

Estado da publicação: O preprint não foi publicado em outro meio.

A segunda barreira: gênero, classe e diversidade social no recrutamento de deputadas federais no Brasil (1998-2022)

Adriano Codato

<https://doi.org/10.1590/SciELOPreprints.16970>

Submetido em: 2026-07-20

Postado em: 2026-07-20 (versão 1)

(AAAA-MM-DD)

A moderação deste preprint recebeu o(s) endosso(s) de:

- Nilton Sainz (ORCID: <https://orcid.org/0000-0002-3957-2714>)

A segunda barreira: gênero, classe e diversidade social no recrutamento de deputadas federais no Brasil (1998-2022)*

The Second Barrier: Gender, Class, and Social Diversity in the Recruitment of Women to Brazil's Chamber of Deputies (1998-2022)

Adriano Codato¹ <https://orcid.org/0000-0002-5015-4273>

Universidade Federal do Paraná (UFPR); Instituto Nacional de Ciência e Tecnologia Representação e Legitimidade Democrática (INCT ReDem)

*Paper apresentado no seminário “30 anos de cotas eleitorais de gênero no Brasil” realizado no dia 28 de abril de 2026 na Câmara dos Deputados, Brasília – DF.

Resumo

O crescimento da representação feminina costuma ser avaliado pelo número de cadeiras conquistadas. Este artigo desloca o foco de “quantas mulheres” para “quais mulheres” e investiga se a expansão numérica das deputadas federais brasileiras foi acompanhada por maior *diversidade social*, mensurada a partir das posições ocupacionais declaradas ao Tribunal Superior Eleitoral. O universo reúne os 3.591 eleitos para a Câmara dos Deputados nas sete eleições de 1998 a 2022, entre os quais 380 mulheres. A análise principal utiliza 3.342 registros (356 mulheres e 2.986 homens), classificados em 12 categorias sociais; são excluídos apenas 249 casos registrados como ‘aposentado’, ‘não informado’ ou ‘outros’. A média não ponderada de Simpson das bancadas femininas caiu de 0,862 para 0,633 (-26,5%), mas a tendência monotônica não é estatisticamente corroborada (Kendall $\tau = -0,429$; p exato bilateral = 0,239). A inclinação de Sen é -0,0308 por eleição (IC95% [-0,0845; 0,0166]). Simpson, Shannon corrigido e rarefação não apoiam de modo consistente a hipótese de menor diversidade feminina. tampouco há intensificação temporal da diferença ou maior diversidade das deputadas da esquerda. Nenhum dos quatro testes principais permanece significativo após Holm. Os dados mostram expansão numérica e queda descritiva da diversidade social, mas não confirmam de modo robusto uma segunda barreira. O desenho observacional não identifica mecanismos causais.

Palavras-chave: representação política, gênero; classe social, recrutamento parlamentar, diversidade social, origem ocupacional, cotas eleitorais, Brasil.

Abstract

The growth of women's political representation is usually assessed by the number of seats won. This article shifts the focus from “how many women” to “which women,” asking whether the numerical expansion of women in Brazil's Chamber of Deputies was accompanied by greater social diversity, operationalized through the occupations candidates self-declared to the Superior Electoral Court. The study population comprises all 3,591 representatives elected across the seven general elections from 1998 to 2022, of whom 380 are women. The primary analysis uses 3,342 records (356 women and 2,986 men) classified into 12 social categories; only 249 cases coded as “retired,” “not reported,” or “other”

¹ Corresponding Author:

adrianocodato@gmail.com

<https://orcid.org/0000-0002-5015-4273>

Universidade Federal do Paraná

Departamento de Ciência Política

Rua General Carneiro, 460, sala 514

Curitiba - Paraná - Brasil

are excluded. The unweighted mean Simpson diversity of women's delegations fell from 0.862 to 0.633 (−26.5%), but the monotonic trend is not statistically supported (Kendall's $\tau = -0.429$; exact two-sided $p = 0.239$). Sen's slope is -0.0308 per election (95% CI $[-0.0845, 0.0166]$). The Simpson index, bias-corrected Shannon entropy, and rarefaction do not consistently support lower diversity among women, and there is no evidence that the gender gap widens over time or that women elected on the left are more socially diverse. None of the four primary tests remains significant after Holm correction. The data show numerical expansion alongside a descriptive decline in social diversity, but do not robustly confirm a “second barrier.” The observational design precludes causal identification.

Keywords: political representation, gender, social class, legislative recruitment, social diversity, occupational background, electoral quotas, Brazil.

1. Introdução

A eleição de 2022 produziu a maior bancada feminina da história da Câmara dos Deputados: 91 mulheres, ou 17,7% das 513 cadeiras. Em 1998, primeira eleição federal realizada sob a legislação nacional de cotas de candidaturas, haviam sido eleitas 29 deputadas. O crescimento foi de 214% em 24 anos. A expansão é politicamente relevante, mas a contagem de cadeiras responde apenas a uma dimensão da representação descritiva. Ela informa quantas mulheres chegaram ao Parlamento, não quais trajetórias sociais foram admitidas nem se o conjunto das eleitas se tornou internamente mais heterogêneo.

Essa distinção importa porque *grupos historicamente excluídos não são socialmente homogêneos*. A presença de mulheres pode aumentar enquanto permanecem sub-representadas mulheres oriundas do trabalho manual, da agricultura, dos serviços de baixa remuneração ou de outras posições com menor disponibilidade de recursos econômicos, redes partidárias e capital político. O problema não é substituir gênero por classe, mas observar sua articulação no recrutamento. Se a seleção partidária exige credenciais cada vez mais raras, a abertura de uma porta institucional pode coexistir com o estreitamento das origens sociais de quem consegue atravessá-la.

O artigo propõe o conceito de *segunda barreira* para nomear essa possibilidade.

A primeira barreira é a *desigualdade de gênero* que reduz a entrada de mulheres na competição e no exercício de mandatos. A segunda corresponde a uma seleção social dentro do próprio conjunto feminino: entre as mulheres potencialmente elegíveis, perfis dotados de maior capital econômico, educacional, familiar, ocupacional ou político teriam maior vantagem. O conceito é uma hipótese de mecanismo, não uma conclusão derivada automaticamente dos índices de diversidade. Os dados utilizados aqui observam o resultado agregado do recrutamento; não acompanham decisões de dirigentes, oferta de candidaturas, distribuição intrapartidária de recursos ou conversão de votos em cadeiras.

Empiricamente, o estudo pergunta: *a expansão numérica das deputadas federais entre 1998 e 2022 foi acompanhada por diversificação das posições sociais medida pela variável “ocupação”?*

Três desdobramentos orientam a análise.

- a) Primeiro, testa-se se a diversidade social das bancadas femininas apresenta tendência temporal.
- b) Segundo, compara-se a diversidade social de mulheres e homens dentro do mesmo bloco ideológico e da mesma eleição, distinguindo dominância de categorias frequentes e amplitude do repertório de posições sociais.

- c) Terceiro, avalia-se se as bancadas femininas dos partidos de esquerda são socialmente mais diversas do que as da direita.

A contribuição do paper é dupla. Substantivamente, o artigo amplia a agenda brasileira sobre cotas, partidos e carreiras políticas ao tratar a composição social interna das bancadas femininas como variável dependente. Metodologicamente, adapta medidas de diversidade oriundas da ecologia e da teoria da informação, explicita o nível de agregação das categorias do TSE e submete as conclusões a testes de classificação, tamanho dos grupos e dependência temporal.

2. Representação, recrutamento e composição social

2.1 Da presença à diversidade interna

A teoria da representação distingue dimensões formais, descritivas e substantivas (Pitkin, 1967). A representação descritiva pergunta em que medida os representantes se parecem socialmente com os representados; a substantiva, se atuam em favor de seus interesses. Phillips (1995) mostrou que a “política de presença” não pode ser dissolvida numa política de ideias: experiências e posições sociais influenciam quais problemas entram na agenda, como são interpretados e quem é reconhecido como interlocutor legítimo. A literatura sobre cotas traduziu esse argumento em reformas destinadas a aumentar a presença de mulheres nos espaços decisórios (Krook, 2009).

“Presença”, contudo, não é uma variável unidimensional. A categoria “mulheres” inclui diferenças de classe, raça, região, religião, geração e trajetória profissional. Uma abordagem interseccional alerta que desigualdades não operam apenas de forma aditiva e que a inclusão de membros socialmente privilegiados de um grupo não elimina a exclusão de suas parcelas subalternizadas (Crenshaw, 1989). Por isso, um aumento da proporção de mulheres pode coexistir com estabilidade ou redução da diversidade interna do grupo eleito.

No Brasil, a Lei nº 9.100/1995 reservou 20% das candidaturas nas eleições municipais de 1996 para mulheres. A Lei nº 9.504/1997 estendeu a regra às eleições proporcionais, com percentual transitório de 25% em 1998 e mínimo de 30% nos pleitos seguintes. A Lei nº 12.034/2009 substituiu a simples reserva pela obrigação de preenchimento. Essas mudanças definem o contexto institucional da série, mas não constituem, por si, uma estratégia de identificação causal.

A literatura documenta que regras formais são mediadas por partidos, financiamento, incumbência, redes políticas e seleção de candidaturas (Miguel, 2008; Rezende, 2017; Sacchet & Speck, 2012; Wylie, 2018; Wylie & dos Santos, 2016). A passagem de 29 deputadas em 1998 para 91 em 2022 registra a expansão da presença durante o período das cotas. Porém, não demonstra quanto desse crescimento foi produzido por cada reforma nem permite, isoladamente, avaliar a amplitude social das trajetórias incorporadas. Estudos interseccionais mostram, ademais, que medidas dirigidas a um eixo de desigualdade podem ter efeitos distintos entre os subgrupos atravessados por outros eixos (Hughes, 2011).

