Nv14 8/5/26

Analysis of Three Metrics of Partisan Bias

John F Nagle

Abstract: Assuming that partisan bias should be avoided in redistricting, it is important that it be based on a firm normative principle of fairness and that it be accurately measurable. A traditional metric, called seats bias, is based on the principle of equal seat share for equal vote share. When its implementation is suitably modified to take into account competitive districts, accuracy at the 1% level is achieved by using results from many previous elections, and the difference in partisan bias between two plans is even better determined. However, seats bias fails to detect cracking in unbalanced states. Proportionality has often been considered to be a desirable norm for representation even though it is generally unachievable over significant ranges of the vote in the single member district system. Nevertheless, it is appealing to measure bias as a deviation from proportionality, but just at the projected average partisan vote. However, this paper argues that even this modified proportionality is not a good metric of bias because it fails to take into account the effect of responsiveness in the single member district system. As is well known, responsiveness is also an important consideration, and three different forms of it play roles in considerations of partisan advantage for states with dominant parties. To take into account non-proportional responsiveness, as well as to focus on the average partisan vote, a metric based on the historical cubic law is also considered. This paper presents analyses of these three metrics for the 2022 US congressional plans of states with more than one district.

1. Introduction

Electoral maps that favor the voters of one party over voters of another party are by now clearly understood to be undesirable by fair minded people. Accordingly, redistricting legislation ought to consider partisan fairness since legislative bodies, especially Congress, have become so rigorously partisan. SCOTUS has abdicated oversight of partisan bias to the states[1], so any action has to take place in state election law (Cervas et al., 2022). Although redistricting for congress in many states in the USA has recently been weaponized to maximize partisan advantage, in the hope that fairness will eventually prevail, there remains the long-term issue of how to actually measure partisan bias.

There are at least two sources of partisan bias. Bias can arise from the political geography of the state, such as the higher concentrations of Democrats in cities than those of Republicans in rural areas. The traditional districting rules of a state, such as compactness and not splitting cities and counties can exacerbate that bias. While the neutral criteria constrain the worst gerrymanders, it is also claimed that they make it nearly impossible to obtain partisan fairness in many states (Keena et al., 2021, Rodden, 2019, Nagle, 2019). Then, there is intentional gerrymandering to achieve partisan advantage above whatever "natural" bias can be attributed to the other causes (Barton and Eguia, 2024, Magleby and McDonald, 2025a); it is often important for challenging maps in state courts to establish intentional gerrymandering. Disentangling intentional gerrymandering from unintentional partisan bias is difficult and beyond the scope of this paper. Instead, we assume in this paper that total partisan bias, whatever its provenance, should be minimized in choosing new maps,[2] and that election law should be written to enable this.

There are many metrics for partisan bias, and which should be used has been rather unclear and even contentious. This has weakened the case for including partisan bias in election law, instead falling back on the traditional so-called "neutral" criteria.[3] Even when the neutral criteria

[1] Rucho v. Common Cause, 139 S. Ct. 2484 (2019).

[2] As mentioned by Katz et al. (2020), "…the absence of intentional gerrymandering is not the same as fairness."

[3] An interesting new criterion even works against partisan fairness in half the states (DeFord et al., 2022)

permit a fair map, they allow maps that cover a fairly wide range of partisan bias.[4] In that case, the map that minimizes partisan bias could be chosen, while still satisfying other criteria, but for that it is still necessary to have a good metric.

The many metrics that have been proposed[5] represent different perspectives of partisan bias. It has been suggested that five metrics correlated well with each other and, rather than choosing just one of them, that a composite of fairness metrics would provide such a single number. (Nagle and Ramsay, 2021) However, that composite metric is not simple to calculate and some of the included metrics could not be applied to plans in some states. Furthermore, even when courts have expressed interest in measuring partisan bias, they have tended to call for a single, simple metric.[6] so this paper focusses on comparing single metrics based on the percentage of seat share that can be estimated to ensue from redistricting plans, although it ultimately recognizes that multiple approaches are necessary.

The first metric this paper focusses on is the time-honored seats bias (SB) metric, not only because it is a traditional metric, but because it is based on the strong normative principle of equal seats for equal votes. Namely, a fair plan should make it most likely that parties that obtain equal vote shares would obtain the same number of seats.[7] When there are two dominant parties that together obtain all the seats, it is convenient to concentrate on the two-party vote share.[8] Then the fairness principle, succinctly stated, is half the seats for half the votes. The seats bias metric of party A is calculated by how much the seats percentage of party A differs from 50% when its vote is 50%. Equation (1) expresses SB as a percentage,

$$SB = S(50\%) - 50\% \ , \qquad (1)$$

where S(50%) is the percentage of seats at 50% of the two-party vote.[9] Of course, elections

[4] Goedert et al. (2024) have recently employed "hill-climbing" simulations that exhibit the "stealth gerrymandering" effect with respect to compactness.

[5] The popular on-line tool DRA2020 (Bradlee, 2020) reports values from other worthy metrics in its Advanced section, notably mean-median (McDonald and Best 2015), the efficiency gap (McGhee 2014,) and declination (Warrington 2018). A relatively new quadratic rule metric (Barton,2022) has merit, although it does not actually ensue from the cited appealing cut and choose districting method.

[6] Davis v Bandemer 478 U.S. 109 (1986) , at 125 n9. Justice White, Stevens and Powell concurring.

[7] It is noteworthy that this principle is not limited to two parties.

[8] This paper is confined to current USA states. As these are dominated by two parties, for convenience we use the two-party vote such that the GOP and DEM vote percentages add to 100%.

[9] In a two-party state, all bias metrics, including SB, simply have opposite signs for the two parties, so it suffices to only consider metrics for one party, which is chosen to be the GOP in this paper.

almost never have equal vote, so the historical concern with this metric has been the necessity of having to shift the vote so as to be able to draw the seats S versus vote V curve (S(V)) to the 50% point using data from an election where V may be far from 50%. It will be shown that the SB metric provides highly consistent values of bias over many past elections that span a wide vote range. That is to say, the average values of bias, so obtained, have quite small statistical uncertainties, thereby showing that it is a robust metric for how much the seat share differs from the vote share at 50% vote share. Nevertheless, accuracy doesn't necessarily guarantee reliability, and SB becomes problematic for some, although not all, highly unbalanced states.[10]

One of the metrics that was found by (Nagle and Ramsay, 2021) not to correlate well with the others is proportionality. Its normative principle is that when a percentage V_A of voters vote for party A, then party A should obtain the same percentage of the seats $S_A = V_A$. This way, even though some A voters will live in districts represented by other parties, A voters as a group have the same average representational power, defined as S_A/V_A, as voters for a different party B that have representational power S_B/V_B. Assuming this principle leads to proportionality which is represented by the seats-votes function S(V) = V. The proportionality metric of partisan bias (P) is then given as the difference between the estimated seat percentage S(V) and the vote percentage V,

$$P(V) = S(V) - V. \qquad (2)$$

However, a voter's representational power clearly decreases abruptly when the voter's party loses a majority of the seats, so the normative principle behind proportionality is rather weak.

Proportionality has been achieved by being baked into various voting systems.[11] However, proportionality is clearly not achievable in the single member district voting system in the United States (Grofman, 1982). Parties with 1% of the vote in a two party state cannot possibly have any district with more than 50% of its supporters when there are fewer than 50 districts, even if all those voters comprise 100% of the population in one geographical area. More realistic political geography makes it likely that a political minority will not obtain any representation in a

[10] Unbalanced states are defined as states whose voters have a considerably higher preference for one party than the other.

[11] The list systems and the mixed member proportional system used in other countries are notable examples (Farrell, 2011).

two party state even with 25% of the vote. An important paper definitively and rigorously showed that it was mathematically impossible to draw a single district that would lean Republican in the decade before the MA map was drawn even though the vote was nearly 40% Republican (Duchin et al., 2019). That is because the partisan voter preference for MA is geographically relatively homogeneous outside the Boston area. Likewise, it is important to consider a hypothetical state that has completely homogeneous partisan preference geographically for two major parties. For such a state, all districts, no matter how drawn, would have the same partisan preference and the seats-votes S(V) curve would necessarily have zero seats for a party when it has V less than 50% and all the seats when V is greater than 50%.[12] Empirical studies (Goedert, 2014, Goedert et al., 2024, Barton, 2022) have shown that "partisan-neutral maps rarely give seats proportional to vote" (Rodden and Weighill, 2022).

Even though it is not generally possible to achieve proportionality over a substantial range of the vote, proportional outcomes have remained appealing, so a fall-back goal is to draw maps that can be estimated to most closely achieve proportionality at an historic value of the statewide vote estimated from previous elections. Proportionality in this modified form is consistent with Justice O'Connors characterization in *Davis v. Bandemer* (1986)[13] that "the plurality opinion ultimately rests on a political preference for proportionality - not an outright claim that proportional results are required, but a conviction that the greater the departure from proportionality, the more suspect an apportionment plan becomes", and to a 2015 amendment to the Ohio state constitution that the "statewide proportion of districts … shall correspond closely to the statewide preferences of the voters of Ohio."[14] The modified metric of bias mP is then defined by how far the seat share differs from the vote share when the vote share is the estimated historic vote share V_h

$$mP = S(V_h) - V_h, \qquad (3)$$

[12] Of course, even in such a state statistical fluctuations would make actual outcomes more smoothly varying near 50% vote than the most probable projection represented by this winner take all S(V) function, but the ensuing S(V) curve looks more like an abrupt jump at V=50% than like the proportional S(V) = V function.

[13] Davis v. Bandemer, 478 U.S, 109 (1986).

[14] Article XI, Sections 6(B) of the Ohio Constitution.

It is far less demanding of a plan to require that it make mP equal to zero because it only has to do this for one estimated voter preference V_h rather than make the entire seats/votes curve S(V) equal to V for all possible votes. Proportionality in a similar form has been clearly described and advocated (Duchin and Schoenbach, 2023).[15] Recent papers in these pages have essentially approved the mP metric (Ramsay, 2023, Gordon and Yntiso, 2024). Importantly, the popular and very useful online redistricting tool, DRA2020 (Bradlee, 2020), prioritizes and advocates mP over SB and many of the other metrics it calculates. Even though modified proportionality mP would therefore appear to be an attractive and accessible metric for measuring partisan bias, we present arguments that it is a poor one for the single member district system used in the USA. Furthermore, reformers who would employ it in an unbalanced state would disadvantage that state's power in the USA congress relative to states that do not employ it.

