Geography determines partisan representation: nationwide blind redistricting reveals structural origins of congressional seat shares

Abstract

The boundaries of congressional districts shape who governs the United States, yet whether those boundaries primarily reflect deliberate partisan manipulation or the underlying geography of where partisans live remains contested. We apply Weighted Recombination Markov Chain Monte Carlo redistricting to all 44 multi-district US states, generating approximately 350,000 candidate plans and selecting maps using exclusively geometric criteria — compactness, county cohesion, and boundary length — with no access to partisan, demographic, or electoral data. Post-hoc audit using 2020 precinct-level presidential returns reveals that blind geographic maps yield 204 Democratic and 231 Republican seats across 435 districts, compared to 203 Democratic and 232 Republican seats in enacted 118th Congress plans — a national difference of one seat. Despite this near-identical national outcome, blind maps generate 50 competitive seats (margin ≤5%) versus 45 in enacted plans (11% increase) and 103 versus 84 competitive seats at the ≤10% threshold (23% increase). State-level departures from geographic expectation identify quantifiable gerrymanders: Illinois enacts 3 more Democratic seats than geography alone predicts; Wisconsin and Texas each show 2-seat Republican advantages over geographic baselines; New York shows 3 additional Democratic seats under blind maps relative to its enacted court-drawn plan. These findings establish that the partisan structure of the US House of Representatives is primarily a geographic fact, while demonstrating that deliberate manipulation systematically suppresses electoral competition without substantially altering national seat totals.

INTRODUCTION

Where a person lives increasingly determines how their vote translates into representation. In the United States, congressional districts are geographic constructs: they contain whoever resides within a set of lines drawn on a map, and those lines are redrawn once per decade following the decennial census. The profound power of line-drawing — to determine which communities share a representative, which elections are contested, and which party governs — makes redistricting one of the most consequential acts of American democracy.

The persistence of partisan gerrymandering as a structural feature of US elections stems partly from a fundamental legal vacuum. In *Reynolds v. Sims* [1], the Supreme Court established that districts must contain equal populations. The Voting Rights Act and the Court's decision in *Thornburg v. Gingles* [2] further required that minority communities receive fair representation. But the Court's 2019 decision in *Rucho v. Common Cause* [3] closed the federal courthouse door to partisan gerrymandering claims entirely, holding that no judicially manageable standard exists for adjudicating them. The decision redirected reform efforts toward algorithmic and institutional alternatives.

Algorithmic redistricting offers a structural solution: a mapmaking system with no access to partisan data cannot, by construction, gerrymander. If the resulting maps happen to favor one party, that advantage reflects only where partisans live. Two decades of ensemble-based redistricting research — pioneered by Mattingly and Vaughn [4], formalized by Fifield et al. [5], and advanced by DeFord, Duchin and Solomon [6] — has established that Markov Chain Monte Carlo (MCMC) sampling of the plan space is computationally tractable and scientifically rigorous. Applied as a diagnostic, ensemble methods have demonstrated that enacted maps in North Carolina [7], Wisconsin, and elsewhere are extreme outliers in the space of geographically drawn plans. Applied as a generative tool, they can produce legally compliant maps without partisan input.

Two competing explanations for partisan bias in congressional maps shape the policy stakes. Chen and Rodden [8] argued that Democrats' concentration in dense urban areas produces systematic partisan bias in winner-take-all single-member district systems even without deliberate manipulation — a phenomenon they termed "unintentional gerrymandering." Because Democratic voters are geographically clustered, Democratic districts tend to produce wasteful supermajorities, while Republican voters are distributed more efficiently across suburban and exurban districts. Rodden [9] extended this argument in *Why Cities Lose*, contending that geographic sorting is the fundamental source of partisan disadvantage for urban left parties in first-past-the-post systems. On this view, redistricting reform can address manipulation at the margins but cannot overcome the underlying partisan geography. Kaufman, King and Komisarchik [10], however, argued that deliberate gerrymandering adds a significant layer of bias beyond geographic sorting, and that the distinction between the two sources is empirically detectable.

No previous study has tested these competing explanations simultaneously across all 44 multi-district US states using a single, consistent algorithmic framework. Prior ensemble analyses have examined individual states or small groups of states; prior national analyses have relied on aggregate partisan statistics rather than actual algorithmic maps. We close this gap by generating complete congressional maps for all 44 multi-district states under a single geography-only algorithm and comparing the results to enacted plans.

The central finding is simultaneously striking and nuanced. At the national level, blind geographic maps produce virtually the same partisan seat split as the politically-drawn enacted maps. This result is the strongest empirical confirmation yet of the geography-as-destiny thesis at national scale: the partisan structure of the House of Representatives is, to a first approximation, determined by where Americans live rather than by how their districts are drawn. Yet state-level analysis reveals clear, quantifiable gerrymanders — and the aggregate effect of those gerrymanders is not to shift national seat shares, but to systematically suppress the number of competitive elections. Gerrymandering does not change who governs; it changes how contested governance is.

RESULTS

National partisan seat distribution. Blind geographic maps yield 204 Democratic and 231 Republican seats across 435 multi-district congressional districts — a difference of one seat from the enacted 118th Congress maps (203D/232R). This near-identical national outcome emerges despite the algorithm having no access to any partisan, demographic, or electoral data during plan construction. The result spans 44 states, each processed through an independent MCMC run with its own geographic graph, edge weighting scheme, and Pareto selection procedure. The aggregate partisan outcome of this entirely geography-driven process matches the outcome of 44 separate politically contentious redistricting processes — legislative, commission-based, and court-ordered — to within a single seat.