2.2 Partidos como *gatekeepers* e a hipótese da segunda barreira

O modelo de oferta e demanda de Norris e Lovenduski (1995) situa o recrutamento na interseção entre *pessoas dispostas e capazes de concorrer* e *seletores partidários* que avaliam sua viabilidade. As duas faces são *socialmente estruturadas*. Recursos para desenvolver ambição, disponibilidade de tempo,

redes e experiência não se distribuem de modo uniforme. Tampouco são neutros os critérios pelos quais dirigentes reconhecem uma candidatura como “segura”. Instituições informais e redes homossociais podem favorecer perfis já familiares às elites partidárias (Bjarnegård, 2013; Bjarnegård & Kenny, 2015).

O caso brasileiro reúne características que tornam esse problema particularmente relevante: competição centrada em candidatos, listas abertas, campanhas territorializadas, desigualdade de financiamento e organizações partidárias marcadas por assimetrias de gênero. Estudos mostram que recursos financeiros se associam fortemente ao sucesso eleitoral e que mulheres historicamente receberam menos recursos que homens (Sacchet, 2018; Sacchet & Speck, 2012). Trajetória política prévia, incumbência, capital familiar e inserção partidária também funcionam como credenciais importantes (Araújo & Borges, 2013; Miguel et al., 2015). Wylie (2018) descreve os perfis limitados disponíveis às mulheres que conseguem avançar em instituições simultaneamente frágeis e estruturadas por desigualdades de gênero.

A *profissionalização política* oferece uma explicação concorrente. A consolidação de carreiras de tempo integral pode valorizar mandatos prévios, posições partidárias e recursos acumulados, elevando barreiras para pessoas externas aos circuitos políticos (Best & Cotta, 2000). Se esse processo atravessa homens e mulheres e ocorre simultaneamente às reformas de gênero, uma série de eleitos não consegue separar seus efeitos. A segunda barreira deve, portanto, ser tratada como hipótese sobre o mecanismo que gera a composição observada, e não como efeito já atribuído às cotas ou à profissionalização.

A perspectiva de classe complementa esse diagnóstico. Ocupações anteriores à política condensam, de forma imperfeita, diferenças de escolaridade, renda, autonomia, prestígio e controle de tempo. Legisladores latino-americanos são socialmente muito mais seletos do que as populações que representam e a origem de classe pode afetar atitudes e prioridades políticas mesmo sob disciplina partidária (Carnes & Lupu, 2015; Schwindt-Bayer, 2011). No Brasil, a classificação de ocupações é um problema metodológico próprio: rótulos administrativos precisam ser transformados em posições comparáveis, preservando a distinção entre inserção ocupacional, status e carreira política (Codato et al., 2014).

O conceito de fechamento social ajuda a interpretar a seleção. Em Weber (1978), grupos podem restringir oportunidades e monopolizar posições. Parkin (1979) distingue estratégias de exclusão pelas quais coletividades privilegiadas controlam recursos escassos. Aplicada ao recrutamento feminino, a segunda barreira sugere que a correção parcial de uma exclusão de gênero pode deixar intacta, ou tornar mais visível, a exclusão por origem social. A expectativa empírica não é que todas as deputadas pertençam a uma única elite ocupacional, mas que a variedade de posições de origem se estreite quando partidos privilegiam candidatas já dotadas de capital político e social.

2.3 Ideologia partidária e hipóteses

Partidos de esquerda mantiveram historicamente vínculos mais fortes com sindicatos, movimentos sociais e organizações populares. Em princípio, essas redes poderiam ampliar o recrutamento de trabalhadoras, professoras, estudantes, agentes religiosos, militantes e lideranças associativas, produzindo bancadas femininas socialmente mais heterogêneas. Partidos de direita, por sua vez, poderiam recrutar proporcionalmente mais empresárias, profissionais liberais e mulheres

dotadas de recursos privados. Essas expectativas são plausíveis, mas não determinísticas, pois *profissionalização, incumbência, tamanho do partido e competição eleitoral* atravessam ambos os blocos.

As hipóteses são:

H1: concentração temporal. A diversidade social das bancadas femininas *diminui* ao longo das eleições de 1998 a 2022. O teste avalia a tendência; a explicação por profissionalização ou cotas permanece uma interpretação a ser testada por desenho causal.

H2a: assimetria de gênero. Dentro do mesmo bloco e eleição, *as bancadas femininas são menos socialmente diversas do que as masculinas*. A hipótese é examinada separadamente com Simpson, mais sensível à dominância de categorias frequentes, e Shannon, mais sensível à riqueza e às categorias raras, mas também ao tamanho desigual dos grupos.

H2b: intensificação da assimetria. A diferença entre diversidade feminina e masculina torna-se *progressivamente mais negativa ao longo do tempo*.

H3: diferenciação ideológica. *As bancadas femininas da esquerda são socialmente mais diversas do que as da direita*.

As quatro hipóteses possuem direção substantiva, mas a análise não foi pré-registrada. Por isso, os testes principais são bilaterais e exatos; *p-valores* unilaterais são informados apenas como complemento na direção originalmente prevista. A inferência de H2a é agregada por eleição, reduzindo os 14 pares bloco-eleição a sete diferenças médias independentes entre pleitos. Como verificação de multiplicidade, aplica-se Holm aos quatro testes principais (ver o Apêndice estatístico).

3. Dados e métodos

3.1 Universo, fontes e unidades de análise

O universo contém todos os eleitos para a Câmara dos Deputados nas sete eleições gerais de 1998, 2002, 2006, 2010, 2014, 2018 e 2022. Como a Câmara possui 513 cadeiras, são 3.591 registros de eleitos. Desses, 380 correspondem a mulheres: 29 em 1998, 42 em 2002, 45 em 2006, 45 em 2010, 51 em 2014, 77 em 2018 e 91 em 2022. Os dados eleitorais e as ocupações declaradas provêm dos registros públicos do Tribunal Superior Eleitoral (TSE).

A observação individual é usada para construir distribuições agregadas por sexo, bloco ideológico e eleição. A principal unidade comparativa é o bloco-eleição: dois blocos, esquerda e direita, observados em sete eleições, totalizando 14 unidades. A classificação partidária é derivada de pesquisas com especialistas que posicionam as legendas brasileiras no eixo esquerda-direita (Bolognesi et al., 2023, 2025). A dicotomização favorece a comparabilidade e preserva um tamanho mínimo das bancadas, mas elimina variação entre partidos do mesmo bloco e *constitui uma limitação substantiva*.

Para H1, calcula-se, em cada eleição, a média não ponderada dos índices das bancadas femininas de esquerda e direita. Assim, ambos os blocos recebem o mesmo peso, independentemente do número de deputadas. A análise principal utiliza os registros atribuídos a uma das 12 categorias sociais e exclui apenas aposentado, não informada e outros. As versões com as 15 categorias, com nove categorias estritamente ocupacionais e com macrocategorias agregadas são

reportadas como sensibilidades. Os resultados descrevem a trajetória média dos dois blocos, e não a distribuição agregada de todas as deputadas.

Para H2a, cada bancada feminina é pareada à bancada masculina do mesmo bloco e eleição, formando 14 contrastes descritivos. Como os dois blocos do mesmo pleito compartilham contexto e não constituem réplicas independentes, a inferência principal usa, em cada eleição, a média das duas diferenças $D_F - D_M$, resultando em sete pares temporais. H2b testa a tendência dessas sete diferenças médias.

Para H3, comparam-se os sete índices femininos da esquerda aos sete índices femininos da direita.

3.2 Classificação das posições sociais

O campo de ocupação declarado ao TSE é usado como indicador imperfeito de posição social. A variável-base da análise harmoniza as 116 denominações originais de ocupação original da TSE em 15 categorias. A análise principal preserva 12 delas, sem novas fusões: trabalhadores manuais e agrícolas; empresários e gestores; advogados; outros profissionais liberais; políticos de carreira; assessores políticos; funcionários públicos; forças policiais; professores de todos os níveis; sacerdotes e religiosos; estudantes; e trabalho doméstico não remunerado, registrado pelo TSE como dona de casa. A mesma regra é aplicada a mulheres e homens.

São excluídos somente “aposentado”, “não informada” e “outros” porque essas respostas não permitem identificar com segurança uma posição social de origem. A exclusão alcança 249 casos e preserva 3.342 observações: 356 mulheres e 2.986 homens. As taxas são semelhantes entre mulheres (24/380; 6,3%) e homens (225/3.211; 7,0%). O banco contém 24 reclassificações externas feitas apenas para mulheres e nenhuma para homens; por isso, *ocupacao_paper* é usada somente em auditoria, evitando mensuração assimétrica por gênero. Como sensibilidades, estimam-se modelos com todas as 15 categorias, com nove categorias estritamente ocupacionais e com agregações em nove e seis macrogrupos (ver o Apêndice B. Correspondência entre as classificações das posições ocupacionais).

Ocupação continua sendo uma *proxy* de posição social, não uma medida direta de classe.

3.3 Medidas de diversidade

Simplesmente contar quantas categorias aparecem numa bancada mede riqueza, mas ignora sua distribuição. Uma bancada concentrada em uma posição e com poucos casos dispersos em outras pode ter a mesma riqueza de uma bancada equilibrada. Para captar variedade e distribuição, empregam-se índices que respondem de maneira diferente às categorias frequentes e raras. A preservação de 12 categorias reduz o achatamento produzido por macrocategorias, mas aumenta a importância de verificar o efeito de posições pouco frequentes.

O índice principal é a versão de Simpson interpretada como probabilidade de encontro entre categorias, isto é, a probabilidade de duas pessoas sorteadas sem reposição pertencerem a posições sociais distintas (Hurlbert, 1971; Simpson, 1949):

$$D = N/(N - 1) \times [1 - \sum(p_i^2)],$$

em que N é o tamanho da bancada e p_i é a proporção da categoria i . O fator $N/(N - 1)$ ajusta a probabilidade para o sorteio de duas pessoas sem reposição dentro da própria bancada; como os dados abrangem o universo de eleitos, ele não corrige erro amostral. D varia de 0 a 1: valores maiores indicam maior diversidade social e valores menores, maior concentração. Por elevar proporções ao quadrado, Simpson responde sobretudo à dominância de categorias frequentes. Seu nível absoluto cresce mecanicamente quando a classificação admite maior número de categorias, razão pela qual valores obtidos em esquemas diferentes não são comparados como se estivessem na mesma escala substantiva.

