

Gasto em Redes Sociais, Incumbência e Dispersão Territorial do Voto nas Eleições Legislativas Brasileiras[†]

Marco Antonio Faganello*

Fundação Getulio Vargas (FGV)

28 de agosto de 2026

Resumo

Apesar de a literatura experimental mostrar efeito persuasivo próximo de zero para o gasto de campanha em redes sociais, candidatos brasileiros a cargos legislativos ampliaram exponencialmente esse investimento entre 2014 e 2022. Usando um desenho canônico de diferenças-em-diferenças na margem extensiva sobre um painel de recandidatos e uma regressão descontínua (RDD) multi-cadeira para sistemas proporcionais de lista aberta, este artigo isola dois momentos causais do gasto digital de campanha: a decisão de começar a investir e a transição para a incumbência. Os resultados mostram que candidatos que passam a investir em redes sociais, vindo de gasto zero, dispersam mais o voto territorialmente na eleição seguinte, com efeito concentrado no impulsionamento pago e ausente em gastos com sites. Candidatos que vencem uma cadeira por margem estreita reduzem, por outro lado, o peso proporcional do digital no orçamento seguinte, sem cortar o valor investido em termos absolutos. Essa diluição devolve a composição da campanha ao padrão tradicional associado à incumbência. Os achados sugerem que o gasto digital funciona, no Brasil, menos como instrumento de persuasão e mais como tecnologia de expansão territorial para quem ainda não tem rede local – um recurso que perde peso relativo assim que essa rede é conquistada.

Palavras-chave: gasto de campanha; redes sociais; dispersão eleitoral; incumbência; regressão descontínua; eleições brasileiras.

*Pesquisador Pós-Doutor Associado junto ao Centro de Política e Economia do Setor Público (CEPESP), Fundação Getulio Vargas. Contato: marcofaga@gmail.com

[†]Título de trabalho, sujeito a revisão. Comentários bem-vindos.

1 Introdução

A conclusão mais recorrente da literatura experimental sobre a eficácia da propaganda eleitoral é que ela não funciona: em meta-análise de 49 experimentos de campo conduzidos em eleições gerais americanas, robusta tanto a canvassing presencial quanto a mídia de massa, Kalla and Broockman (2018) mostram que o efeito persuasivo médio de contato de campanha é estatisticamente indistinguível de zero. O achado se replica no ambiente digital, onde a expectativa de um efeito diferente seria, a priori, ao menos plausível: Broockman and Green (2014) não encontram efeito de anúncios de display no Facebook sobre reconhecimento ou favorabilidade de candidato, e Coppock et al. (2022), testando anúncios patrocinados no Facebook e no Instagram durante as eleições de meio de mandato de 2018, com mais de um milhão de exposições registradas, estimam um efeito de $-0,04$ ponto percentual sobre o voto. Ou seja, um efeito estatisticamente nulo. Se propaganda paga, em qualquer meio, não muda votos, a sabedoria convencional deveria prever estagnação ou recuo do investimento das campanhas em publicidade digital. No entanto, os dados mostram o contrário.

Os dados categorizados de prestação de contas do Tribunal Superior Eleitoral (TSE) revelam o tamanho do descompasso: entre 2012 e 2024, a proporção de candidatos a prefeito que declaram algum gasto em redes sociais, impulsionamento pago e WhatsApp sobe de 13,1% para 48,4%, e o peso desse gasto sobre o orçamento total sobe de 0,66% para 6,52%. A trajetória é muito parecida com a observada entre candidatos legislativos no mesmo período. Esse aparente paradoxo – orçamento crescente, em praticamente todo cargo eletivo, para um instrumento de efeito persuasivo nulo – não é resolvido pela literatura existente. A pergunta da literatura de efeitos mínimos se refere ao volume e à relação entre o aumento do contato de campanha e o número de votos que um candidato recebe. Ora, se o gasto digital não compra volume adicional de voto e mesmo assim segue crescendo de forma generalizada, o que ele está comprando? Provavelmente, não mais eleitores. O gasto pode estar comprando alcance territorial: votos em municípios onde o candidato não tinha estrutura local anteriormente.

A quase totalidade da evidência experimental de efeito nulo vem de eleições majoritárias em distritos uninominais, nas quais a única pergunta possível é quanto se ganha no total, nunca onde esse voto está. O sistema brasileiro, ao contrário, elege deputados federais e estaduais por lista aberta proporcional, em circunscrições amplas e heterogêneas que coincidem com estados inteiros, onde cada candidato compete por voto pessoal não apenas contra adversários de outros partidos, mas contra os próprios colegas de legenda. O problema estratégico do candidato a deputado não é apenas quantos votos conseguir, mas onde consegui-los, e é exatamente essa dimensão territorial, ausente da competição binária de uma disputa majoritária, que a literatura de efeitos mínimos não tem como capturar.

Este artigo testa se o gasto digital de campanha, no contexto brasileiro, funciona como tecnologia de expansão territorial, permitindo que candidatos convertam investimento em redes sociais em alcance eleitoral fora de seus redutos históricos, inclusive por meio da formação de comunidades digitais que transcendem fronteiras físicas (Bor, 2014) – em vez de, ou além de, persuasão. Responder a essa pergunta exige isolar dois momentos causais distintos do ciclo de decisão de uma campanha, e atribuímos a cada um uma estratégia de identificação própria. O primeiro isola a decisão de adotar o digital: usando um painel de candidatos que disputam a mesma cadeira em eleições consecutivas, comparamos recandidatos que passam a investir em redes sociais, vindo de gasto zero na eleição anterior, com recandidatos que seguem sem investir, em um desenho canônico de diferenças-em-diferenças de dois períodos, na margem extensiva. O segundo isola a consequência de vencer: usando a generalização de Torres (2023) do desenho de regressão descontínua (RDD) para sistemas eleitorais multi-cadeira, comparamos candidatos que ficaram na fronteira exata entre conquistar ou perder uma vaga pelo método do maior quociente (D’Hondt), uma descontinuidade quase-experimental na definição de quem se torna incumbente.

Os dois desenhos, mesmo estimados de forma inteiramente independente sobre amostras e fontes de variação que não se sobrepõem, revelam um padrão consistente. Candidatos a deputado estadual que começam a gastar em redes sociais, vindo de zero, tendem a dispersar mais o voto entre municípios na eleição seguinte do que candidatos que seguem sem investir, efeito concentrado especificamente em impulsionamento pago e WhatsApp e ausente quando o gasto é em sites e hospedagem. Candidatos que vencem uma cadeira por margem mínima, por sua vez, passam a gastar, na eleição seguinte, uma fatia proporcionalmente menor de um orçamento maior em redes sociais: a queda chega a oito pontos percentuais entre deputados federais, ainda que o valor absoluto investido em digital não caia. Ou seja: as duas pontas do mesmo ciclo aparecem em dois momentos distintos da carreira do candidato. Quem ainda não tem rede local investe em digital e se dispersa territorialmente; quem já venceu, e presumivelmente já dispõe dessa rede, devolve o orçamento de campanha, em termos proporcionais, ao padrão tradicional associado à incumbência.

O artigo dialoga de perto com três linhas de pesquisa, a saber: 1) a literatura americana sobre geo-targeting digital; 2) o índice de concentração eleitoral que fornece a variável dependente do primeiro desenho; e 3) a literatura brasileira pré-digital sobre categorias de gasto. Quanto à primeira, Fowler et al. (2021) mostram que candidatos alocam mais verba para o Facebook quanto menor a congruência entre seu distrito e o mercado de cobertura da TV local, e Haenschen (2023) encontram que anúncios digitais só afetam comparecimento em contextos competitivos e em redutos de oposição sem estrutura presencial viável, mecanismo análogo ao que este artigo testa para o caso brasileiro, mas nunca estimado com desenho

causal fora de sistemas eleitorais majoritários e distritais. Quanto à segunda, Avelino et al. (2016) documentam que, entre 1994 e 2014, candidatos eleitos são sistematicamente mais dispersos territorialmente do que os não eleitos, mas seu desenho é descritivo e os autores convocam explicitamente pesquisa futura sobre os mecanismos que permitem essa dispersão. Quanto à terceira, Speck and Mancuso (2017) estabelecem, para 2014, que despesas tradicionais superam despesas modernas em retorno eleitoral e que incumbentes investem menos em mídia moderna mesmo vencendo muito mais. Em que pese o avanço dessa literatura, ela permanece descritiva, de corte transversal; o desenho de RDD deste artigo testa causalmente, pela primeira vez, o componente especificamente digital dessa relação.

A resposta importa além do caso brasileiro. Sistemas proporcionais de lista aberta, em que o voto pessoal compete com o voto de legenda e a circunscrição frequentemente ultrapassa qualquer mercado de mídia coerente, são regra, não exceção, em boa parte da América Latina e em segmentos crescentes da Europa. Se o gasto digital funciona nesses sistemas como substituto de rede local para quem não a tem, a expansão do impulsionamento pago tem implicação direta para o debate sobre regulação de propaganda digital em eleições: restringir o instrumento sem entender sua função pode fechar precisamente o canal que hoje permite a desafiantes territorialmente desconectados competir. O risco é reverter, na prática, um dos poucos mecanismos equalizadores aqui identificados.

O artigo segue a seguinte estrutura: a Seção 2 revisa a literatura sobre efeitos mínimos, equalização digital e dispersão territorial do voto, e deriva as previsões testáveis dos dois desenhos. A Seção 3 descreve os dados: são 29 milhões de registros de despesa eleitoral do TSE classificados por rede neural em 23 categorias (2002–2024), fundidos à votação por município via CEPESPDData. A Seção 4 detalha as duas estratégias de identificação. A Seção 5 apresenta os resultados de cada desenho, cada um acompanhado de seus testes de validade e da respectiva bateria de robustez. Por fim, a Seção 6 discute os achados à luz do debate entre equalização e normalização, nomeia as limitações do artigo e finaliza com a conclusão.

2 Teoria e Hipóteses

O mecanismo central deste artigo pode ser resumido numa frase: o gasto digital funciona como substituto de rede local. Candidatos sem estrutura física (comitês, cabos eleitorais, presença histórica num determinado território) usam o baixo custo das redes sociais para alcançar eleitores que a campanha presencial não alcançaria de outra forma. Esse rótulo reaparece ao longo desta seção e do restante do artigo, no lugar de repetir a descrição completa a cada menção.