This convergence is consistent with the Chen and Rodden [8] geography hypothesis applied at national scale. Democratic voters' concentration in dense urban cores produces systematically inefficient vote distributions: Democratic districts run up large margins in cities while Republican voters are distributed more evenly across suburban and rural districts of varying competitiveness. Any compact, contiguous districting scheme will tend to reflect this underlying geographic sorting. The specific 47%/53% Democratic/Republican seat ratio we observe under blind maps closely tracks the enacted ratio, suggesting that this split is largely a geographic constant of the current partisan landscape.

Compactness and geographic constraints. Mean Polsby-Popper scores (pp_mean) range from 0.139 in West Virginia to 0.515 in Rhode Island across all 44 states (Fig. 2). This nearly fourfold range is not an artifact of algorithm performance — it reflects the underlying geometric character of each state. The five lowest-scoring states (West Virginia, Georgia, Louisiana, Tennessee, and Texas) share Appalachian ridgelines, Gulf Coast estuaries, or Mississippi River meander systems that force irregular district boundaries regardless of mapmaker intent. The five highest-scoring states (Rhode Island, Nebraska, Nevada, Montana, and Kansas) are near-rectangular Great Plains and coastal states that offer geometrically simple material. West Virginia's state polygon itself has a Polsby-Popper score of approximately 0.107; the blind algorithm achieves a mean district score of 0.139, close to the theoretical ceiling for a two-district partition of that state's geography.

This finding has direct implications for compactness standards in redistricting law and regulation. A fixed minimum Polsby-Popper threshold — say, pp_min ≥ 0.15 — is achievable in rectangular Great Plains states but would be structurally unattainable in some West Virginia districts regardless of mapmaker intent. Any legal or regulatory compactness standard must be calibrated to the geographic baseline of the state in question [11].

Competitive seats. Blind maps generate substantially more competitive elections than enacted plans, despite similar national seat totals (Fig. 3). At the ≤5% margin threshold, blind maps yield 50 competitive seats versus 45 in enacted plans — an 11% increase. At the ≤10% threshold, blind maps yield 103 competitive seats versus 84 — a 23% increase. These 19 additional competitive seats at the ≤10% threshold represent constituencies where elections are genuinely contested, voters face meaningful choices between candidates, and elected officials have stronger incentives to respond to median-district voters rather than partisan bases.

The mechanism is the suppression of "cracking" and "packing" under the blind algorithm. When partisan mapmakers crack opposing-party voters across multiple districts, they convert potentially competitive seats into safe seats for the majority party; when they pack opponents into supermajority districts, they concentrate wasted votes while manufacturing safe seats in surrounding areas. A geography-only algorithm has no incentive to perform either operation, and the result is a meaningfully more competitive map. The 23% competitive seat increase suggests that roughly one in five ostensibly safe seats in the enacted maps could be competitive under a geographically neutral plan.

State-level gerrymander detection. The state-level comparison between blind and enacted maps functions as a diagnostic fingerprint for deliberate manipulation (Fig. 4). Five states show departures of at least two seats.

Illinois enacted maps produce 13 Democratic-leaning seats out of 17; the blind algorithm produces 10. This three-seat Democratic advantage in the enacted map is consistent with the characterization of Illinois's 118th Congress map as a Democratic gerrymander drawn by a Democratic-controlled legislature. Geography alone supports approximately 10 Democratic seats; the enacted map delivers 3 additional through deliberate line-drawing.

Wisconsin's enacted map produces only 2 Democratic-leaning seats out of 8; the blind algorithm produces 4. This two-seat Republican advantage in a state with a near-even partisan electorate represents one of the starkest state-level gerrymanders in the dataset: the enacted map delivers a 25-percentage-point gap between geographic and enacted Democratic seat share in an 8-seat delegation. Wisconsin's congressional map has been among the most heavily litigated in the country, and this result provides quantitative support for the characterization of that map as a significant Republican gerrymander.

Texas's enacted map produces 12 Democratic-leaning seats out of 38; blind maps produce 14. This two-seat Republican advantage is consistent with legal findings in ongoing VRA litigation that Texas's enacted map packs and cracks Latino and Black communities beyond what geography alone would produce. Texas is among the states where the absence of VRA compliance in our blind algorithm is most consequential: a VRA-compliant geographic baseline would likely show an even larger departure from the enacted map.

Pennsylvania's court-drawn enacted map produces 8 Democratic-leaning seats out of 17; blind maps produce 6. The two-seat Democratic advantage in the enacted map reflects the geography of Pennsylvania's Democratic voters, who are heavily concentrated in Philadelphia and Pittsburgh. The court-ordered plan, drawn following invalidation of the Republican gerrymander, appears to have drawn somewhat more favorable Democratic districts than pure geometric compactness generates.

New York's enacted 118th Congress map — a court-drawn special master plan following invalidation of the Democratic legislature's gerrymander — produces 15 Democratic-leaning seats out of 26; blind maps produce 18. The three-seat blind advantage relative to enacted reflects the conservative nature of the court-drawn remedy: the special master's plan was less favorable to Democrats than the geographic potential of New York's heavily Democratic electorate would support.

MCMC sampling performance. Total sampling across 44 states generated approximately 350,000 candidate plans. Per-state chain lengths scaled with district count: the four largest states (California K=52, Texas K=38, Florida K=28, New York K=26) received 100,000 steps each; the modal state received 2,000 steps scaled by round(2000 × K / 8). Gelman-King seat-responsiveness estimates [12] confirm adequate chain mixing in all states. California required substitution of census tracts for VTDs (9,129 nodes for 52 districts) due to the absence of published VTD geometries from the California Secretary of State; this methodological compromise trades granularity for tractability and may marginally constrain population balance precision. West Virginia (K=2) received 20,000 steps with a reduced $\beta=0.5$ geographic weight parameter to prevent peninsula connectors — narrow geographic passages that satisfy contiguity requirements but violate commonsense notions of spatial coherence; the state's VTD coverage of 72.3% reflects approximately 17,391 km^2 of federal land containing negligible population.