Notwithstanding our concerns about proportionality, it is nevertheless desirable to focus on the historic partisan vote. But because responsiveness is so non-proportional, a third metric is considered that assumes the historical cubic "law" that fit old data (Kendall and Stuart, 1950, Gudgin and Taylor, 1978) and that has also recently been reported to provide a good fit to 22 states (Wise, 2026). This has a normative seats-votes curve

$$S_3(V) = 100/(1+((100-V)/V)^3) \quad . \quad (4)$$

Figure 1 compares $S_3(V)$ to proportionality and the efficiency gap (Stephanopoulos and McGhee, 2015). The cubic $S_C(V)$ has responsiveness $\rho_{50} = dS_3/dV = 3$ at V = 50% compared to 1 for proportionality and 2 for the efficiency gap. The corresponding bias, to be designated the 3B bias, is then calculated as the difference between estimated seats S(V) and the cubic $S_3(V)$ in Eq. (4) at the estimated historical average vote V_h,

$$3B = S(V_h) - S_C(V_h) \quad . \quad (5)$$

The shortcoming of the 3B metric is that it lacks a firm normative principle since each state generally would be expected to have its own level of responsiveness (Tufte, 1973). However, it will be shown that the 3B values agree much better with SB than with mP for many states, indicating that the level of responsiveness is generally closer to cubic than to simple

[15] A minor difference is that Duchin and Schoenbach (2023) obtain the proportionality bias for several elections and then effectively average rather than averaging the elections first as for the mP bias in this paper.

proportionality. Importantly, it will be shown that 3B has plausible differences from SB for those states that are problematic for the SB metric. As has been noted many times in the literature, no one metric is likely to work well in all states, so consideration of several metrics is appropriate.

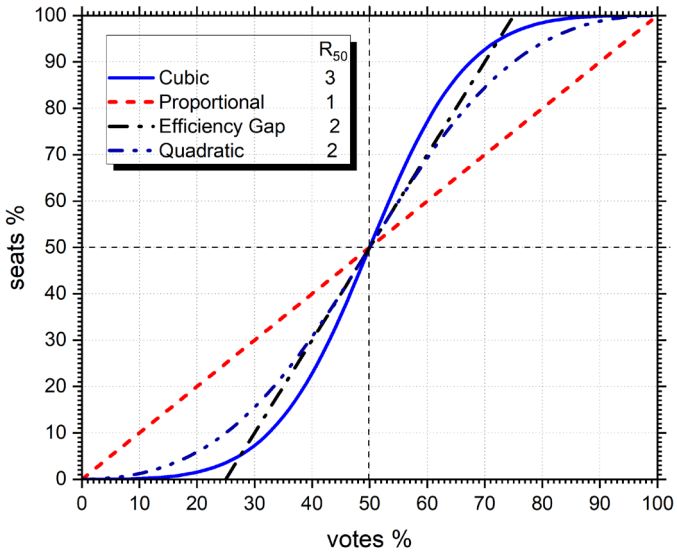

Fig. 1. Ideal cubic seats-votes curve compared to proportionality, the efficiency gap (EG) and the quadratic bilogit; the latter has the same responsiveness R_{50} at 50% vote as the EG and a more gradual approach to 0% and 100% seats.

The outline of this paper follows. The next section 2 reviews the methodology in DRA2020 (Bradlee, 2020) that we use. This is followed by section 3 that focuses on a single state and a single plan, the VA 2022 Congressional plan. This in-depth example illustrates the underlying issues, particularly the role of responsiveness. Section 4 tabulates values of the bias metrics and various responsiveness metrics for the 44 USA states with more than one district. Then, section 5 focuses on the SB metric, and it emphasizes that SB has been made into a precise metric in DRA2020. This is followed by section 6 that critiques the mP metric. It has long been known (Tufte, 1973) that responsiveness affects what ultimately constitutes partisan fairness in states with a dominant party, so the next section 7 considers two additional metrics of responsiveness that lead to differentiating three broad types of plans. While SB is satisfactory for type A plans, 3B is more appropriate for type B plans, and type C plans are basically those that are unresponsive. Section 8 recapitulates the results and discusses them more broadly. Section 9 recommends how partisan bias might be included in election law. Conclusions are listed in Section 10.

For convenience, the following glossary collects some of the main terms used in this paper.

- SB is seats bias defined as GOP seats in excess of 50% for equal vote for the two parties, as in Eq. (1).
- P is proportionality bias defined as the difference between % GOP seats S and % GOP vote V, as in Eq. (2).
- mP is proportionality P just at the estimated average vote V, as in Eq. (3).
- 3B is cubic bias defined as GOP seats in excess of the cubic law (Eq. (4)) at the estimated average vote V, as in Eq. (5).
- S(V) is the seats-votes curve defined as seats S for vote V, *viz.* the Methods section.
- $R_{40\text{-}60}$ is a broad responsiveness defined as [S(60%)-S(40%)]/20%.
- R_{50} is responsiveness narrowly defined as the slope dS/dV in the S(V) curve just at V = 50%.
- R_V is responsiveness narrowly defined as the slope dS/dV in the S(V) curve just at the most likely vote V.

2. Methods

Results in this paper are obtained using Dave's Redistricting App (DRA2020 (Bradlee, 2020)), which will be abbreviated as DRA. DRA has the following four features of primary importance for this study. First, it applies the two-party vote in a past statewide election to each precinct to obtain the estimated voter preference in each precinct and thus in each district, which is a standard practice (Magleby and McDonald, 2025a). Statewide elections are deemed preferable to district elections because (i) they remove the issue of assigning votes to uncontested districts and (ii) the voter preferences of all the precincts that will form a new district are estimated by the same election.[16]

The second feature in DRA is an important modification of the traditional way to estimate seats. Of course, in an actual election, there is one winner in each district, so the number of seats is precisely determined by winner take all (WTA) in each district. However, this is a poor way to estimate seats before an actual election. Indeed, it has been shown that WTA seat assignment leads to severe contradictions for the SB metric in contrived examples (Nagle, 2015) and

[16] Congressional elections likely estimate the voter preference of two precincts with the same underlying preference differently if they are in different districts with different candidates.

(DeFord and Veomett, 2025), and to unrealistic discontinuities in the seats-votes curve that is especially egregious for states with a small number of districts. Appendix A elaborates on this feature. The problem with WTA can most obviously be seen by considering a district where the two-party voter preference is estimated to be 50%. Prior to any subsequent election, such a district should clearly be evaluated as half a seat for either party. A district that differs from 50% by a small amount should also have fractional seats because future district votes are variable.[17] DRA assigns fractional seats to all districts using a standard probability function that gradually[18] approaches winner take all as the voter preference moves away from 50%.[19,20]

The third DRA feature estimates a seats-votes curve S(V) for all V for a statewide election by shifting the estimated voter preference in each district. Seats-votes curves have long been recognized to be central for analysis of partisan fairness; in particular, they are necessary for the SB metric since statewide elections generally differ from 50% vote. Of course, different statewide elections determine different S(V) curves. Importantly, averaging these for a suite of elections provides a better estimate of S(V) along with statistical uncertainties. The traditional uniform shift method (Tufte,1973) shifts the vote by the same percentage in each district. DRA employs proportional shift (Nagle, 2015) (Katz et al., 2020) (Wilson and Grofman, 2022)[21] of those voters whose votes have to shift in order to shift the overall statewide vote. The motivation for proportional shift is to avoid having shifted districts with more than 100% or less than 0% voters in a party, something that occurs for large uniform shifts, but it makes only small differences for typical shifts of interest in this paper.[22]

---

[17] It may also be noted that Gelman and King (1994) accomplished variability by use of a random vote in the JudgeIt tool (Gelman et al., 2012).

[18] It is often supposed that all districts in the 45-55% preference range are competitive, but assigning all of these the same fraction of estimated seats makes the implausible assumption that there is an abrupt shift from just below to just above the values of 45% and 55% preference. Similarly to the issue of WTA versus fractional seats, sharp cutoffs like these should be avoided in quantitative estimations.

[19] A methodological improvement of the current DRA fractional seat calculation would be to use the suite of elections to estimate the variability of each district, but this is beyond the scope of this paper.

[20] It may also be noted that a party that wishes to maximize its seat share should also use fractional seats instead of winner take all because a map that has many districts slightly leaning to a party looks unrealistically too favorable to that party using winner take all. This is recognized qualitatively in some pieces in the press; fractional seats takes this into account quantitatively.

[21] This paper refers to a flawed method as proportional shift and then uses a new name for what the other two references call proportional shift.

[22] Subtleties involving different turnout in different districts is discussed in Appendix B.

Fourth, DRA provides a composite election formed by aggregating six past statewide elections. The average vote of this composite election is then an estimate of the historically most likely vote. The mP metric is then obtained as proportionality at this composite vote. However, DRA also provides individual statewide elections that this paper makes use of in order to obtain statistical uncertainties for the SB metric, as well as to focus on close elections for which all three metrics agree.[23]

3. An in-depth example

To gain insight, the 2022 congressional plan for the state of Virginia is considered in detail by applying the results of ten statewide elections to the plan's districts.[24] Examining the results from such a suite of elections provides a more comprehensive view of a plan's characteristics than using only one election or a single composite of several elections. Figure 2 shows the seats shares for the ten elections $e = 1$ to 10 versus their vote shares V_e. These vote shares span a range from 49% to 58%, so simply connecting these ten points would provide a reasonable estimate of the seats votes curve S(V) over this vote range. DRA also provides an S(V) curve from any election; the solid line in Fig. 2 is the S(V) derived from the DRA composite election. Figure 2 also shows the proportionality ideal and the 3B ideal for comparison. The DRA S(V) curve clearly has an overall slope, i.e. responsiveness to voters, closer to the 3B ideal than to proportionality. Ideal proportionality, shown by the bold dashed straight line, would switch 2% of the seats if the overall vote swings by 2%. In contrast, the seats/votes S(V) curve shown for VA in Fig. 2 would switch 5.1% of the seats when the vote swings between 52% and 54%.[25] Another major comparison between the ideal curves in Fig. 2 and the DRA S(V) curve is that the latter has fewer than 50% of the seats for 50% of the vote, indicating a seats bias in favor of the GOP for the composite election.

[23] An intriguing possibility to avoid having to shift the vote to 50% for the SB metric would be to form a composite election by adding together two or more elections weighted such that the composite vote is 50%.

[24] The VA elections were those available in the DRA database (Bradlee, 2020), president 2016 and 2020 (P16 & P20), senate 2018 and 2020 (S18 & S20), governor, lt. governor, attorney general 2017 and 2021 (G17,G21,LG17,LG21,AG17,AG21).

[25] This is even greater than the responsiveness of the ideal efficiency gap (EG) which idealizes changing 4% of the seats for a 2% vote swing, and it is closer to the 6% seat swing of the 3B ideal.