O debate entre equalização¹ e normalização² digital ainda não tem resposta consolidada, mas quatro estudos recentes convergem em direções específicas, a saber: 1) o próprio mapeamento do debate, que alerta contra estudar redes sociais isoladamente; 2) evidência de que desafiantes gastam proporcionalmente mais em digital do que incumbentes; 3) evidência de que esse padrão se repete em sistemas de lista partidária; e 4) uma distinção de função entre mídia tradicional e digital que ajuda a resolver o paradoxo da introdução. Quanto ao primeiro, em revisão de 49 estudos, Severin-Nielsen (2023) sustenta que boa parte da literatura comete um erro metodológico ao estudar o uso de redes sociais isoladamente, sem controlar a presença do candidato na mídia tradicional, e nota que estudos sobre o resultado eleitoral (outcome) são bem mais raros do que estudos sobre o comportamento de campanha (output), lacuna que este artigo, focado em resultado, ajuda a preencher. Quanto ao segundo, Vanden Eynde and Maddens (2022) mostram, na Colômbia, que desafiantes gastam proporcionalmente mais em digital do que incumbentes, um efeito que desaparece quando o gasto é em rádio ou televisão; o tamanho do distrito eleitoral não prediz essa decisão, só a infraestrutura local de internet, o que refuta a hipótese de que o digital compensaria, por si só, a baixa densidade populacional. Quanto ao terceiro, Enli and Skogerbø (2013) encontram, num sistema de lista partidária comparável ao brasileiro, que desafiantes adotam redes sociais mais cedo do que incumbentes, motivados por marketing pessoal e mobilização, não por diálogo democrático. Quanto ao quarto, Bode et al. (2016) descrevem uma divergência de audiência: anúncios pagos de televisão mobilizam recursos para persuadir eleitores indecisos, com tom negativo, enquanto o conteúdo digital mobiliza a própria base partidária, com tom positivo. Se o digital serve à mobilização e não à persuasão, o paradoxo que abre este artigo deixa de ser paradoxo: o gasto cresce porque cumpre uma função eleitoral diferente da que a literatura de efeitos mínimos testa.

A distinção entre função de persuasão e função de mobilização, que Bode et al. (2016) documentam para o caso americano, ajuda a revisitar com mais profundidade a literatura de efeitos mínimos já citada na introdução. Kalla and Broockman (2018), Broockman and Green (2014) e Coppock et al. (2022) medem todos a mesma coisa: se o contato de campanha muda a decisão de um eleitor que, de outra forma, votaria diferente – a definição estrita de persuasão. Nenhum desses estudos testa se o gasto muda a distribuição espacial do voto entre eleitores que já estavam inclinados a apoiar o candidato, e é exatamente essa lacuna que Fowler et al. (2021) começam a preencher: quanto maior a precisão do targeting digital, mais a mensagem se desloca de conteúdo persuasivo para conteúdo mobilizador. Efeito

¹A ideia de que a internet e as redes sociais nivelam o campo de jogo eleitoral, permitindo que candidatos com menos recursos superem desvantagens estruturais frente a rivais mais bem financiados.

²A contrate-se de que as desigualdades já existentes fora da internet (recursos financeiros, tempo de TV) se reproduzem também no ambiente digital, mantendo a vantagem de quem já tem mais recursos.

persuasivo nulo, em agregado, e efeito de mobilização territorial positivo não são resultados contraditórios: são respostas a perguntas diferentes, sobre variáveis dependentes diferentes.

O segundo passo da lógica causal explica como a mobilização digital se converte especificamente em alcance territorial. Em análise qualitativa de campanhas americanas, Bor (2014) descreve como equipes de campanha usam redes sociais, deliberadamente, para engajar eleitores em localizações geográficas distantes, construindo o que a autora chama de comunidade virtual geograficamente distribuída, que se converte em ativismo de base fora do distrito físico do candidato. Fowler et al. (2021) oferecem a versão quantitativa do mesmo mecanismo: quanto menor a congruência entre o distrito eleitoral de um candidato e a área de cobertura do mercado de mídia de televisão, maior a fatia do orçamento alocada para o Facebook, para evitar o desperdício de impressões em eleitores fora da jurisdição, e essa realocação desloca, mais uma vez, a mensagem de persuasão para mobilização. Haenschen (2023), em experimento com mais de 870 mil eleitores, não encontra efeito agregado de anúncios digitais sobre o comparecimento às urnas, mas encontra um aumento de 1,66 ponto percentual especificamente em distritos competitivos e em redutos de oposição sem estrutura de campanha presencial: evidência direta de que o digital substitui, em vez de complementar, a campanha física em território negligenciado. É esse mecanismo, gasto digital como substituto de estrutura territorial ausente, que este artigo testa para o caso brasileiro.

Esse mecanismo importa mais para quem ainda não tem rede local do que para quem já tem, e a literatura brasileira pré-digital já documentava essa assimetria antes mesmo de o gasto digital se tornar relevante. Para as eleições legislativas de 2014, Speck and Mancuso (2017) estimam, por regressão logística, que um ponto percentual a mais de gasto em mídia tradicional multiplica por 3,0 a chance de eleição (razão de chances, $RC = 3,0$), e o mesmo ponto percentual em mídia moderna multiplica a chance por 1,78. No entanto, já ocupar um mandato multiplica a chance de eleição por 19,8: incumbentes vencem por um caminho que não depende de atualizar sua estratégia de mídia, e por isso investem pouco nela. O mesmo padrão aparece em contextos institucionais bem diferentes do brasileiro: na Colômbia, Vanden Eynde and Maddens (2022) encontram que a vantagem digital dos desafiantes desaparece quando o gasto é em rádio ou televisão; na Noruega, Enli and Skogerbø (2013) encontram que incumbentes adotam redes sociais mais devagar do que desafiantes. Três sistemas eleitorais distintos – lista aberta no Brasil, sistema misto na Colômbia, lista partidária na Noruega – convergem na mesma direção: o digital é ferramenta de quem ainda não tem vantagem estrutural.

A dispersão territorial do voto, o mecanismo por trás da nossa primeira hipótese (H1), precisa de uma forma de medição antes de virar hipótese testável. Avelino et al. (2016) propõem o Índice G, adaptado do índice de concentração industrial de Ellison and Glaeser

(1997) para a competição eleitoral, e mostram que, entre 1994 e 2014, candidatos eleitos são sistematicamente mais dispersos territorialmente do que os não eleitos, com o G médio dos eleitos caindo de 0,125 em 1994 para 0,091 em 2014. Silva and Davidian (2013) complementam o índice G com duas medidas que dizem onde, não só quanto: o Quociente Locacional (LQ), que compara a participação de um candidato num município com sua participação na circunscrição inteira, e o Cluster Horizontal (HC), que pondera essa razão pelo peso eleitoral do próprio município. Juntos, G, LQ e HC operacionalizam o alcance geográfico distante que Bor (2014), Fowler et al. (2021) e Haenschen (2023) descrevem qualitativa e experimentalmente: G mede se o candidato se dispersou; LQ e HC dizem para onde.

A lógica dos parágrafos anteriores converge numa previsão testável, batizada aqui de H1 (Hipótese da Expansão): candidatos que passam a investir em redes sociais, vindo de gasto zero na eleição anterior, dispersam mais o voto territorialmente na eleição seguinte do que candidatos que seguem sem investir. A previsão é sobre a margem extensiva, ou seja, a decisão de adotar o digital e não sobre a intensidade contínua do investimento. É precisamente o primeiro real gasto que compra o acesso à infraestrutura de segmentação que permite alcançar eleitores fora do reduto histórico; gastar mais depois de já ter essa infraestrutura montada é uma decisão de natureza diferente, sobre a qual a teoria aqui desenvolvida não faz previsão.

A mesma lógica, lida ao contrário, produz uma segunda previsão: H2 (Hipótese da Reversão). Se o gasto digital compensa a ausência de rede local, conquistar essa rede deveria reduzir a necessidade do substituto. A vitória numa eleição legislativa é, ela mesma, evidência de que o candidato construiu, ou herdou, uma base eleitoral capaz de garantir uma cadeira; um recém-eleito tem, na eleição seguinte, menos motivo estrutural para investir proporcionalmente em digital do que um candidato que ainda não venceu. H2 prevê, portanto, que a vitória de uma cadeira legislativa reduz a composição digital do orçamento de campanha na eleição seguinte, mesmo que o orçamento total não caia: a previsão é sobre como o dinheiro se distribui entre categorias de gasto, não sobre quanto dinheiro entra.

H1 e H2 não são a mesma hipótese testada duas vezes; são duas faces da mesma moeda, observadas em momentos diferentes da carreira de um candidato. H1 descreve o momento em que a rede local ainda não existe e o digital a substitui; H2 descreve o momento em que a rede local acaba de ser conquistada e o digital deixa de ser necessário na mesma proporção. Testar as duas juntas, e não cada uma isoladamente, é o que permite distinguir um mecanismo consistente ao longo do ciclo eleitoral de dois achados pontuais e desconectados: se apenas H1 se sustentasse, poderíamos suspeitar de uma tendência geral de dispersão do voto ao longo do tempo, sem relação com o digital; se apenas H2 se sustentasse, poderíamos suspeitar de um efeito genérico de incumbência sobre composição de gasto, sem relação com território. As duas hipóteses, sustentadas em conjunto e testadas com desenhos causais que não compartilham

amostra nem fonte de variação, dão à interpretação de substituto de rede local um grau de credibilidade que nenhuma delas teria isoladamente.

A hipótese de dispersão territorial não nasce no vácuo: a literatura brasileira pré-digital já documentava heterogeneidade geográfica no retorno do gasto de campanha, décadas antes de o digital virar categoria relevante de despesa. Para deputados estaduais do Ceará em 2010, Arraes et al. (2017) estimam que cada R\$100 mil gastos equivalem a 36.810 votos adicionais no estado como um todo, mas apenas 19.180 na região metropolitana de Fortaleza e 14.680 no interior: o mesmo volume de dinheiro rende retornos eleitorais bem diferentes dependendo de onde é gasto. Para prefeitos em 2012, Speck and Cervi (2016) decompõem o sucesso eleitoral num triângulo causal em que o dinheiro (coeficiente de 0,420) supera a memória eleitoral (0,375) e o tempo de rádio e TV (0,164), e mostram que o efeito do dinheiro diminui à medida que o município cresce. Para prefeitos em 2016, ainda antes da consolidação das redes sociais, Sampaio (2020) encontram que despesas tradicionais superam despesas modernas em todos os cinco portes de município pesquisados, com a criação de páginas na internet respondendo por apenas 0,46% do gasto total. Para prefeitos em 2020, Souza and Freitas (2024) encontram um cenário já bem diferente: o gasto digital soma R\$53,96 milhões, 3,25% do total declarado, 12^a posição entre 41 categorias de despesa – ainda um padrão de normalização, concentrado em candidatos homens, brancos e mais escolarizados, mas já não mais marginal. Em oito anos, o digital saiu de quase irrelevante para uma das dez categorias de despesa mais importantes do país; a heterogeneidade geográfica de retorno que Arraes et al. (2017) documentam para o gasto total, uma década antes de o digital existir como categoria declarável, é o precedente direto da hipótese que este artigo testa para o gasto especificamente digital.

