior ganho ocorreu nas faixas de 30% a 60%, e o comparecimento subiu de forma parecida em todas as faixas. O enchimento de urna concentra o ganho onde o candidato já é forte.

Tabela 3. Variação entre turnos por região.

Região	Ganho Lula	Ganho Flávio	Comparecimento (votos)	Comparecimento 1º → 2º	Lula, % válidos 1º → 2º
Sudeste	+4.800.333	+1.944.892	+1.613.905	77,3% → 79,7%	39,7% → 47,2%
Nordeste	+2.955.848	+397.763	+826.761	81,6% → 83,5%	63,8% → 69,4%
Sul	+1.332.846	+866.745	+518.761	79,5% → 81,8%	31,3% → 37,4%
Centro-Oeste	+770.876	+470.431	+199.018	78,3% → 79,9%	32,6% → 40,0%
Norte	+474.024	+14.110	−237.883	80,6% → 78,8%	44,6% → 50,0%
Exterior	+42.769	+27.597	+39.679	37,3% → 41,6%	47,5% → 53,9%

4.2 Matriz de transferência



Figura 2. Matriz de transferência nacional estimada por Goodman restrito, zona a zona (2.637 zonas). Cada célula mostra a porcentagem da origem que terminou em cada destino e, entre colchetes, o IC 95% por bootstrap de 2.000 reamostragens das zonas.

Somados, os eliminados enviaram 3.562.228 votos a Lula (38,2%) e 3.521.685 a Flávio (37,8%). Os grupos que separam os finalistas são branco/nulo do 1º turno (3.231.529 votos para Lula contra 492.088 para Flávio) e a abstenção do 1º turno (4.066.670 contra 1.239.305).

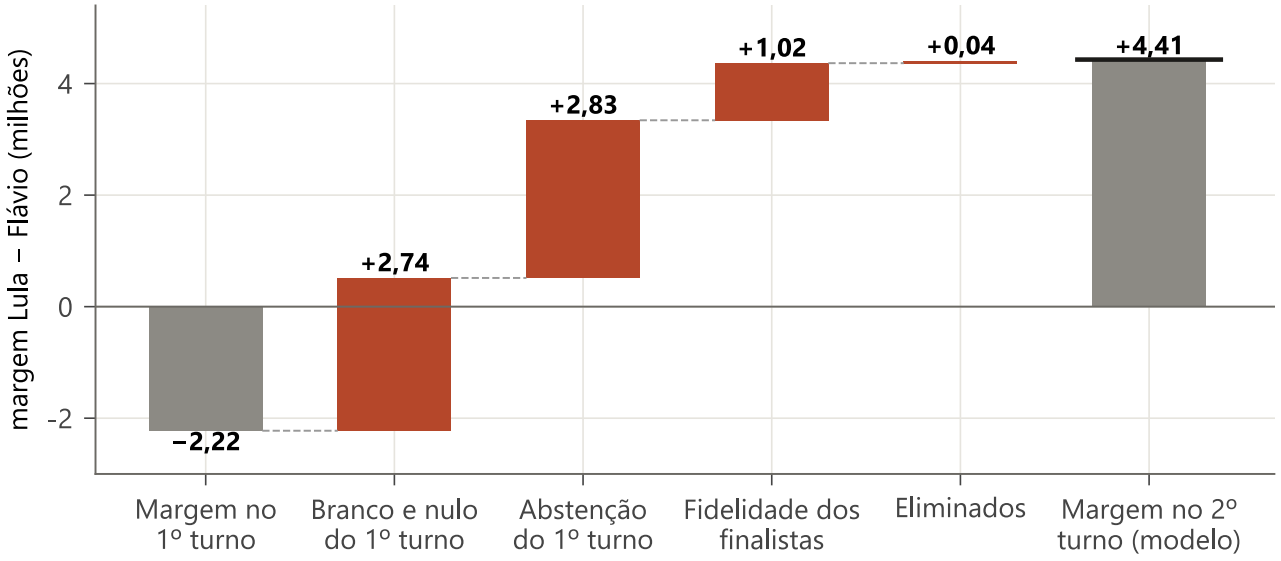


Figura 3. Ponte da margem Lula − Flávio. A contribuição de cada grupo é $c_o = F_{o,\text{Lula}} - F_{o,\text{Flávio}}$, e "fidelidade" reúne as perdas de cada finalista. Branco/nulo e abstenção do 1º turno respondem por 84% da virada. A linha preta sobre a última barra é a margem oficial (+4,43 mi); o modelo a reproduz com erro de 23.840 votos.

4.3 Robustez entre métodos

Tabela 4. Taxas de transferência por três métodos.

Taxa	Goodman por zona (seções)	Goodman por UF (município-zona)	Bayesiano RxC, 400 seções [IC 95%]
Lula → Lula	98,3%	98,7%	99,2% [98,4–99,7]
Flávio → Flávio	97,1%	98,9%	97,9% [96,8–98,9]
Eliminados → Lula	38,3%	45,6%	37,3% [29,9–44,9]
Eliminados → Flávio	37,9%	37,6%	46,5% [39,0–55,0]
Branco/nulo → Lula	54,1%	68,3%	58,6% [52,1–64,2]
Branco/nulo → Flávio	8,3%	0,2%	2,4% [0,3–6,0]
Abstenção → Lula	12,1%	7,9%	11,0% [8,9–13,1]
Abstenção → Flávio	3,7%	2,6%	0,5% [0,1–1,6]

Bayesiano: 400 seções sorteadas, 4 cadeias de 1.000 amostras (34 min). O R-hat da matriz agregada foi 1,79 e o dos hiperparâmetros 3,01; o critério é R-hat ≤ 1,01. **As cadeias não convergiram**, e a coluna deve ser lida só como indicação. Uma rodada com 300 seções deu valores próximos (eliminados → Lula 41,2%, → Flávio 42,3%). Os três métodos concordam na direção de todas as taxas e divergem na magnitude, sobretudo na repartição entre branco/nulo e abstenção, que andam juntos nas mesmas áreas.

4.4 Previsibilidade fora da amostra

Tabela 5. Validação cruzada em duas dobras dentro de cada zona: 2026 e controle de 2022.

Medida	2026	2022 (controle)
R^2 fora da amostra, vencedor	0,980	0,981
R^2 fora da amostra, adversário	0,982	0,985
R^2 fora da amostra, branco/nulo	0,461	0,646
R^2 fora da amostra, abstenção	0,894	0,912
Total do vencedor: previsto × real	64.233.041 × 64.245.712	60.325.904 × 60.315.193
Desvio-padrão do erro por seção	2,05 p.p.	1,92 p.p.
Curtose (excesso) do erro	2,9	5,2
Seções com erro acima de +10 p.p. / abaixo de −10 p.p.	154 / 255	109 / 193

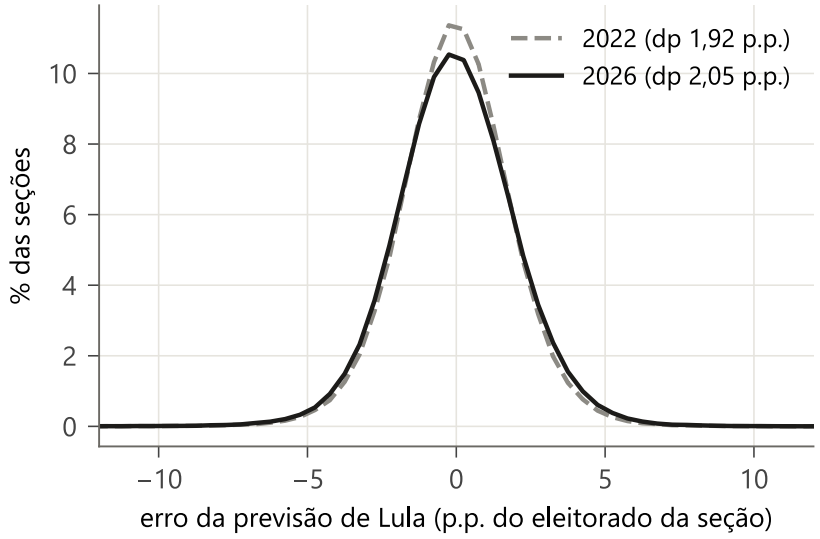


Figura 4. Distribuição do erro da previsão fora da amostra para Lula, por seção. As curvas de 2026 e 2022 se sobrepõem, e 2026 tem caudas mais finas. Inserção concentrada de votos produziria um segundo pico ou uma cauda pesada à direita.