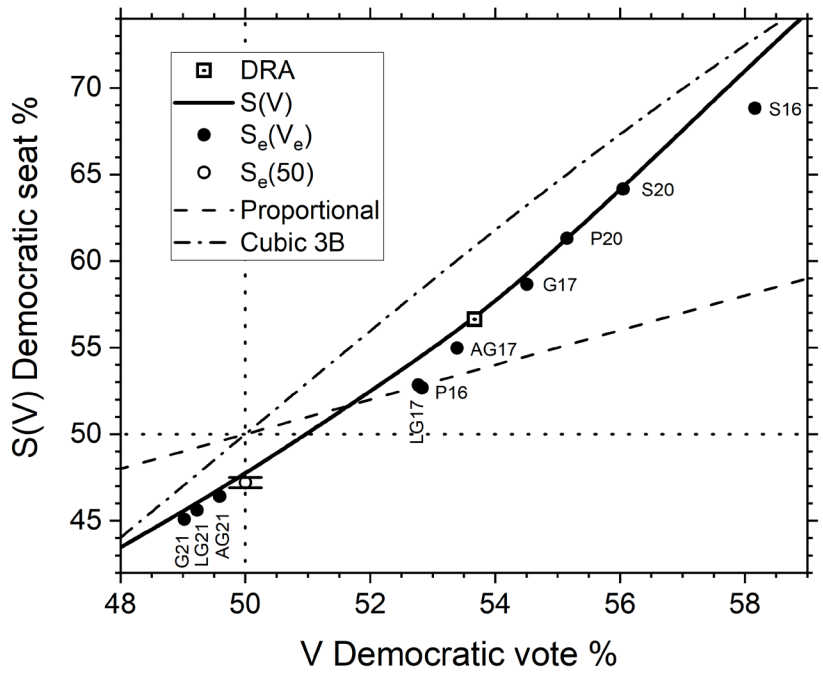

Fig. 2. The solid circles show the results $S_e(V_e)$ using ten statewide elections (e=1,…,10) for the Virginia congressional plan of 2022. The solid curve shows S(V) produced from the DRA composite election (open square). The open circle at V = 50% is the average seat % obtained from all ten elections. Proportionality is the dashed line and the dash-dot line is the 3B ideal.

Figure 3 takes a closer look at the partisan bias metrics by comparing the results for each of the elections e. Each value of P_e in Fig. 3 is the negative difference between the filled circle for election e in Fig. 2 and the proportionality line at the same V_e.[26] Similarly, each value of $3B_e$ in Fig. 3 is the negative difference between the filled circle in Fig. 2 and the cubic 3B line at the same V_e. In contrast, each SB_e value is obtained at V=50% from the $S_e(V)$ curve generated by DRA.

As would be expected, each election gives different values. However, even though there is a 9% range of votes over these ten elections, the values of SB are nearly the same with an average of 2.3% and a standard deviation of only 0.9%, which is a 0.3% 'standard error of the mean'; this is the uncertainty in the average value.[27] These results are consistent with SB being a precise metric. The $3B_e$ bias values in Fig. 3 are mostly somewhat larger than the SB_e values, with a

[26] Using V – S(V) instead of S(V) – V (Eq. 1) conforms to the DRA sign convention that the GOP is favored by a positive value and the Democrats are favored by a negative value.

[27] The size of the uncertainty is about the same in the S(V) curves at other V. Although DRA does not show uncertainties, they were shown for other states by Nagle and Ramsay (2021).

4.2% average. Importantly, there are no trends in the SB_e or the $3B_e$ values over the 9% range of vote shares. In contrast, the P_e values vary systematically from plus 4% GOP bias for the G21 election that had 49% Democratic vote to minus 11% for the S18 election that had 58% Democratic vote.

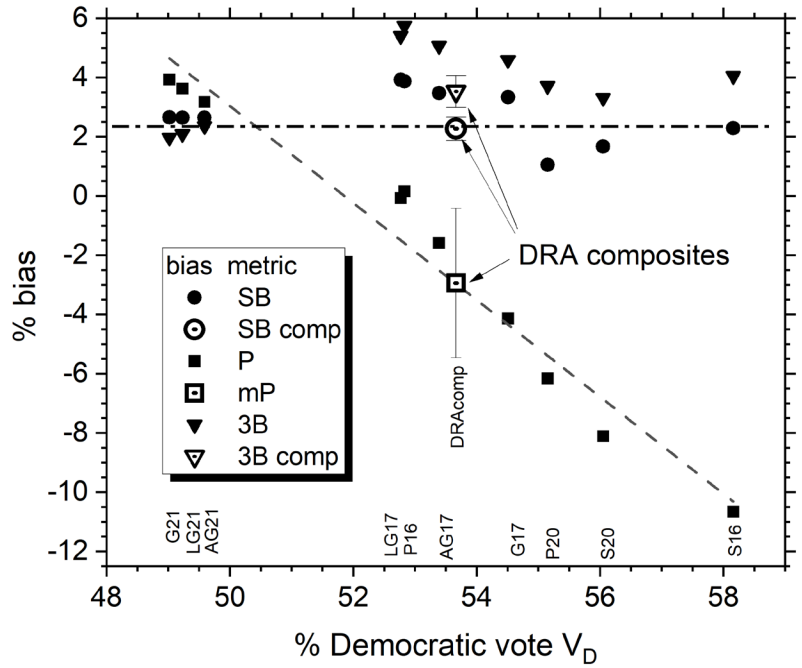

Fig. 3. Seats bias SB (circles), proportionality bias P (squares), and cubic bias 3B (downward triangles) along the vertical axis for the VA congressional plan of 2022 versus the statewide votes. The ten statewide elections are shown by closed symbols with the statewide election identified along the votes axis as in footnote 25. The three biases at the DRA composite vote have open symbols. The dashed line is a linear fit to the P data and the dash-dot line is the SB average.

Such wide and systematic variations imply that proportionality does not measure a valid bias because the bias of a plan is an intrinsic property of the plan that shouldn't vary with swings in actual elections. The mP modification alleviates this somewhat by taking an average over elections which is shown in Fig. 3 as the mP open square, but the standard error of the mean shown as the error bar on the open square is quite large, consistent with differences that would occur if different elections were chosen to compile a composite election.[28] Importantly, the mP

[28] The DRA composite only uses six of the ten elections in Fig. 2. Other analysts often use only one or two elections, usually the latest presidential elections. Using only the P20 election, mP would report even more bias favoring Democrats in Fig. 2 than the DRA composite.

values differ substantially from the SB and 3B values, even reversing the bias from favoring the GOP to favoring the Democrats.

The reason that proportionality varies so much and so consistently with vote shares is due to the VA plan having a feature that is often considered to be desirable, namely, that the plan is more responsive to voters than proportionality. When the S(V) curve is more responsive than proportionality, then that curve systematically obtains values of the proportionality metric that become more negative as V increases because it equals the difference V – S(V) between the proportionality ideal and the S(V) curve.[27] Of course, SB_e, $3B_e$ and P_e agree well for the elections that had the vote closest to 50% because P and 3B at 50% vote are identical to SB. However, the DRA composite mP value in Fig. 3 deviates from that because mP uses an historic average vote V_h, which is greater than 50%. Then, mP makes VA look biased in favor of Democrats. As we will show shortly, mP makes some other states look more biased in favor of the GOP.

4. Tabulated results for many states

Although each of the 44 states with more than one congressional district deserves an in-depth analysis like the one for Virginia in the previous section, that would require a very long paper.[29] Instead, we present a table that more succinctly provides overall perspectives as well as comparisons of the plans of the individual states. Listing the states in order of the vote in Table 1 enables seeing important correlations. Specifically, for unbalanced state that have greater than 55% Democratic two-party historic vote share, the mP metric claims that all 12 plans are more strongly biased in favor of Democrats than the SB metric, and the 3B values are generally closer to the SB values. For the 17 unbalanced states that have greater than 55% GOP vote, the mP metric has large biases in favor of the GOP, which are mostly much greater than the SB and 3B values. Importantly, this feature of the mP metric, assigning large bias to majority parties, is non-partisan.

State	#CDs	D vote	SB	mP	3B	$R_{40\text{-}60}$	R_V	R_{50}	type
HI	2	68.3	-0.1±0.0	-31.7	-9.1	4.9	0.0	9.9	B
NY	26	64.8	3.6±0.6	-20.0	1.4	2.2	2.2	1.7	A
CA	52	64.2	-2.8±0.3	-19.7	1.3	3.2	1.5	3.5	A

[29] We urge those interested in a particular state to view the DRA seats-votes curves and rank-votes graphs for different statewide elections.

MD	8	62.2	3.9±0.8	-16.9	2.6	2.5	2.7	2.3	A
RI	2	61.9	-1.6±0.1	-37.1	-17.9	4.8	0.7	7.1	B
MA	9	61.4	7.5±0.6	-34.7	-16.0	4.1	1.6	5.8	B
IL	17	58.2	-3.0±1.3	-20.1	-5.3	3.0	1.7	3.3	A
WA	10	58.2	7.4±0.7	-10.0	4.8	3.3	3.0	3.8	A
CT	5	58.1	-1.0±0.3	-34.5	-19.9	4.8	2.8	7.3	B
OR	6	57.5	2.2±0.4	-19.7	-6.0	3.3	2.5	4.2	B
NM	3	56.1	1.0±0.2	-36.2	-24.7	4.9	3.3	9.6	B
NJ	12	56.0	-4.3±0.5	-18.5	-7.2	3.3	2.7	4.5	B
CO	8	54.5	6.2±0.5	-2.1	6.6	2.8	2.6	3.0	A
MN	8	54.5	8.7±0.8	0.8	9.5	2.8	2.9	2.7	A
VA	11	53.7	2.3±0.3	-3.0	4.2	2.7	2.8	2.3	A
PA	17	52.5	3.2±0.9	-1.5	3.5	2.8	2.7	3.1	A
MI	13	51.9	4.5±0.6	-0.2	3.6	3.1	3.5	3.5	A
NV	4	51.5	-11.1±0.7	-16.7	-13.7	4.6	4.2	4.3	A
WI	8	50.7	13.9±0.9	12.7	14.1	2.9	3.1	3.6	A
ME	2	50.6	0.9±0.2	-0.2	1.0	4.4	2.8	2.7	A
NC	14	49.4	0.3±0.1	1.2	0.0	3.2	2.6	2.7	A
AZ	9	48.9	5.9±0.4	9.1	6.9	3.4	3.5	4.3	A
GA	14	48.0	13.4±0.3	13.0	9.0	2.4	0.7	1.0	C
FL	28	48.4	7.8±0.2	10.8	9.0	3.5	2.3	3.8	A
NH	2	46.5	-0.0±0.0	26.4	19.5	4.9	7.0	9.6	B
OH	15	46.4	7.9±0.5	17.6	10.5	3.5	3.4	4.2	A
TX	38	46.2	10.8±0.1	9.8	2.4	2.4	0.5	1.2	C
IA	4	45.0	-6.2±0.6	22.4	12.8	4.6	6.6	6.0	B
MT	2	43.6	1.1±0.2	27.8	15.8	4.7	4.5	5.8	B
IN	9	43.5	17.4±0.4	21.2	9.0	3.7	0.4	5.6	BC
SC	7	43.3	11.6±0.7	27.5	14.9	3.9	0.8	6.8	BC
MS	4	43.2	19.7±0.5	18.1	5.5	3.2	0.1	0.8	C
MO	8	42.8	14.9±0.6	16.6	3.3	2.7	0.6	1.8	C
KS	4	41.9	5.1±1.1	28.5	13.9	4.5	3.1	6.1	B
KY	6	40.5	12.0±0.7	19.9	3.4	3.5	1.4	4.3	B
AL	7	40.0	14.4±0.5	25.6	8.5	3.7	0.1	6.2	BC
LA	6	39.1	23.7±0.4	22.4	4.2	3.7	0.0	3.7	C
TN	9	38.7	9.7±0.9	26.4	7.8	3.9	0.6	5.3	BC
NE	3	37.5	-9.3±0.4	25.3	5.6	3.3	3.1	2.6	A
AR	4	35.2	2.8±0.4	34.7	13.3	4.7	0.3	7.4	B
ID	2	33.9	-0.5±0.1	33.8	11.8	4.8	0.0	8.3	B
OK	5	33.6	0.2±0.3	33.5	11.4	4.7	0.1	6.7	B
WV	2	33.5	0.0±0.1	33.5	11.3	4.9	0.0	10.0	B
UT	4	32.9	0.1±0.1	32.9	10.5	4.9	0.0	9.1	B
USA	429	51.9	5.1±0.5	0.9	4.1	3.2	1.9	3.7	

Table 1. Metric values for 2022 congressional plans of states. State names are abbreviated in the first column, with the number of congressional districts in the second column, arranged in decreasing order of the DRA 2016-2020 composite vote in the third column. Column 4 shows the mean values of the seats bias SB with standard errors of the mean, column 5 shows mP bias and column 6 shows 3B. The seventh column shows the broad responsiveness metric $R_{40\text{-}60}$.