H1 (Expansão) e H2 (Reversão) são as duas previsões testáveis que orientam o restante deste artigo. A seção seguinte descreve os dados e as duas estratégias de identificação usadas para testá-las; a seção de resultados retoma cada uma pelo nome, na mesma ordem em que aparecem aqui.

3 Dados

Os dois desenhos deste artigo usam uma base de despesa eleitoral que foi classificada de forma automática e inédita, por rede neural, a partir dos registros de prestação de contas do TSE.³ Compilamos, padronizamos e categorizamos um universo de 29 milhões de registros de despesas eleitorais de candidatos, nas 12 eleições ocorridas entre 2002 e 2024, no âmbito do Centro de Política e Economia do Setor Público (CEPESP) da Fundação Getulio Vargas

³Para mais informações e acesso à base: https://github.com/marcofaga/bert_gastos_candidatos_tse.

(FGV); a base resultante foi incorporada ao CEPESPData⁴, o repositório público de dados eleitorais do centro. A categorização foi construída com um modelo de linguagem baseado em BERT, pré-treinado em português (BERTimbau), aplicado ao próprio corpus de despesas (Task-Adaptive Pretraining) e refinado em quatro rodadas de aprendizado ativo, classificando cada descrição de gasto em uma de 23 categorias⁵ informadas pela literatura especializada em despesas de campanha e refinadas indutivamente, a partir de exploração e rotulação manual do próprio corpus.

Identificamos o gasto digital a partir de duas subcategorias: "Internet – redes sociais, impulsionamento e WhatsApp"(1,09% dos registros) e "Internet – sites, blogs e hospedagem"(0,26%). As duas categorias captam fenômenos de natureza distinta: a primeira cobre o gasto com alcance pago e segmentado (anúncios patrocinados, impulsionamento de conteúdo, disparo em massa via WhatsApp), o mecanismo de geo-targeting central à teoria deste artigo. A segunda é infraestrutura estática de presença digital, sem o componente de segmentação de público que sustenta a hipótese de expansão territorial. Esta é também a mais antiga das duas, já em queda desde o início da série, enquanto o impulsionamento pago só decola a partir de 2018.

Nenhum estudo do eixo de literatura brasileira revisado na seção anterior cobre escala comparável. Speck and Mancuso (2017) recodificam manualmente as categorias oficiais do TSE para um único ano (2014), restrito a deputados; Sampaio (2020) cobrem apenas o ano de 2016 e cargo de prefeito; Souza and Freitas (2024), o único trabalho nacional com o gasto digital como objeto central, também se limitam a um único ano (2020). A base deste artigo cobre as quatro eleições do seu escopo (2010, 2014, 2018 e 2022) com uma classificação automatizada e validada. É, ao que sabemos, a primeira a sustentar um teste causal do gasto digital ao longo de uma série multieleitoral brasileira.

Antes de qualquer desenho causal, os próprios dados de despesa já mostram o padrão que motiva os dois estudos. A Figura 1 traça a composição percentual do gasto de campanha por categoria, para Deputado Estadual e Federal, entre 2010 e 2022: Tradicional segue dominante o período inteiro (58% a 84% do gasto, a depender do ano e do cargo), mas o Digital (a fatia no topo de cada barra) cresce de forma visível e monotônica a partir de 2014, tanto para Estadual quanto Federal. O gasto digital médio salta de menos de 1% do orçamento em 2010 (0,66% em Estadual, 0,89% em Federal) para 7,37% em Estadual e 7,99% em Federal em 2022, mas a mediana entre candidatos permanece em zero até a própria eleição de 2022 – sinal, já nesta etapa puramente descritiva, de que gastar em digital é adoção recente e ainda minoritária, concentrada numa fração dos candidatos, não intensificação de um padrão já

⁴Cf. <https://cepesp.fgv.br/data>

⁵A lista completa das categorias que aparecem na amostra deste artigo, com a participação de cada uma no gasto total declarado, está na Tabela 6 do Apêndice.

disseminado.

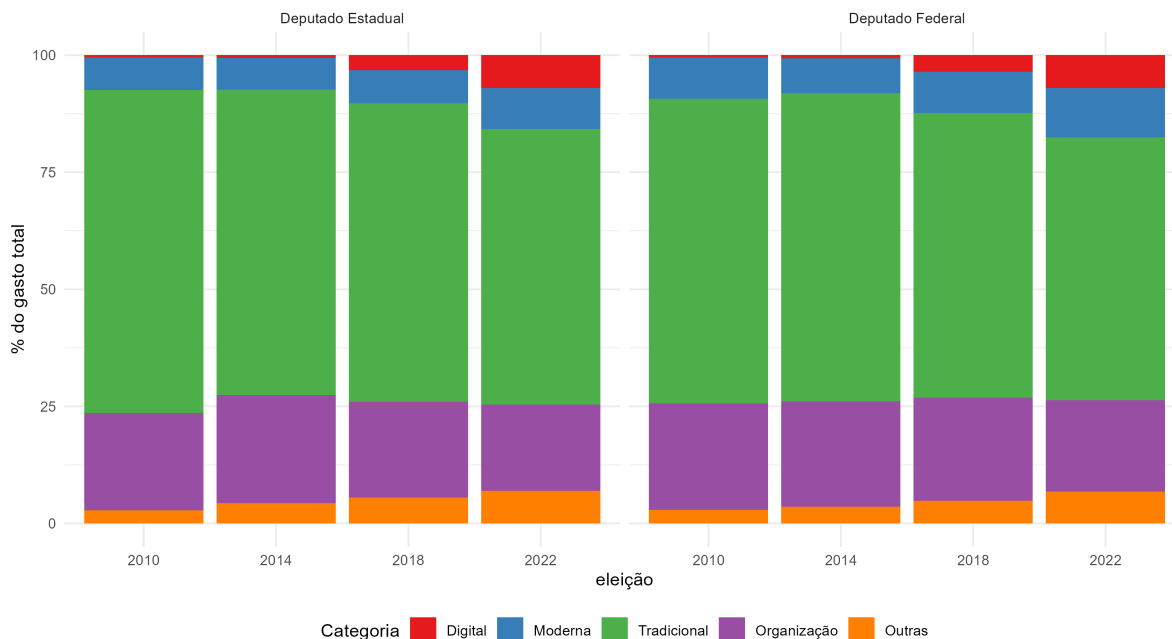


Figura 1: Composição percentual do gasto de campanha por categoria, Deputado Estadual e Federal (2010–2022). Script: `scripts/desc01_gasto_por_tipo.R`.

Esse padrão bruto, no qual o digital está em expansão real, mas ainda é minoritário, é o pano de fundo empírico para o teste formal da seção seguinte.

4 Desenho de Pesquisa

Os dois desenhos causais deste artigo isolam ameaças de identificação distintas. O primeiro isola a decisão de adotar o digital, eliminando heterogeneidade fixa entre candidatos (carisma, rede local histórica, reconhecimento de nome) que confundiria qualquer comparação puramente observacional; é o desenho que sustenta H1 (Expansão). O segundo isola a incumbência de seus determinantes de seleção: quem vence já era, antes da eleição, sistematicamente mais viável. O desenho restringe a comparação à fronteira exata de conquista de uma cadeira para isolar exatamente essa seleção; é o que sustenta H2 (Reversão).

4.1 Desenho 1: diferenças-em-diferenças na margem extensiva

O primeiro desenho compara candidatos que passam a gastar em redes sociais e WhatsApp, vindo de gasto zero na eleição anterior (tratamento: inicia o gasto em digital), com candidatos que seguem gastando zero em digital nas duas eleições (controle), dentro do mesmo painel

de recandidatos. O tratamento é definido na margem extensiva – uma variável binária, não a intensidade contínua do gasto – estimado por um desenho canônico de diferenças-em-diferenças de dois períodos e dois grupos (Card and Krueger, 1994; Lechner, 2011), sem exposição ao viés de comparação de tratamento escalonado que afeta desenhos com múltiplos períodos e múltiplos momentos de adoção (Goodman-Bacon and Marcus, 2020). Gastar zero em digital está longe de ser um comportamento raro que inviabilizaria o grupo de controle: é o comportamento modal do painel (a mediana de porcentagem do gasto digital é zero em quase todo ano e cargo), e a proporção de candidatos com algum gasto digital sobe de 16,31% (Estadual) e 23,02% (Federal) em 2010 para 47,97% (Estadual) e 57,53% (Federal) em 2022. A comparação intra-candidato, antes e depois da adoção, elimina a heterogeneidade fixa entre candidatos já discutida acima, sem precisar medi-la diretamente.

Formalmente, essa heterogeneidade fixa entra na equação em nível como um efeito não observado do candidato:

$$G_{i,t} = \alpha_i + \beta D_{i,t} + \gamma X_{i,t} + \lambda_t + \varepsilon_{i,t}$$

onde α_i absorve carisma, rede local histórica e reconhecimento de nome, e $D_{i,t}$ é o indicador de gasto digital positivo. Diferenciar entre o par de eleições consecutivas elimina α_i por construção:

$$\Delta G_i = \beta \Delta D_i + \gamma \Delta X_i + \Delta \lambda_t + \Delta \varepsilon_i$$

Como a amostra é restrita a candidatos que gastavam zero em $t - 1$, ΔD_i assume só dois valores: ficou em zero (controle) ou passou a gastar (tratado). A regressão de ΔG_i sobre esse indicador binário é a forma em painel do estimador de diferenças-em-diferenças (Angrist and Pischke, 2009, cap. 5). A especificação estimada também acrescenta o controle de escala e o efeito fixo de UF, com erro-padrão clusterizado por candidato, já que recandidatos com mais de uma transição de eleição contribuem com mais de um par ΔG_i ao painel. Sob o pressuposto de tendência paralela, β identifica o efeito médio do tratamento sobre os tratados (ATT).