As análises de sensibilidade usam a entropia de Shannon (Shannon, 1948a, 1948b), sua correção de Miller-Madow para grupos pequenos e rarefação:

$$H' = -\sum p_i \ln(p_i).$$

Shannon aumenta com a riqueza e a equitabilidade e atribui relativamente mais informação às categorias raras. Como seu estimador bruto tende a crescer com N quando há categorias raras, calcula-se também $H_{MM} = H' + (K - 1)/(2N)$, em que K é o número de categorias observadas.

Para H2a, as bancadas masculinas também são rarefeitas ao tamanho da bancada feminina correspondente por 20.000 amostras hipergeométricas sem reposição. A comparação entre Shannon bruto, Miller-Madow e rarefação permite distinguir diferenças de composição de diferenças induzidas pelo tamanho muito desigual das bancadas. Nenhum índice, isoladamente, identifica quais ocupações explicam um contraste; isso requer inspeção composicional.

3.4 Estratégia estatística

Com apenas sete eleições, testes paramétricos baseados em normalidade seriam frágeis. H1 e H2b usam Mann-Kendall, que compara a ordem de todos os pares temporais e resume a direção por τ (Kendall, 1975; Mann, 1945). Os p -valores principais são bilaterais e exatos por permutação das sete observações; a aproximação normal com correção de continuidade é apresentada como sensibilidade. A magnitude de H1 é estimada pela inclinação de Sen (Sen, 1968), reportada por eleição e por ano, com intervalo de 95%.

H2a utiliza o teste exato de postos sinalizados de Wilcoxon sobre as sete diferenças médias por eleição. Os 14 contrastes bloco-eleição são apresentados descritivamente, mas não tratados como 14 réplicas independentes. H3 também usa Wilcoxon exato pareado por pleito; Mann-Whitney permanece apenas como sensibilidade que ignora o pareamento temporal (Mann & Whitney, 1947; Wilcoxon, 1945). As alternativas principais são bilaterais porque não houve pré-registro.

A inferência não se reduz ao limiar de 0,05: direção, magnitude, intervalo da inclinação, coerência entre índices, estabilidade diante das classificações alternativas e dependência entre testes orientam a interpretação. A correção de Holm é aplicada, como sensibilidade, aos p -valores bilaterais exatos de H1, H2a, H2b e H3. Com $n = 7$, qualquer resultado é necessariamente discreto e sensível a poucas eleições.

Tabela 1. Hipóteses, unidades e testes da diversidade social

Hipótese	Relação estimada	Unidade / n	Teste e alternativa
H1	Tendência da diversidade social feminina	eleição / 7	Mann-Kendall bilateral exato; inclinação de Sen com IC95%
H2a	Diferença entre diversidade social feminina e masculina	eleição / 7 pares	Wilcoxon pareado bilateral; Shannon corrigido e rarefação
H2b	Tendência de D_F – D_M	eleição / 7	Mann-Kendall bilateral exato
H3	Diversidade social feminina da esquerda versus direita	eleição / 7 pares	Wilcoxon pareado bilateral exato; Mann-Whitney como sensibilidade

Fonte: elaboração do autor.

4. Resultados

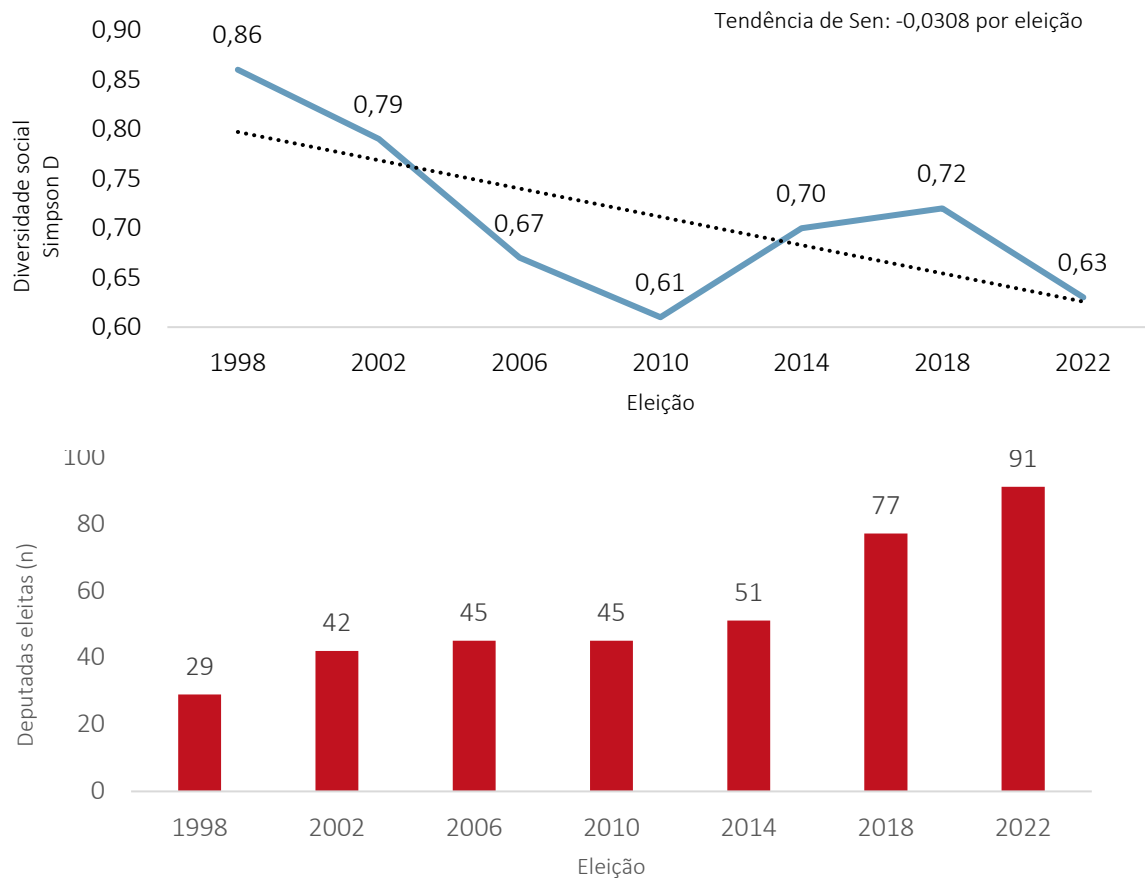
4.1 Concentração temporal das bancadas femininas

A Figura 1 mostra a expansão de 29 para 91 deputadas e a queda descritiva do índice médio de Simpson de 0,862 para 0,633 (–26,5%). A série com 12 categorias é 0,862; 0,788; 0,668; 0,614; 0,700; 0,718; e 0,633. Há queda acentuada até 2010, recuperação em 2014-2018 e novo recuo em 2022; portanto, o padrão não é uma descida contínua.

O Mann-Kendall não corrobora uma tendência monotônica no teste principal: $\tau = -0,429$ e p exato bilateral = 0,239 (aproximação normal com correção de continuidade: $p = 0,230$). A inclinação de Sen é $-0,0308$ por eleição, ou $-0,0077$ por ano, mas seu IC95% $[-0,0845; 0,0166]$ inclui zero. *H1, portanto, não é corroborada, embora a magnitude descritiva da queda mereça investigação.*

Shannon bruto conduz a uma tendência ainda mais fraca ($\tau = -0,238$; p exato = 0,562); Miller-Madow produz $\tau = -0,429$ e $p = 0,239$. A queda descritiva aparece nas classificações alternativas, mas sua significância varia: com todas as 15 categorias, $p = 0,136$; com nove categorias estritamente ocupacionais, $p = 0,239$; com seis macrocategorias, $p = 0,136$; e com nove macrogrupos que preservam estudantes, sacerdotes e trabalho doméstico, $p = 0,069$. Nenhuma dessas especificações sustenta um achado robusto após considerar multiplicidade. Na análise principal, o p ajustado de Holm para H1 é 0,717.

Figura 1. Diversidade social e número de deputadas federais, 1998-2022.



Fonte: elaboração do autor com dados do Tribunal Superior Eleitoral. Nota: Categorias sociais excluídas: aposentado, não informada e outros excluídos.

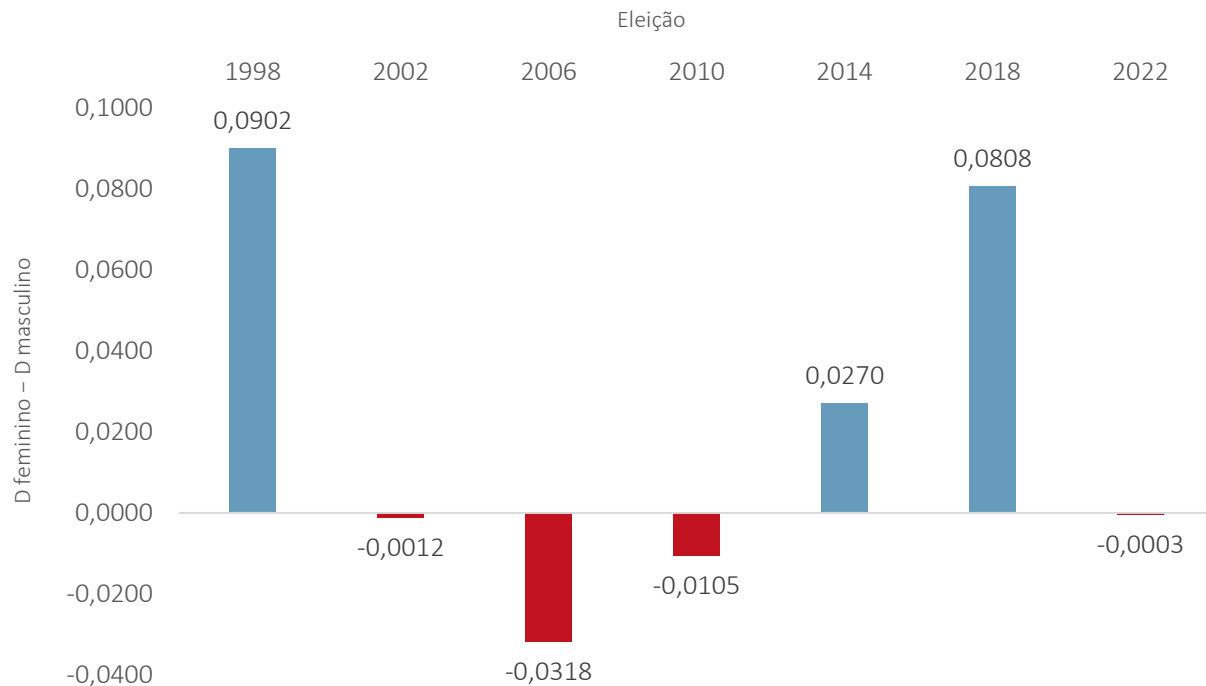
4.2 Assimetria entre bancadas femininas e masculinas

H2a não é corroborada por Simpson. As mulheres apresentam D maior em dez dos 14 contrastes bloco-eleição, mas apenas três das sete diferenças médias por pleito são positivas. A diferença média é +0,0220 e a mediana, -0,0003. No teste principal agregado por eleição, $W^+ = 17$, p bilateral exato = 0,688 e p unilateral na direção de mulheres menos diversas = 0,711. Não há evidência de assimetria na direção prevista.