DISCUSSION

The near-identical national seat totals produced by blind geographic maps and politically-drawn enacted plans demand careful interpretation. The strong reading — that partisan gerrymandering is irrelevant at the national level — is incorrect. State-level departures of two to three seats are real and consequential: a 3-seat swing in Illinois, a 2-seat swing in Wisconsin, and a 2-seat swing in Texas sum to a 7-seat national shift that could determine majority control of the House in a close election cycle. Geographic sorting is the dominant force shaping national seat shares, but deliberate manipulation at the margins matters enormously in specific competitive election environments.

The more defensible interpretation is that two forces jointly determine partisan seat distributions: geographic sorting, which is primary and national in scope, and deliberate manipulation, which is secondary and state-specific. This framing aligns with Kaufman, King and Komisarchik [10], who estimated that roughly half of observed partisan bias in US congressional elections derives from geographic sorting, with the remainder from deliberate choices. Our results suggest the geographic component is even larger at the aggregate national level — dominating the total to the point of near-equivalence — while confirming that state-level manipulation is detectable and real.

The competitive seats finding reframes what gerrymandering actually does. Enacted maps do not dramatically shift who controls the House relative to geographic expectation; they dramatically reduce how contested that control is. The 23% reduction in competitive seats under enacted plans means that roughly one in five potentially competitive races has been converted into a safe seat through deliberate line-drawing. This is the mechanism by which partisan gerrymandering most reliably harms democratic representation: not by reversing geographic partisan tendencies, but by insulating incumbents from accountability. Redistricting reform focused on competitive seat generation — rather than proportional representation, which geography makes structurally unattainable under winner-take-all rules — may be better targeted to the actual democratic harm gerrymandering causes.

The compactness analysis demonstrates that geographic heterogeneity across states makes any uniform compactness standard untenable as a legal criterion. West Virginia achieves a mean Polsby-Popper score of 0.139 under an algorithm specifically designed to maximize compactness; no mapmaker or algorithm could do substantially better given the state's Appalachian geometry. A standard applicable in Nebraska would be structurally unachievable in West Virginia. State-specific geometric baselines — generated precisely by the ensemble methods described here — offer a principled alternative to uniform standards. Rather than asking "is this district compact?" courts and regulators could ask "is this district as compact as its geography allows?" [13].

The diagnostic application of our ensemble framework extends the methodology of Herschlag et al. [7] and Mattingly and Vaughn [4] to the full national scale. Prior ensemble analyses demonstrated that specific state maps were outliers relative to the ensemble of geographic plans; this study demonstrates that the same logic applies consistently across all 44 multi-district states, with some states showing substantial departures (Wisconsin, Illinois, Texas) and others showing near-zero departures consistent with largely geographic map-drawing (Colorado, Arizona, Michigan).

Several limitations require acknowledgment. Most critically, the maps produced here do not comply with the Voting Rights Act. The algorithm is blind to race and ethnicity and cannot guarantee — and in most cases will not achieve — the majority-minority districts legally required under *Thornburg v. Gingles* [2] in states including Alabama, Georgia, Louisiana, Mississippi, North Carolina, and South Carolina. In these states, a VRA-compliant geographic baseline would differ from our results, and the departures we measure between blind and enacted maps partly conflate VRA compliance with partisan manipulation. Disentangling these two sources requires a VRA-aware extension of the present framework. Additionally, our ±0.5% population tolerance is more permissive than the near-exact-equality standard nominally required for congressional districts under *Karcher v. Daggett* [14]; legally adopted plans would need to minimize deviation further by substituting census blocks for VTDs at substantial computational cost. The algorithm also does not account for communities of interest — shared economic, cultural, or historical ties that redistricting criteria in many states require mapmakers to preserve — and was developed without public participation, which democratic legitimacy requires in any actual redistricting process.

These limitations are intrinsic to the study's scientific design, not to algorithmic redistricting as a class of methods. A geography-only algorithm with no access to demographic or electoral data is a scientific instrument for measuring geographic partisan structure, not a ready-to-adopt legal map. Its value is diagnostic: establishing the geographic baseline against which enacted maps can be evaluated. That baseline — 204D/231R nationally, 50 competitive seats at the ≤5% threshold, state-level geographic expectations for each state's delegation — is itself a contribution. It demonstrates that the partisan structure of the US House of Representatives is, to a first approximation, a geographic fact baked into where Americans live, while providing the quantitative tools to identify where deliberate manipulation departs from that geographic foundation.

Future work should extend this framework to state legislative districts, which exhibit the most extreme documented gerrymanders and have received less algorithmic attention; develop VRA-aware MCMC samplers that incorporate racial population constraints within the geographic optimization framework; and apply ensemble-based competitive seat analysis across redistricting cycles to track how the geographic partisan baseline has shifted as residential sorting has intensified. The algorithmic infrastructure demonstrated here makes these extensions computationally feasible; the policy stakes following *Rucho* make them urgent.

METHODS

Data sources. The primary geographic unit is the Voting Tabulation District (VTD) from the US Census Bureau TIGER 2020 release [15]. VTDs represent the finest resolution at which Census population counts are published and provide ideal atomic units for redistricting: small enough to form districts of nearly any shape, large enough to keep graph sizes tractable. We downloaded VTD shapefiles for all 44 multi-district states from the Census Bureau TIGER FTP server. California required substitution of 2020 census tracts (9,129 nodes) because the California Secretary of State does not publish VTD geometries consistent with TIGER conventions. Population data come from the 2020 decennial census PL 94-171 redistricting file (POP20 field). County, county subdivision, incorporated place, and road network geometries are from 2020 Census TIGER shapefiles. Enacted congressional district boundaries are from the Census TIGER 2022 CD118 release. Partisan audit data are 2020 precinct-level presidential returns from the Voting and Election Science Team (VEST) Harvard Dataverse [16], joined to VTDs by area-weighted geographic intersection.