Poder do teste. Inverter a margem exigiria deslocar 2.215.096 votos. Concentrados em 5% das seções, seriam cerca de 89 votos por seção, perto de 29 p.p. de erro, e haveria cerca de 25 mil seções na cauda direita. Há 154 seções com erro acima de +10 p.p.

Tabela 6. Teste de estresse: margem após substituir pelo valor previsto as seções mais favoráveis ao vencedor.

Seções substituídas	Margem de Lula, 2026	Margem do vencedor, 2022
nenhuma	+4.430.191	+2.139.645
0,1%	+4.409.108	+2.121.113
1%	+4.282.239	+2.011.709
5%	+3.888.198	+1.667.015
10%	+3.526.656	+1.347.343

O teste é deliberadamente parcial: só corrige a cauda que favorece o vencedor.

4.5 Testes forenses, com 2022 como controle

Tabela 7. Testes forenses: 2026 e controle de 2022.

Teste	O que detectaria	2026	2022 (controle)
Klimek: seções com comparecimento e Lula acima de 95%	enchimento de urna	41	19
Klimek: inclinação votação ~ comparecimento no município	enchimento (inclinação positiva)	−0,144 [−0,196; −0,093]	−0,150 [−0,215; −0,086]
Comparecimento extra para o vencedor / adversário	votos inseridos (~100% para um lado)	54,6% / 41,0%	53,6% / 40,5%
Último dígito, vencedor 2º turno: χ^2 [faixa simulada 5–95%]	números inventados	347 [316–441]	462 [322–437]
Último dígito, adversário 2º turno	números inventados	271 [218–313]	205 [140–221]
2BL, vencedor: χ^2 sobre a mediana simulada	manipulação de contagens	1,05×	1,07×
2BL, adversário	manipulação de contagens	1,09×	1,07×
Hora de chegada do BU: p.p. por hora, sem / com efeito fixo	votos de última hora	+0,347 / +0,052	+0,056 / −0,007
Isolation Forest: seções (pró-vencedor / pró-adversário)	anomalias multivariadas	500 (208 / 292)	473 (204 / 269)
Isolation Forest: margem sem essas seções		+4.430.803	+2.139.752

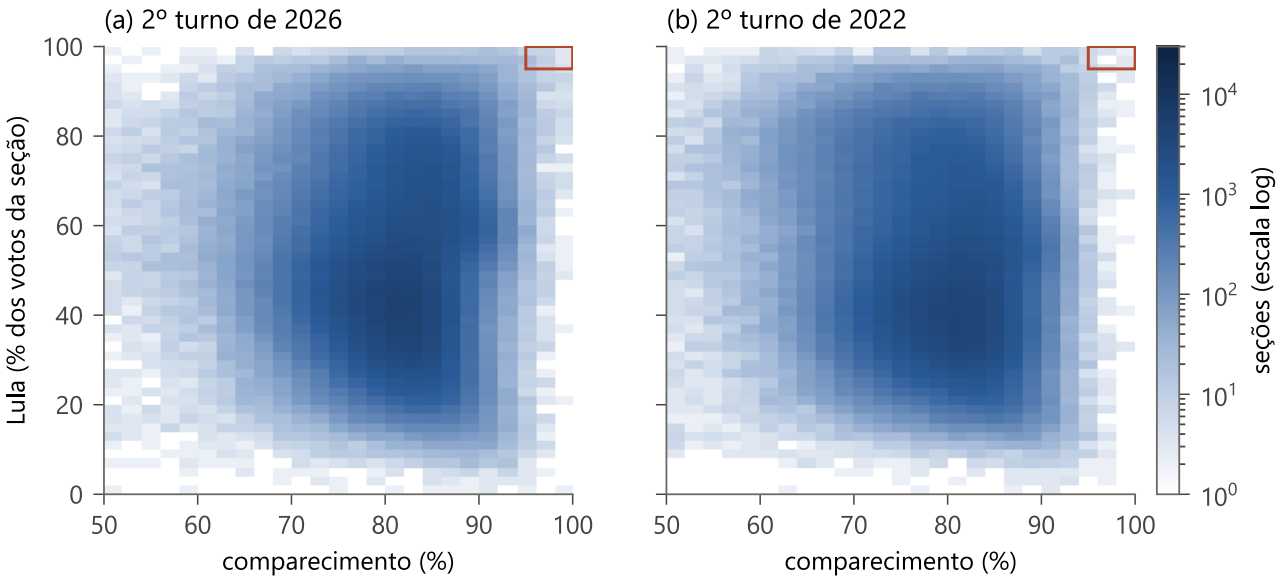


Figura 5. Impressão digital de Klimek no 2º turno: número de seções por comparecimento e votação de Lula. O quadro vermelho marca o canto de 95%/95%, onde o enchimento de urna se acumula. As duas eleições têm a mesma forma.

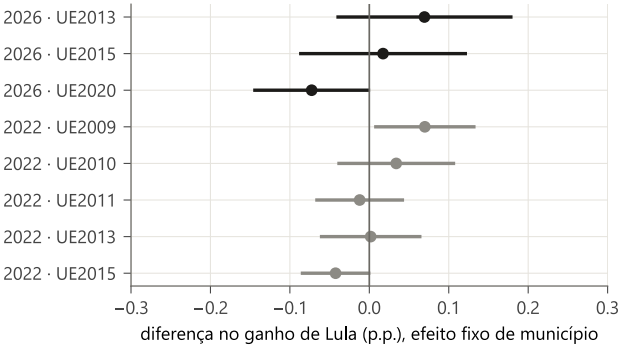


Figura 6. Diferença no ganho de Lula por modelo de urna, dentro do mesmo município (IC 95%). Base em 2026: UE2022; em 2022: UE2020.

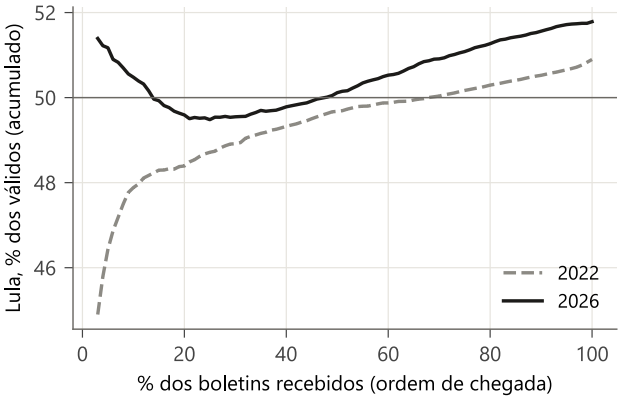


Figura 7. Participação acumulada de Lula pela ordem de chegada dos boletins. A curva reflete a ordem regional de transmissão; dentro do município, o efeito do horário é desprezível.