Column 8 shows the DRA responsiveness R_V, which is the differential dS/dV slope of the S(V) curve, at the DRA composite V. The ninth column shows the slope R_{50} of the S(V) curve at V=50%. The letters in the final column roughly divide the states into three types defined in section 7. The last USA row is the CD weighted average of the 44 state plans.

The example of VA in the previous section suggests that high responsiveness helps to understand the problematic behavior of the mP metric. For responsiveness Table 1 begins with a metric that looks at a broad range of vote share,

$$R_{40\text{-}60} = (S(60)\text{-}S(40))/20, \qquad (6)$$

where S(60) is the seat share percentage at 60% vote and S(40) is the percentage at 40% vote. An important finding in Table 1 is that the high responsiveness of the Virginia plan is more the rule than an exception. The broad responsiveness $R_{40\text{-}60}$ is much larger for all 44 state plans than the value of 1 that proportionality idealizes. This is shown graphically in Fig. 4. The average over all states is 3.7. This is also considerably larger than the value of 2 for the ideal efficiency gap (McGhee, 2014) and it is even larger than the value of three[30] of the classical cube law (Kendall and Stuart, 1950). Smaller states tend to have larger values of $R_{40\text{-}60}$ than more populous states but even the average weighted by the number of CDs is 3.2.[31] Figure 4 also indicates that $R_{40\text{-}60}$ is about the same for deeply red GOP states as for deeply blue Democratic states and as for more balanced purple states. Therefore, the large broad $R_{40\text{-}60}$ responsiveness appears to be a feature of the single member district system independently of partisanship. Importantly, $R_{40\text{-}60}$ also appears to be independent of who draws the lines, as shown in Table 2.

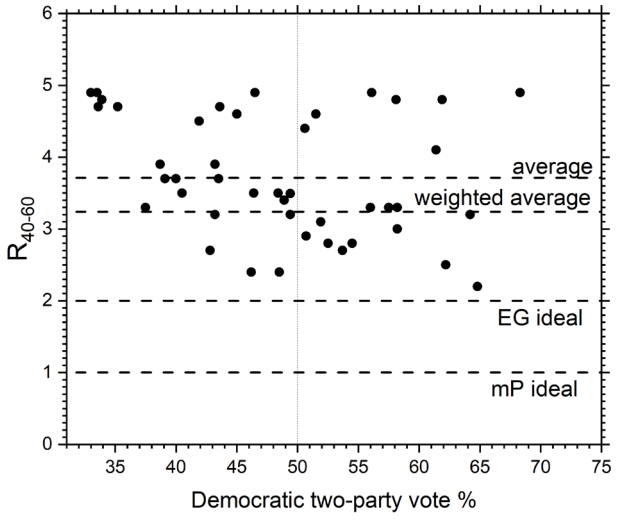

Fig. 4. Broad responsiveness $R_{40\text{-}60}$ extracted using DRA for congressional plans as of 11/2024 for all states with more than one congressional district. Seats values used to compute $R_{40\text{-}60}$ were obtained from the DRA seats-votes graphs using its 2016-2020 composite statewide data bases.

[30] Although the ideal 3B curve has a slope R_V of 3 at V = 0.5, its $R_{40\text{-}60}$ is only 2.72 as can be seen in Fig. 1,

[31] Goedert et al. (2024) have noted a responsiveness of 2.6 from simulations of the 34 most populous states.

Control	# states	average	# districts	average
Indy	8	3.5	107	3.2
GOP	18	3.8	176	3.3
Dems	8	3.7	75	3.1
Split	10	3.6	71	3.2

Table 2. Average $R_{40\text{-}60}$ for four types of commissions identified in DRA as independent commissions, GOP controlled, Democratic controlled and split control. Number of states and the averages by states are shown in columns 2 and 3, respectively and number of CDs and the averages by CDs are shown in columns 4 and 5, respectively. Further details can be found in (Magleby and McDonald, 2025a).

Because seats-votes curves are not straight lines, responsiveness can be measured at different vote shares. Table 1 also shows narrow responsiveness metrics R_V and R_{59} that will be important in the analysis in section 7.

5. Seats bias

Importantly, there is close agreement of the S(V) curves drawn from different elections, as was shown by (Nagle and Ramsay, 2021)), even when including elections with V far from 50%, as in Fig. 2. This is further quantified in Table 1 by the small ± uncertainties in SB that were determined from the $S_e(50\%)$ from the many elections *e*. Indeed, using the suite of elections in DRA we find here that the average standard error of the mean for all 2022 state plans in Table 1 is 0.45% and is only greater than 1% for IL (1.27%) and KS (1.07%). These results show that DRA methodology has overcome the criticism that the traditional SB metric is not robust.

Better yet, comparison of two plans using seats bias can be even more precise. For example, when the Pennsylvania Supreme Court called for plans to choose from, there was a major contender called the Gressman Math/Sciences (GMS) plan. The PA2022 enacted plan has a DRA composite seats bias of 3.2% in Table 1 and the GMS plan had 1.9%, only 1.3% smaller. With 0.9% SEM uncertainty from Table 1, the two plans could be judged to have insignificant difference in bias within statistical uncertainty,[32] but this ignores the correlations that occur

[32] The court did not explicitly consider partisan bias but made its decision based on county splits and not splitting the city of Pittsburgh, two criteria that make it more difficult to minimize partisan bias (Nagle, 2019). The split in 2022 was 9D and 8R, so it was widely believed that the plan was fair or even tilted toward the Democrats, even though the average D vote for governor and senator was 55%. The average of the presidential and senatorial vote in 2024 was 49.4%% and the split was 7D and 10R

within the suite of elections. Table 3 shows a better way to compare the two plans. The bias is compared election by election and then the differences are averaged over the suite of elections. This gives a difference of 1.1% bias (3.5% - 2.4%), but the uncertainty in this difference is only 0.3%, which means that the 1.3% difference in bias is statistically highly significant.[33] In passing, it would seem that this type of comparative analysis should be used for any metric.[34]

Election	PA2022	GMS	Δ
P16	4.7	2.8	1.9
S16	7.5	5.5	2.0
AG16	5.7	3.5	2.2
T16	7.2	5.5	1.7
Aud16	8.9	9.7	-0.8
S18	4.8	3.3	1.5
G18	2.8	2.1	0.6
P20	0.9	0.3	0.6
AG20	0.7	-0.7	1.3
Aud20	1.3	-0.5	1.8
T20	0.8	-0.8	1.5
S22	0.4	-0.2	0.6
G22	-0.2	0.2	-0.4
Averages	3.5±0.9	2.4±0.9	1.1±0.3
DRAcomp	3.2	1.9	1.3

Table 3. Seats bias for the PA2022 enacted plan compared to the GMS plan for the statewide elections in the first column identified by year and office (P-president; S-senate; G-governor; AG-attorney general; T-treasurer; AUD-auditor). The differences are given in the Δ column. Averages with standard error of the mean are given in the penultimate row. The last row gives the DRA composite values, which only averaged over P16, P20, S16, S18, G18 and AG20.

By arranging the elections by year, Table 3 also indicates that bias favoring the GOP in PA has been decreasing with time. This is largely uncorrelated with shifts in the overall vote and

compared to 7.6 D seats using the methods in this paper. Both 2022 and 2024 outcomes are consistent with what SB estimated.

[33] A Condorcet winner may not emerge when comparing a suite of maps, in which case any of the maps in the Condorcet indeterminate set would be suitable with regard to partisan bias.

[34] Note that this kind of detailed statistical analysis cannot be performed for mP or 3B because they give only one value of bias for each plan, so it is not possible to determine if any obtained difference is significant. However, it can be applied just to proportionality and just to the cubic ideal. For both, the mean difference between the two plans is 0.11% with a standard error of the mean of 0.53%, so, in contrast to SB, neither proportionality nor 3B detect a significant difference in these two plans. It may also be noted that for both plans the average 3B bias is near 4%, closer to the SB bias than the average proportionality bias which is near -1.5% for both plans.

instead reflects shifts in voter preference between districts.[35] A refinement would be to fit the trend with time to extrapolate the bias to the time of the first election under the plan. However, that is not necessary to compare the two plans in Table 3 because both have the same temporal trend.

6. Critique of the mP metric

An important way to evaluate bias examines elections that were close to 50% vote (Chen and Rodden, 2013), as this is firmly based on the normative principle of equal seats for equal votes. Table 4 first compares the metrics for close elections for selected states. As expected, SB, mP and 3B have nearly the same values for close elections because these metrics are identical by their definitions when V = 50%. The second row for each state shows the value for the composite election, so the 'Prop' entry in the 'comp' row is the value of mP. The SB and 3B values agree reasonably well with each other and also with their values for the close elections. The outliers are the mP values in the comp rows and the Prop column. The mP value purports that the plan favors the majority party more than what is obtained by SB and by 3B for both elections and by proportionality at the close election. In NC, AZ, and FL, mP would blame the GOP for having too much bias in its favor. In the other states mP would tilt the bias toward the Democrats, erroneously attributing Democratic bias in VA, PA, and MI and attributing too little GOP bias in MN. These flawed mP results are accounted for by the seats-votes curve being considerably more responsive than proportionality as shown by the final column in Table 4.