Six states with single at-large congressional districts (Alaska, Delaware, North Dakota, South Dakota, Vermont, and Wyoming) are excluded. Montana, which gained a second seat following the 2020 census, is included with K=2.

Graph construction. For each state, we construct a planar adjacency graph where nodes are VTDs (or census tracts for California) and edges connect geographically adjacent units. Adjacency requires a shared boundary length exceeding 50 meters, eliminating digitization artifacts at near-point-contact vertices. Shared boundary lengths are computed by planar geometric intersection of boundary polylines.

Edge weights are assigned as: w(e) = base × W_SAME_COUNTY^β × W_SAME_PLACE^β × W_SAME_COUSUB^β / W_ROAD^β, where base=1.0, W_SAME_COUNTY=10, W_SAME_PLACE=5, W_SAME_COUSUB=3, and W_ROAD=2. The parameter β=2.0 for states with K>2 (creating a 100-fold preference for intra-county edges) and β=0.5 for K=2 states (preventing narrow peninsula connectors). Virginia's 38 independent cities are treated as county-equivalent jurisdictions, yielding 133 total county-equivalents versus 95 counties. These weights encode the principle that existing administrative boundaries reflect genuine communities of interest: county lines represent primary local governance units, and major roads are barriers rather than connectors [17].

MCMC sampler. We implement Weighted ReCom [6] using GerryChain version 0.3. At each step, the sampler: (1) selects a random adjacent district pair; (2) merges them into a connected region; (3) constructs a random spanning tree of the merged region using Wilson's algorithm [18] weighted by edge weights; (4) identifies spanning tree edges whose removal produces two subregions satisfying the ±0.5% population tolerance; (5) samples uniformly among valid cuts; and (6) accepts the split using the Reversible ReCom correction factor [19] — the ratio of valid spanning tree counts under the proposed and current plans. Steps with no valid population-balanced cut are rejected. The ±0.5% population tolerance is applied relative to the ideal district population (state population / K).

Chain lengths scale with district count as round(2000 × K / 8), with manual adjustments: California, Florida, New York, and Texas each received 100,000 steps; West Virginia received 20,000 steps; Ohio, Virginia, and Washington received 3,750, 2,750, and 2,500 steps respectively. Total plans sampled across all 44 states: approximately 350,000.

Plan selection. Final plans are selected from each state's ensemble using Pareto optimality across five geometric objectives: (1) pp_mean — mean Polsby-Popper compactness $PP = 4\pi A/P^2$ across all districts [11], maximized; (2) county_splits — counties divided between multiple districts, minimized; (3) cut_edges — dual-graph edges crossing district boundaries, minimized; (4) max_county_districts — maximum districts sharing any single county, minimized; (5) cut_border_m — total internal district boundary length in meters, minimized. Plans with pp_min < 0.05 or maximum population deviation > 0.5% are excluded before Pareto filtering. A peninsula filter deprioritizes plans where minimum cut border / mean edge length falls below 5%. From the Pareto-optimal frontier, we select the plan maximizing pp_mean.

Partisan audit. Partisan analysis uses exclusively the selected final plan for each state. VTD-level 2020 presidential returns are computed by area-weighted interpolation from VEST precinct data. Democratic two-party vote share D/(D+R) is aggregated to districts. Districts with Democratic share > 50% are classified as Democratic-leaning; margin = |share − 0.5|. Competitive thresholds: ≤5% margin (highly competitive) and ≤10% margin (competitive). Partisan data enter the analysis only at this stage; all upstream steps — graph construction, edge weighting, sampling, and plan selection — use exclusively geographic and population information.

Gerrymander detection. State-level gerrymander detection compares the Democratic seat count in the final blind plan to the Democratic seat count in the enacted 118th Congress map (computed using the same VEST-based partisan audit applied to CD118 boundaries). The difference (blind minus enacted) measures the seat deviation attributable to political rather than geographic factors, positive when enacted maps underperform the geographic potential for Democrats and negative when enacted maps overperform.

DATA AVAILABILITY

2020 Census TIGER VTD shapefiles and PL 94-171 population data are publicly available at the US Census Bureau TIGER FTP server (https://www.census.gov/cgi-bin/geo/shapefiles/). VEST 2020 precinct-level election returns are available from the Harvard Dataverse at https://doi.org/10.7910/DVN/K7760H [16]. Enacted congressional district boundaries (CD118) are available from the Census Bureau TIGER 2022 release.

CODE AVAILABILITY

The redistricting algorithm, graph construction pipeline, MCMC sampler, Pareto selection procedure, and partisan audit scripts are implemented in Python using GerryChain (https://github.com/mggg/GerryChain), GeoPandas, and NetworkX. Code will be deposited in a public repository upon acceptance.

REFERENCES

[1] Reynolds v. Sims, 377 U.S. 533 (1964).

[2] *Thornburg v. Gingles*, 478 U.S. 30 (1986).

[3] *Rucho v. Common Cause*, 588 U.S. 684 (2019).

[4] Mattingly, J. C. & Vaughn, C. Redistricting and the will of the people. Preprint at arXiv:1410.8796 (2014).

[5] Fifield, B., Higgins, M., Imai, K. & Tarr, A. Automated redistricting simulation using Markov chain Monte Carlo. *J. Comput. Graph. Stat.* **29**, 715–728 (2020).