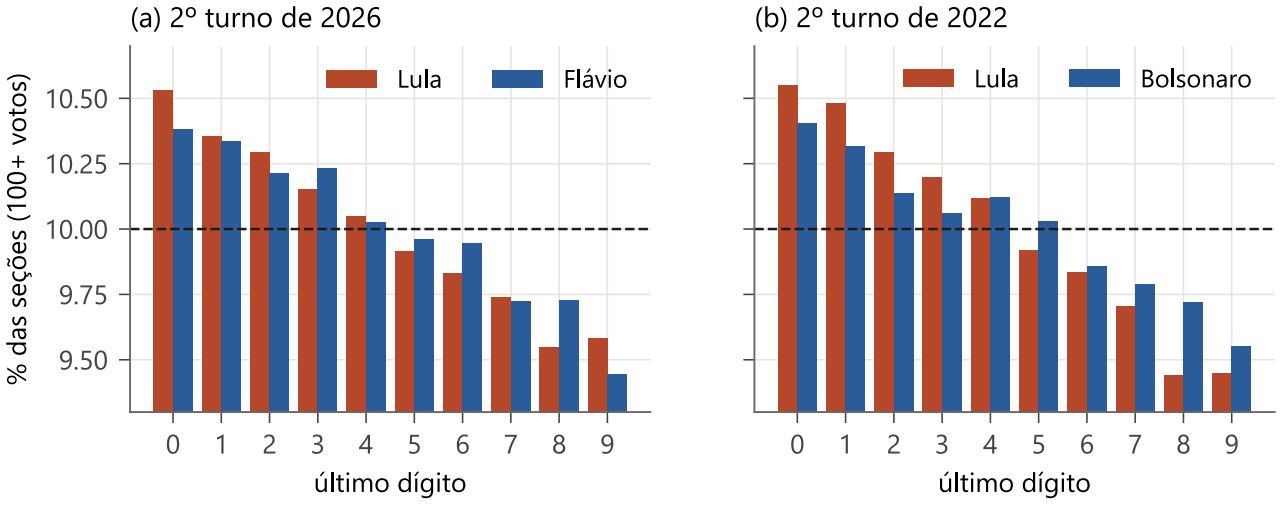


Figura 8. Frequência do último dígito dos votos por seção, 2º turno, seções com 100 ou mais votos. A linha tracejada marca 10%. Nos quatro casos de 2026 (1º e 2º turno), o qui-quadrado fica dentro da faixa de 5% a 95% das contagens simuladas sem fraude.

Testes múltiplos. Dos sete testes de irregularidade com p-valor (Apêndice C), dois ficam perto de 0,05 sem correção: o modelo UE2020 e a hora de chegada. Nenhum continua significativo após a correção de Holm, e os dois têm magnitude desprezível (menos de 0,1 p.p.). Em 2022 aparece o mesmo número de resultados nesse limite.

4.6 Comparação histórica, contrafactual e decomposição

Tabela 8. Do 1º ao 2º turno, 2014–2026 (Goodman por UF sobre município-zona; vencedor × perdedor).

Medida	2014	2018	2022	2026
Finalistas	Dilma × Aécio	Bolsonaro × Haddad	Lula × Bolsonaro	Lula × Flávio
Líder do 1º turno venceu?	sim	sim	sim	**não**
Ganho do vencedor	+11,23 mi	+8,52 mi	+3,09 mi	+10,38 mi
Ganho do perdedor	+16,14 mi	+15,70 mi	+7,13 mi	+3,72 mi
Votos dos eliminados	25,86 mi	26,43 mi	9,90 mi	9,32 mi
Variação de comparecimento	−2,44 mi	−1,43 mi	+0,57 mi	+2,96 mi
Variação de branco e nulo	−3,96 mi	+0,78 mi	+0,25 mi	−1,82 mi
Eliminados → vencedor / perdedor	33% / 55%	30% / 53%	38% / 42%	46% / 38%
Branco e nulo → vencedor / perdedor	42% / 8%	6% / 28%	10% / 33%	68% / 0%
Abstenção → vencedor / perdedor	0,2% / 3,2%	0,3% / 0,8%	1,1% / 3,6%	7,9% / 2,6%
R^2 fora da amostra (vencedor)	0,977	0,996	0,996	0,994

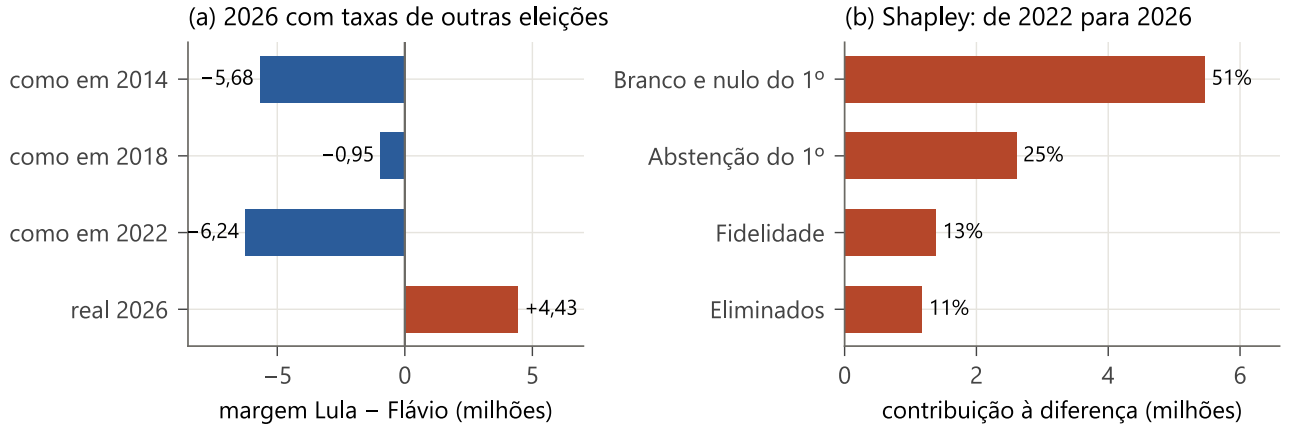


Figura 9. (a) Margem de Lula se o eleitorado de 2026 tivesse se movido com as taxas de cada eleição. (b) Decomposição de Shapley da diferença entre o contrafactual de 2022 (−6,24 mi) e a margem do modelo (+4,40 mi).

Com as taxas de 2014, 2018 ou 2022, Flávio teria vencido por 5,68 mi, 0,95 mi e 6,24 mi, respectivamente. A virada exigiu um comportamento diferente das três eleições anteriores. A diferença em relação a 2022 se decompõe em branco/nulo do 1º turno (51%), abstenção do 1º turno (25%), fidelidade dos finalistas (13%) e eliminados (11%). O padrão de branco/nulo do 1º turno favorecendo o candidato do PT já apareceu em 2014 e 2018, com intensidade menor.

4.7 Taxa de base de viradas

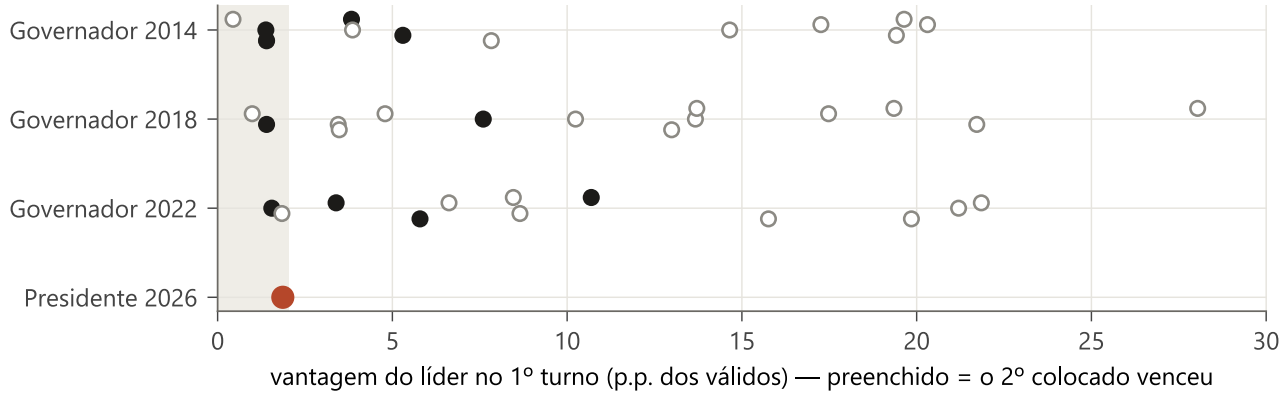


Figura 10. Disputas de governador com 2º turno em 2014, 2018 e 2022 (39 disputas), pela vantagem do líder no 1º turno. Ponto preenchido indica que o segundo colocado venceu. A faixa sombreada marca vantagem até 2 p.p.; o ponto vermelho é a disputa presidencial de 2026.