State	election	V	SB	Prop mP	3B	R_V
MN	Aud22	50.2	7.1	6.9	7.3	2.2
MN	comp	54.5	8.7	0.8	9.5	2.9
VA	AG21	49.6	2.7	3.2	2.4	2.3
VA	comp	53.7	2.3	-3.0	4.2	2.8
PA	P16	49.6	4.7	5.4	4.6	2.8
PA	comp	52.5	3.2	-1.5	3.5	2.7
MI	P16	49.9	5.9	6.2	6.0	3.3

[35] There is a similar trend in some other states, but not all; a detailed analysis of the durability of bias is beyond the scope of this paper.

MI	comp	51.9	4.5	-0.2	3.6	3.5
NC	AG20	50.1	0.2	0	0.2	2.6
NC	comp	49.4	O.3	1.2	0.0	2.6
AZ	P20	50.2	3.9	3.5	3.9	3.8
AZ	comp	48.9	5.9	9.1	6.9	3.5
FL	S18	49.9	7.1	7.3	7.1	2.9
FL	comp	48.4	7.8	10.8	7.6	2.3

Table 4. Comparison of metrics obtained from close elections for the 2022 congressional plans with those of the DRA 2016-2020 composite elections. There are two rows for each state identified in the first column. The second column identifies the election used, P (president), G (governor), S (senate), AG (attorney general) or AU (auditor) and the year for the close elections or comp for the DRA composite. Subsequent columns give the vote V and the seats bias SB. The column labelled Prop gives mP for the composite election and simple proportionality for the close election. The penultimate column gives the cubic bias 3B and the final column gives responsiveness R_V at the vote V.

Turning from the nearly balanced states in Table 4, it has been emphasized in section 4 that modified proportionality reports numbers for bias that tend to blame the majority party for partisan bias. Naively, this could be attributed to majority parties always intentionally gerrymandering, but the analysis in this paper ascribes much of it to responsiveness being generally greater than proportionality in the single member district system.

Another problem with the mP metric is that it depends so sensitively on how one decides what is the historic vote. Although we have chosen this to be the DRA composite vote for the period 2016-2020, one could choose different composites or even just one favorite election, so this makes mP an intrinsically non-robust metric.[36]

Furthermore, the mP metric tends to lead to anti-majoritarian outcomes that favor one party.[37] To see why, consider Fig. 5. This figure shows S(V) for two hypothetical plans that both have responsiveness $R_V = 3$. Let's consider what occurs when the average state vote is 53%.[38] The plan shown by the dash-dot line has the same seat % as proportionality at V = 53%, so its

[36] The example in footnote 28 gives differences of 3%. Also, the 3B metric is not so sensitive to this because responsiveness is much closer to the cubic ideal.

[37] Less than half the seats for more than half the two-party vote defines anti-majoritarian, which is universally agreed to vitiate democracy.

[38] States in Table 1 with average vote for the majority party in the range 52-55% are CO, MN, VA, PA, GA, FL, NH, OH, TX and IA.

mP value of zero declares it to be fair. However, vote swings of 5% are normal. When the vote swings to between 50% and 52%, there is an anti-majoritarian range, in which the vote percentage exceeds 50% but the seat percentage is less than 50%. Importantly, the vote V in Fig. 5 can be either V_{DEM} or V_{GOP}, so this type of anti-majoritarianism can occur against either party. Therefore, this flaw in the mP metric is non-partisan. Table 1 indicates that states where this flaw is most likely to occur are CO, MN, VA, PA and MI because these states all have votes that differ a few percent from 50% and values of mP and SB that have similar values as in Fig. 5.[39]

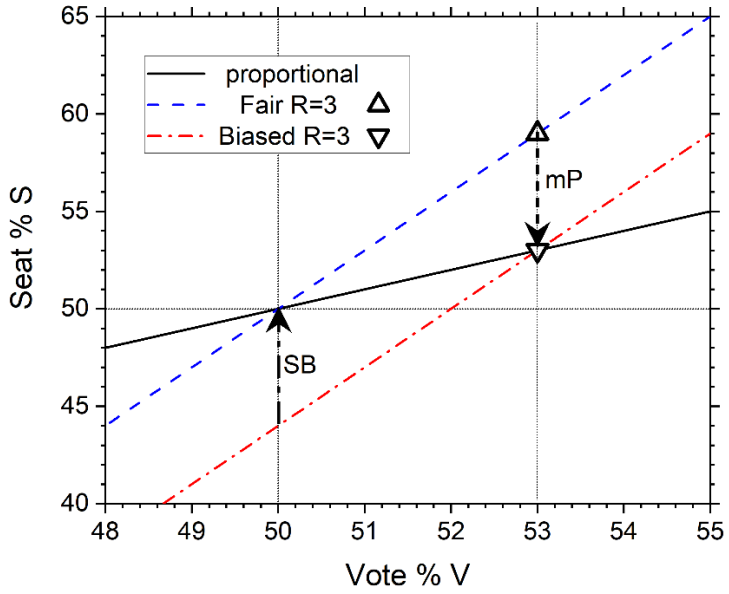

Fig. 5. Illustration of anti-majoritarianism when the mP metric claims fairness. The seats-votes curve shown by the dot-dash line is deemed fair by mP when the vote is 53%, but it becomes anti-majoritarian when the vote swings to between 50 and 52% vote, Its +6% seats bias is shown by the vertical SB arrow. The seats-votes curve shown by the dashed line has 50% seats for 50% vote but it is deemed by the mP metric to be biased by - 6% as shown by the arrow labelled mP.

Interestingly, if the group of states with party X majority try to minimize the mP value in their plans and the states with party Y majority try to minimize the SB value, then party X will have put itself at a disadvantage for control of Congress. For the example in Fig. 5, this disadvantage is the seats bias SB. Reformers who like proportionality might keep this in mind.

[39] Of course, anti-majoritarian outcomes are even more likely to occur in balanced states when mP is not small and SB is not zero. In Fig. 5 this occurs for the biased dash-dot S(V) function when the average minority party vote is less than 50% and the swing vote takes it to between 50% and 52%. States most likely to have anti-majoritarian tendencies favoring the GOP are WI, NC, AZ, GA, FL, OH and TX, while plans in NV and IA have Democratic anti-majoritarian tendencies.

7. Responsiveness and Types of Plans

This section will continue to emphasize the importance of responsiveness in assessing partisan fairness. However, it now becomes appropriate to consider narrow responsiveness which is the slope dS/dV in the seats-votes S(V) curve because the differences in votes are much smaller than the broad 20% difference in the $R_{40\text{-}60}$ responsiveness metric. Of course, the slope of the S(V) curve depends upon the vote V where the slope is evaluated, so it is valuable to consider more than one vote share; we consider R_V, which is the slope of the S(V) curve right at the historic V_h, and R_{50}, which is the slope at V=50%. These new responsiveness metrics will be instrumental in classifying the different types of plans.

State plans will be classified as types A, B or C depending upon their responsiveness. The states in Table 4 provide examples of type A plans. Other type A states that did not have close elections are listed in Table 1. They have values of responsiveness between 1.7 and 4 as shown in the last column of Table 4 and in the R_{50} column of Table 1. The values of R_V persist in type A plans over large ranges of V in the S(V) curve until the curve eventually flattens (R_V goes towards zero) as V approaches either 0 or 100% as it always does for plans in the single member district system; this feature is illustrated by VA in Fig. 6. It is the claim of this paper that the SB metric is valid for type A plans.

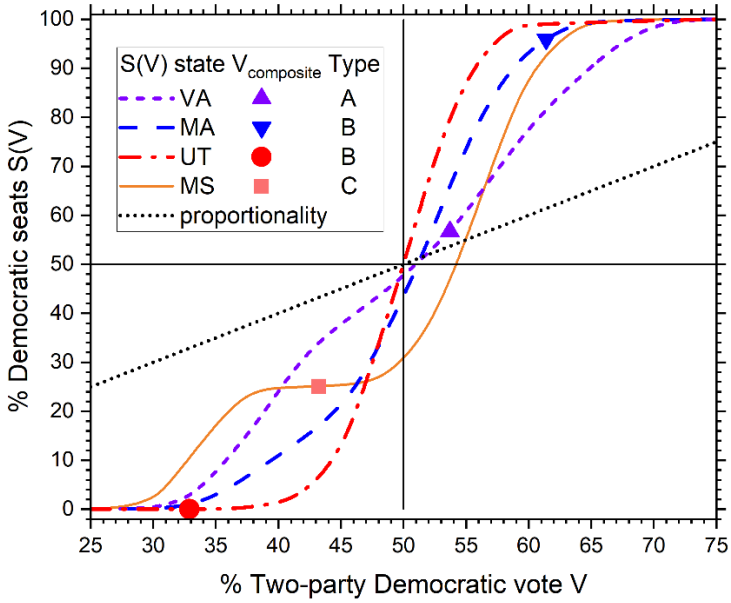

Fig. 6. Seats-votes curves. VA and MS represent types A and C, respectively. MA and UT represent type B with opposite dominant parties. The solid symbols show the historic DRA composite votes V_h.

Plans of B type have two distinguishing features, both of which are necessary for a type B designation. One is that they are unbalanced as roughly indicated by a composite vote less than 45% or more than 55%. The second is a high value of responsiveness R_{50}, taken to be greater than 4 for the type B assignments shown in Table 1.[40] This combination raises an important concern. In a 60-40 vote share plan, the majority party would obtain essentially all the seats if it drew a map with 60-40 voter preferences in all districts. The S(V) curves for two type B states, MA and UT are shown in Fig. 6.

The 2022 congressional Utah plan is an especially interesting type B example. By cracking Democratically leaning Salt Lake county into all of Utah's four districts in the 2022 plan, all districts had between 63% and 71 % GOP preference, virtually guaranteeing all four seats for the GOP. In this case cracking is not detected by the SB metric whose value of 0.1 in Table 1 declares this plan to be essentially fair; this comes about because shifting the vote to 50% makes all the districts competitive. However, cracking is detected in Utah by the very high value of 9 for the R_{50} metric. This responsiveness metric reports that the voter preference in all the districts has been made relatively uniform; that makes each district highly likely to vote for the GOP at the historic UT Democratic vote share of 32.9%.

Further failure of the SB metric that can occur when the vote is unbalanced statewide is revealed by the UT 2026 plan. District 1 is entirely within Salt Lake County and has a composite preference[41] of 55.2% that DRA estimates as 0.9 Democratic seats. In contrast to the 2022 plan which had SB = 0, SB is 17.8% favoring the GOP even though it would not get as many seats as the 2022 plan. The reason for this disparity is that shifting the vote to 50% makes district 1 highly packed with 72% Democratic voter preference while the other three districts remain highly GOP preferred. This would-be packing gerrymander is a caution that shifting the vote to obtain the SB metric, while accurate, can be misleading.

[40] Table 1 lists 21 type B plans. Although this is almost half the 44 states with more than one district, most type B states have fewer than the average number of ten districts, so only 25% of the congressional districts are in type B plans.

[41] Using the DRA 2012-2020 composite as in Table 1 for UT.