Shannon bruto apresenta diferença feminina negativa em seis das sete médias eleitorais ($W^+ = 3$; p bilateral = 0,078; p unilateral = 0,039). Como a análise não foi pré-registrada e Shannon bruto é sensível ao tamanho e às categorias raras, o resultado exige cautela. Miller-Madow não confirma a diferença ($W^+ = 8$; p bilateral = 0,375), e a rarefação das bancadas masculinas ao tamanho das femininas elimina o contraste ($W^+ = 14$; $p = 1,000$). O conjunto dos estimadores, portanto, não sustenta menor diversidade feminina de forma robusta.

A Figura 2 mostra a diferença de Simpson entre mulheres e homens, promediada entre os blocos. Os valores são +0,0902; -0,0012; -0,0318; -0,0105; +0,0270; +0,0808; e -0,0003. A alternância de sinais e os valores próximos de zero em três eleições não indicam deterioração contínua.

Figura 2. Assimetria de gênero: diferença média de diversidade social entre mulheres e homens, por eleição, 1998-2022.



Fonte: elaboração do autor com dados do Tribunal Superior Eleitoral. Nota: 12 categorias sociais; aposentado, não informada e outros excluídos.

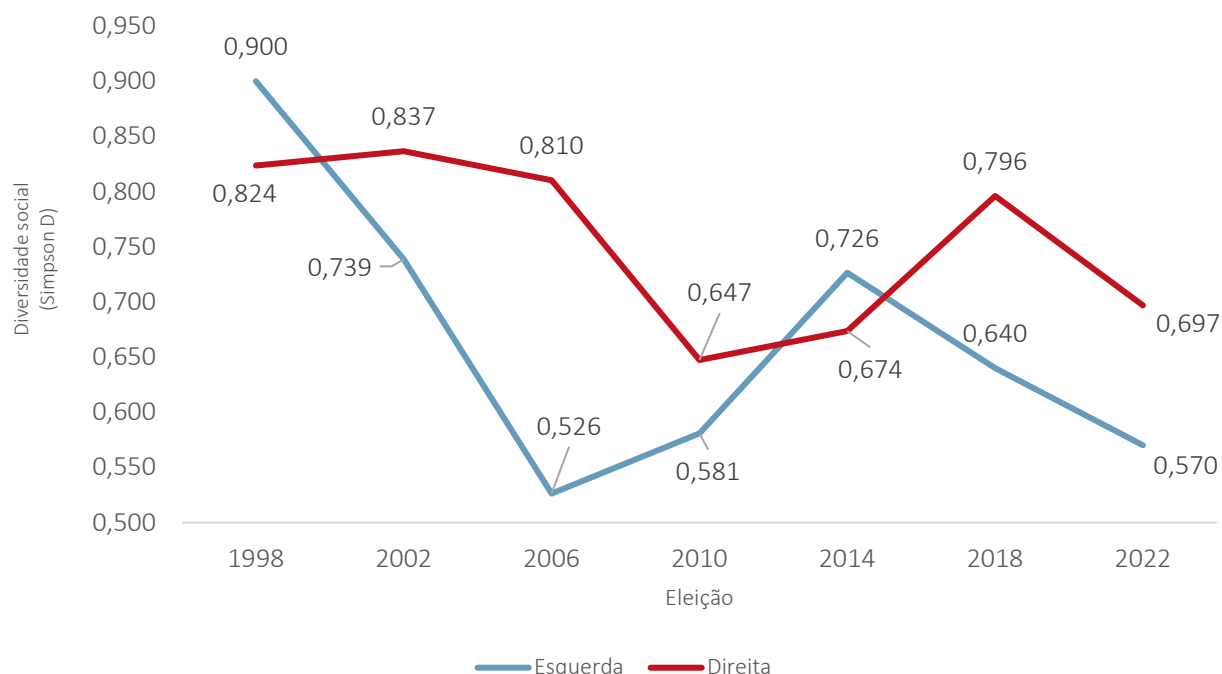
H2b não é corroborada: o Mann-Kendall da diferença $D_F - D_M$ resulta em $\tau = +0,048$ e p exato bilateral = 1,000. A especificação agregada de todas as pessoas por sexo em cada eleição produz o mesmo $\tau = +0,048$ e $p = 1,000$. Não há tendência monotônica de ampliação da assimetria.

Assim, H2 não recebe apoio consistente. Simpson não mostra diversidade feminina menor; o sinal do Shannon bruto desaparece com Miller-Madow e rarefação, e a diferença entre sexos não se torna progressivamente mais negativa. *O conceito de segunda barreira deve ser investigado com frequências específicas de posições sociais e com o funil completo de recrutamento, não inferido apenas de índices agregados.*

4.3 Ideologia e diversidade social

A Figura 3 compara as bancadas femininas de esquerda e direita. A direita apresenta D maior em cinco eleições e a esquerda em duas. As médias são 0,669 para a esquerda e 0,755 para a direita, direção descritiva oposta a H3.

Figura 3. Diversidade social das bancadas femininas por bloco, 1998-2022 (12 categorias).



Fonte: elaboração do autor com dados do Tribunal Superior Eleitoral. Nota: 12 categorias sociais; aposentado, não informada e outros excluídos.

O Wilcoxon exato pareado por eleição não confirma maior diversidade à esquerda: $W^+ = 4$, p unilateral na direção de H3 = 0,961 e p bilateral = 0,109. O Mann-Whitney bilateral também não é significativo ($U = 13$; $p = 0,165$). Algumas agregações mais grossas fortalecem a diferença em favor da direita, direção oposta a H3, mas nenhuma permanece significativa após Holm; na especificação principal, p ajustado = 0,438.

Entre as 356 deputadas incluídas, políticos de carreira representam 57,7% das eleitas da esquerda (82/142) e 40,7% das da direita (87/214). O contraste ajuda a explicar a diversidade média menor à esquerda, mas é descritivo e não identifica seleção partidária, oferta de candidaturas ou causalidade.

Tabela 2. Síntese dos resultados

Hipótese	Estatística principal	Corroborada	Interpretação restrita
H1: concentração temporal	$\tau = -0,429$; p exato = 0,239; Sen = $-0,0308$ /eleição; IC95% $[-0,0845; 0,0166]$	Não	D cai de 0,862 para 0,633; p de Holm = 0,717
H2a: mulheres menos diversas, Simpson	$W^* = 17$; p bilateral = 0,688; p unilateral na direção de H2a = 0,711; Holm = 1,000	Não	Diferença média +0,0220; mediana $-0,0003$
H2a: entropia corrigida/rarefação	Shannon bruto: $p = 0,078$; Miller-Madow: $W^* = 8$, $p = 0,375$; rarefação: $W^* = 14$, $p = 1,000$	Não	O sinal do Shannon bruto não resiste às correções de tamanho
H2b: intensificação temporal	$\tau = +0,048$; p exato = 1,000	Não	Sem tendência monotônica da diferença entre sexos
H3: esquerda mais diversa	$W^* = 4$; p unilateral = 0,961; p bilateral = 0,109; Holm = 0,438	Não	Média maior à direita: 0,755 versus 0,669

Fonte: elaboração do autor. Nota: p-valores principais bilaterais e exatos; H2a agregado por eleição; análise principal com 12 categorias sociais e exclusão de aposentado, não informada e outros. A correção de Holm é apresentada como sensibilidade.

5. Discussão

O resultado mais robusto é o crescimento da bancada feminina. A diversidade social medida por Simpson apresenta queda descritiva de 26,5%, mas sua tendência não é estatisticamente corroborada na classificação principal com 12 categorias. A direção negativa reaparece nas classificações alternativas, enquanto a força inferencial varia com o grau de agregação. O padrão merece investigação, mas não deve ser apresentado como tendência estatística estabelecida.

Isso não equivale a afirmar que as cotas fracassaram.

As cotas de candidaturas enfrentam a sub-representação de gênero, não equalizam automaticamente classe ou posição social, e o desenho não estima seu efeito causal. A formulação defensável é mais restrita: *durante o período das cotas, a presença feminina aumentou, enquanto a diversidade social caiu em termos descritivos, com incerteza estatística e alguma sensibilidade ao nível de agregação.*

A auditoria esclarece a comparação por gênero. O resultado de menor Shannon feminino na versão anterior combinava entropia bruta sensível a N, 14 pares tratados como mais independentes do que são e 24 reclassificações externas feitas apenas para mulheres. Com ocupação categorizada em 12 classes, aplicada simetricamente e inferência agregada por eleição, o Shannon bruto ainda sugere menor diversidade feminina, mas Miller-Madow e rarefação não confirmam o contraste.

Decisões de codificação e tamanho dos grupos são, portanto, parte da incerteza substantiva.

O conceito de *segunda barreira* organiza uma agenda de pesquisa. Seu teste exigiria decompor o recrutamento em etapas.

- Primeiro, comparar a distribuição ocupacional da população elegível à de filiados e aspirantes.
- Segundo, estimar quem obtém candidatura formal.
- Terceiro, medir a distribuição de recursos partidários, tempo de propaganda e posições competitivas.
- Quarto, separar seleção partidária de resposta eleitoral.