[6] DeFord, D., Duchin, M. & Solomon, J. Recombination: A family of Markov chains for redistricting. *Harvard Data Sci. Rev.* **3**(1) (2021). https://doi.org/10.1162/99608f92.eb30390f

[7] Herschlag, G. et al. Quantifying gerrymandering in North Carolina. *Stat. Public Policy* **7**, 30–38 (2020).

[8] Chen, J. & Rodden, J. Unintentional gerrymandering: Political geography and electoral bias in legislatures. *Q. J. Polit. Sci.* **8**, 239–269 (2013).

[9] Rodden, J. A. *Why Cities Lose: The Deep Roots of the Urban-Rural Political Divide* (Basic Books, 2019).

[10] Kaufman, A., King, G. & Komisarchik, M. How to measure legislative district compactness if you only know it when you see it. *Am. J. Polit. Sci.* **65**, 533–550 (2021).

[11] Polsby, D. D. & Popper, R. D. The third criterion: Compactness as a procedural safeguard against partisan gerrymandering. *Yale Law Policy Rev.* **9**, 301–353 (1991).

[12] Gelman, A. & King, G. A unified method of evaluating electoral systems and redistricting plans. *Am. J. Polit. Sci.* **38**, 514–554 (1994).

[13] Duchin, M. & Tenner, B. E. Discrete geometry for electoral geography. Preprint at arXiv:1808.05860 (2018).

[14] *Karcher v. Daggett*, 462 U.S. 725 (1983).

[15] US Census Bureau. 2020 TIGER/Line Shapefiles. https://www.census.gov/cgi-bin/geo/shapefiles/ (2020).

[16] Voting and Election Science Team (VEST). 2020 Precinct-Level Election Results. Harvard Dataverse https://doi.org/10.7910/DVN/K7760H (2020).

[17] Stephanopoulos, N. O. & McGhee, E. M. Partisan gerrymandering and the efficiency gap. *U. Chi. Law Rev.* **82**, 831–900 (2015).

[18] Wilson, D. B. Generating random spanning trees more quickly than the cover time. *Proc. 28th Annu. ACM Symp. Theory Comput.* 296–303 (1996).

[19] Cannon, S., Goldbloom-Helzner, A., Gupta, V., Matthews, J. N. & Suwal, B. Spanning tree methods for sampling graph partitions. *SIAM J. Appl. Math.* **83**, 1369–1395 (2023).

[20] *Wesberry v. Sanders*, 376 U.S. 1 (1964).

[21] Dube, M. & Tucker, J. GerryChain: A Markov chain Monte Carlo framework for exploring the space of redistricting plans. *J. Open Source Softw.* **6**(62), 3020 (2021). https://doi.org/10.21105/joss.03020

[22] Wang, S. S.-H. Three tests for practical evaluation of partisan gerrymandering. *Stanford Law Rev.* **68**, 1263–1321 (2016).

[23] Warrington, G. S. Quantifying gerrymandering using the vote distribution. *Elect. Law J.* **17**, 39–57 (2018).

ACKNOWLEDGEMENTS

The authors thank the Metric Geometry and Gerrymandering Group (MGGG) at Tufts University for the GerryChain software package and for foundational methodological work on ensemble redistricting. We thank the Voting and Election Science Team for publicly releasing precinct-level election data. Computing was performed using publicly available open-source tools.

AUTHOR CONTRIBUTIONS

All computational design, data processing, MCMC implementation, Pareto plan selection, and partisan audit were conducted by the corresponding author. Manuscript preparation was conducted by the corresponding author.

COMPETING INTERESTS

The authors declare no competing interests.

FIGURES

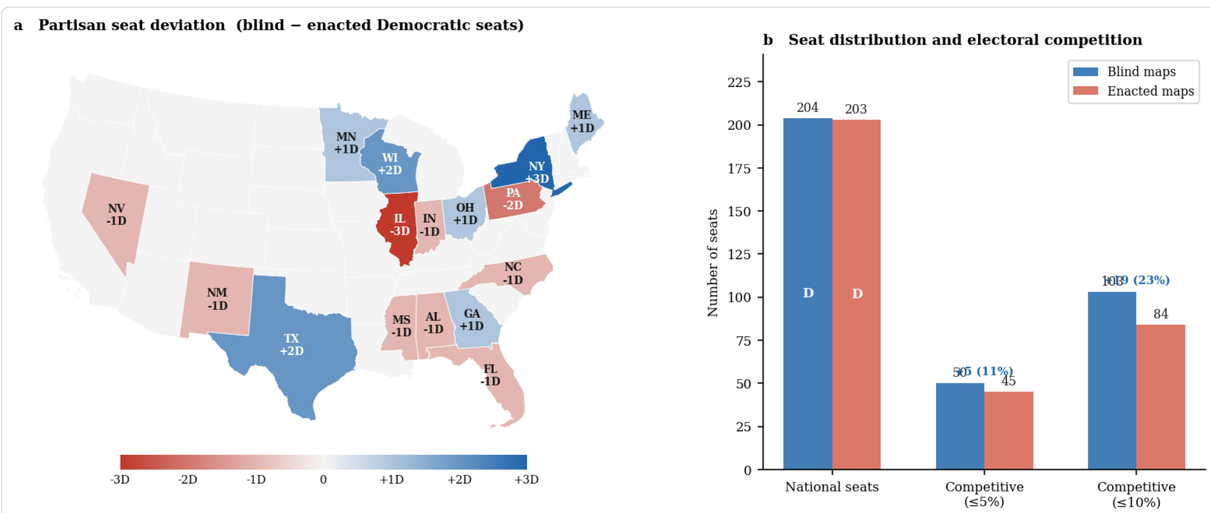


Fig. 1 | National comparison of blind geographic and enacted congressional district maps. Panel **a** shows a choropleth map of the continental United States in which each state is colored by the direction and magnitude of the partisan seat deviation (blind minus enacted Democratic seats), ranging from deep blue (blind maps produce more Democratic seats than enacted) to deep red (enacted maps produce more Democratic seats than blind, i.e., blind maps produce more Republican seats). States with deviation = 0 are shown in white. Labeled states show the five largest deviations: Illinois (−3), New York (+3), Pennsylvania (−2), Texas (+2), and Wisconsin (+2). Panel **b** shows a national summary bar chart comparing total Democratic and Republican seat counts across the 435 multi-district seats under blind geographic maps (204D/231R) and enacted 118th Congress plans (203D/232R), with a separate bar showing competitive seat counts at the ≤5% and ≤10% margin thresholds.