It therefore behooves us to consider the mP and 3B metrics for Utah because they do not shift the vote. The distinction between these two is whether proportional or cubic responsiveness is more appropriate for Utah. More generally, one would also consider other levels of responsiveness than the level of 1 for mP and the level of 3 for 3B (Tufte, 1973. Levels of responsiveness can be obtained by looking at state house and senate plans; unfortunately, the DRA rank-votes graphs for Utah show fewer competitive districts near the composite vote than in other vote ranges; that is consistent with drawing safe districts for incumbents, and that reduces R_V. However, other vote ranges are consistent with values of responsiveness close to 3 for both the Utah state house and the state senate. This favors using the 3B metric rather than the mP metric. The 3B metric gives 10.5% bias for the 2022 plan (Table 1), which is a substantial GOP bias that makes more sense than either the flawed SB or mP metrics.[42]

An equally interesting type B plan is the Massachusetts 2022 plan. Figure 6 shows that its S(V) curve is also considerably steeper than the VA type A curve. As already mentioned in the introduction, the relative geographical political homogeneity of MA prevents drawing any district that leans to the GOP at the most likely statewide vote, so it can hardly be designated as intentionally biased against the GOP, in agreement with the rather different analysis of the similar 2012 MA plan (McDonald et al., 2018). Interestingly, the MA map has a seats bias of 7.5% favoring the GOP. This comes about because district MA7 is packed with 82% Democrats. Importantly, it is consistent that seven of the nine MA districts favor the GOP for the close V=49% G14 election.[43] It is also appropriate to consider the 3B metric for Massachusetts. That declares MA to be biased by – 16%, strongly favoring Democrats. But this assumes the cubic level of responsiveness whereas the introduction notes that an even higher level of responsiveness would be expected for a state that has voter preferences that are relatively the same geographically. Replacing the 3 in Eq. 4 by a 6 yields a curve that would declare no bias in MA when substituted into Eq. 5.

[42] Interestingly, 3B for the 2026 plan has a bias -12,0% that just as substantially favors the Democrats as the 2022 plan favors the GOP. To be 3B fair, district 1 would need to be drawn more competitively with 49.3% Democratic voter preference, and that would substantially increase responsiveness at the composite vote. It is easy to draw such a district entirely within Salt Lake county.

[43] This G14 election gives a rank-votes graph (see DRA advanced section) similar to the 2026 Utah plan when both are shifted to 50% vote.

The UT and MA cases emphasize that appropriate auxiliary tests should be applied to type B states that are flagged by a very high responsiveness R_{50} at V = 50%.

Some state plans are classified as type C in Table 5. These plans have small responsiveness R_V at the historic (DRA composite) vote, but that alone does not make them type C. As illustrated in Fig. 6, S(V) curves in the single member district system flatten out at both high and low values of V as seats S approaches either of its extreme values of 0 or 100%, and this automatically gives a small value of R_V if the composite vote places S in one of those two regions. For example, the UT composite 32% vote places V where S(V) is flattened near 0 in Fig. 6. In contrast, S(V) for the composite vote for Mississippi in Fig. 6 has an additional flattening between those at the low and high vote regions. We will classify state plans that have S(V) curves like the MS curve in Fig. 6 as type C. These S(V) curves look quite different from those of type A, represented in Fig. 6 by VA, and from those of type B, represented in Fig. 6 by UT and MA. Some type C unbalanced states are exceptional in having a value of SB nearly the same or even a bit larger than mP in Table 1. This comes about for MS in Fig. 6 because the flattening of the S(V) curve for vote between the composite vote and 50% makes the difference in seats between the composite vote and the 50% vote smaller than the corresponding proportionality difference. Importantly, curve flattening with small responsiveness means that the districts are less competitive and safer for both parties at the historically most likely vote share, so type C plans are additionally undesirable compared to type A plans that may be strongly biased, but are still responsive.[44] Some plans like AL and SC are typed as both B and C in Table 1 because they have both high R_{50} and low R_V.

8. Discussion

The firmest normative principle for partisan fairness is equal seat share for equal vote share. Furthermore, all three metrics considered in this paper agree on equal seats for equal votes. Accordingly, evaluating plans using elections that had close to 50% two-party vote share should play a central role in evaluating the bias of a plan (Chen and Rodden, 2013). Table 4 lists recent close elections and confirms that seats bias, proportionality, and cubic bias have nearly the same

[44] A final possible type of S(V) curve would be both unresponsive and have small seats bias. This does not occur in any current congressional plans, but the NY legislature 2022 plan rejected by the court *vide infra* leans in this direction.

values of partisan bias. We believe that these normatively supported values are close to the true partisan bias for the state plans in Table 4.

However, many states have not had close elections and then it is necessary to use elections that are not close. Even when there is a close election, the mP and 3B metrics require using elections with a more likely two-party vote share that can deviate considerably from 50%. Then, the choice of the partisan bias metric becomes important because even the three considered in this paper do not generally give the same bias value when the vote share is not exactly 50%, and the differences grow as the vote share difference grows. Table 4 shows values of bias using the DRA composite/historic vote share. Importantly, the close elections in Table 4 provide a crucial test for the three metrics considered in this paper. A valid metric should have nearly the same value of bias for a plan as the value that is obtained from a close election. Table 4 shows that the SB and the 3B metrics pass this test and the mP metric fails it. To rationalize the mP metric would require asserting that bias depends on vote share, but bias should be an innate property of a plan and not be contingent on swings in the vote.[45]

The failure of the mP metric is simply related to responsiveness being considerably greater than proportionality, and this is the reason that the 3B metric is usually better than mP. The SB metric utilizes the full seats-votes curve, which includes detailed responsiveness, to extrapolate to the 50% two-party vote share. Importantly, the bias values obtained from the SB metric are nearly the same for elections with different overall vote shares, consistent with SB measuring an innate property of the plan. Also, the use of many elections allows the calculation of estimated uncertainty in the seats bias metric. However, the Utah example in section 8 shows that SB fails to detect cracking in strongly unbalanced states, so 3B is a better metric than SB for type B plans.

Of course, each state is different with its own political geography (Rodden and Weighill, 2022), and a redistricting commission might look at other partisan bias metrics not focused on in this paper and, especially, at the rank and seats figures in DRA. Nevertheless, groupings into types may be helpful to discern broadly common feature*s*. In particular, the separation of some of the plans into types B and C emphasizes the importance of responsiveness in addition to

[45] *A priori*, any statewide election should be equally applicable for evaluating a plan, although local distortions could appear that are correlated to where the candidates live.

metrics of partisan bias. The trade-off between these two criteria is illustrated by comparing metrics for the current court enabled NY plan and the rejected legislative plan that have been carefully discussed (Magleby and McDonald, 2025b). Using the DRA 2016-2020 composite, the 2022 enacted plan had more seats bias, 2.9% versus 1.5% for a statistically significant difference of 0.4 seats. However, the court was more concerned with responsiveness. That was only $R_V = 1.0$ in the rejected plan (same as proportionality, interestingly) whereas R_V increased to 2.2 in the court enabled plan.

While responsiveness is important, it is germane to consider different metrics for it. The broad responsiveness $R_{40\text{-}60}$ is useful to illustrate that the USA single member district system is far more responsive than the normative proportionality ideal of R = 1, whoever draws the maps. The narrow responsiveness R_V helps diagnose unresponsive plans with mostly safe districts, the type C states in Table 1. A large value of the narrow R_{50} helps diagnose whether cracking voters of a minority party might be occurring in unbalanced state plans of type B.

Figure 7 rearranges the states into types to facilitate comparison of the values of the SB, mP and 3B metrics of partisan bias. Starting with type A states all metrics agree for the highly balanced plans (WI, ME and NC), as they must. With the exception of Nebraska, SB and 3B values are also largely in agreement for the other type A states in contrast to the mP values. The mP differences are generally larger the farther the average vote differs from 50%. We explain the difference as due to responsiveness being significantly greater than proportionality. When SB and 3B claim that a state is biased in favor of the dominant party, mP claims it is even more biased, as for the three GOP leaning type A states and the five most unbalanced DEM type A states in Fig. 7. For the less unbalanced states CO, MN, VA, PA and MI, mP claims the plans are nearly fair but SB and 3B claim GOP bias.

Measuring partisan bias is considerably more complex for type B state plans, as is indicated by the disagreement of all three metrics in Fig. 7. However, the extreme mP values are easily discounted due to high responsiveness, and the SB values do not detect cracking and tend to give values that are too fair. The concern with the 3B metric is that it assumes an ideal level of responsiveness, but that level increases for states, like MA, that have a more homogeneous distribution of voters. The analysis of the UT and MA cases in section 7 emphasizes that

appropriate auxiliary tests should be applied to type B states that are flagged by high responsiveness that could be due to cracking (UT), but not necessarily (MA).

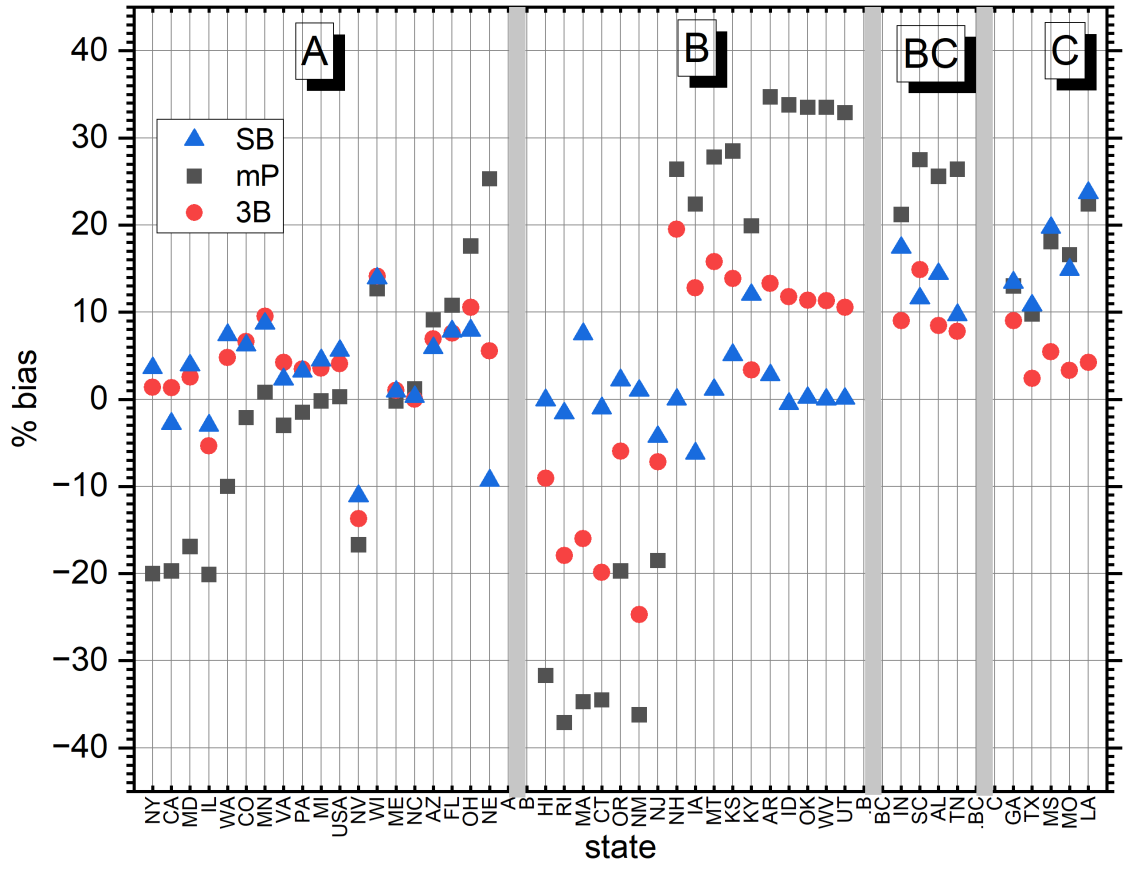

Fig. 7. State plans are grouped into types A, B, BC and C on the horizontal axis. Within each group the Democratic vote decreases from left to right. The vertical axis shows % bias for SB (triangles), mP (squares) and 3B (circles),

Figure 7 has a separate section for four unbalanced states whose plans have both type B and type C character. All metrics agree that these plans are biased in favor of the GOP. As usual, the mP metric assigns more bias than SB or CB. In contrast, for the five type C states in Fig. 7, SB and mP values agree, whereas the 3B values are lower. More important than partisan bias for these plans is their undesirable feature of very low R_V responsiveness, indicative of incumbency protection.