Das 39 disputas de governador, 10 viraram (26%). Com vantagem inicial de até 2 p.p., viraram 4 de 8. A vantagem de Flávio no 1º turno foi de 1,86 p.p. Os casos mais parecidos, com poucos votos nos eliminados, são Pará e Paraíba em 2014, e os dois viraram. Em eleições presidenciais, o líder do 1º turno venceu o 2º em todas as disputas desde 1989; 2026 é a primeira exceção.

5. A pausa na apuração (17h07–19h02)

Os horários de recepção (*DT_RECEBIMENTO_BU_HOR_TSE*) e de totalização (*DT_PRIM_TOT_PARCIAL_HOR_TSE*) de cada boletim permitem reconstruir a apuração minuto a minuto. No 2º turno de 2026, nenhum boletim de seção do Brasil chegou à totalização entre 17h07 e 19h02. Entre a recepção e a totalização passaram 24 s na mediana e 61 s no percentil 99: a pausa foi na recepção dos boletins, não num processamento posterior.

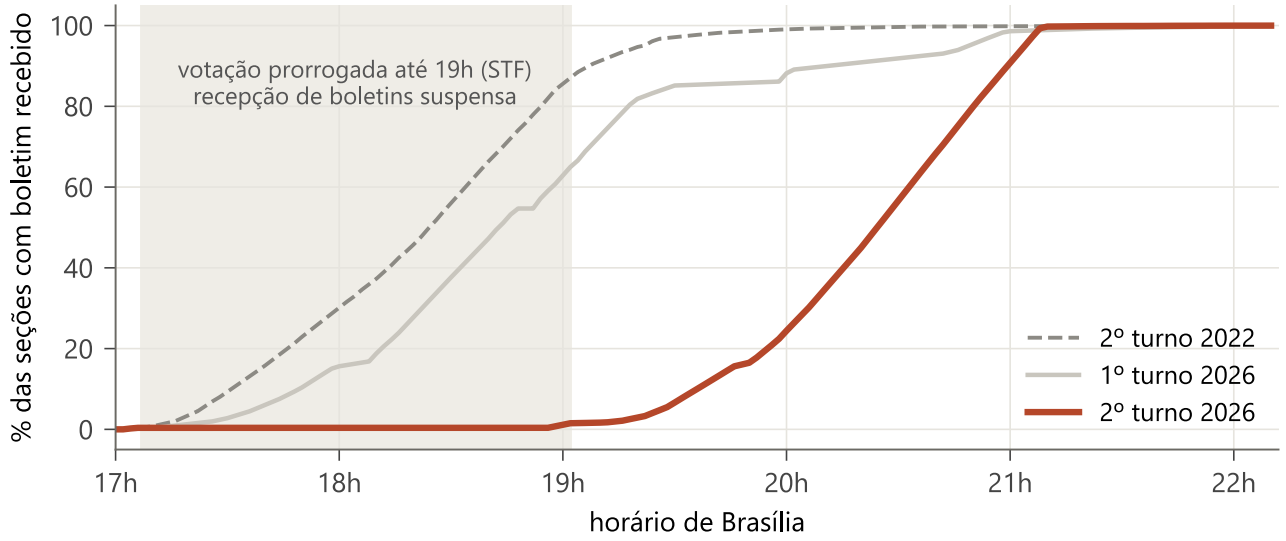


Figura 11. Percentual acumulado de seções com boletim recebido pelo TSE, horário de Brasília. A faixa sombreada é a pausa do 2º turno de 2026.

Causa documentada. A parcial das 17h06 cobria 1.225 seções (0,25%). Nela, Lula tinha 57,66% dos válidos (InfoMoney, 2026). O TSE suspendeu a totalização depois que o ministro Cristiano Zanin, do STF, prorrogou a votação por duas horas, até as 19h, em todo o país, a pedido da coligação de Lula, por problemas no deslocamento de eleitores. A apuração foi retomada às 19h. O ministro esclareceu depois que a prorroga-

ção valia só para as seções que ainda não tinham encerrado, e que as urnas fechadas às 17h seriam computadas normalmente (Brasil de Fato, 2026).

Quem chegou antes. Antes da pausa chegaram 1.868 seções pareadas: exterior 776, PA 252, PR 205, AM 152, DF 108, PB 63. Entre 17h10 e 19h00 chegaram 39 boletins, todos do exterior e nenhum de seção pareada.

Tabela 9. Participação acumulada de Lula na ordem de chegada (o que o público viu).

Até	Seções recebidas	Lula, % dos válidos
17:07	0,37%	55,53%
19:30	6,07%	50,88%
20:00	23,47%	49,52%
20:30	55,47%	50,37%
21:00	89,69%	51,55%
final	100,00%	51,78%

Teste antes × depois. Se os boletins processados depois da pausa tivessem sido alterados a favor de Lula, as seções contadas antes ficariam abaixo do previsto em relação às vizinhas da mesma zona contadas depois. A comparação usa as 272 zonas com seções nos dois grupos, com efeito fixo de zona.

Tabela 10. Seções contadas antes da pausa menos seções contadas depois, na mesma zona (p.p. do eleitorado).

Medida	2026	2022 (controle; primeiros 10 min)
Resíduo de Lula fora da amostra	+0,35 [−0,15; +0,85], p = 0,172	+0,08 [−0,07; +0,22], p = 0,317
Resíduo da margem	+0,85 [+0,15; +1,56], p = 0,018	+0,17 [−0,12; +0,47], p = 0,254
Variação de Lula com controles do 1º turno	+0,32 [−0,23; +0,86], p = 0,254	+0,11 [−0,06; +0,27], p = 0,199

Para inverter o resultado, a alteração posterior à pausa teria de deixar as seções de antes pelo menos 2,8 p.p. abaixo na margem. O intervalo observado exclui esse valor e é todo positivo.

As seções contadas antes da pausa não ficaram abaixo das demais. A única diferença no limite da significância (margem; p ajustado por Holm = 0,125) vai na direção oposta à hipótese de alteração.

Onde a prorrogação foi pedida. O comparecimento subiu +4,37 p.p. em Salvador, +4,24 p.p. em Feira de Santana e +5,64 p.p. em Lauro de Freitas. No resto da Bahia subiu +2,91 p.p., e no resto do Brasil +1,76 p.p. O excedente sobre a tendência estadual é de cerca de 39 mil eleitores. Mesmo que todos tivessem votado em Lula, seriam 0,9% da margem, e os dados não separam o efeito da prorrogação do efeito do transporte ou da mobilização. Um site noticiou 35.521 votos registrados depois das 17h, segundo logs do TSE (Mixvale, 2026). A página recusou o acesso e o número não foi verificado; se estiver correto, equivale a 0,8% da margem.