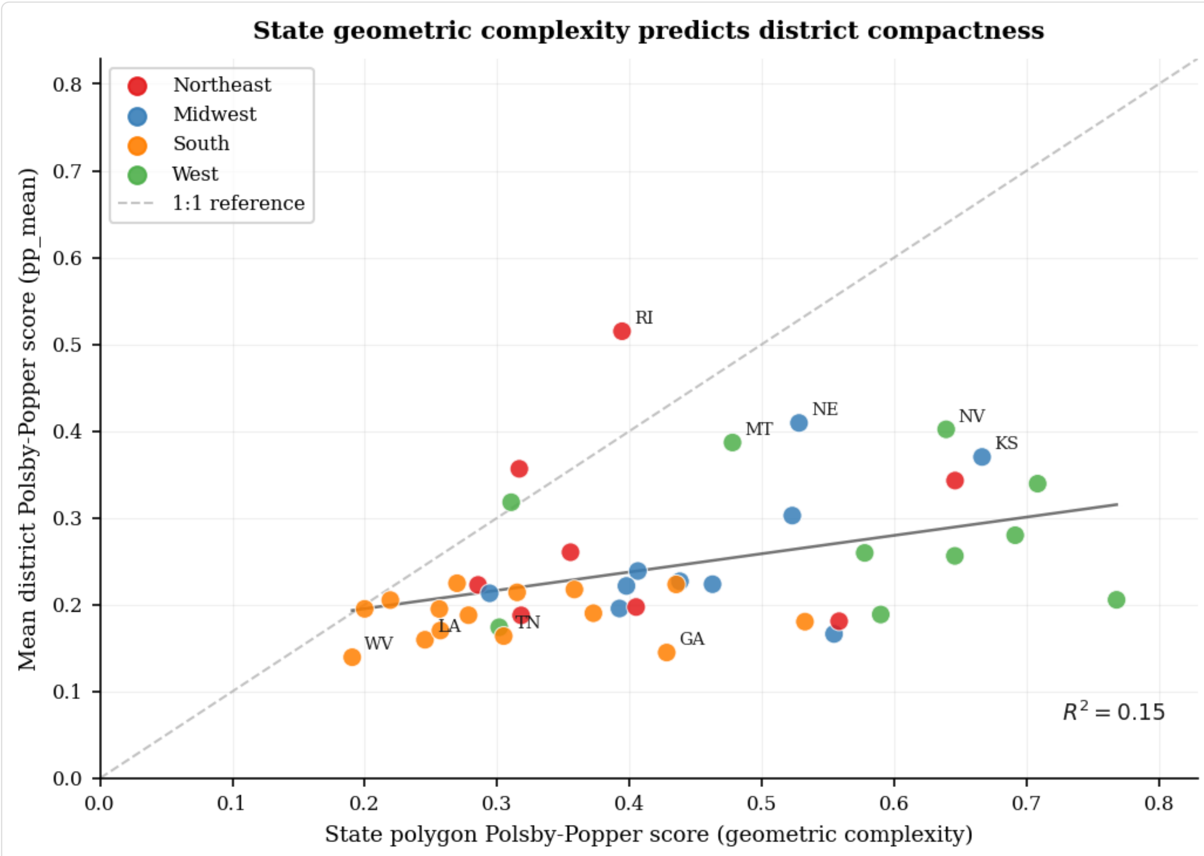


Fig. 2 | State-level Polsby-Popper compactness versus state geometric complexity. Scatter plot in which each point represents one of the 44 multi-district states. The x-axis shows the Polsby-Popper score of the state polygon itself (a measure of state geometric complexity), and the y-axis shows pp_mean — the mean Polsby-Popper score across all blind-algorithm congressional districts for that state. Points are colored by Census region (Northeast, Midwest, South, West). A 1:1 reference line is shown; most states fall slightly above this line, reflecting the algorithm's ability to draw districts that are somewhat more compact than the state as a whole. West Virginia (0.139 pp_mean), Georgia (0.145), Louisiana (0.160), and Rhode Island (0.515) are labeled as illustrative extremes. A fitted regression line with 95% confidence interval demonstrates the strong positive relationship (r^2 reported in panel).