A variável resposta é o Índice G de Avelino et al. (2016)⁶ dos candidatos a Deputado Estadual e Federal no período correspondente. O índice G possui a seguinte formulação:

$$G_i = \sum_m \left(\frac{V_{im}}{V_i} - \frac{V_m}{V} \right)^2$$

onde V_{im} é o voto do candidato i no município m , V_i seu voto total na circunscrição, e V_m/V a participação do município no eleitorado da circunscrição inteira. O índice só faz sentido em

⁶Validamos o índice contra os valores publicados na Tabela 1 de Avelino et al. (2016): para Deputado Federal, o G médio dos eleitos em 2010 é 0,093 (publicado: $\approx 0,095$) e em 2014 é 0,091 (publicado: 0,091, exato).

circunscrições multi-município com competição por voto pessoal, como no caso do legislativo brasileiro. O reduto de cada candidato (diagnóstico espacial complementar, via os índices LQ e HC de Silva and Davidian (2013)) é identificado pelo município de LQ máximo com piso de dez votos, evitando redutos espúrios; 96,5% dos candidatos têm reduto identificado por esse critério. A interpretação do índice G é direta: quanto maior o índice, mais concentrada é a votação do candidato nos municípios; quanto menor, mais dispersa.

4.2 Desenho 2: RDD multi-cadeira (Torres, 2023)

O segundo desenho testa a outra ponta do ciclo: o que acontece com o gasto digital quando um candidato conquista uma cadeira. Regressão descontínua (RDD) exige uma *running variable* contínua com um corte único que separa quem ganha de quem perde, mas o sistema proporcional de lista aberta não tem esse corte: existe uma zona de sobreposição em que um candidato pode conquistar cadeira com menos voto pessoal do que outro que perdeu, o que quebra a monotonicidade que o RDD de corte único exige. Torres (2023) resolve isso generalizando a *running variable* para sistemas multi-cadeira via distância mínima (resolvida por programação linear) até a fronteira entre duas alocações de cadeira consecutivas no método do maior quociente (D'Hondt), o critério usado no Brasil para distribuir cadeiras proporcionais entre partidos e coligações.⁷

Usamos uma amostra nacional com eleições para Deputado Federal e Estadual, nos anos 2010, 2014, 2018 e 2022; com a identificação de 3.648 candidatos na fronteira entre conquistar e perder uma cadeira. A variável resposta é a composição digital do orçamento na eleição seguinte, Y_i ; restrita a quem recandidatou. A amostra final soma 1.675 recandidatos (533 Federal e 1.142 Estadual).

Os dois desenhos diferem, por fim, no uso de controles. O Desenho 1 inclui em toda especificação o Δ do gasto total entre eleições, que isola composição de escala: sem ele, um candidato em ascensão que aumenta o orçamento inteiro, não especificamente a fatia digital, poderia ser confundido com adoção do digital em si; a especificação mais exigente acrescenta efeito fixo de UF, absorvendo choques estaduais comuns a todos os candidatos de uma circunscrição. O Desenho 2 é estimado sem covariáveis, como é padrão em RDD: a identificação vem da continuidade na fronteira, não do ajuste por observáveis, e a comparabilidade dos dois lados do corte é verificada diretamente pelos testes de validade em vez de imposta por controle.

Com os dois desenhos especificados e suas ameaças de identificação endereçadas, a seção seguinte apresenta os resultados que sustentam H1 (Expansão) e H2 (Reversão).

⁷A construção passo a passo, com exemplo numérico, está documentada no Apêndice (Seção A.5).

5 Resultados

5.1 H1: adoção digital e dispersão territorial

Testamos aqui a hipótese da expansão territorial (H1): candidatos que passam a investir em redes sociais, vindo de gasto zero, dispersam mais o voto territorialmente na eleição seguinte do que candidatos que seguem sem investir? O desenho é o de diferenças-em-diferenças na margem extensiva a partir de um painel de recandidatos restrito a quem gastava zero em digital em $t - 1$, comparando quem passa a gastar (tratado, $\Delta D_i = 1$) com quem segue em zero (controle). A amostra soma 5.490 recandidatos a Deputado Federal e Estadual, 4.061 em Estadual e 1.429 em Federal.

O resultado principal se mostrou significativo apenas para o cargo de Deputado Estadual. A adoção do gasto digital reduz o Índice G em 0,0150 (Tabela 1) – uma queda de aproximadamente 5,6% sobre a média do índice nesse cargo (G médio em Estadual = 0,270). Em Federal, o coeficiente é próximo de zero e não significativo (-0,0030 sem efeito fixo de UF; -0,0029 com efeito fixo; $p = 0,64$).

Um teste ingênuo poderia esperar o oposto: se o digital funciona como substituto de rede local, o cargo já mais disperso deveria ser onde esse substituto tem mais espaço para operar, e Deputado Federal é sistematicamente mais disperso do que Deputado Estadual, mesmo antes de qualquer gasto digital entrar na conta (G médio 0,220 contra 0,270; entre eleitos, a diferença quase dobra, 0,083 contra 0,145). É exatamente o contrário que os dados mostram: o efeito aparece só no cargo menos disperso (Estadual) e é nulo no mais disperso (Federal).

Nossa hipótese é que a dispersão elevada de Deputado Federal já é pré-condição estrutural da própria candidatura. Deputado Federal exige um quociente eleitoral (votos por cadeira) em média 2,49 vezes maior que Deputado Estadual, no mesmo estado e eleição, um padrão sem uma única exceção nas 104 combinações UF-eleição observadas entre 2010 e 2022. Um candidato federal minimamente viável já precisa de alcance territorial amplo antes de gastar um real em redes sociais; a dispersão chega por outros canais, como estrutura partidária e cobertura de mídia, e pelo próprio filtro de viabilidade que barra quem não tem esse alcance. É por isso que o digital não move o número em Federal: a variância que sobraria para ele explorar já é menor ali (teste F, $p < 0,001$). Em Deputado Estadual, por outro lado, um reduto único já garante uma cadeira, a dispersão é opção, não pré-condição, e é onde o digital, quando adotado, faz a diferença que a teoria prevê.

Tabela 1: H1 – efeito da adoção digital sobre a dispersão territorial

	Pooled	Estadual	Federal
Sem efeito fixo de UF	-0,011** (0,003)	-0,015*** (0,004)	-0,003 (0,006)
Com efeito fixo de UF	-0,010** (0,003)	-0,013*** (0,004)	-0,003 (0,007)
Observações	5.490	4.061	1.429

Variável dependente: ΔG_i . Coeficiente de ΔD_i , controlando $\Delta \log(\text{gasto total})_i$. Erro-padrão clusterizado por candidato entre parênteses. * $p < 0,1$; ** $p < 0,05$; *** $p < 0,01$. Fonte: CEPESPData/TSE. Script: `scripts/inferencia01_fd_h1_h2.R`.

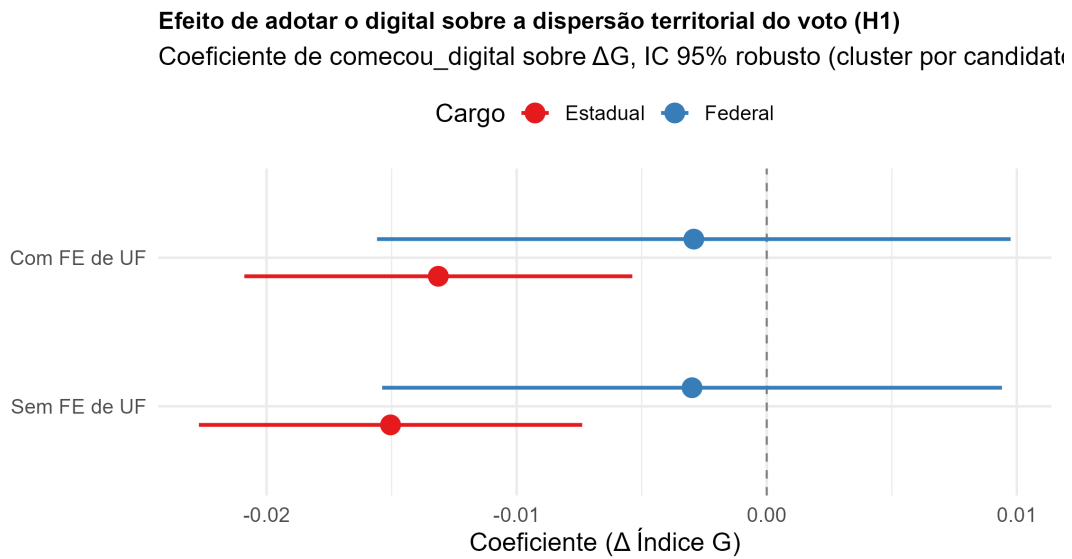


Figura 2: H1 – coeficiente de ΔD_i sobre ΔG_i , por cargo, com e sem efeito fixo de UF (IC 95% robusto). Script: `scripts/desc12_coefplot_h1b.R`.

A decomposição por subcategoria de gasto liga o resultado ao mecanismo de geo-targeting: em Estadual, o efeito está concentrado em impulsionamento pago e WhatsApp (Tabela 2) e ausente em sites e hospedagem. Isso condiz com nossa hipótese uma vez que o gasto em sites é uma infraestrutura estática e passiva, sem segmentação de público. Em Federal, as duas subcategorias são nulas, consistente com o efeito agregado também nulo nesse cargo. Os dados mostram que, para deputados estaduais, o gasto em publicidade digital segmentada está associado a uma maior dispersão territorial do voto para além do reduto local.

Tabela 2: H1 – decomposição por subcategoria de gasto digital

	Estadual	Federal
Redes sociais/WhatsApp	-0,0181*** (0,0041)	-0,0089 (0,0066)
Sites/hospedagem	-0,0071 (0,0057)	0,0056 (0,0081)
Observações	4.061	1.429

Variável dependente: ΔG_i . Cada célula é uma regressão separada, mesma amostra restrita a $p_{\text{digital},0} = 0$ da análise principal (Tabela 1), controlando $\Delta \log(\text{gasto total})_i$; erro-padrão clusterizado por candidato entre parênteses. * $p < 0,1$; ** $p < 0,05$; *** $p < 0,01$. Fonte: CEPESPDData/TSE.