Turning to the overall US Congress, the last row in Table 1 averages the metrics for the congressional districts of the 44 states that have more than one CD. The average $R_V = 1.9$ is smaller than the broad $R_{40\text{-}60}$ responsiveness because the vote in highly unbalanced states, like MA and UT or in type C states, locates their R_V on flattened places in their S(V) curves; the

ensuing small R_V values lower the national average. The average R_{50} is larger than $R_{40\text{-}60}$ because the S(V) curves are generally steeper in their middle. Most interesting is that mP only has 0.9% national bias whereas the national seats bias metric SB has 5.1% and 3B has 4.1%. The SB value yields a GOP average of 0.551x429 = 236.2 seats in these 429 districts when the national vote is 50%, to which 3.8 seats could be added for the six states with only one district, for a total of 240 GOP seats and 195 Democratic seats. The number of Democratic seats rises to 211 for their national two-party vote of 51.9% in Table 1 using the national average responsiveness $R_V = 1.9$.[46] It would appear from these SB results that Democrats would have needed their national average vote to have increased to 52.8% to have obtained control of the House.[47] Importantly, bias estimates are not predictions of outcomes because they only take into account the voter preference of districts from statewide elections and do not take into account district incumbency and amount of campaign finances that are specific to actual elections; however, these should not be considered in the map drawing process even if they could be. Nevertheless, it is encouraging for the methodology in this paper that the estimate of 211 DEM seats for the 2022 plans evaluated with the 2016-2020 election data is so close to the actual 213 DEM seat outcome of 2022. This is consistent with specific contingencies averaging out in the 429 Congressional districts.

One might suppose that even a 10% seats bias is insignificant in a state with only 5 districts on the grounds that this produces a 3 to 2 seat split, only different by one seat and a frequent estimated outcome with 0% bias. The fallacy of such thinking is that small seat differences in each state add up to the significant differences for congress obtained in the previous paragraph.

Let us finally return to considering normative principles. The normative principle of equal seats for equal votes underlying the seats bias metric seems unassailable but it becomes problematic to apply it in strongly unbalanced states that have no elections with nearly equal votes. The introduction describes a possible normative principle behind proportionality as empowering voters of like mind equally. However, in the highly polarized US two party system,

[46] This is nearly the same as what Goedert (2014) reported for the previous decade.

[47] Note that this two-party national vote is district averaged. This deliberately ignores different voter turnout in different districts, quantified as turnout bias (McDonald, 2009), as irrelevant for determining seats, so this national vote is different from typical national vote tallies that would be used in the National Popular Vote for president proposal.

effective empowerment of voters of like mind is not the seat percentage divided by the voter percentage. Instead, actual empowerment in the US congress becomes rather more like nearly complete for voters of the party in majority and closer to zero for voters in minority. Therefore, there does not seem to be a valid normative principle that underlies proportionality. The cubic 3B metric is not based on a normative principle. It simply assumes a particular level of responsiveness three times greater than proportionality; that level has overall empirical support, but it should be fine-tuned for each state, but that is beyond the scope of this paper.[48]

9. Recommendations for Election Law

While we believe that partisan fairness should be an important feature in states' redistricting laws, such laws should be sufficiently general to allow a redistricting commission to adopt the metrics that are pertinent to the contingencies of its state.[49] This approach follows from the analysis in this paper which affirms that there is no one-size fits-all metric of partisan bias. Importantly, our recommendation does not preclude consideration of future improved metrics for determining bias.[47] But in any case, proportionality is especially to be avoided in law,[50] because responsiveness in the single member district system is far from proportional and this leads to distortion even when a state's partisan balance is only a few percent from even. The traditional seats bias is more reliable for many of the more populous states, and it is a very precise metric for comparing plans, but it should be used with caution in highly unbalanced states where the cubic bias or some variant of it may be more appropriate. Furthermore, responsiveness (*aka* competitiveness) should be an inherent part of the process of adopting redistricting plans. However, when there has been a close statewide election, that alone provides a very strong and direct assessment of partisan bias,[51] so a redistricting body might primarily consider that, thereby simplifying its deliberations.

10. Conclusions

[48] An improvement could estimate each state's intrinsic responsiveness and use it rather than the number 3 in Eq. 4 to calculate bias in Eq. 5. An ensemble of simulated plans could be used to estimate intrinsic responsiveness.

[49] For example, Pennsylvania Senate Bill 131 Session 2025 introduced March 26, 2025 § 19 (a)(3) "A redistricting plan may not provide an advantage to any political party. An advantage to a political party shall be determined by using accepted measures of partisan fairness."

[50] For example, the previous 2024 PA House Bill 1776 prohibited "disproportionate advantage", which could have been construed as requiring proportionality.

[51] Even if there isn't a close election, footnote 23 suggests a way to create a close election when there is an election that has more than 50% vote and one that has less than 50% vote.

1. Equal seats for equal votes is a firm normative principle for partisan fairness.
2. Imbalance of seats estimated from close statewide elections, if available, provide the best estimate of partisan bias.
3. mP is a poor metric that overly assigns bias to majority parties.
4. SB is a good metric for type A plans.
5. 3B is a better metric for some type B plans.
6. Type C plans should be rejected due to low R_V responsiveness.
7, High R_{50} responsiveness identifies plans that may have cracked voters.
8. The single member district system is incompatible with simple proportionality.
9. Assigning partisan bias to plans in unbalanced states remains challenging and requires careful attention to the contingencies of each state.

Appendix A: Critique of Winner-Take-All estimates of bias

Preliminarily, it is necessary to clarify the conflicting use of common terms. In this paper partisan bias is used to refer to the general thing to be measured, recognizing that there are different metrics to do so. However, (Katz et al., 2020) used the term partisan bias (PB) for one specific metric that looks at the difference between the number of seats that party A obtains at its vote V and the number of seats that party B obtains when its vote is V.[52] Although PB is equal to SB when the vote is 50%, they generally diverge when it is not. A recent paper (DeFord and Veomett, 2025) in these pages referred to (Katz et al., 2020) and used the term PB, but they actually calculated SB.[53] There was, however, a major difference from DRA in that (DeFord and Veomett, 2025) used winner-take-all (WTA) to calculate the number of seats, so that will be called SB_{WTA} here.[54]

Deford and Veomett (2025) strongly criticized the seats bias metric using toy examples. Here we use those same examples to illustrate the importance of assigning fractional seats to competitive districts rather than by WTA, which assigns 0 or 1 seats depending upon whether the

[52] PB is often called the symmetry metric and is sometimes notated as β in the literature.
[53] The percentage of districts with votes greater than the average vote in the definition of PB (DeFord and Veomett, 2025,6) is the same as S(50) in Eq. 1 upon shifting the vote, and the percentage with votes less than the average vote is 100 – S(V). The equivalence is also apparent in their Fig. 1.
[54] There is also the minor difference that DRA uses proportional shift. The discussion in this appendix uses the uniform shift used by (DeFord and Veomett, 2025).

two-party vote is less than or greater than 50%. Table A1 displays those toy examples although here the district preferences are uniformly shifted to 50%, because that is where the seats bias metric is calculated.[55] Comparing the percentages in the last two columns illustrates how much difference there is using WTA versus fractional seats. This raises the concern that the SB_{WTA} values are artificially large because competitive districts are not properly taken into account. As noted in the main text, it makes no sense to assign a district with 50+x% partisan preference entirely to one party when x is small. Specifically, case 1 has 9 party A seats using WTA but this does not take into account that all those districts are competitive.[56] In contrast DRA realistically estimates a smaller fraction of these 9 seats for party A.

Districts	1	2-5	6-9	10	SB_{WTA}	SB
Case 1	37	51	51	55	40	6.8
Case 2	21	51	55	55	40	18.6
Case 3	13	53	55	55	40	25.7
Case 4	27	51	51	65	40	6.6
Case 5	39	49	49	69	-40	-6.6
Case 6	48.8	48.8*	50.8	50.8	10	0.0

Table A1: Partisan preference percentages for party A at 50% statewide average for six toy models, each with ten districts, that were considered by (DeFord and Veomett, 2025) as described in the text. The SB_{WTA} column gives the (DeFord and Veomett, 2025) percentages that used WTA and the SB columns shows the values calculated as in DRA. *Note that for case 6 district 5 has 50.8 preference.

Cases 1-3 in Table A1 are from Table 1 in (DeFord and Veomett, 2025), where those cases are described as the results of different elections. If so, then the large differences in the values of SB would contradict the results in the present paper that obtain closely similar values of SB for

[55] Compared to (Deford and Veomett, 2025), cases 1-3 are shifted by +2%, cases 4 and 5 by -10% and case 6 by -0.2%.

[56] It's not necessary to consider fractional seats when there are the same effective number for each party, but overall balance does not generally occur in real maps and the toy examples appear to have been chosen not only not to have this redeeming feature, but to exacerbate the opposite.

suites of actual elections. However, those three toy elections are highly unlikely to be mutually representative of the distribution of actual elections. It is highly unlikely, given the case 1 election, that any subsequent election, such as case 2, would swing district 1 by 16% from its case 1 value while districts 6-9 all shift by 4% in the opposite direction. Somewhat differently stated, at least two of these elections would be drawn from highly improbable tails of the universe of elections in which only the vote region close to one of the elections would be probable. Basically, the concern in this paragraph is that one should not draw general inferences by selectively picking outliers.