Se a diversidade já cair entre *aspirantes* e *indicadas*, o mecanismo se localiza dentro dos partidos, predominantemente; se cair entre candidatas e eleitas, recursos de campanha, incumbência e comportamento eleitoral são as variáveis que então ganham importância.

O resultado sobre ideologia também recomenda cautela.

A direita tem média descritiva maior, e esse sinal aparece em diferentes classificações, mas não resiste ao ajuste por multiplicidade. A concentração de políticos de carreira entre deputadas da esquerda é uma pista composicional, não uma demonstração de homogeneização programática. *Análises por partido, família ideológica e grau de “programaticidade” são necessárias.*

Cinco *limitações* merecem destaque.

- a) A primeira é o *n temporal pequeno*. Sete eleições permitem testes de postos, mas reduzem o poder estatístico e tornam os resultados sensíveis a observações individuais.
- b) A segunda é a *ocupação autodeclarada*. Ela pode funcionar como identidade eleitoral, conter categorias vagas e ocultar mudanças de carreira. Por isso, ‘aposentado’, ‘não informada’ e ‘outros’ são excluídos; estudantes, sacerdotes e trabalho doméstico são mantidos porque identificam posições sociais, embora não sejam ocupações convencionais.
- c) A terceira é a *classificação em 12 grupos sociais*. Preservar mais detalhes reduz o achatamento das seis macrocategorias, mas introduz grupos muito raros, como assessores políticos e trabalho doméstico não remunerado. A exclusão de 249 casos e as 24 reclassificações externas exclusivamente femininas mostram que toda decisão classificatória precisa ser documentada e testada.
- d) A quarta é a *inferência sobre classe*. Ocupação ou posição declarada *não mede diretamente classe, renda, escolaridade ou prestígio*, e ainda pode ocultar heterogeneidade dentro das 12 categorias. Médicos, engenheiros e outras profissões permanecem reunidos em outros profissionais liberais, por exemplo. *A maior desagregação melhora a fidelidade, mas não elimina os limites do indicador.*

A próxima etapa deve combinar ocupação com escolaridade, patrimônio, raça, idade, região, capital familiar e experiência política.

Há ainda questões estatísticas.

As inferências de H2a foram agregadas por eleição para não tratar os dois blocos do mesmo pleito como réplicas independentes, mas restam apenas sete pontos temporais. Os testes não modelam mudanças estruturais, autocorrelação complexa ou incerteza de classificação. A presença

de categorias raras torna especialmente importantes Miller-Madow e rarefação. **Extensões importantes seriam *bootstrap*, modelos hierárquicos e validação cega das categorias.**

6. Conclusão

Entre 1998 e 2022, o Brasil elegeu mais deputadas federais. Na análise principal com 12 categorias sociais, Simpson cai de 0,862 para 0,633, mas Mann-Kendall não confirma tendência monotônica ($\tau = -0,429$; p exato = 0,239) e o IC95% da inclinação de Sen $[-0,0845; 0,0166]$ inclui zero. *A queda relativa de 26,5% é substantivamente relevante, porém não constitui evidência estatística conclusiva.*

A comparação por gênero não apoia diversidade feminina menor de forma consistente após codificação simétrica, agregação por eleição, correção de entropia e rarefação. A diferença entre sexos não se intensifica monotonicamente, e a expectativa de maior diversidade das deputadas da esquerda não é confirmada. Nenhum dos quatro testes principais permanece significativo após Holm.

A agenda deve continuar avançando de “quantas mulheres” para “quais mulheres”, mas os dados atuais não confirmam uma segunda barreira social robusta. O conceito permanece uma hipótese teórica a ser testada com mensuração simétrica, categorias socialmente informativas, dados sobre etapas anteriores à eleição e maior número de unidades.

O próximo passo é observar o recrutamento *antes* da eleição, ampliar o universo para *outros níveis de mandato* e decompor *ocupação, classe, raça e trajetória política*. Só então será possível distinguir se a concentração nasce da oferta desigual de aspirantes, das escolhas partidárias, da distribuição de recursos ou da competição eleitoral.

Autoria/Authorship

Adriano Codato (adrianocodato@gmail.com) (<https://orcid.org/0000-0002-5015-4273>) é professor de Ciência Política na Universidade Federal do Paraná e Editor-chefe da *Revista de Sociologia e Política*. É pesquisador do CNPq e coordena o INCT Representação e Legitimidade Democrática (ReDem).

Declaração de contribuição dos autores/Author Contributions (CRediT)

Adriano Codato: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing - original draft, Writing - review & editing.

Declaração de conflito de interesse/Conflict-of-interest statement

O autor declara que não possui quaisquer conflitos de interesse, financeiros ou pessoais que possam ter influenciado, direta ou indiretamente, a elaboração, análise e publicação deste trabalho.

Declaração de disponibilidade de dados da pesquisa/Data and Replication Materials Availability

Todo o conjunto de dados de apoio aos resultados deste estudo foi disponibilizado em **Harvard Dataverse** e pode ser acessado em Codato, Adriano. 2026. “Replication Data for: The Second Barrier: Gender, Class, and Social Diversity in the Recruitment of Women to Brazil's Chamber of Deputies (1998–2022).” Harvard Dataverse. <https://doi.org/10.7910/DVN/AIWPWC>.

Os registros eleitorais brutos são públicos e estão disponíveis no Portal de Dados Abertos do Tribunal Superior Eleitoral (<https://dadosabertos.tse.jus.br/>).

Declaração de uso de IA/Generative Artificial Intelligence Use Statement

O autor declara ter utilizado o ChatGPT 5.6 (OpenAI) como apoio à realização e à conferência dos cálculos estatísticos, à revisão linguística e editorial do manuscrito e à revisão, formatação e padronização das referências bibliográficas. Todas as saídas foram verificadas pelo autor, que assume responsabilidade integral pela análise, pela interpretação dos resultados e pelo conteúdo final.

Funding/Financiamento

Este trabalho foi financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), no âmbito do Instituto Nacional de Ciência e Tecnologia em Representação e Legitimidade Democrática (INCT ReDem), processo nº 406649/2022-7, e pela Fundação Araucária de Apoio ao Desenvolvimento Científico e Tecnológico do Estado do Paraná (Projeto nº 115/2024 PDI).

Apêndice A. Detalhamento dos procedimentos estatísticos

Este apêndice apresenta a lógica dos índices e testes utilizados no artigo, as razões para sua escolha e os cuidados adotados na interpretação. A estratégia estatística foi definida principalmente por três características dos dados:

- a) pequeno número de eleições,
- b) a dependência entre observações realizadas no mesmo pleito e
- c) a grande diferença de tamanho entre as bancadas femininas e masculinas.

Embora o banco contenha todos os 3.591 deputados federais eleitos entre 1998 e 2022, a unidade inferencial não é o indivíduo. As hipóteses dizem respeito a tendências e diferenças entre bancadas ao longo de sete eleições. Portanto, utilizar milhares de registros como se fossem observações independentes produziria uma falsa impressão de precisão (pseudorreplificação).

Para H1 e H2b, a unidade temporal é a eleição, com $n = 7$.

Para H2a e H3, as comparações também são resumidas ou pareadas por eleição.

Os 14 grupos formados pelo cruzamento entre bloco ideológico e pleito são utilizados para produzir os índices e apresentar contrastes descritivos, mas os dois blocos observados na mesma eleição não são tratados como réplicas completamente independentes.

A.1 Índice de diversidade de Simpson

$$D = \frac{N}{N-1} \left(1 - \sum_{i=1}^K p_i^2 \right) \quad (A1)$$

Em que: D = índice de diversidade de Simpson corrigido para o tamanho do grupo; N = número total de integrantes da bancada analisada; K = número de categorias ocupacionais consideradas; p_i = proporção da bancada pertencente à categoria ocupacional i ; i = identificador da categoria, variando de 1 a K ; Σ = soma dos valores de todas as categorias; p_i^2 = quadrado da proporção da categoria i , que dá maior peso às categorias mais frequentes.

O índice pode ser interpretado como a probabilidade de que duas pessoas sorteadas sem reposição (depois de sortear a primeira pessoa, ela não volta para o grupo antes do segundo sorteio) na mesma bancada pertençam a categorias diferentes. Quanto maior o D , maior a *diversidade*. Um valor próximo de zero indica forte concentração em uma ou poucas categorias.

A multiplicação por $N/(N - 1)$ ajusta o índice para um sorteio sem reposição. Esse fator não representa uma correção do erro amostral, pois o banco abrange o universo dos eleitos. Ele apenas torna a medida compatível com a interpretação probabilística de selecionar duas pessoas distintas dentro da bancada.

O índice de Simpson foi escolhido como medida principal porque responde principalmente à *dominância das categorias mais frequentes*. Categorias com poucos casos afetam relativamente pouco o resultado, enquanto o crescimento de uma categoria majoritária reduz sensivelmente a diversidade.

Essa propriedade é adequada para investigar se as bancadas estão se tornando concentradas em posições sociais dominantes, como **empresários** ou profissionais liberais.

Uma limitação importante é que o nível absoluto de Simpson depende do número de categorias permitido pela classificação. Uma divisão em 12 categorias tende a produzir valores maiores do que uma divisão em seis **grandes** categorias **abrangentes**. Por isso, valores obtidos com classificações diferentes não devem ser comparados diretamente como se estivessem na mesma escala substantiva.

As classificações alternativas são usadas para verificar a estabilidade das conclusões, não para comparar os níveis absolutos dos índices.

A.2 Entropia de Shannon

$$H' = - \sum_{i=1}^K p_i \ln(p_i) \quad (A2)$$

Em que: H' = entropia de Shannon observada no grupo; p_i = proporção de integrantes pertencentes à categoria ocupacional i ; $\ln$ = logaritmo natural; i = identificador da categoria ocupacional; K = número de categorias ocupacionais observadas; Σ = soma dos termos calculados para todas as categorias.