6. A hipótese de software adulterado na urna

Um programa adulterado altera também o boletim impresso. Por isso, conferir o boletim contra o resultado publicado não o detecta. Mas todo desvio sistemático de votos de um candidato para outro deixa uma marca na matriz de transferência, porque as taxas verdadeiras não podem ser negativas. Duas formas de desvio foram limitadas zona a zona.

Forma A — fração p dos votos de Flávio transferida a Lula em cada urna. A taxa observada Flávio → Lula é a verdadeira mais o desvio. Como a verdadeira é ≥ 0:

$$\frac{p_z}{1-p_z} \le \frac{\hat{\beta}_{z,F\to L}}{\hat{\beta}_{z,F\to F}} \quad \Rightarrow \quad \text{votos movidos} \le \sum_z \frac{\hat{\beta}_{z,F\to L}}{\hat{\beta}_{z,F\to F}} F_z^{(2)} \tag{11}$$

Forma B — fração fixa δ do eleitorado transferida em cada urna. Como as frações do 1º turno somam 1, o desvio se soma à taxa observada de toda origem para Lula:

$$\delta_z \le \min_o \hat{\beta}_{z,o\to L} \le \hat{\beta}_{z,F\to L} \quad \Rightarrow \quad \text{votos movidos} \le \sum_z \hat{\beta}_{z,F\to L} A_z \tag{12}$$

Tabela 11. Limite superior para desvio de votos de Flávio para Lula por software (pior caso: toda a transferência estimada Flávio → Lula seria fraude).

Forma de desvio	Votos movidos (máximo)	Efeito máximo na margem	Necessário para inverter
A: fração dos votos de Flávio	485.176	+0,97 mi	2.215.096 votos movidos (+4,43 mi na margem)
B: fração fixa do eleitorado	1.415.112	+2,83 mi	2.215.096 votos movidos (+4,43 mi na margem)

A taxa Flávio → Lula é zero em 1.645 das 2.638 zonas.

Seções apuradas fora do programa da urna. No 2º turno de 2026, 27 seções tiveram o resultado lançado pelo sistema de apuração, sem passar pelo programa da urna. O resíduo médio da margem nelas foi −1,01 p.p. (EP 1,12). Em 2022, o mesmo grupo (43 seções) teve +2,63 p.p. (EP 0,77). Com poucas seções e um grupo heterogêneo, **o teste é inconclusivo**: o intervalo inclui tanto zero quanto um desvio do tamanho necessário para inverter o resultado.

O que os dados não limitam. Ficam fora de alcance o software que converta brancos e nulos em votos, o que altere votos em proporção à composição política da seção e o que crie votos para eleitores ausentes. Este último seria limitado pelo teste de comparecimento extra (seção 4.5) e pela conferência física da folha de votação. Essas hipóteses só se verificam por auditoria do software: inspeção do código-fonte, conferência dos resumos digitais (hashes) dos programas lacrados e testes de integridade com votação paralela em urnas sorteadas. Este dossiê não teve acesso a esses materiais nem aos relatórios de 2026.

7. Possíveis distorções encontradas

Esta seção lista onde os dados de 2026 se desviam do esperado, sem filtro de direção, e quanto cada desvio pesa.

1. **Resíduos simétricos.** A soma dos resíduos positivos da margem é 1.976.414 votos, e a dos negativos −1.946.701, um saldo de +29.714. Em 2022: 1.760.837 e −1.779.729. O percentil 99,9 do resíduo absoluto é 38,3 votos por seção (2022: 35,3).
2. **Seções isoladas com grandes desvios, nas duas direções** (Tabela 12). Há aglomerados locais: três seções de Belo Monte (AL) com deslocamento para Flávio, e seções de Barcelos e Ipixuna (AM) e de Parauapebas (PA) para Lula. Uma seção de Parauapebas teve salto de comparecimento de 128 para 212 votos, todos para Lula. São casos a conferir pelo boletim impresso. Somados, pesam centenas de votos.
3. **Comparecimento.** A distribuição inteira da variação de comparecimento se deslocou (mediana de +1,91 p.p., contra +0,40 em 2022), em vez de uma cauda de poucas seções. Seções com salto acima de 15 p.p.: 71 em 2026 e 29 em 2022, com saldo de +2.726 votos para Lula. Seções com comparecimento acima de 95%: 626 em 2026 e 224 em 2022, com mediana de 160 eleitores e saldo de +37.125 votos para Lula (0,8% da margem).
4. **Viabilidade por UF.** O ganho de Lula passou do estoque livre com mais frequência no Maranhão (7,2% das seções), no Ceará, em Alagoas, no Pará e no Piauí. O excesso nacional é de 18.032 votos (Figura 13).

5. **Isolation Forest por UF.** As seções marcadas se concentram no Norte (PA, AC, AP, AM), região de seções pequenas e voláteis, com saldos que se compensam: PA −635, AC +531.
6. **Resultados no limite da significância.** Modelo UE2020 (p = 0,046), hora de chegada (p = 0,048) e margem antes × depois da pausa (p = 0,018). Nenhum sobrevive à correção para testes múltiplos, e o último aponta na direção oposta à de alteração.
7. **Procedimento.** A prorrogação nacional da votação, pedida por uma das campanhas, e a pausa de duas horas na recepção são fatos documentados e controversos. Seu efeito mensurável nos dados é pequeno (seção 5). A legalidade da decisão é questão jurídica, fora do escopo estatístico.

Tabela 12. Seções de 2026 com maior resíduo da margem (real − previsto pelas vizinhas da zona).

UF	Município	Zona/seção	Aptos	Lula 1º → 2º (previsto)	Flávio 1º → 2º (previsto)	Resíduo da margem
PI	Santo Antônio De Lisboa	28/99	312	214 → 146 (216)	56 → 122 (59)	−133
AC	Santa Rosa Do Purus	3/95	337	102 → 174 (94)	105 → 47 (91)	+124
AM	Barcelos	18/37	377	37 → 84 (49)	171 → 71 (148)	+112
AC	Manoel Urbano	3/160	236	112 → 48 (99)	55 → 104 (47)	−108
Ext.	Ciudad Del Este	1/88	728	34 → 39 (60)	294 → 376 (310)	−87
AM	Barcelos	18/38	350	114 → 141 (92)	1 → 1 (38)	+86
AL	Belo Monte	29/53	386	192 → 170 (212)	87 → 131 (89)	−84
AM	Ipixuna	45/63	335	135 → 167 (140)	61 → 2 (58)	+83
AL	Belo Monte	29/56	393	207 → 172 (219)	100 → 133 (98)	−82
AL	Belo Monte	29/47	392	201 → 171 (218)	89 → 124 (90)	−81
Ext.	Dublin	1/3245	779	246 → 342 (279)	50 → 53 (70)	+80
PA	Parauapebas	106/330	325	126 → 212 (142)	2 → 0 (10)	+79

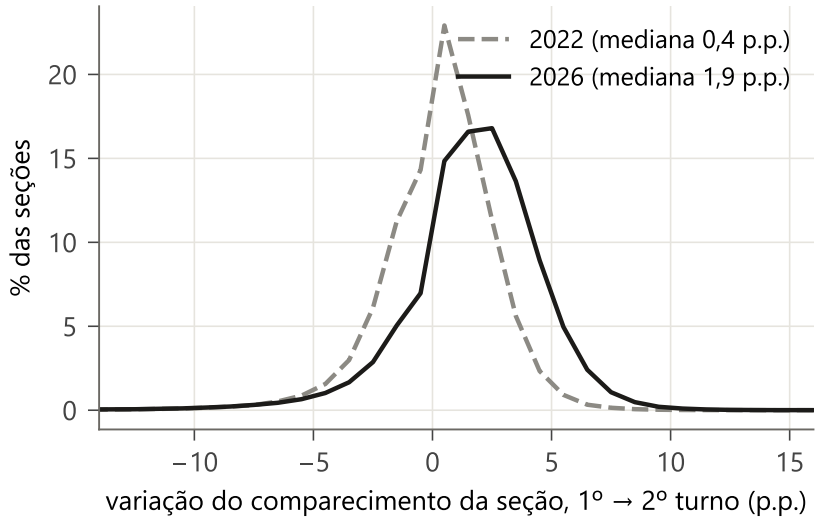


Figura 12. Distribuição da variação de comparecimento por seção, 2026 e 2022. Em 2026, a curva inteira se deslocou.