Nevertheless, it is interesting to consider these examples, not as different elections, but as coming from different maps with the same election. Case 1 is clearly biased in favor of party A because district 1 is packed with the other party. Even though party A has no completely safe district, the preferences in the other nine districts provide it with more than half the seats. Case 2 packs district 1 even more, thereby providing more preference in districts 6-9, so SB is understandably greater than for case 1. Case 3 continues this packing process with the expected further increase in SB. In contrast, there is no difference in the SB_{WTA} values for these three cases. While this supports the claim of (DeFord and Veomett, 2025) that SB_{WTA} is a poor metric, the failure is due to using WTA rather than the basic concept of seats bias.

Cases 4 and 5 come from Table 1 of (DeFord and Veomett, 2025). Districts 1 and 10 are safe districts with fractional seats 0 and 1 respectively both before and after shifting, so differences come from districts 2-9. The signs of SB and SB_{WTA} are consistent with the different leanings of those districts from case 4 to case 5, but SB_{WTA} unrealistically assumes that estimates of preferences guarantee outcomes in competitive districts. These two cases were presented (DeFord and Veomett, 2025) as examples for how SB_{WTA} can give unreliable results for an unbalanced state with 60% overall preference for party A. For 60% vote both SB and SB_{WTA} give 9 seats to party A for both case 4 and case 5, so, even though the SB values are much smaller, their opposite sign does indicate a failure of the SB metric. Importantly, the ρ_{50} responsivity for both cases is a high 7.9 so this would be a type B state which the present paper agrees is problematic for the SB metric.[57]

[57] The 3B metric favors case 5 over case 4 but declares both to be biased in favor of party A.

Case 6 comes from the example that has vote share 50.2% in Table 3 of (DeFord and Veomett, 2025). At that vote share districts 1-4 have party A preference 49% and districts 5-10 have preference 51% so WTA gives 6 seats to party A, and shifting down to 50% overall preference doesn't change that as shown in Table A1. However, the shift gives less preference to party A in the six districts 5-10 and more preference to the other party in the four districts 1-4; this small shift happens to result in very close to 5 seats for each party, so SB is very close to zero instead of the 10% obtained by SB_{WTA}. Furthermore, at 50.2% overall vote, there are only 5.2% fractional seats[58] for party A because all the districts are quite competitive. Case 6 also has a high responsiveness of 9.7, but this wouldn't raise the concern in the text because, at 50.2% overall preference, this would not be a type B state.

We agree with (DeFord and Veomett, 2025) that SB_{WTA} is a poor and imprecise metric for partisan bias, but we assert that SB is a precise metric that is good for type A states because it takes competitive districts into account.

Appendix B. Two options for treating turnout bias.

To draw the S(V) curve, DRA2020 shifts from the statewide two-party vote V. An alternative would be to shift from the average of the districts' two party vote V_{alt}. Turnout Bias occurs when these votes differ, $TB = V - V_{alt}$. Most often, TB is negative which corresponds to more voters in districts with smaller two-party vote, *i.e.* greater turnout in GOP districts. Then, using V, whether using proportional shift or uniform shift, SB appears to favor the Democrats more than if one used V_{alt}. The case of states with only two districts illustrates this difference. For example, RI in Table 1 has SB = - 1.6% from DRA. This comes about because district 1 has fewer voters than district 2 and v_1 is greater than v_2, so after shifting to a statewide V'=50, v_1' is more favorable to Democrats than v_2' is to the GOP (i.e., $| v_1'-50|$ is greater than $| v_2'-50|$), so SB is negative, favoring Democrats. The alternative of using V_{alt} gives SB = 0 ($|v_1'-50| = | v_2'-50|$). It can be argued that the alternative is better because it conforms to the normative ideal of equal representation for equal populations that is diluted by turnout bias, which, for example, can reflect disparities in voting age population.

[58] This is not 5.0% because of the asterisk regarding district 5 in Table A1.

EndNote placeholder for footnote references. (Rodden, 2019),(McDonald and Best, 2015),(Warrington, 2018),(Keena et al., 2021),(Farrell, 2011),(Goedert et al., 2024), (Goedert, 2014), (King and Browning, 1987), (McDonald, 2009),(Barton, 2022),(Barton and Eguia, 2024), (Kendall and Stuart, 1950), (Gelman and King, 1994),(Gelman et al., 2012), (DeFord et al., 2022)

References

BARTON, J. T. 2022. Fairness in plurality systems with implications for detecting partisan gerrymandering. *Mathematical Social Sciences,* 117**,** 69-90.

BARTON, J. T. & EGUIA, J. X. 2024. A decomposition of partisan advantage in electoral district maps. *Electoral Studies,* 92**,** 102871.

BRADLEE, D. 2020. *Dave's Redistricting* [Online]. Available: https://davesredistricting.org [Accessed].

CERVAS, J., GROFMAN, B. & MATSUDA, S. 2022. The Role of State Courts in Constraining Partisan Gerrymandering in Congressional Elections. *UNHL Rev.,* 21**,** 421.

CHEN, J. W. & RODDEN, J. 2013. Unintentional Gerrymandering: Political Geography and Electoral Bias in Legislatures. *Quarterly Journal of Political Science,* 8**,** 239-269.

DEFORD, D. & VEOMETT, E. 2025. Bounds and bugs: The limits of symmetry metrics to detect partisan gerrymandering. *Election Law Journal: Rules, Politics, and Policy,* ?**,** 1-27.

DEFORD, D. R., EUBANK, N. & RODDEN, J. 2022. Partisan Dislocation: A Precinct-Level Measure of Representation and Gerrymandering. *Political Analysis,* 30**,** 403-425.

DUCHIN, M., GLADKOVA, T., HENNINGER-VOSS, E., KLINGENSMITH, B., NEWMAN, H. & WHEELEN, H. 2019. Locating the Representational Baseline: Republicans in Massachusetts. *Election Law Journal,* 18**,** 388-401.

DUCHIN, M. & SCHOENBACH, G. Redistricting for proportionality. The Forum, 2023. De Gruyter, 371-393.

FARRELL, D. M. 2011. *Electoral Systems: A Comparative Introduction,,* London, Red Globe Press, Springer.

GELMAN, A. & KING, G. 1994. A Unified Method of Evaluating Electoral Systems and Redistricting Plans. *American Journal of Political Science,* 38**,** 514-554.

GELMAN, A., KING, G. & THOMAS, A. C. 2012. JudgeIt II: A Program for Evaluating Electoral Systems and Redistricting Plans. *http://gking.harvard.edu/judgeit*.

GOEDERT, N. 2014. Gerrymandering or geography? How Democrats won the popular vote but lost the Congress in 2012. *Research and Politics,* 1**,** 1-8.

GOEDERT, N., HILDEBRAND, R., TRAVIS, L. & PIERSON, M. 2024. Asymmetries in Potential for Partisan Gerrymandering. *Legislative Studies Quarterly*.

GORDON, S. C. & YNTISO, S. 2024. Base Rate Neglect and the Diagnosis of Partisan Gerrymanders. *Election Law Journal: Rules, Politics, and Policy,* 23**,** 193-210.

GROFMAN, B. 1982. For Single Member Districts Random is Not Equal. *Representation and Redistricting Issues," Lexington Books*.

GUDGIN, G. & TAYLOR, P. J. 1978. *Seats, Votes, and the Spatial Organization of Elections,* London, Pin Limited.

KATZ, J. N., KING, G. & ROSENBLATT, E. 2020. Theoretical Foundations and Empirical Evaluations of Partisan Fairness in District-Based Democracies. *American Political Science Review,* 114**,** 164-178.

KEENA, A., LATNER, M., MCGANN, A. J. M. & SMITH, C. A. 2021. *Gerrymandering the states: Partisanship, race, and the transformation of American federalism*, Cambridge University Press.

KENDALL, M. G. & STUART, A. 1950. The law of cubic proportion in election results. *British Journal of Sociology,* 1**,** 183-197.

KING, G. & BROWNING, R. X. 1987. Democratic Representation and Partisan Bias in Congressional Elections. *American Political Science Review,* 81**,** 1251-1273.
MAGLEBY, D. B. & MCDONALD, M. D. 2025a. Assessing Gerrymandering after the 2020 Census. *Election Law Journal: Rules, Politics, and Policy,* 24**,** 388-399.
MAGLEBY, D. B. & MCDONALD, M. D. 2025b. The New York Congressional Gerrymander: A Social Science and Policy Lesson. *Polity,* 57**,** 000-000.
MCDONALD, M. D. 2009. The Arithmetic of Electoral Bias, with Applications to U.S. House Elections. *APSA 2009 Toronto Meeting Paper. Available at SSRN*.
MCDONALD, M. D. & BEST, R. E. 2015. Unfair Partisan Gerrymanders in Politics and Law: A Diagnostic Applied to Six Cases. *Election Law Journal,* 14**,** 312-330.
MCDONALD, M. D., MAGLEBY, D. B., KRASNO, J., DONAHUE, S. J. & BEST, R. 2018. Making a Case for Two Paths Forward in Light of Gill v. Whitford. *Election Law Journal,* 17**,** 315-327.
MCGHEE, E. 2014. Measuring Partisan Bias in Single-Member District Electoral Systems. *Legislative Studies Quarterly,* 39**,** 55-85.
NAGLE, J. F. 2015. Measures of Partisan Bias for Legislating Fair Elections. *Election Law Journal,* 14**,** 346-360.
NAGLE, J. F. 2019. What Criteria Should Be Used for Redistricting Reform? *Election Law Journal,* 18**,** 63-77.
NAGLE, J. F. & RAMSAY, A. 2021. On measuring two-party partisan bias in unbalanced states. *Election Law Journal: Rules, Politics, and Policy,* 20**,** 116-138.
RAMSAY, A. 2023. Estimating Seats–Votes Partisan Advantage. *Election Law Journal: Rules, Politics, and Policy,* 22**,** 67-79.
RODDEN, J. & WEIGHILL, T. 2022. Political geography and representation: A case study of districting in Pennsylvania. *Political Geometry: Rethinking Redistricting in the US with Math, Law, and Everything In Between.* Springer.
RODDEN, J. A. 2019. *Why Cities Lose,* New York, N.Y., Hachette Book Group.
STEPHANOPOULOS, N. O. & MCGHEE, E. M. 2015. Partisan Gerrymandering and the Efficiency Gap. *University of Chicago Law Review,* 82**,** 831-900.
TUFTE, E. R. 1973. The relationship between seats and votes in two-party systems. *American Political Science Review,* 67**,** 540-554.
WARRINGTON, G. S. 2018. Quantifying Gerrymandering Using the Vote Distribution. *Election Law Journal,* 17**,** 39-57.
WILSON, M. C. & GROFMAN, B. N. 2022. Models of inter-election change in partisan vote share. *Journal of Theoretical Politics,* 34**,** 481-498.
WISE, G. M. 2026. One Person, One District: A democratically derived standard of maximum disproportionality. *Election Law Journal: Rules, Politics, and Policy,* 25**,** 32-55.