A bateria de robustez sustenta a leitura causal, em dois testes (Tabela 7, Apêndice). O teste de pré-tendência mostra que o ΔG do par de eleições anterior não difere por status de adoção do par seguinte em Estadual ($p = 0,43$): candidatos que estão prestes a adotar o digital não vinham, antes disso, numa trajetória de dispersão territorial diferente da de quem segue sem adotar. Em Federal, no entanto, esse mesmo teste é marginalmente significativo ($p = 0,05$). O teste de placebo, aplicando a mesma especificação à categoria Moderna (rádio, TV, anúncio impresso, não digital), não encontra efeito em nenhum dos dois cargos (Estadual $p = 0,52$; Federal $p = 0,88$) – o resultado principal não é, portanto, artefato de alguma tendência geral de campanhas em expansão que afetaria qualquer categoria de mídia paga. Um controle de Δ do gasto total, presente em toda especificação, isola a mudança de composição do orçamento do efeito de escala geral da campanha. Quebrado por par de eleição, o efeito em Estadual se concentra no par mais recente (2018→2022, coeficiente -0,0180, $p < 0,05$), consistente com a maturação do impulsionamento pago como categoria de gasto (Seção 3).

H1 estabelece, portanto, que a decisão de adotar o digital dispersa o voto territorialmente em Deputado Estadual, e que esse efeito é específico à categoria com componente de segmentação paga. O Estudo seguinte testa a outra ponta do mesmo ciclo: o que acontece com a composição do gasto quando o candidato já conquistou a rede local via incumbência.

5.2 H2: incumbência e composição digital do gasto

H1 isola a decisão de adotar o digital, mas não testa o outro lado do ciclo: o que acontece com a composição do gasto quando o candidato já tem a rede local que o digital viria substituir. A margem de vitória, ao contrário da decisão de investir, não é escolha do candidato – é quase-experimental na fronteira exata de conquista de uma cadeira, o que permite isolar a incumbência de seus determinantes de seleção.

O desenho é o RDD multi-cadeira de Torres (2023) descrito na Seção 4. Temos aqui uma amostra nacional de 3.648 candidatos na fronteira entre conquistar e perder uma cadeira, restrita a quem recandidatou. São 1.675 recandidatos, sendo 1.142 em Estadual e 533 em Federal. O outcome é a composição digital do orçamento na eleição seguinte, Y_i .

Vencer a cadeira por margem mínima reduz Y_i em 3,4 pontos percentuais no pooled (Tabela 3), concentrado em Deputado Federal: 8,3 pontos percentuais. Em Estadual, o efeito, ao contrário de H1, não é distinguível de zero. Para efeito de escala, o gasto digital médio da população de candidatos gira em torno de 7 a 8% do orçamento em 2022 (Seção 3) – a queda de 8,3 pontos em Federal representa a maior parte dessa fatia. A decomposição por subcategoria mostra o mesmo padrão de H1 (Tabela 4): o efeito está concentrado em redes sociais/WhatsApp. O gasto dessa categoria cai 3,0 pontos percentuais na versão pooled e 8,1 pontos percentuais entre deputados federais. Novamente é nulo em Estadual e ausente em sites/hospedagem em todos os recortes.

Tabela 3: H2 – efeito de vencer a cadeira sobre a composição digital do gasto

	Pooled	Federal	Estadual
Coefficiente	-0,034**	-0,083***	-0,018
EP (convencional)	(0,013)	(0,025)	(0,015)
p (robusto)	0,023	0,001	0,280
N	1.675	533	1.142

Variável dependente: Y_i (composição digital do orçamento em t_1). `rdrobust`, bandwidth MSE-ótimo, erro-padrão clusterizado por candidato. * $p < 0,1$; ** $p < 0,05$; *** $p < 0,01$ (baseado no p robusto). Fonte: CEPESPDData/TSE.

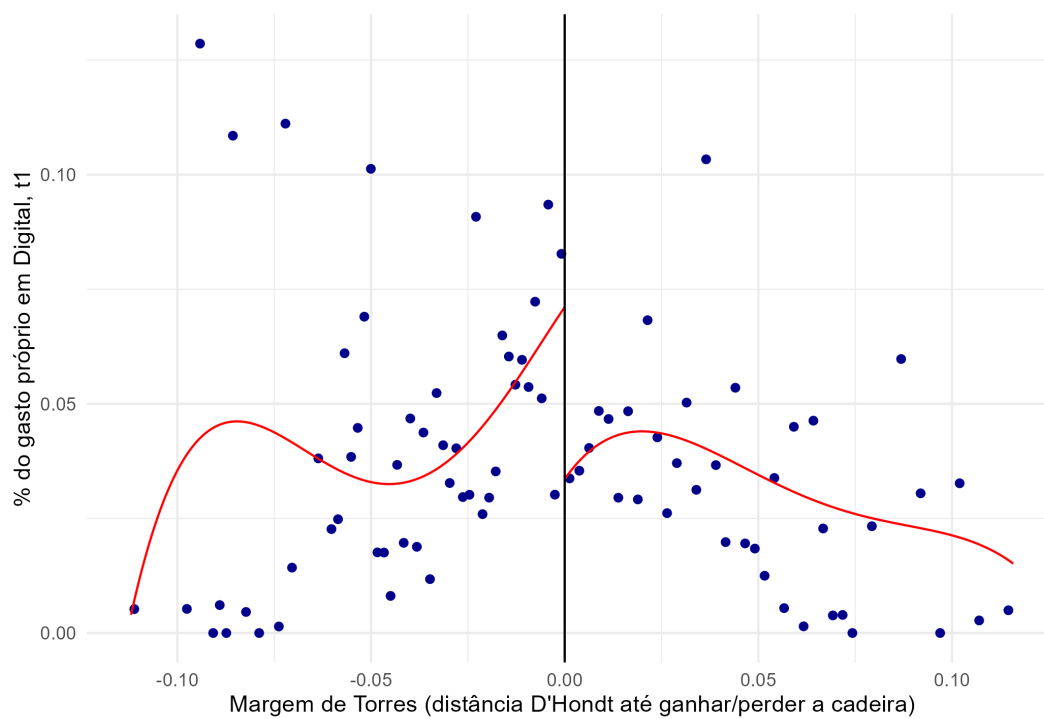


Figura 3: $H2 - Y_i$ no corte da margem de Torres, pooled nacional.

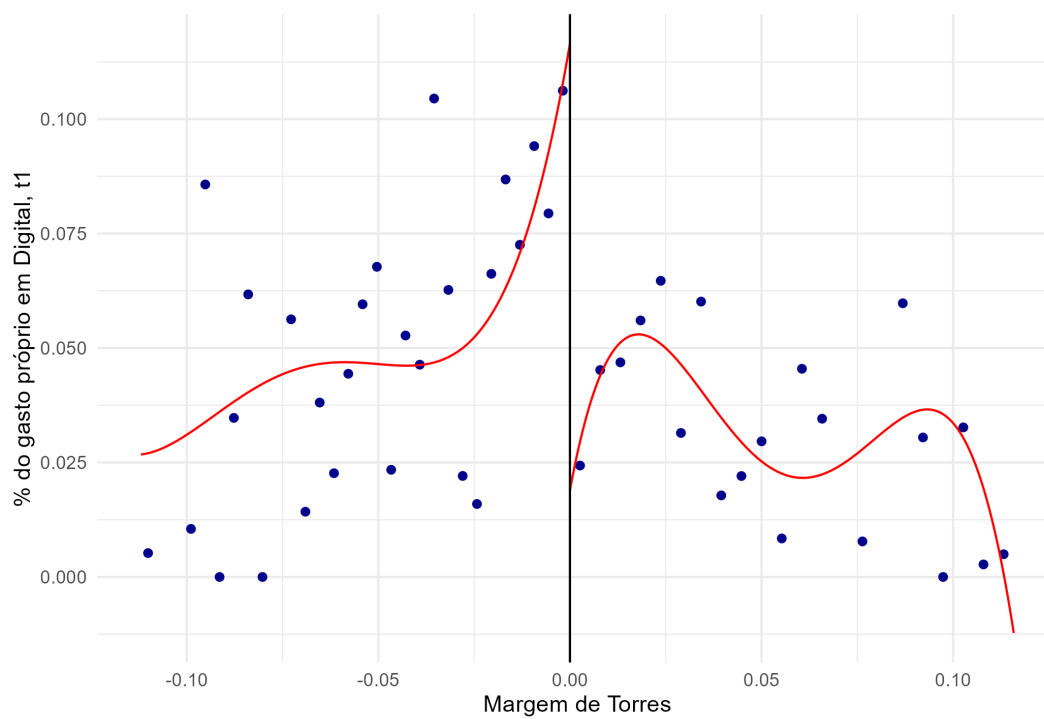


Figura 4: $H2 - Y_i$ no corte da margem de Torres, Deputado Federal (onde o efeito se concentra). Script: `scripts/bd38_rdd_p_digital_torres_brasil.R`.

Tabela 4: H2 – decomposição por subcategoria de gasto digital

	Pooled	Federal	Estadual
Redes sociais/WhatsApp	-3,0** (1,2)	-8,1*** (2,5)	-1,4 (1,5)
Sites/hospedagem	-0,4 (0,3)	-0,3 (0,6)	-0,4 (0,3)
N	1.675	533	1.142

Variável dependente: subcategoria de Y_i , em pontos percentuais do orçamento em t_1 . Coeficiente convencional (`rdrobust`), erro-padrão entre parênteses; significância baseada no p robusto. Bandwidth MSE-ótimo, erro-padrão clusterizado por candidato. * $p < 0,1$; ** $p < 0,05$; *** $p < 0,01$. Fonte: CEPESPDData/TSE. Script: `scripts/bd38_rdd_p_digital_torres_brasil.R`.

A validade do desenho passa por dois testes. O teste de manipulação de densidade de McCrary (2008), na versão generalizada de Cattaneo et al. (2020), sobre todos os 3.648 candidatos na fronteira, não rejeita a continuidade da densidade no corte em nenhum dos três recortes (pooled $T = -1,214$, $p = 0,225$; Federal $T = -1,157$, $p = 0,247$; Estadual $T = -0,914$, $p = 0,361$). Não há sinal de que partidos ajustam coligação para empurrar um candidato de um lado ao outro do corte. O segundo teste verifica o balanceamento de covariáveis pré-tratamento no corte, prática recomendada para qualquer RDD de corte eleitoral (Lee and Lemieux, 2010): nenhuma das três covariáveis testadas – composição digital, gasto total e incumbência, todas medidas em t_0 – rejeita continuidade em nenhum dos três recortes (p robusto entre 0,127 e 0,865, Tabela 8). A figura do primeiro teste está no Apêndice (Figura 5). O balanceamento na própria composição digital do gasto em t_0 é o teste mais direto contra a ameaça de que o resultado acima já existisse antes da vitória.