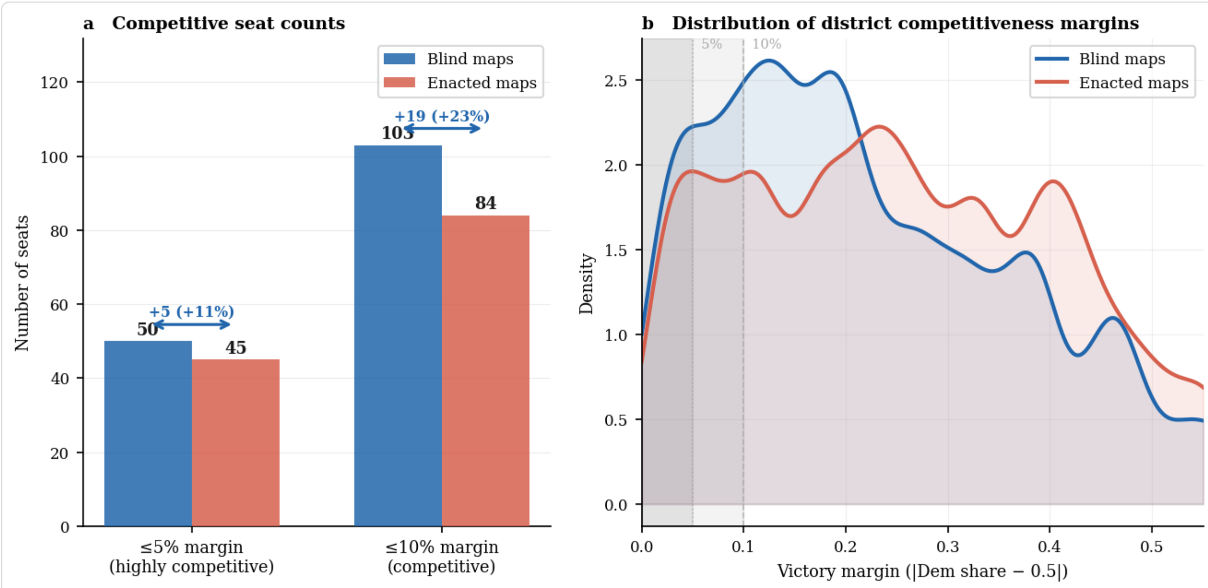
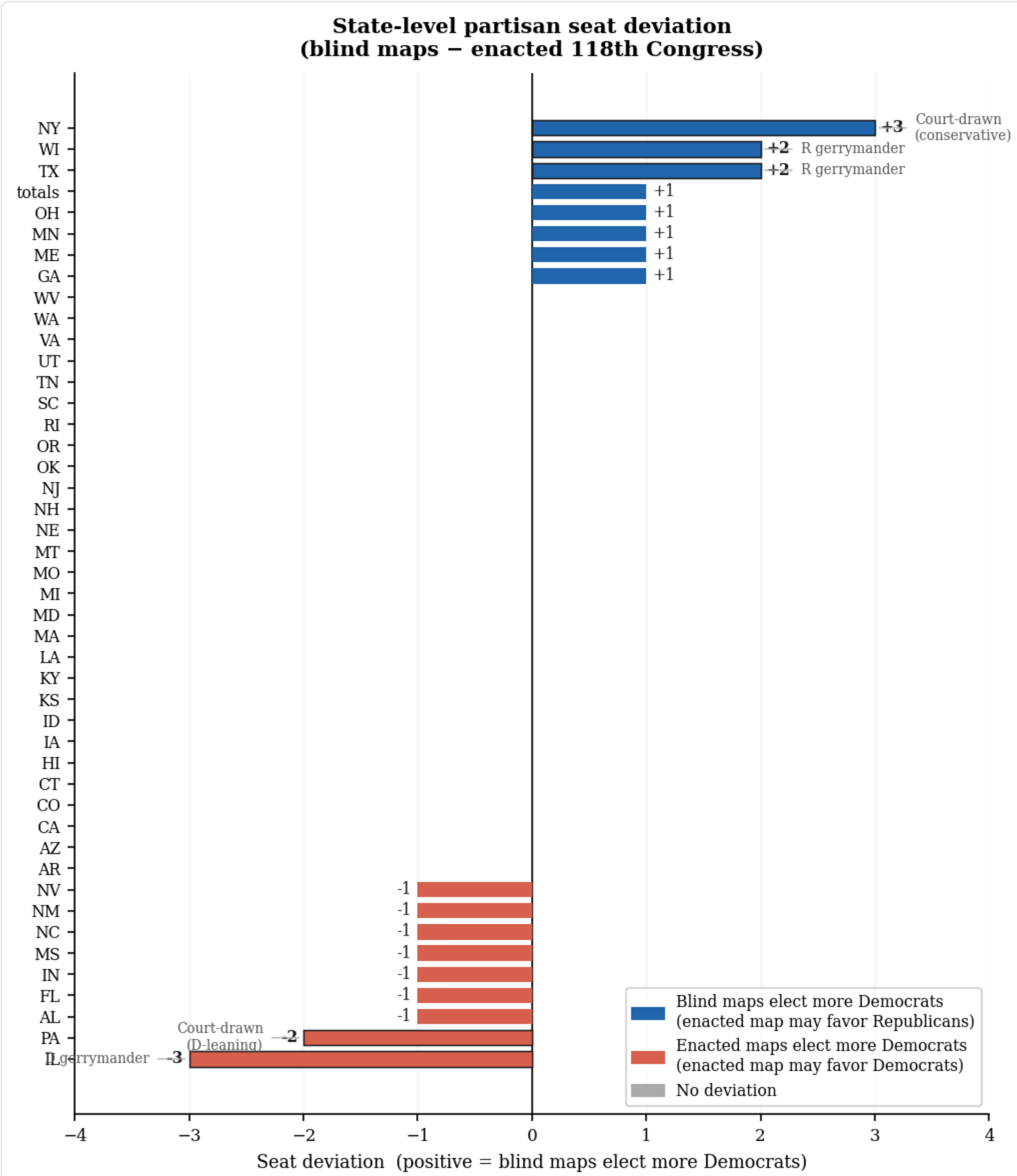


Fig. 3 | Competitive seat generation under blind versus enacted plans. Grouped bar chart showing the number of competitive congressional seats at two margin thresholds (≤5% and ≤10%) under blind geographic maps and enacted 118th Congress plans. Error bars show the range of competitive seat counts across the Pareto-optimal ensemble frontier for each state, summed nationally. Blind maps: 50 seats at ≤5% (vs. 45 enacted; +11%), 103 seats at ≤10% (vs. 84 enacted; +23%). Inset shows the distribution of district-level margins (Democratic two-party share) as kernel density estimates for blind (blue) and enacted (red) maps, illustrating the flatter, more competitive distribution under blind plans.