Shannon combina duas dimensões: a quantidade de categorias observadas, chamada “riqueza”, e o grau de equilíbrio entre elas, chamado “equitabilidade”. Diferentemente de Simpson, Shannon atribui relativamente mais importância às categorias raras.

Essa característica é útil porque uma bancada pode apresentar pouca concentração nas categorias dominantes, mas ainda ter uma variedade limitada de posições sociais minoritárias. Simpson e Shannon observam, portanto, aspectos relacionados, mas não idênticos, da diversidade.

O problema é que Shannon também é mais sensível ao tamanho do grupo.

Uma bancada grande tem maior probabilidade de conter categorias raras simplesmente porque possui mais integrantes. Como as bancadas masculinas são muito maiores que as femininas, uma comparação direta poderia confundir diversidade social com tamanho do grupo.

Por essa razão, Shannon bruto é tratado como um *indicador de sensibilidade*, e não como evidência suficiente por si só.

A.3 Correção de Miller-Madow

$$H_{MM} = H' + \frac{K-1}{2N} \quad (A3)$$

Em que: H_{MM} = entropia de Shannon após a correção de Miller-Madow; H' = entropia de Shannon bruta; K = número de categorias com pelo menos um caso no grupo; N = número total de integrantes do grupo; $(K-1)/(2N)$ = termo acrescentado para reduzir o viés associado a grupos pequenos.

Em amostras ou grupos pequenos, algumas categorias podem não aparecer, mesmo que façam parte da população de interesse. Isso faz com que Shannon bruto tenda a subestimar a diversidade. A correção de Miller-Madow acrescenta um termo proporcional ao número de categorias e inversamente proporcional ao tamanho do grupo.

A correção era necessária porque algumas bancadas femininas, especialmente nas primeiras eleições e dentro de cada bloco ideológico, possuem poucos integrantes. Em 1998, por exemplo, havia apenas 29 deputadas no total, divididas entre esquerda e direita. Nessa situação, a ausência de uma categoria pode decorrer do pequeno tamanho da bancada, e não necessariamente de um processo substantivo de fechamento social.

Miller-Madow não elimina toda a incerteza produzida por grupos pequenos, mas oferece uma verificação mais adequada do que a comparação baseada apenas em Shannon bruto.

A.4 Rarefação

A rarefação oferece outra forma de controlar a diferença entre os tamanhos das bancadas femininas e masculinas. Em cada combinação de eleição e bloco ideológico, a bancada masculina é reduzida artificialmente ao tamanho da bancada feminina correspondente.

O procedimento é realizado por amostragem hipergeométrica multivariada sem reposição:

- a) registra-se o número de mulheres na bancada;
- b) sorteia-se, sem reposição, o mesmo número de homens na bancada masculina correspondente;
- c) calcula-se Shannon para o grupo masculino sorteado;
- d) repete-se o procedimento 20.000 vezes;
- e) utiliza-se a média dos 20.000 valores como diversidade masculina rarefeita.

A distribuição hipergeométrica é apropriada porque o sorteio ocorre sem reposição dentro de uma população finita, a bancada masculina daquele bloco e eleição. Cada sorteio preserva a composição ocupacional observada, mas iguala o número de pessoas disponíveis para a comparação.

A rarefação responde à seguinte pergunta: *qual seria a diversidade esperada dos homens se sua bancada tivesse exatamente o mesmo tamanho que a bancada das mulheres?*

Esse procedimento era necessário porque uma bancada de 200 homens tem muito mais oportunidade de conter ocupações raras do que uma bancada de 20 mulheres. Se a diferença desaparecer após a rarefação, isso indica que o resultado original provavelmente estava associado ao tamanho desigual dos grupos, e não necessariamente a uma seleção social mais restritiva das mulheres.

A.5 Teste de Mann-Kendall

$$S = \sum_{i < j} \text{sign}(y_j - y_i) \quad (\text{A4})$$

Em que: S = estatística de tendência do teste de Mann-Kendall; y_i = valor do índice na eleição situada na posição temporal i ; y_j = valor do índice na eleição posterior, situada na posição temporal j ; $i < j$ = comparação entre uma eleição anterior e uma eleição posterior; $\text{sign}(\cdot)$ = função que retorna +1, 0 ou -1 conforme a diferença seja positiva, nula ou negativa; Σ = soma dos resultados de todas as comparações entre pares de eleições.

$$\tau = \frac{S}{n(n-1)/2} \quad (\text{A5})$$

Em que: τ = coeficiente de associação ordinal de Kendall; S = estatística definida na Equação A4; n = número de observações temporais; neste artigo, sete eleições; $n(n-1)/2$ = número total de pares de eleições que podem ser comparados. *A expressão corresponde ao caso sem empates, que é o observado nas séries principais do artigo.*

Cada par de eleições é comparado. Se o valor posterior for maior, soma-se +1; se for menor, soma-se -1; se houver igualdade, soma-se zero. O resultado é transformado em uma medida padronizada de associação ordinal.

Valores negativos de τ indicam tendência decrescente; valores positivos indicam tendência crescente; valores próximos de zero indicam ausência de ordenamento temporal claro.

O teste foi usado em H1 para verificar se a diversidade social das bancadas femininas diminui ao longo das eleições.

Em H2b, ele avalia se a diferença $D_F - D_M$, entre diversidade feminina e masculina, se torna progressivamente mais negativa.

Mann-Kendall foi escolhido porque não exige normalidade dos dados, não pressupõe uma relação linear e é adequado para séries curtas. Ele não pergunta se o primeiro e o último valores são diferentes. Pergunta se os valores apresentam uma ordenação temporal relativamente consistente.

Essa distinção é importante.

Uma série pode terminar muito abaixo do ponto inicial, mas apresentar quedas, recuperações e novos recuos. Nesse caso, a variação entre os extremos pode ser substantivamente grande sem constituir uma tendência monotônica estatisticamente bem definida.

Como existem apenas sete eleições, o *p-valor* foi calculado de maneira exata, considerando todas as $7! = 5.040$ permutações possíveis da série observada. O procedimento compara o valor de S encontrado com a distribuição que seria obtida se a ordem temporal dos sete valores fosse permutada sob a hipótese nula de ausência de tendência.

A.6 Inclinação de Sen

$$\hat{\beta}_{\text{Sen}} = \text{mediana}_{i < j} \left(\frac{y_j - y_i}{j - i} \right) \quad (\text{A6})$$

Em que: $\hat{\beta}_{\text{Sen}}$ = inclinação estimada pelo método de Sen; o acento circunflexo indica que se trata de uma estimativa; y_i = valor do índice na eleição situada na posição i ; y_j = valor do índice na eleição situada na posição posterior j ; $j - i$ = distância, em intervalos eleitorais, entre as duas observações; **mediana** = valor central de todas as inclinações calculadas para os pares em que $i < j$.

A inclinação é obtida pela mediana de todas as inclinações calculadas entre pares de observações.

A mediana torna o estimador menos sensível a valores extremos do que uma regressão linear convencional. Como as eleições estão separadas por intervalos regulares de quatro anos, a inclinação é apresentada por eleição e, complementarmente, dividida por quatro para obter uma aproximação anual.

O intervalo de confiança de 95% mostra a incerteza em torno da magnitude estimada. Se o intervalo inclui zero, os dados permanecem compatíveis tanto com uma redução como com estabilidade ou pequena elevação.

A inclinação de Sen era necessária porque um *p-valor*, isoladamente, não informa se a mudança é substantivamente pequena ou grande. No artigo, a combinação de Mann-Kendall e Sen permite distinguir três aspectos:

- a) a direção da tendência;
- b) sua regularidade temporal;
- c) o tamanho estimado da mudança.

A.7 Teste de postos sinalizados de Wilcoxon

$$d_t = x_t - y_t \quad (\text{A7})$$

Em que: d_t = diferença entre os dois grupos comparados na eleição t ; x_t = valor do índice no primeiro grupo da comparação, na eleição t ; y_t = valor do índice no segundo grupo da comparação, na mesma eleição; t = eleição usada como unidade de pareamento.

$$W^+ = \sum_{d_t > 0} R(|d_t|) \quad (A8)$$

Em que: W^+ = soma dos postos associados às diferenças positivas; d_t = diferença pareada na eleição t ; $|d_t|$ = valor absoluto da diferença, isto é, sua magnitude sem considerar o sinal; $R(|d_t|)$ = posto atribuído à magnitude da diferença entre todas as diferenças não nulas; $\Sigma \text{ com } d_t > 0$ = soma restrita às eleições em que a diferença é positiva.

As diferenças iguais a zero são retiradas. As diferenças restantes são ordenadas segundo seus valores absolutos, recebem postos e recuperam seus sinais. A estatística W^+ corresponde à soma dos postos associados às diferenças positivas.

Sob a hipótese nula, sinais positivos e negativos devem distribuir-se de maneira compatível com a ausência de diferença sistemática. Com m diferenças não nulas, a distribuição exata é obtida examinando as 2^m combinações possíveis de sinais.

Em H2a, inicialmente são calculadas as diferenças entre mulheres e homens em cada combinação de bloco e eleição:

$$d_{bt} = D_{F, bt} - D_{M, bt} \quad (A9)$$

Em que: d_{bt} = diferença entre a diversidade feminina e a masculina no bloco b e na eleição t ; $D_{F, bt}$ = índice de diversidade da bancada feminina no bloco b e na eleição t ; $D_{M, bt}$ = índice de diversidade da bancada masculina no mesmo bloco e eleição; F = bancada feminina; M = bancada masculina; b = bloco ideológico: esquerda ou direita; t = eleição.

Esses 14 contrastes são apresentados descritivamente. Para a inferência principal, entretanto, calcula-se em cada eleição a média das diferenças da esquerda e da direita:

$$\bar{d}_t = \frac{d_{\text{esquerda}, t} + d_{\text{direita}, t}}{2} \quad (A10)$$

Em que: $\bar{d}_t$ = média das diferenças entre mulheres e homens na eleição t ; $d_{\text{esquerda}, t}$ = diferença entre mulheres e homens no bloco de esquerda, na eleição t ; $d_{\text{direita}, t}$ = diferença correspondente no bloco de direita; 2 = número de blocos ideológicos incluídos na média.