Os dois estudos, estimados de forma inteiramente independente sobre amostras e fontes de variação que não se sobrepõem, descrevem as duas pontas do mesmo ciclo: quem ainda não tem rede local adota o digital e se dispersa territorialmente (H1, Estadual); quem já venceu e, presumivelmente, já construiu essa rede devolve o orçamento, em termos proporcionais, ao padrão tradicional associado à incumbência (H2, mais forte em Federal).

6 Discussão e Conclusão

Este artigo abriu com uma pergunta simples, motivada por um paradoxo que a literatura experimental não resolve: se o gasto de campanha em redes sociais não muda quantos votos um candidato recebe, o que ele está comprando, afinal? Os dois desenhos causais analisados aqui respondem, de forma independente e concordante, à mesma pergunta: o gasto digital não compra votos adicionais, compra alcance territorial. Quem ainda não tem uma rede local investe em redes sociais e passa a dispersar mais o voto entre municípios na eleição seguinte;

quem já venceu uma cadeira e, presumivelmente, já construiu essa rede devolve na eleição seguinte parte proporcional do orçamento de campanha ao padrão tradicional associado à incumbência. Estimados sobre amostras e fontes de variação que não se sobrepõem, H1 e H2 descrevem, cada um a seu tempo, o mesmo mecanismo: o gasto digital funciona, no Brasil, como substituto de rede local.

Uma limitação merece registro explícito. O mecanismo territorial que sustenta a interpretação dos dois achados – geo-targeting, no sentido de Fowler et al. (2021) e Haenschen (2023) – não é observado diretamente: não temos acesso a dados geolocalizados de anúncio nem à segmentação de público que cada campanha efetivamente usou, e inferimos o mecanismo do padrão agregado de gasto e votação, não de registro publicitário.

O que esses achados dizem sobre o debate entre equalização e normalização digital, revisitado na Seção 2? A leitura aqui sustentada é deliberadamente estreita: os achados apontam para uma equalização territorial, não de desempenho. Os dois desenhos testam se o digital muda a distribuição espacial do voto (H1) e a composição do orçamento de campanha (H2) – nunca se o digital aumenta a probabilidade de vitória ou o volume total de votos de um desafiante, pergunta à qual a própria literatura de efeitos mínimos já responde negativamente (Seção 1). O digital não muda quantos votos um desafiante recebe; muda onde esses votos estão, permitindo que candidatos sem rede local preexistente cheguem a eleitores que a campanha presencial não alcançaria. É uma forma de equalização mais modesta do que uma hipótese ampla de equalização de desempenho, mas também mais precisa: reconcilia o efeito persuasivo nulo da literatura experimental com o crescimento real do investimento, ao localizar o benefício do digital numa dimensão, a distribuição espacial do voto, que a maior parte dessa literatura não mede.

A implicação prática decorre diretamente dessa leitura territorial. Propostas de regulação de propaganda digital em eleições – restrição a impulsionamento pago, limites de gasto por categoria, controle mais rígido sobre disparo em massa via WhatsApp – tendem a discutir o instrumento pelo prisma da desinformação e da manipulação de voto. No entanto, os achados aqui sugerem que o mesmo instrumento cumpre, para uma parte dos candidatos, uma função distinta: substituir a estrutura territorial que desafiante sem rede local historicamente não têm. Restringir esse instrumento sem diferenciar sua função pode fechar, na prática, um dos poucos canais que hoje permite a candidatos territorialmente desconectados competir por voto fora do próprio reduto, reduzindo o risco de manipulação ao custo de reduzir também a dispersão territorial da própria competição eleitoral. O desenho de qualquer regulação futura ganharia precisão ao distinguir esses dois efeitos, em vez de tratá-los como a mesma coisa.

Resta fechar o paradoxo que abriu este artigo: por que o investimento em redes sociais cresce de forma tão consistente se o efeito persuasivo médio, testado repetidamente por de-

senhos experimentais robustos, é nulo? A resposta que os dois estudos sugerem é que o gasto cresce porque cumpre uma função diferente daquela que a literatura de efeitos mínimos testa. Persuasão pergunta se o contato de campanha muda o voto de quem, de outra forma, votaria diferente; os achados aqui reportados são sobre onde o voto de quem já estava inclinado a apoiar o candidato se distribui territorialmente. Não há contradição entre um efeito agregado nulo sobre a primeira pergunta e um efeito real sobre a segunda – são variáveis dependentes diferentes, e a literatura de efeitos mínimos nunca teve a segunda em seu escopo. O paradoxo se desfaz, portanto, não porque a evidência experimental esteja errada, mas porque ela responde a uma pergunta mais estreita do que a que orienta o comportamento real das campanhas.

Se há uma frase que resume o que este artigo encontrou, é a que já nomeia o mecanismo central: o gasto digital funciona, na competição legislativa brasileira, como substituto de rede local. É o recurso de quem ainda não tem reduto – e, assim que o reduto é conquistado, um recurso que perde peso relativo diante das mesmas ferramentas tradicionais que garantiram a vitória. Nesse sentido bem específico, digital sem reduto talvez seja também digital com prazo de validade.

Referências

- Angrist, J. D. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.
- Arraes, R., Amorim, O., and Simonassi, A. (2017). Despesas de Campanha e Sucesso Eleitoral nos Pleitos Legislativos Brasileiros*. *Dados*, 60:1059–1093.
- Avelino, G., Biderman, C., and Silva, G. P. d. (2016). A Concentração Eleitoral no Brasil (1994-2014). *Dados*, 59(4):1091–1125.
- Bode, L., Lassen, D., Kim, Y., Shah, D., Franklin Fowler, E., Ridout, T., and Franz, M. (2016). Coherent campaigns? Campaign broadcast and social messaging. *Online Information Review*, 40:580–594.
- Bor, S. E. (2014). Using Social Network Sites to Improve Communication Between Political Campaigns and Citizens in the 2012 Election. *American Behavioral Scientist*, 58(9):1195–1213.
- Broockman, D. E. and Green, D. P. (2014). Do Online Advertisements Increase Political Candidates' Name Recognition or Favorability? Evidence from Randomized Field Experiments. *Political Behavior*, 36(2):263–289.

- Card, D. and Krueger, A. B. (1994). Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. *American Economic Review*, 84(4):772–793.
- Cattaneo, M. D., Jansson, M., and Ma, X. (2020). Simple Local Polynomial Density Estimators. *Journal of the American Statistical Association*, 115(531):1449–1455.
- Coppock, A., Green, D. P., and Porter, E. (2022). Does digital advertising affect vote choice? Evidence from a randomized field experiment. *Research & Politics*, 9(1):20531680221076901.
- Ellison, G. and Glaeser, E. L. (1997). Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach. *Journal of Political Economy*, 105(5):889–927.
- Enli, G. S. and Skogerbø, E. (2013). Personalized Campaigns in Party-Centred Politics: Twitter and Facebook as arenas for political communication. *Information, Communication & Society*, 16(5):757–774. _eprint: <https://doi.org/10.1080/1369118X.2013.782330>.
- Fowler, E. F., Franz, M. M., Martin, G. J., Peskowitz, Z., and Ridout, T. N. (2021). Political Advertising Online and Offline. *American Political Science Review*, 115(1):130–149.
- Goodman-Bacon, A. and Marcus, J. (2020). Using Difference-in-Differences to Identify Causal Effects of COVID-19 Policies. *Survey Research Methods*, 14(2):153–158.
- Haenschen, K. (2023). The Conditional Effects of Microtargeted Facebook Advertisements on Voter Turnout. *Political Behavior*, 45(4):1661–1681.
- Kalla, J. L. and Broockman, D. E. (2018). The Minimal Persuasive Effects of Campaign Contact in General Elections: Evidence from 49 Field Experiments. *American Political Science Review*, 112(1):148–166.
- Lechner, M. (2011). The Estimation of Causal Effects by Difference-in-Difference Methods. *Foundations and Trends in Econometrics*, 4(3):165–224.
- Lee, D. S. (2008). Randomized experiments from non-random selection in U.S. House elections. *Journal of Econometrics*, 142(2):675–697.
- Lee, D. S. and Lemieux, T. (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, 48(2):281–355.
- McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714.

- Sampaio, D. (2020). Campanhas Tradicionais ou Modernas? Estratégias de Gastos nas Eleições Municipais de 2016. *Revista Brasileira de Ciências Sociais*, 36:e3610511.
- Severin-Nielsen, M. K. (2023). Politicians’ social media usage in a hybrid media environment: A scoping review of the literature between 2008–2022. *Nordicom Review*, 44(2):172–193.
- Silva, G. P. d. and Davidian, A. (2013). Identification of Areas of Vote Concentration: Evidences from Brazil. *Brazilian Political Science Review*, 7(2):141–155.
- Souza, B. M. and Freitas, A. S. K. L. (2024). Eleições pós-web: os gastos com campanha online em destaque nas eleições municipais de 2020 no Brasil. *Em Tese*, 21:01–34.
- Speck, B. W. and Cervi, E. U. (2016). Dinheiro, Tempo e Memória Eleitoral: Os Mecanismos que Levam ao Voto nas Eleições para Prefeito em 2012. *Dados*, 59:53–90.
- Speck, B. W. and Mancuso, W. P. (2017). “Street fighters” e “media stars”: estratégias de campanha e sua eficácia nas eleições brasileiras de 2014. *Cadernos Adenauer*, XVIII.
- Torres, S. (2023). Close Elections Regression Discontinuity Designs in Multi-seat Systems. Technical Report 23, Universidade de los Andes – CEDE. SSRN 2536286.
- Vanden Eynde, G. and Maddens, B. (2022). Explaining digital campaign expenses: The case of the 2018 legislative elections in Colombia. *Journal of Information Technology & Politics*, 19(3):302–315. _eprint: <https://doi.org/10.1080/19331681.2021.1973643>.

A Apêndice

A.1 Apêndice de dados: definição de variáveis

A Tabela 5 documenta a correspondência entre a notação matemática usada no corpo do artigo e as variáveis efetivamente construídas na base de dados, para fins de reprodutibilidade.