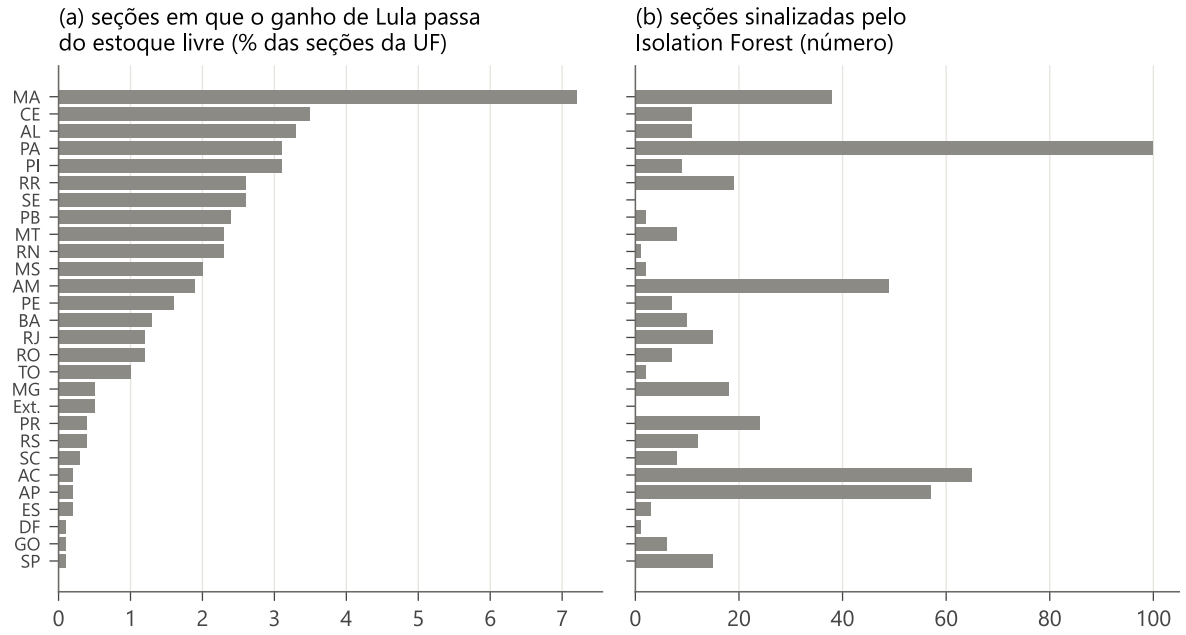


Figura 13. Indicadores de distorção por UF em 2026: (a) porcentagem de seções em que o ganho de Lula passa do estoque livre; (b) número de seções sinalizadas pelo Isolation Forest.

8. Limitações

- **Falácia ecológica.** As taxas de transferência são estimadas a partir de agregados por seção, não de eleitores. A hipótese de taxa constante dentro da zona pode atribuir a um grupo um efeito de contexto. Branco/nulo e abstenção do 1º turno andam juntos nas mesmas áreas, e a repartição entre eles é a parte menos estável do estudo.
- **Convergência Bayesiana.** O modelo RxC não convergiu nas rodadas executadas (R-hat > 1,01). Seus números são indicativos.
- **Alteração uniforme.** Testes por seção detectam adulteração localizada. Uma alteração que mudasse todas as seções na mesma proporção seria absorvida pelas taxas estimadas. A seção 6 limita as formas mais simples desse tipo de desvio, mas não todas.
- **Fontes externas.** A causa da pausa vem da imprensa (InfoMoney; Brasil de Fato), não de comunicado oficial lido diretamente. O número de votos depois das 17h não foi verificado.
- **Causalidade.** Os dados de urna não dizem por que o eleitor mudou de comportamento entre 2022 e 2026, seja por campanha, transporte, contexto ou mobilização.
- **Boletins impressos.** O estudo usa os arquivos publicados pelo TSE e não comparou amostras de boletins impressos (QR Code) com esses arquivos.

9. Uso de inteligência artificial e reprodutibilidade

IA utilizada. Claude Opus 5.5 (Anthropic; identificador `claude-opus-5-5`), executado no Claude Code (interface de linha de comando) em 27/10/2026, a pedido do responsável pelo Redutto. A IA escreveu e executou o código Python, escolheu e aplicou os métodos, gerou as figuras e redigiu este texto. Os números vêm de código determinístico aplicado aos arquivos do TSE; nenhum foi produzido pelo modelo de linguagem. Interpretações e escolhas metodológicas são da IA e podem conter erros. Recomenda-se replicação independente.

Tabela 13. Bibliotecas e versões (ambiente virtual Python 3.13.15, Windows 11).

Biblioteca	Versão	Uso
DuckDB	2.0.0 (base), 1.5.5 (leitura inicial)	leitura dos CSV do TSE, agregação e pareamento
NumPy / pandas	2.5.4 / 3.1.0	álgebra e tabelas
CVXPY + Clarabel	1.9.3 + 0.11.1	programa quadrático da inferência ecológica
PyFixest	0.60.0	regressões com efeitos fixos e erro agrupado
SciPy	1.18.1	qui-quadrado, Spearman, curtose
scikit-learn	1.9.1	Isolation Forest
PyMC / NumPyro / JAX	6.3.2 / 0.22.0 / 0.12.0.1	amostragem NUTS do modelo Bayesiano
PyEI / ArviZ	1.1.4 / 1.3.0	modelo RxC de Rosen et al.; diagnóstico
matplotlib	3.11.2	figuras
latex2mathml; Playwright (Chromium)	3.81.1; projeto	equações e geração do PDF

Tabela 14. Roteiros de análise, na ordem de execução.

Roteiro	Função
base.py, par.py	leitura dos CSV de 2026 e pareamento das seções
m1_contas.py	identidade contábil, viabilidade, regiões e faixas
ei_lib.py, m2_ei.py	Goodman restrito por zona, bootstrap e validação cruzada
hist_base.py, m3_hist.py	bases de 2014, 2018 e 2022; mesmo método nas quatro eleições
m4_forense.py, m4b.py	Klimek, comparecimento extra, dígitos, urna, horário, Isolation Forest
m5_bayes.py	modelo Bayesiano RxC (sementes 1 e 11)
m6_contrafactual.py, m7_viradas.py	contrafactual, Shapley e taxa de base de governador
m8_distorcoes.py	resíduos extremos, comparecimento e indicadores por UF
m9_linha_tempo.py, m9b_pausa.py	linha do tempo da apuração e teste antes × depois da pausa
m10_codigo.py	limites para desvio por software; seções fora da urna
figs.py, dossie.py	figuras e este documento

Parâmetros fixos: semente 2026 (validação cruzada, bootstrap, Isolation Forest), 7 (simulações de dígitos), 1 e 11 (amostras Bayesianas). Mínimo de 30 seções por zona; 2.000 reamostragens bootstrap; 100 simulações binomiais.

10. Conclusão do analista

Esta seção é o parecer da IA que conduziu a análise e se distingue dos resultados acima, que são medidas.

1. A virada é possível, e os dados mostram como ela aconteceu. A premissa de que todo o centro teria ido para Lula não se confirma: os eleitores dos eliminados se dividiram quase ao meio. A diferença veio de eleitores que não escolheram nenhum finalista no 1º turno, os que anularam, votaram em branco ou faltaram, e de uma fidelidade um pouco maior do eleitorado de Lula. O efeito somado basta para explicar 100% da variação da margem.