O teste de Wilcoxon é então aplicado às sete diferenças médias eleitorais. Essa agregação evita tratar esquerda e direita da mesma eleição como duas réplicas totalmente independentes, pois ambas estão sujeitas ao mesmo contexto eleitoral, às mesmas regras e a choques políticos nacionais.

Em H3, o pareamento ocorre diretamente por eleição:

$$d_t = D_{\text{esquerda}, t} - D_{\text{direita}, t} \quad (A11)$$

Em que: d_t = diferença entre a diversidade das deputadas de esquerda e de direita na eleição t ;
 $D_{\text{esquerda},t}$ = índice de diversidade da bancada feminina de esquerda na eleição t ; **$D_{\text{direita},t}$** = índice de diversidade da bancada feminina de direita na mesma eleição; t = eleição usada para formar o par.

Assim, a diversidade das deputadas da esquerda em 1998 é comparada à diversidade das deputadas da direita em 1998, e assim sucessivamente.

Wilcoxon foi escolhido porque não exige normalidade das diferenças e é apropriado para um número pequeno de pares. Sua interpretação como teste de deslocamento de localização é mais direta quando a distribuição das diferenças é aproximadamente simétrica. Com apenas sete pares, essa condição não pode ser avaliada com precisão, razão pela qual direção, magnitude e resultados de outras especificações também são considerados.

A.8 Testes bilaterais e unilaterais

As hipóteses possuem direções substantivas definidas: queda da diversidade, menor diversidade feminina e maior diversidade à esquerda. Entretanto, a análise não foi pré-registrada antes da inspeção dos dados.

Por essa razão, os testes principais são bilaterais. O teste bilateral considera como relevantes tanto diferenças na direção esperada quanto diferenças na direção contrária. Essa escolha é mais conservadora e reduz o risco de selecionar retrospectivamente a direção que favorece a hipótese.

Os *p-valores* unilaterais são apresentados apenas como complemento, sempre na direção prevista pela hipótese. Quando o resultado aparece na direção contrária, isso é explicitado no texto, sem reinterpretar posteriormente a hipótese.

A.9 Mann-Whitney como análise de sensibilidade

Na avaliação de H_3 , também é apresentado o teste de Mann-Whitney. Esse teste compara a ordenação dos sete valores da esquerda com a dos sete valores da direita, tratando os dois conjuntos como grupos independentes.

Ele é menos adequado como teste principal porque ignora o pareamento por eleição. A esquerda de 1998 está mais diretamente relacionada à direita de 1998 do que à direita de 2022, pois os dois primeiros valores compartilham o mesmo contexto eleitoral.

Mann-Whitney é utilizado apenas como sensibilidade. A concordância entre ele e o Wilcoxon aumenta a confiança na interpretação; uma discordância sugere que o resultado depende da forma como a estrutura temporal é tratada.

A.10 Correção de Holm para comparações múltiplas

O artigo possui quatro testes principais:

- a) **H1: tendência da diversidade feminina;**

- b) H2a: diferença entre mulheres e homens;**
- c) H2b: tendência da diferença entre os sexos;**
- d) H3: diferença entre esquerda e direita.**

Realizar vários testes aumenta a probabilidade de encontrar pelo menos um resultado aparentemente significativo apenas pela multiplicidade das comparações. Para controlar esse risco, aplica-se a *correção de Holm* aos quatro p-valores principais.

O procedimento ordena os p-valores do menor para o maior. O menor é comparado a um critério corrigido pelo número total de testes; os seguintes são avaliados com critérios progressivamente menos restritivos. A correção controla a probabilidade de ao menos uma rejeição falsa no conjunto de hipóteses, conhecida como taxa de erro familiar.

Holm foi escolhido porque controla comparações múltiplas sem ser tão conservador quanto a correção de Bonferroni aplicada de maneira uniforme. Os resultados sem correção continuam sendo apresentados, mas conclusões substantivas robustas deveriam sobreviver ao ajuste.

A.11 Análises de sensibilidade da classificação

A classificação das posições sociais constitui uma decisão de mensuração, não uma etapa neutra.

A análise principal utiliza 12 categorias e exclui somente “aposentado”, “não informada” e “outros”. Para verificar se as conclusões dependem dessa escolha, são estimadas especificações alternativas:

- a) todas as 15 categorias de *ocupacao_categ_1*;
- b) nove categorias estritamente ocupacionais;
- c) nove macrogrupos sociais;
- d) seis macrocategorias ocupacionais.

Essas análises não têm o objetivo de escolher a classificação que produz o menor *p-valor*. Elas verificam se direção, magnitude e inferência permanecem estáveis quando se altera o grau de desagregação.

A queda descritiva da diversidade feminina aparece nas diferentes classificações, mas a força estatística varia. Essa variação é substantivamente informativa: mostra que parte da inferência depende de como posições raras, respostas ambíguas e fusões entre ocupações são tratadas.

A.12 Interpretação dos p-valores e limites da inferência

Os *p-valores* exatos utilizados no artigo não representam a probabilidade de a hipótese nula ser verdadeira. Eles indicam quão incomum seria obter uma estatística tão extrema quanto a observada sob o procedimento de permutação ou atribuição de sinais definido pela hipótese nula.

Um *p-valor* elevado não demonstra ausência de mudança. Significa que *os dados disponíveis não permitem distinguir com segurança o padrão observado da variação compatível com a hipótese nula*.

Essa cautela é especialmente importante porque existem somente sete eleições. Com $n = 7$, os p-valores exatos assumem um conjunto limitado de valores possíveis, os intervalos são amplos e poucos pleitos podem alterar substancialmente o resultado. O pequeno número de pontos temporais reduz o poder estatístico para identificar tendências moderadas.

O fato de o banco conter o universo dos eleitos elimina o erro decorrente de selecionar apenas uma amostra de parlamentares. Não elimina, porém, a incerteza sobre a regularidade temporal do padrão, sua sensibilidade à classificação nem sua generalização para outros períodos e níveis de representação.

Finalmente, os testes identificam associações e ordenamentos, não mecanismos causais.

Eles não permitem atribuir a mudança observada às *cotas*, à *profissionalização política* ou às *decisões partidárias*. A identificação desses mecanismos exigiria dados sobre as etapas anteriores do recrutamento, como *filiação*, *disposição para concorrer*, *seleção de candidaturas*, *distribuição de recursos* e *chances eleitorais*.

Quadro A1. Correspondência entre hipóteses e procedimentos

Hipótese ou problema	Procedimento principal	Finalidade
Medir diversidade	Simpson corrigido	Captar concentração nas categorias mais frequentes
Verificar categorias raras	Shannon	Captar riqueza e equitabilidade
Corrigir grupos pequenos	Miller–Madow	Reduzir o viés de Shannon associado a N pequeno
Igualar bancadas masculinas e femininas	Rarefação	Separar diversidade de diferença de tamanho
H1: tendência temporal da diversidade feminina	Mann–Kendall e inclinação de Sen	Testar tendência e estimar sua magnitude
H2a: diferença entre mulheres-homens	Wilcoxon exato sobre sete diferenças eleitorais	Comparar mulheres e homens preservando o pareamento
H2b: tendência da diferença entre mulheres-homens	Mann–Kendall	Testar se a diferença entre os sexos se intensifica
H3: diferença esquerda-direita	Wilcoxon pareado	Comparar esquerda e direita dentro da mesma eleição
Sensibilidade de H3	Mann–Whitney	Avaliar o resultado ignorando o pareamento temporal
Quatro hipóteses principais	Correção de Holm	Controlar o risco de significância produzido por testes múltiplos

Fonte: elaboração do autor.

Apêndice B. Correspondência entre as classificações das posições ocupacionais

Este apêndice apresenta a correspondência entre as diferentes classificações das posições ocupacionais empregadas no artigo. A análise principal utiliza 12 categorias sociais. As classificações com 15, nove e seis categorias são utilizadas como análises de sensibilidade, permitindo avaliar se os resultados dependem das exclusões realizadas ou do grau de agregação das categorias. Há duas especificações com nove categorias: uma mantém categorias estritamente ocupacionais sem novas fusões; a outra reúne as 12 posições sociais em nove macrogrupos.

Quadro B1. Correspondência entre as classificações ocupacionais utilizadas na análise

Classificação original (15 categorias)	Principal (12 categorias sociais)	9 categorias estritamente ocupacionais	9 macrogrupos sociais	6 macrocategorias ocupacionais
Trabalhadores	Trabalhadores manuais e agrícolas	Trabalhadores manuais e agrícolas	Trabalhadores manuais e agrícolas	Trabalhadores manuais e agrícolas
Empresários	Empresários e gestores	Empresários e gestores	Empresários e gestores	Empresários e gestores
Advogado	Advogados	Advogados	Profissionais liberais	Profissionais liberais
Outros profissionais liberais	Outros profissionais liberais	Outros profissionais liberais	Profissionais liberais	Profissionais liberais
Político de carreira	Políticos de carreira	Políticos de carreira	Políticos de carreira	Políticos de carreira
Assessores políticos	Assessores políticos	Assessores políticos	Políticos de carreira	Políticos de carreira
Funcionários públicos	Funcionários públicos	Funcionários públicos	Serviço público e forças policiais	Serviço público e forças policiais
Forças policiais	Forças policiais	Forças policiais	Serviço público e forças policiais	Serviço público e forças policiais
Professores de todos os níveis	Professores de todos os níveis	Professores de todos os níveis	Professores e pesquisadores	Professores e pesquisadores
Sacerdotes	Sacerdotes e religiosos	Excluída	Sacerdotes e religiosos	Excluída
Estudante	Estudantes	Excluída	Estudantes	Excluída
Dona de casa	Trabalho doméstico não remunerado	Excluída	Trabalho doméstico não remunerado	Excluída
Aposentado	Excluída	Excluída	Excluída	Excluída
Não informada	Excluída	Excluída	Excluída	Excluída
Outros	Excluída	Excluída	Excluída	Excluída

Nota: a classificação com 15 categorias reproduz integralmente `ocupacao_categ_1`. A análise principal preserva 12 posições sociais e exclui somente aposentado, não informada e outros. A classificação estritamente ocupacional também exclui sacerdotes, estudantes e trabalho doméstico não remunerado. Os macrogrupos resultam da fusão de categorias ocupacionais próximas.