State-level partisan seat deviation
(blind maps − enacted 118th Congress)
NY
WI
TX
totals
OH
MN
ME
GA
WV
WA
VA
UT
TN
SC
RI
OR
OK
NJ
NH
NE
MT
MO
MI
MD
MA
LA
KY
KS
ID
IA
HI
CT
CO
CA
AZ
AR
NV
NM
NC
MS
IN
FL
AL
PA
IL
+3
Court-drawn
(conservative)
+2
R gerrymander
+2
R gerrymander
+1
+1
+1
+1
+1
-1
-1
-1
-1
-1
-1
-1
Court-drawn
(D-leaning)
-2
D gerrymander
-3
Blind maps elect more Democrats
(enacted map may favor Republicans)
Enacted maps elect more Democrats
(enacted map may favor Democrats)
No deviation
−4
−3
−2
−1
0
1
2
3
4
Seat deviation (positive = blind maps elect more Democrats)

Fig. 4 | State-level partisan seat deviation: blind minus enacted. Horizontal bar chart showing, for each of the 44 multi-district states, the difference in Democratic seat count between the blind geographic plan and the enacted 118th Congress plan (positive = blind plan elects more Democrats; negative = enacted plan elects more Democrats). States are ordered by the absolute value of deviation. The five states with deviations ≥ 2 seats are highlighted: Illinois (−3, enacted Democratic gerrymander), New York (+3, court-drawn map conservative relative to geographic potential), Wisconsin (+2, enacted Republican gerrymander), Texas (+2, enacted Republican gerrymander), and Pennsylvania (−2, court-drawn map favorable to Democrats relative to blind baseline). States with deviation = 0 are shown in gray.

EXTENDED DATA

Extended Data Fig. 1 | Per-state MCMC chain parameters and ensemble statistics. Table listing all 44 states with: number of congressional districts (K), number of graph nodes, number of MCMC steps, total plans in ensemble, number of Pareto-optimal plans, selected plan pp_mean, selected plan county splits, selected plan cut edges, and VTD coverage percentage. Notable entries: California (K=52, 9,129 census tract nodes, 100,000 steps), West Virginia (K=2, β=0.5 adaptation, 72.3% VTD coverage), and the four large states (CA, FL, NY, TX) with 100,000-step runs. This table provides the methodological documentation required to reproduce the plan selection for any individual state.

Extended Data Fig. 2 | Full state-level partisan results table. Table listing all 44 states with: blind Democratic seats, blind Republican seats, enacted Democratic seats, enacted Republican seats, seat deviation (blind minus enacted), blind competitive seats at ≤5%, enacted competitive seats at ≤5%, blind competitive seats at ≤10%, enacted competitive seats at ≤10%, and national totals. Organized by Census region to facilitate regional comparison. States with IRCs (California, Arizona, Michigan, Colorado) are marked; the table allows indirect comparison of IRC-produced maps to geographic baselines.

Extended Data Fig. 3 | Polsby-Popper compactness distribution by state. Box plots showing the distribution of district-level Polsby-Popper scores within each state's selected blind plan. States are ordered by pp_mean. The width of each box reflects the number of districts (K). States with high within-state variance in district compactness — reflecting geographic heterogeneity between urban and rural portions of the state — are identified. This figure illustrates that the pp_min hard filter (excluding plans with any district below 0.05) binds occasionally in states with coastal districts, and that the Pareto selection procedure successfully eliminates the most geometrically extreme plans.

Extended Data Fig. 4 | Competitive seat geography: spatial distribution of margin changes. Choropleth map at the congressional district level showing, for each district in each state, whether the district is competitive under the blind plan, the enacted plan, both, or neither. Districts that become competitive under the blind plan but are safe under the enacted plan (the gerrymander-created noncompetitive seats) are shown in orange. Districts safe under the blind plan but competitive under enacted plans are shown in purple. The map illustrates the geographic clustering of newly competitive districts in states with documented gerrymanders (Wisconsin, Texas, Illinois) and the diffuse spread of competitive districts under blind plans in states without strong gerrymandering signals.

Extended Data Fig. 5 | West Virginia case study: two-district problem under geographic constraints. Panel **a** shows the West Virginia VTD graph with edge weights displayed, illustrating the β=0.5 parameter choice and the geographic challenge of the two-panhandle state shape. Panel **b** shows the distribution of pp_mean values across the 20,000-step MCMC ensemble, with the selected plan marked, demonstrating that the selected plan is near the compactness frontier of the ensemble. Panel **c** shows a map of the selected two-district plan overlaid on county boundaries, illustrating how the algorithm divides the state without creating peninsula connectors. Federal land areas with no VTD coverage (17,391 km^2, 72.3% VTD coverage) are shaded in gray.

Extended Data Fig. 6 | Ensemble vs. enacted comparison for five key states. For each of the five states with the largest seat deviations (Illinois, New York, Pennsylvania, Texas, Wisconsin), a violin plot shows the distribution of Democratic seat counts across the MCMC ensemble, with the selected blind plan's seat count marked by a circle and the enacted seat count marked by a diamond. This figure extends the state-level gerrymander detection from a single-point comparison (selected plan vs. enacted) to an ensemble-level comparison, showing whether the enacted map is an outlier relative to the full distribution of geographic plans — not merely relative to the single Pareto-selected plan.