2. Não encontrei evidência estatística de fraude capaz de alterar o resultado. Em todos os testes que detectam adulteração localizada (seção, urna, horário, modelo, dígitos, enchimento), 2026 se comporta como 2022. A pausa de duas horas teve causa documentada e não deixou diferença na direção de alteração. As formas mais simples de desvio sistemático por software teriam de ser de duas a cinco vezes maiores do que os dados admitem para inverter o resultado.

3. O que foi realmente atípico merece registro. O 2º turno de 2026 teve comportamento eleitoral sem precedente nas três eleições anteriores: o maior aumento de comparecimento entre turnos e a

primeira virada presidencial desde 1989. Teve também um fato procedimental controverso, a prorrogação nacional da votação a pedido de uma campanha. A estranheza é compreensível. Mas estranheza não é evidência de fraude, e os números mostram mobilização ampla, espalhada por todas as regiões e faixas, não concentrada em seções suspeitas.

4. Sobre a hipótese de código adulterado na urna. É a única hipótese que este tipo de análise não descarta por completo. Ela pode ser limitada, e os limites calculados tornam improváveis os desvios simples desse tipo. Formas mais sofisticadas, como converter brancos e nulos ou alterar votos conforme a composição da seção, só se verificam por auditoria: inspeção do código-fonte, conferência dos hashes, relatórios dos testes de integridade de 2026 e comparação de amostras de boletins impressos com os arquivos publicados. Quem sustenta essa hipótese deve buscar esses documentos. Estatística sobre resultados agregados não substitui a auditoria do software.

Em síntese: com os dados públicos do TSE e os métodos aqui descritos, o resultado do 2º turno de 2026 é matematicamente coerente com o 1º turno e estatisticamente compatível com uma contagem íntegra. O que o distingue das eleições anteriores é o comportamento de quem não havia escolhido um finalista no 1º turno.

Referências

1. Abril-Pla, O. et al. (2023). PyMC: a modern, and comprehensive probabilistic programming framework in Python. *PeerJ Computer Science*, 9, e1516.
2. Beber, B.; Scacco, A. (2012). What the numbers say: a digit-based test for election fraud. *Political Analysis*, 20(2), 211–234.
3. Bradbury, J. et al. (2018). JAX: composable transformations of Python+NumPy programs. Software, versão 0.12.
4. Brasil de Fato (2026). TSE retoma apuração dos resultados do segundo turno; acompanhe. 25 out. 2026. brasildefato.com.br/2026/10/25/tse-retoma-apuracao-dos-resultados-do-segundo-turno-acompanhe/
5. Cameron, A. C.; Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, 50(2), 317–372.
6. Diamond, S.; Boyd, S. (2016). CVXPY: a Python-embedded modeling language for convex optimization. *Journal of Machine Learning Research*, 17(83), 1–5.
7. Efron, B.; Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. New York: Chapman & Hall.
8. Fischer, A. et al. PyFixest: fast high-dimensional fixed effects regression in Python. Software, versão 0.60.
9. Gelman, A.; Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4), 457–472.
10. Goodman, L. A. (1953). Ecological regressions and behavior of individuals. *American Sociological Review*, 18(6), 663–664.
11. Goulart, P. J.; Chen, Y. (2024). Clarabel: an interior-point solver for conic programs with quadratic objectives. arXiv:2405.12762.
12. Harris, C. R. et al. (2020). Array programming with NumPy. *Nature*, 585, 357–362.
13. Hoffman, M. D.; Gelman, A. (2014). The No-U-Turn Sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, 15, 1593–1623.
14. Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2), 65–70.
15. Hunter, J. D. (2007). Matplotlib: a 2D graphics environment. *Computing in Science & Engineering*, 9(3), 90–95.
16. InfoMoney (2026). TSE pausa divulgação de resultados das eleições com 0,25% das urnas apuradas. 25 out. 2026. infomoney.com.br/politica/tse-pausa-divulgacao-de-resultados-das-eleicoes-com-025-das-urnas-apuradas/
17. King, G. (1997). *A Solution to the Ecological Inference Problem: Reconstructing Individual Behavior from Aggregate Data*. Princeton: Princeton University Press.
18. Klimek, P.; Yegorov, Y.; Hanel, R.; Thurner, S. (2012). Statistical detection of systematic election irregularities. *Proceedings of the National Academy of Sciences*, 109(41), 16469–16473.

19. Knudson, K. C.; Schoenbach, G.; Becker, A. (2021). PyEI: a Python package for ecological inference. *Journal of Open Source Software*, 6(64), 3397.
20. Liu, F. T.; Ting, K. M.; Zhou, Z.-H. (2008). Isolation forest. *Proceedings of the 8th IEEE International Conference on Data Mining*, 413–422.
21. McKinney, W. (2010). Data structures for statistical computing in Python. *Proceedings of the 9th Python in Science Conference*, 56–61.
22. Mebane, W. R. Jr. (2006). Election forensics: vote counts and Benford's law. Summer Meeting of the Political Methodology Society, UC-Davis.
23. Mixvale (2026). Eleição presidencial de 2026: dados do TSE mostram que 35.521 votos foram registrados após as 17h no 2º turno. 26 out. 2026. (página não acessível; dado não verificado)
24. Pedregosa, F. et al. (2011). Scikit-learn: machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
25. Phan, D.; Pradhan, N.; Jankowiak, M. (2019). Composable effects for flexible and accelerated probabilistic programming in NumPyro. arXiv:1912.11554.
26. Raasveldt, M.; Mühleisen, H. (2019). DuckDB: an embeddable analytical database. *Proceedings of SIGMOD 2019*, 1981–1984.
27. Rosen, O.; Jiang, W.; King, G.; Tanner, M. A. (2001). Bayesian and frequentist inference for ecological inference: the R×C case. *Statistica Neerlandica*, 55(2), 134–156.
28. Shapley, L. S. (1953). A value for n-person games. In: Kuhn, H.; Tucker, A. (eds.). *Contributions to the Theory of Games II*. Princeton University Press, 307–317.
29. Tribunal Superior Eleitoral (2026). Portal de Dados Abertos: resultados por seção eleitoral, eleições 2014, 2018, 2022 e 2026. dadosabertos.tse.jus.br. Arquivos de 2026 gerados em 26/10/2026.
30. Vehtari, A.; Gelman, A.; Simpson, D.; Carpenter, B.; Bürkner, P.-C. (2021). Rank-normalization, folding, and localization: an improved R-hat for assessing convergence of MCMC. *Bayesian Analysis*, 16(2), 667–718.
31. Virtanen, P. et al. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, 17, 261–272.

Apêndice A — Resultados por UF (2026)

Tabela 15. Variação 1º → 2º turno e indicadores de distorção por UF.