Tabela 5: Definição de variáveis

Símbolo	Variável	Descrição	Script
$D_{i,t}$	<code>p_digital > 0</code>	Indicador de gasto digital positivo (redes sociais/WhatsApp ou sites/hospedagem) na eleição t	<code>bd04_painel_candidato_eleicao.R</code>
ΔD_i	<code>comecou_digital</code>	Indicador de adoção: <code>p_digital</code> = 0 em $t-1$ e > 0 em t , no painel de recandidatos	<code>inferencia01_fd_h1_h2.R</code>
$G_{i,t}$	<code>G</code>	Índice de dispersão territorial do voto (Avelino et al., 2016)	<code>bd06_indice_g_lq.R</code>
$\Delta \log(\text{gasto total})_i$	<code>delta_log_gasto</code>	Varição do logaritmo do gasto total do candidato (<code>val_total_def</code>) entre o par de eleições	<code>inferencia01_fd_h1_h2.R</code>
–	<code>sg_uf</code>	Efeito fixo de Unidade da Federação	<code>inferencia01_fd_h1_h2.R</code>
–	<code>cd_cargo_padrao</code>	Cargo (Deputado Federal ou Estadual); usado para restringir/dividir a amostra	<code>inferencia01_fd_h1_h2.R</code>
Y_i	<code>p_digital_t1</code>	Composição digital do orçamento na eleição seguinte à fronteira de Torres, outcome do Desenho 2	<code>bd38_rdd_p_digital_torres_brasil.R</code>
–	<code>margem_torres</code>	Distância mínima assinada até a fronteira de Torres (2023): <i>running variable</i> do Desenho 2	<code>bd37_ponte_candidato_torres_brasil.R</code>
–	<code>ganhou_t0</code>	Indicador de vitória na fronteira, tratamento do Desenho 2	<code>bd38_rdd_p_digital_torres_brasil.R</code>

Todas as variáveis derivam da base de despesas classificada (Seção 3) fundida com a votação por candidato $\times$ município (CEPESPDData/TSE). Painel de candidato $\times$ eleição salvo em `dados/painel_candidato_eleicao.rds`; construção completa documentada na pipeline `scripts/bdXX`.

A.2 Categorias de despesa

A Tabela 6 lista as categorias de despesa que aparecem na amostra deste artigo (Deputado Federal e Estadual, 2010–2022), com o número de registros e a participação de cada uma no gasto total declarado (valores deflacionados, IPCA). A segunda coluna indica o agrupamento em cinco categorias mais amplas – Digital, Moderna, Tradicional, Organização, Outras –, usado na Figura 1 e nos dois desenhos causais das seções seguintes.

Tabela 6: Categorias de despesa, participação no gasto total (Deputado Federal e Estadual, 2010–2022)

Categoria	Grupo	N	% do gasto total
Materiais impressos e publicidade visual	Tradicional	780.228	29,59
Militância, cabos e panfletagem	Tradicional	2.627.526	17,06
Serviços gerais, pessoal, coordenação, administração de campanha	Organização	1.435.285	16,08
Locação de veículo	Tradicional	227.573	6,50
Combustível e peças automotivas	Tradicional	231.268	3,94
Publicidade, design, marketing, comunicação	Moderna	68.277	3,36
Internet – redes sociais, impulsionamento e WhatsApp	Digital	72.954	2,76
Comitê, imóvel, escritório e materiais de expediente	Organização	175.121	2,58
Carro de som	Tradicional	65.647	2,48
Jurídico	Outras	38.529	2,07
Jingles, locução, vídeos, áudio, audiovisual	Moderna	45.206	1,89
Anúncios em jornal ou revista	Moderna	68.168	1,78
Contabilidade e prestação de contas	Outras	60.333	1,77
Despesas postais	Tradicional	55.043	1,50
Eventos e comícios	Tradicional	40.923	1,44
Transporte e passagens	Organização	47.560	1,28
Rádio e TV	Moderna	26.799	0,99
Alimentação e água	Organização	89.557	0,93
Indefinido	Outras	40.189	0,59
Taxas, impostos, estorno e sobras	Outras	429.054	0,41
Internet – sites, blogs, hospedagem	Digital	17.541	0,41
Pesquisa	Moderna	3.095	0,34
Utilidades públicas	Organização	49.712	0,24

N é o número absoluto de lançamentos de despesa em cada categoria; % do gasto total é a participação da categoria no valor total gasto (R\$, deflacionado), não no número de lançamentos. As duas colunas não são proporcionais entre si, já que o valor médio por lançamento varia por categoria.

Fonte: CEPESData/TSE.

A.3 Testes de validade – Desenho 1

A Tabela 7 reporta os dois testes de validade do Desenho 1 (Seção 4), Estadual e Federal lado a lado. Colunas (1)-(2): pré-tendência – ΔG do par de eleições *anterior* contra o status de adoção do par *atual*; um coeficiente não significativo confirma que quem está prestes a adotar o digital não vinha, antes disso, numa trajetória de dispersão diferente. Colunas (3)-(4): placebo – mesma especificação, mas o tratamento é começar do zero em gasto Moderna (não digital); um coeficiente não significativo confirma que o achado principal não é um artefato de expansão geral de campanha.

Tabela 7: Desenho 1 – testes de pré-tendência e placebo

Cargo	ΔG (par anterior)		ΔG (par atual)	
	Estadual (1)	Federal (2)	Estadual (3)	Federal (4)
Constante	-0,024*** (0,005)	-0,002 (0,010)	-0,013*** (0,004)	-0,003 (0,009)
Começou digital (par atual)	-0,006 (0,007)	-0,024* (0,012)		
$\Delta \log(\text{gasto total})$	-0,001 (0,004)	-0,002 (0,004)	-0,001 (0,002)	0,002 (0,003)
Começou Moderna			0,004 (0,007)	-0,002 (0,011)
Observações	887	262	1.670	564

Erro-padrão clusterizado por candidato entre parênteses. * $p < 0,1$; ** $p < 0,05$; *** $p < 0,01$. Fonte: CEPESP-Data/TSE.

A.4 Testes de validade – Desenho 2

A Figura 5 dá suporte visual ao teste de manipulação de densidade (McCrary, 2008; Cattaneo et al., 2020) na margem de Torres; o segundo teste de validade (balanceamento de covariáveis pré-tratamento) está na Tabela 8, abaixo.

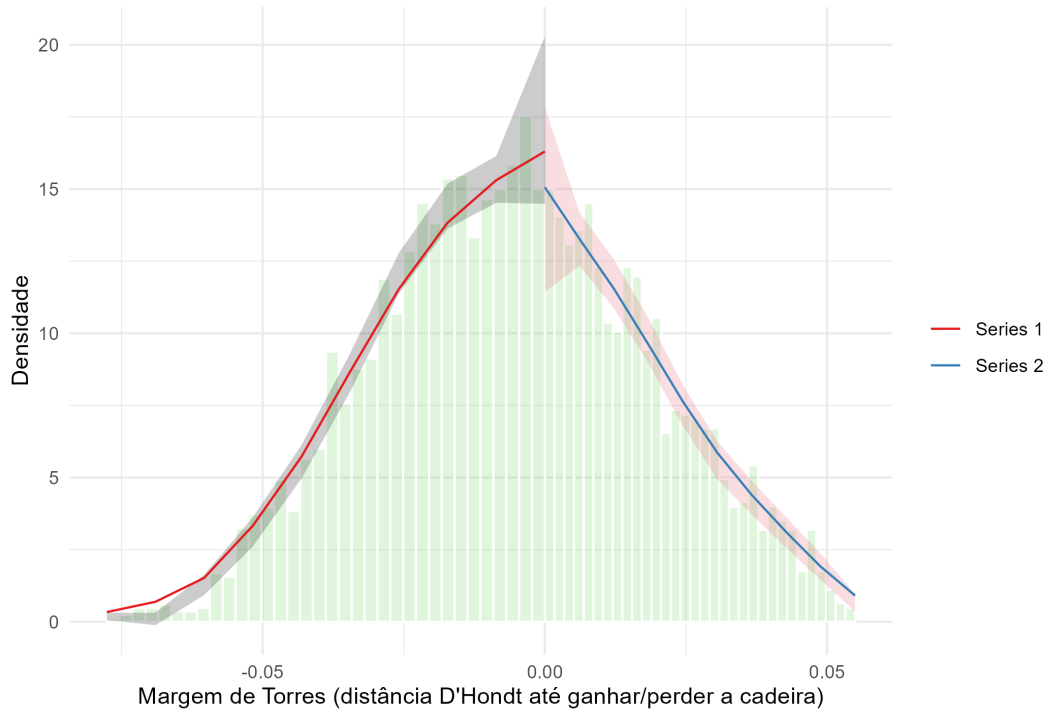


Figura 5: Teste de manipulação (densidade) na margem de Torres, pooled nacional. Script: `scripts/bd40_manipulacao_torres-brasil.R`.

Tabela 8: Balanceamento de covariáveis pré-tratamento no corte da margem de Torres

Covariável (em t_0)	Amostra	Coefficiente	EP (robusto)	p (robusto)	N
Composição digital do gasto	Pooled	-0,007	0,008	0,470	3.568
	Federal	-0,013	0,014	0,407	1.198
	Estadual	0,002	0,011	0,674	2.370
Log do gasto total	Pooled	0,127	0,160	0,308	3.568
	Federal	0,328	0,252	0,127	1.198
	Estadual	0,027	0,176	0,865	2.370
Incumbência prévia	Pooled	0,035	0,051	0,624	2.987
	Federal	0,151	0,107	0,159	1.043
	Estadual	0,025	0,086	0,594	1.944

Cada linha é uma regressão RDD separada (`rdrobust`, bandwidth MSE-ótimo, erro-padrão clusterizado por candidato), com a covariável pré-tratamento no lugar do outcome. Coeficiente e EP na especificação convencional; p na especificação robusta (bias-corrected). Nenhuma covariável rejeita continuidade no corte. Fonte: CEPESPDData/TSE. Script: `scripts/bd52_balanceamento_covariaveis_torres-brasil.R`.

A.5 Exemplo numérico: a fronteira proposta por Torres (2023)

O caso já conhecido: dois partidos. Em eleição majoritária, "ganhar por pouco" é simples: basta calcular a diferença de voto entre o primeiro e o segundo colocado. Qualquer resultado possível é um par de fatias (A, B) que soma 100% entre os dois candidatos, e todo esse espaço de resultados pode ser representado em uma reta com uma escala espelhada ao longo da reta: um extremo é representado por $(100, 0)$, o outro por $(0, 100)$, e a fronteira entre "A ganha/B perde" e "B ganha/A perde" é um único ponto: o empate em 50%-50% (Figura 6). A margem de vitória é a distância do ponto observado até o ponto de empate hipotético.