Quadro B2. Casos analisados e casos excluídos por categorias ocupacionais

Especificação	Casos analisados	Casos excluídos	Uso no paper
15 categorias originais	3.591	0	Sensibilidade com todas as respostas
12 categorias sociais	3.342	249	Análise principal
9 categorias estritamente ocupacionais	3.296	295	Sensibilidade sem posições não ocupacionais
9 macrogrupos sociais	3.342	249	Sensibilidade à agregação, preservando posições sociais
6 macrocategorias ocupacionais	3.296	295	Especificação mais agregada

Nota: A expressão “295 casos excluídos” aplica-se, portanto, às classificações estritamente ocupacionais de nove e seis categorias. Na análise principal com 12 categorias, foram excluídos 249 casos.

Referências

- Araújo, C. (2005). Partidos políticos e gênero: Mediações nas rotas de ingresso das mulheres na representação política. *Revista de Sociologia e Política*, (24), 193-215. <https://doi.org/10.1590/S0104-44782005000100013>
- Araújo, C., & Borges, D. (2013). Trajetórias políticas e chances eleitorais: Analisando o “gênero” das candidaturas em 2010. *Revista de Sociologia e Política*, 21(46), 69-91. <https://doi.org/10.1590/S0104-44782013000200005>
- Best, H., & Cotta, M. (Eds.). (2000). *Parliamentary representatives in Europe 1848-2000: Legislative recruitment and careers in eleven European countries*. Oxford University Press. <https://doi.org/10.1093/oso/9780198297932.001.0001>
- Bjarnegård, E. (2013). *Gender, informal institutions and political recruitment: Explaining male dominance in parliamentary representation*. Palgrave Macmillan. <https://doi.org/10.1057/9781137296740>
- Bjarnegård, E., & Kenny, M. (2015). Revealing the “secret garden”: The informal dimensions of political recruitment. *Politics & Gender*, 11(4), 748-753. <https://doi.org/10.1017/S1743923X15000471>
- Bolognesi, B., Ribeiro, E., & Codato, A. (2023). Uma nova classificação ideológica dos partidos políticos brasileiros. *Dados*, 66(2), e20210164. <https://doi.org/10.1590/dados.2023.66.2.303>
- Bolognesi, B., Ribeiro, E., Codato, A., & Silva, B. F. da. (2025). O desaparecimento do centro ideológico no sistema partidário brasileiro: A classificação mais atualizada dos experts. *Opinião Pública*, 31, e31120. <https://doi.org/10.1590/1807-0191202531120>
- Brasil. (1995). *Lei nº 9.100, de 29 de setembro de 1995: Estabelece normas para a realização das eleições municipais de 3 de outubro de 1996, e dá outras providências*. https://www.planalto.gov.br/ccivil_03/leis/19100.htm
- Brasil. (1997). *Lei nº 9.504, de 30 de setembro de 1997: Estabelece normas para as eleições*. https://www.planalto.gov.br/ccivil_03/leis/19504.htm
- Brasil. (2009). *Lei nº 12.034, de 29 de setembro de 2009: Altera as Leis nº 9.096, de 19 de setembro de 1995, nº 9.504, de 30 de setembro de 1997, e nº 4.737, de 15 de julho de 1965*. https://www.planalto.gov.br/ccivil_03/_ato2007-2010/2009/lei/l12034.htm
- Carnes, N., & Lupu, N. (2015). Rethinking the comparative perspective on class and representation: Evidence from Latin America. *American Journal of Political Science*, 59(1), 1-18. <https://doi.org/10.1111/ajps.12112>
- Codato, A., Costa, L. D., & Massimo, L. (2014). Classificando ocupações prévias à entrada na política: Uma discussão metodológica e um teste empírico. *Opinião Pública*, 20(3), 346-362. <https://doi.org/10.1590/1807-01912014203346>
- Crenshaw, K. (1989). Demarginalizing the intersection of race and sex: A Black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. *University of Chicago Legal Forum*, 1989(1), 139-167. <https://chicagounbound.uchicago.edu/uclf/vol1989/iss1/8>
- Hurlbert, S. H. (1971). The nonconcept of species diversity: A critique and alternative parameters. *Ecology*, 52(4), 577-586. <https://doi.org/10.2307/1934145>

- Hughes, M. M. (2011). Intersectionality, quotas, and minority women's political representation worldwide. *American Political Science Review*, 105(3), 604-620.
<https://doi.org/10.1017/S0003055411000293>
- Kendall, M. G. (1975). *Rank correlation methods* (4th ed.). Charles Griffin.
- Krook, M. L. (2009). *Quotas for women in politics: Gender and candidate selection reform worldwide*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195375671.001.0001>
- Mann, H. B. (1945). Nonparametric tests against trend. *Econometrica*, 13(3), 245-259.
<https://doi.org/10.2307/1907187>
- Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, 18(1), 50-60.
<https://doi.org/10.1214/aoms/1177730491>
- Miguel, L. F. (2008). Political representation and gender in Brazil: Quotas for women and their impact. *Bulletin of Latin American Research*, 27(2), 197-214. <https://doi.org/10.1111/j.1470-9856.2008.00263.x>
- Miguel, L. F., Marques, D., & Machado, C. (2015). Capital familiar e carreira política no Brasil: Gênero, partido e região nas trajetórias para a Câmara dos Deputados. *Dados*, 58(3), 721-747.
<https://doi.org/10.1590/00115258201557>
- Norris, P., & Lovenduski, J. (1995). *Political recruitment: Gender, race and class in the British Parliament*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511598609>
- Parkin, F. (1979). *Marxism and class theory: A bourgeois critique*. Columbia University Press.
- Phillips, A. (1995). *The politics of presence*. Clarendon Press.
- Pitkin, H. F. (1967). *The concept of representation*. University of California Press.
- Rezende, D. L. (2017). Desafios à representação política de mulheres na Câmara dos Deputados. *Revista Estudos Feministas*, 25(3), 1199-1218. <https://doi.org/10.1590/1806-9584.2017v25n3p1199>
- Sacchet, T. (2018). Why gender quotas don't work in Brazil? The role of the electoral system and political finance. *Colombia Internacional*, (95), 25-54.
<https://doi.org/10.7440/colombiaint95.2018.02>
- Sacchet, T., & Speck, B. W. (2012). Financiamento eleitoral, representação política e gênero: Uma análise das eleições de 2006. *Opinião Pública*, 18(1), 177-197. <https://doi.org/10.1590/S0104-62762012000100009>
- Schwindt-Bayer, L. A. (2011). Women who win: Social backgrounds, paths to power, and political ambition in Latin American legislatures. *Politics & Gender*, 7(1), 1-33.
<https://doi.org/10.1017/S1743923X10000541>
- Sen, P. K. (1968). Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association*, 63(324), 1379-1389.
<https://doi.org/10.1080/01621459.1968.10480934>
- Shannon, C. E. (1948a). A mathematical theory of communication. *Bell System Technical Journal*, 27(3), 379-423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>

- Shannon, C. E. (1948b). A mathematical theory of communication. *Bell System Technical Journal*, 27(4), 623-656. <https://doi.org/10.1002/j.1538-7305.1948.tb00917.x>
- Simpson, E. H. (1949). Measurement of diversity. *Nature*, 163, 688. <https://doi.org/10.1038/163688a0>
- Tribunal Superior Eleitoral. (s.d.). *Portal de Dados Abertos do TSE* [Conjuntos de dados de candidaturas e resultados eleitorais, 1998-2022]. Recuperado em 16 de julho de 2026, de <https://dadosabertos.tse.jus.br/>
- Weber, M. (1978). *Economy and society: An outline of interpretive sociology* (G. Roth & C. Wittich, Eds.). University of California Press.
- Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6), 80-83. <https://doi.org/10.2307/3001968>
- Wylie, K. N. (2018). *Party institutionalization and women's representation in Democratic Brazil*. Cambridge University Press. <https://doi.org/10.1017/9781108612722>
- Wylie, K. N., & dos Santos, P. G. (2016). A law on paper only: Electoral rules, parties, and the persistent underrepresentation of women in Brazilian legislatures. *Politics & Gender*, 12(3), 415-442. <https://doi.org/10.1017/S1743923X16000179>

Este preprint foi submetido sob as seguintes condições:

- Os autores declaram que os necessários Termos de Consentimento Livre e Esclarecido de participantes ou pacientes na pesquisa foram obtidos e estão descritos no manuscrito, quando aplicável.
- Os autores declaram que a elaboração do manuscrito seguiu as normas éticas de comunicação científica.
- Os autores declaram que estão cientes que são os únicos responsáveis pelo conteúdo do preprint e que o depósito no SciELO Preprints não significa nenhum compromisso de parte do SciELO, exceto sua preservação e disseminação.
- Os autores declaram que os dados, aplicativos e outros conteúdos subjacentes ao manuscrito estão referenciados.
- O manuscrito depositado está no formato PDF.
- Os autores declaram que a pesquisa que deu origem ao manuscrito seguiu as boas práticas éticas e que as necessárias aprovações de comitês de ética de pesquisa, quando aplicável, estão descritas no manuscrito.
- Os autores declaram que uma vez que um manuscrito é postado no servidor SciELO Preprints, o mesmo só poderá ser retirado mediante pedido à Secretaria Editorial do SciELO Preprints, que afixará um aviso de retratação no seu lugar.
- Os autores concordam que o manuscrito aprovado será disponibilizado sob licença [Creative Commons CC-BY](#).
- O autor submissor declara que as contribuições de todos os autores e declaração de conflito de interesses estão incluídas de maneira explícita e em seções específicas do manuscrito.
- Os autores declaram que o manuscrito não foi depositado e/ou disponibilizado previamente em outro servidor de preprints ou publicado em um periódico.
- Caso o manuscrito esteja em processo de avaliação ou sendo preparado para publicação mas ainda não publicado por um periódico, os autores declaram que receberam autorização do periódico para realizar este depósito.
- O autor submissor declara que todos os autores do manuscrito concordam com a submissão ao SciELO Preprints.