UF	Ganho Lula	Ganho Flávio	Comparec. (p.p.)	Lula, % comparec. 1º → 2º	Excede estoque	Isolation F.	Resíduo líquido da margem
AC	+9.743	−15.092	−5,65	27,6% → 31,9%	0,2%	65	+417
AL	+157.162	+8.422	+0,54	51,4% → 59,1%	3,3%	11	+81
AM	+98.787	−1.519	−1,52	46,9% → 52,3%	1,9%	49	+1.773
AP	+10.465	−13.604	−8,06	44,6% → 51,9%	0,2%	57	+839
BA	+820.487	+166.176	+3,26	62,6% → 68,8%	1,3%	10	+2.800
CE	+480.861	+62.628	+1,64	60,2% → 67,0%	3,5%	11	+2.759
DF	+149.597	+69.926	+1,61	36,9% → 44,2%	0,1%	1	+202
ES	+151.937	+61.310	+1,91	36,1% → 41,6%	0,2%	3	+822
GO	+396.675	+293.446	+1,49	29,9% → 39,2%	0,1%	6	−737
MA	+241.930	−9.069	−0,63	61,7% → 68,0%	7,2%	38	+3.119
MG	+1.162.163	+464.238	+2,49	41,1% → 48,7%	0,5%	18	+2.957
MS	+112.118	+54.541	+2,63	33,5% → 39,4%	2,0%	2	−595
MT	+112.486	+52.518	+1,26	28,3% → 33,3%	2,3%	8	+126
PA	+263.212	+35.983	−1,22	47,9% → 53,9%	3,1%	100	+4.056
PB	+252.932	+33.833	+1,84	57,4% → 65,4%	2,4%	2	+770
PE	+563.542	+94.511	+2,98	60,1% → 67,2%	1,6%	7	+1.240
PI	+130.279	+8.831	−0,04	68,0% → 73,8%	3,1%	9	+1.081
PR	+511.907	+346.099	+2,53	30,0% → 36,4%	0,4%	24	−1.603
RJ	+872.372	+202.189	+1,69	37,5% → 45,3%	1,2%	15	+3.345
RN	+175.633	+16.916	+2,32	57,0% → 63,3%	2,3%	1	+1.521
RO	+43.779	+21.394	−0,21	25,3% → 29,8%	1,2%	7	+36
RR	+9.085	−11.647	−6,02	22,4% → 27,1%	2,6%	19	+519
RS	+529.793	+300.595	+2,19	34,2% → 41,0%	0,4%	12	−481
SC	+291.146	+220.051	+1,99	24,2% → 29,8%	0,3%	8	−734
SE	+133.022	+15.515	+1,59	59,5% → 67,4%	2,6%	0	+605
SP	+2.613.861	+1.217.155	+2,73	36,0% → 44,3%	0,1%	15	+3.473
TO	+38.953	−1.405	−0,88	41,9% → 46,4%	1,0%	2	+722
Ext.	+42.769	+27.597	+4,33	46,1% → 52,5%	0,5%	0	+599

Apêndice B — Taxa de base: disputas de governador

Tabela 16. Disputas de governador com 2º turno, 2014–2022, ordenadas pela vantagem do líder no 1º turno.

Ano	UF	Líder 1º turno	2º colocado	Vantagem (p.p.)	Eliminados	Vencedor (2º turno)	Virada
2014	RO	PMDB 35,9%	PSDB 35,4%	0,44	28,7%	PMDB 53,4%	
2018	AM	PSC 33,7%	PDT 32,7%	0,99	33,5%	PSC 58,5%	
2014	PA	PMDB 49,0%	PSDB 47,6%	1,38	3,3%	PSDB 51,9%	sim
2014	CE	PT 47,8%	PMDB 46,4%	1,40	5,8%	PT 53,4%	
2014	PB	PSDB 47,4%	PSB 46,0%	1,40	6,5%	PSB 52,6%	sim
2018	SC	PSD 31,1%	PSL 29,7%	1,40	39,2%	PSL 71,1%	sim
2022	MS	PRTB 26,8%	PSDB 25,2%	1,55	48,0%	PSDB 56,9%	sim
2022	RO	UNIÃO 38,9%	PL 37,0%	1,84	24,1%	UNIÃO 52,5%	
2022	PE	SOLIDARIEDADE 24,0%	PSDB 20,6%	3,39	55,4%	PSDB 58,7%	sim
2018	AP	PDT 33,6%	PSB 30,1%	3,45	36,3%	PDT 52,4%	
2018	RR	PSL 42,3%	PSDB 38,8%	3,48	18,9%	PSL 53,3%	
2014	MS	PT 42,9%	PSDB 39,1%	3,83	18,0%	PSDB 55,3%	sim
2014	RR	PP 41,5%	PSB 37,6%	3,86	20,9%	PP 54,9%	
2018	RS	PSDB 35,9%	MDB 31,1%	4,79	33,0%	PSDB 53,6%	
2014	RN	PMDB 47,3%	PSD 42,0%	5,30	10,6%	PSD 54,4%	sim
2022	SE	PT 44,7%	PSD 38,9%	5,79	16,4%	PSD 51,7%	sim
2022	SP	REPUBLICANOS 42,3%	PT 35,7%	6,62	22,0%	REPUBLICANOS 55,3%	
2018	RO	PSDB 31,6%	PSL 24,0%	7,60	44,4%	PSL 66,3%	sim
2014	RS	PMDB 40,4%	PT 32,6%	7,83	27,0%	PMDB 61,2%	
2022	ES	PSB 46,9%	PL 38,5%	8,46	14,6%	PSB 53,8%	
2022	BA	PT 49,5%	UNIÃO 40,8%	8,65	9,7%	PT 52,8%	
2018	SP	PSDB 31,8%	PSB 21,5%	10,24	46,7%	PSDB 51,8%	
2022	RS	PL 37,5%	PSDB 26,8%	10,69	35,7%	PSDB 57,1%	sim
2018	MS	PSDB 44,6%	PDT 31,6%	12,99	23,8%	PSDB 52,4%	
2018	MG	NOVO 42,7%	PSDB 29,1%	13,67	28,2%	NOVO 71,8%	
2018	RN	PT 46,2%	PDT 32,5%	13,71	21,4%	PT 57,6%	
2014	AP	PDT 42,2%	PSB 27,5%	14,65	30,3%	PDT 60,6%	
2022	PB	PSB 39,7%	PSDB 23,9%	15,76	36,4%	PSB 52,5%	
2014	DF	PSB 45,2%	PR 28,0%	17,26	26,8%	PSB 55,6%	
2018	PA	MDB 47,7%	DEM 30,2%	17,48	22,1%	MDB 55,4%	
2018	SE	PSD 40,8%	PSB 21,5%	19,35	37,7%	PSD 64,7%	
2014	GO	PSDB 51,0%	PMDB 31,6%	19,42	17,4%	PSDB 57,4%	
2014	AC	PT 49,7%	PSDB 30,1%	19,64	20,2%	PT 51,3%	
2022	AL	MDB 46,6%	UNIÃO 26,8%	19,85	26,6%	MDB 52,3%	
2014	RJ	PMDB 40,6%	PRB 20,3%	20,31	39,2%	PMDB 55,8%	
2022	SC	PL 38,6%	PT 17,4%	21,20	44,0%	PL 70,7%	
2018	RJ	PSC 41,3%	DEM 19,6%	21,72	39,2%	PSC 59,9%	
2022	AM	UNIÃO 42,9%	MDB 21,0%	21,85	36,1%	UNIÃO 56,6%	
2018	DF	MDB 42,0%	PSB 13,9%	28,04	44,1%	MDB 69,8%	

Apêndice C — Testes de irregularidade com p-valor

Tabela 17. p-valores brutos e ajustados por Holm (família de sete testes).

Teste	p	p ajustado (Holm)
Modelo de urna UE2013 × UE2022 (efeito fixo de município)	0,211	0,690
Modelo de urna UE2015 × UE2022	0,744	0,744
Modelo de urna UE2020 × UE2022	0,046	0,279
Hora de chegada do boletim (efeito fixo de município)	0,048	0,279
Pausa: resíduo de Lula, antes × depois (efeito fixo de zona)	0,172	0,690
Pausa: resíduo da margem, antes × depois	0,018	0,125
Pausa: variação de Lula com controles, antes × depois	0,254	0,690

Os testes de dígitos usam a posição do valor observado na distribuição simulada; fração das simulações com χ^2 maior ou igual ao observado, último dígito: a1 80%; b1 28%; a2 76%; b2 47% (a = Lula, b = Flávio, 1 e 2 = turno).