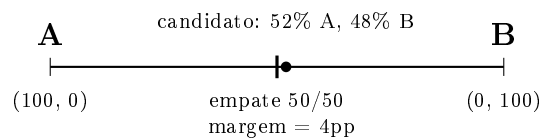


Figura 6: O caso de dois partidos: a fronteira é um ponto só, o empate

Três partidos: a mesma lógica, uma dimensão a mais. Em eleição proporcional com mais de dois partidos, isso não funciona – quem ganha uma cadeira depende de como os votos de todos os partidos se combinam, não de um par isolado de candidatos, e o espaço de resultados possíveis deixa de caber numa reta. O exemplo a seguir, adaptado de Torres (2023), ilustra a lógica com três partidos disputando três cadeiras pelo método do maior quociente (D'Hondt): Partido A com 200 votos, Partido B com 140, Partido C com 70.

D'Hondt divide o voto de cada partido por 1, 2, 3... e aloca as cadeiras aos maiores quocientes da lista combinada (Tabela 9):

Tabela 9: Exemplo numérico – quocientes D'Hondt, 3 partidos, 3 cadeiras

	$\div 1$	$\div 2$	$\div 3$
A (200 votos)	200	100	67
B (140 votos)	140	<u>70</u>	47
C (70 votos)	<u>70</u>	35	23

Os três maiores quocientes (em negrito) decidem a alocação: 200 (A), 140 (B), 100 (A). Os dois quocientes seguintes (sublinhados) empatam em 70.

A alocação resultante é A com duas cadeiras, B com uma, C com nenhuma. A terceira cadeira foi do Partido A, com quociente $A/2 = 100$; o quociente seguinte da fila, o mais próximo de disputar essa cadeira, não é um único rival – $B/2$ e $C/1$ empatam exatamente em 70. O Partido B (que já tem uma cadeira) e o Partido C (que não tem nenhuma) estão

igualmente perto de tomar essa cadeira de A. Essa distância entre o resultado observado e o ponto em que a alocação teria sido diferente é a *running variable* de Torres: a mesma lógica da margem de vitória majoritária, generalizada para o caso em que a cadeira em disputa depende da votação conjunta de todos os partidos – e, como o empate acima já sugere, nem sempre de um único rival.

Lendo um resultado como um ponto no triângulo. Com N partidos, cada resultado possível de eleição pode ser descrito por um ponto em uma figura geométrica de N vértices e $N - 1$ dimensões (*simplex*), onde cada ponto dentro da figura corresponde a uma distribuição possível de votos entre os N partidos. Portanto, no caso acima, a figura é um triângulo bidimensional. Cada vértice do triângulo é, nesse caso, um ponto da figura que representa um extremo hipotético, 100% dos votos para um partido só: (100%, 0%, 0%) no vértice A, (0%, 100%, 0%) no B, (0%, 0%, 100%) no C. O resultado observado é um ponto dentro da figura representando as fatias de cada partido (49%, 34%, 17%). Cada alocação de cadeiras diferente corresponde a uma região desse espaço – todos os resultados de votação que, mantida a mesma regra (D'Hondt), teriam dado exatamente aquela distribuição de cadeiras entre os partidos. A pergunta "qual é a menor mudança que teria alterado o resultado" vira, nesse desenho, uma pergunta geométrica simples de enunciar: qual é o ponto da fronteira entre duas regiões vizinhas mais próximo do resultado observado?

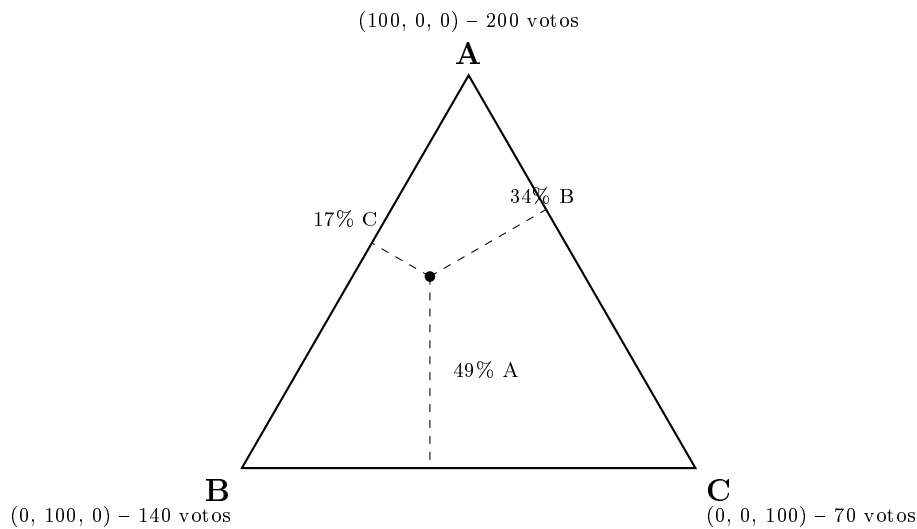


Figura 7: Como ler um ponto: a distância até cada lado recupera a fatia do vértice oposto

Torres (2023) demonstra que essas regiões têm um formato bem-comportado: convexo, sem reentrâncias ou buracos⁸. Na prática, isso significa que não existem atalhos escondidos

⁸Uma região é convexa quando o segmento de reta entre quaisquer dois pontos dela permanece inteiramente dentro dela. Reentrância é uma dobra que faz a borda da região curvar para dentro (como uma lua crescente ou uma estrela);

nem respostas conflitantes, e o ponto exato mais próximo da fronteira pode ser encontrado por métodos padrão de otimização – o mesmo tipo de cálculo usado, por exemplo, em logística para achar a rota mais curta sujeita a restrições. É essa distância mínima entre o resultado eleitoral observado e a fronteira mais próxima que teria mudado a alocação de cadeiras que generaliza a ideia de "margem de vitória" para qualquer sistema proporcional (D'Hondt incluído), preservando a mesma interpretação causal de um RDD tradicional: quanto mais perto da fronteira, mais o resultado se aproxima de um sorteio, e mais segura é a comparação entre quem ficou de um lado e do outro do corte.

Determinação da fronteira. No nosso exemplo, a fronteira nasce do empate da Tabela 9: $B/2$ e $C/1$, ambos em 70, contra $A/2 = 100$. Cada condição de empate – $A=B$, que separa "A mantém a cadeira" de "B toma a cadeira", e $A=2C$, que separa "A mantém" de "C toma" – é uma equação linear nas três fatias. A primeira, $A=B$, não envolve C : em $c = 0$ ela força $a = b = 1/2$, o ponto médio da aresta $A-B$; no outro extremo, fazendo $b = 0$, a mesma equação força $a = 0$ também, sobrando $c = 1$ – o próprio vértice C . A segunda, $A=2C$, não envolve B : em $c = 0$ ela força $a = 0$, sobrando $b = 1$, o próprio vértice B ; na aresta $A-C$ ($b = 0$), resolvendo $A=2C$ junto com $A+C = 1$ dá $c = 1/3$, $a = 2/3$ – um ponto a um terço do caminho entre C e A . As duas retas se cruzam num ponto do triângulo (Figura 8) próximo do resultado observado.

Esse ponto de cruzamento é um **vértice** do mosaico de alocações de cadeiras: perto dali não existe "o" rival mais próximo – existem dois, B e C , ao mesmo tempo, e qual deles de fato toma a cadeira depende da direção específica em que o voto se desloca. Por isso a pergunta certa não é comparar o resultado observado contra um par de partidos escolhido de antemão, mas contra *todas* as regiões de alocação vizinhas ao mesmo tempo – exatamente o que o método geral de Torres (2023) faz, e que uma subtração entre um único par de quocientes, suficiente quando há só uma fronteira em jogo, não capturaria aqui.

Por que não bastaria subtrair o voto bruto. Uma alternativa mais simples pareceria óbvia: pegar o voto pessoal do candidato mais votado entre os que não se elegeram e subtrair do voto pessoal do candidato menos votado entre os que se elegeram. Essa subtração não funciona no sistema brasileiro porque cadeiras são disputadas por **coligações** (ou partidos e federações isolados), não por candidatos isolados – quem decide se uma coligação ganha uma cadeira é o quociente dela inteira (voto agregado dividido pelo número de cadeiras já obtidas), não o voto pessoal de um candidato específico. Isso quebra a comparação direta: um candidato pode ter mais voto pessoal que outro e ainda assim não se eleger, se a coligação

buraco é um vazio interno que a região não cobre (como uma rosquinha). É a ausência dos dois que garante uma resposta única e computável para "qual é o ponto mais próximo da fronteira".

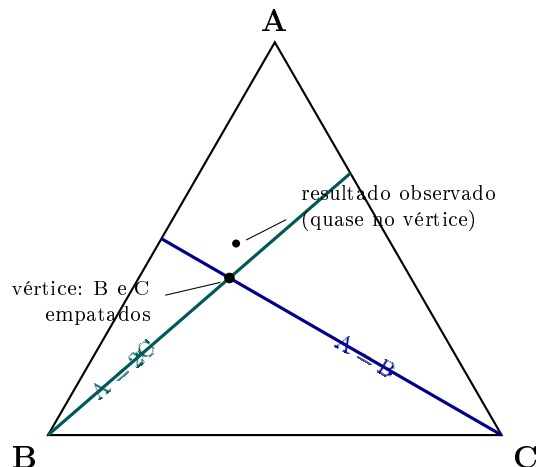


Figura 8: Duas fronteiras (A contra B, A contra C) se encontram no mesmo vértice

dele for mais fraca no total, e o oposto também ocorre. O voto bruto individual, ordenado do maior para o menor, não separa de forma confiável quem ganhou de quem perdeu – é por isso que a margem precisa ser calculada sobre os quocientes de coligação (como fizemos acima), não sobre o ranking de voto pessoal.

Por que isso ainda é um RDD. As retas da Figura 8 não são a *running variable* – são só o objeto geométrico usado para construí-la. A running variable de fato é um único número por observação: a distância mínima, com sinal, entre o resultado observado e o ponto mais próximo dessa fronteira – positiva se o partido ficou do lado de ganhar a cadeira, negativa se ficou do lado de perder, e exatamente zero se o resultado caísse em cima da própria fronteira (o empate). Reduzido a esse único número, o problema geométrico – reta, triângulo, ou o *simplex* de $N - 1$ dimensões do caso geral com N partidos – recupera exatamente a estrutura de um RDD tradicional (Lee, 2008): uma running variable escalar com corte pontual em zero, só que calculado por uma distância geométrica em vez da subtração direta de voto do caso majoritário.