

Implementing partisan symmetry: Problems and paradoxes

**Daryl DeFord¹, Natasha Dhamankar², Moon Duchin²,
Varun Gupta³, Mackenzie McPike², Gabe Schoenbach²,
and Ki Wan Sim²**

¹Washington State University

²Tufts University/Voting Rights Data Institute

³University of Pennsylvania

Abstract

We consider the measures of partisan symmetry proposed for practical use in the political science literature, as clarified and developed in Katz, King, and Rosenblatt 2020. Elementary mathematical manipulation shows the symmetry metrics to have surprising properties that call their meaningfulness into question. To accompany the general analysis, we study measures of partisan symmetry with respect to recent voting patterns in Utah, Texas, and North Carolina, flagging problems in each case. Taken together, these observations should raise major concerns about the available techniques for quantitative scores of partisan symmetry—including the mean-median score, the partisan bias score, and the more general "partisan symmetry standard"—as the decennial redistricting begins.

Keywords: Partisan symmetry, partisan metrics, redistricting, gerrymandering.

1 Introduction

In the political science literature, there is a long legacy of work on partisan gerrymandering, or the act of drawing political boundary lines to secure extra seats for the party that controls the process. One of the questions attracting the most attention has been to measure the degree of partisan advantage secured by a particular choice of redistricting lines. But to counteract partisan gaming requires a baseline notion of partisan fairness—extra seats compared to what baseline?—that has proved elusive. The family of fairness metrics with perhaps the longest pedigree is called *partisan symmetry* scores (Grofman 1983; King and Browning 1987; King et al. 2005; Grofman and King 2007; Gerken et al. 2018; Grofman and Gaddie 2019), which got a conceptual and empirical overview and a timely renewed endorsement in (Katz, King, and Rosenblatt 2020). The partisan symmetry standard is premised on the intuitively appealing fairness notion that the share of representation awarded to one party with its share of the vote *should* also have been secured by the other party, had the vote shares been exchanged. For instance, if Republicans achieve 40% of the seats with 30% of the vote, then it would be deemed fair for the Democrats to also achieve 40% of the seats with 30% of the vote.

At the heart of the symmetry ideal is a commitment to the principle that half of the votes should secure half of the seats. There are several metrics in the symmetry family that derive their logic from this core axiom. The *mean-median metric* is vote-denominated: it produces a signed number that is often described as measuring how far short of half of the votes a party can fall while still securing half the seats. A similar metric, *partisan bias*, is seat-denominated. Given the same input, it is said to measure how much more than half of the seats will be secured with half of the votes. The ideal value of both of these scores is zero. These are two in a large family of partisan metrics that can be described in terms of geometric symmetry of the "seats-votes curve."

The focus in the current work is to show that there are serious obstructions to the practical implementation of symmetry standards. This is of pressing current interest at the time of writing: we are in a period of intense public focus on redistricting reform and on the cusp of a new round of redistricting. In 2018 alone, voter referenda led four states to pass constitutional amendments (CO, MI, MO, OH), and another to write reforms into state law (UT) in anticipation of 2021 redistricting.¹ In Utah, partisan symmetry was adopted as a criterion to be considered by the new independent redistricting commission before plans would be approved.² We sound a note of caution here, showing that the versions of these scores that are realistically useable are eminently gameable by partisan actors and do not have reliable interpretations. To be precise: in each state we studied, the most extreme partisan outcomes for at least one political party are still achievable with a clean bill of health from the full suite of partisan symmetry scores. Furthermore, the signed scores (like mean-median MM and partisan bias PB) make frequent sign errors in terms of partisan advantage.

Utah itself gives strong evidence of the interpretation problems: with respect to recent voting patterns, a good symmetry score can only be achieved by a plan that secures a Republican congressional sweep; what's more, the popular symmetry scores described above flag all possible plans with any Democratic representation as major *Republican* gerrymanders.

We argue that there is at present no workable framework to make good on the idea of partisan symmetry. A manageable symmetry standard requires a swing assumption of some kind because its logic is built on voting counterfactuals (namely, table-turning). But this puts the most tenuous election modeling task, vote prediction, at the center of the methodology. And the symmetry framing requires that the metrics be insensitive to the fundamentally spatial nature of redistricting: there are many ways, not one way, for a vote pattern to shift by a given amount.³⁴

The main contribution of the paper is a Characterization Theorem for partisan symmetry under linear swing that clarifies what is actually measured by the leading symmetry scores. Applying this characterization, we demonstrate that realistic conditions can easily lead to "paradoxes" where one party has an extreme advantage in seats but the other party is flagged as the beneficiary of the gerrymander. We then use recent electoral data from three states to demonstrate the ease of gaming symmetry scores and the prevalence of these paradoxical labelings.

1.1 Literature review

Building the seats-votes curve with available data

We consider an election in a state with k districts and two major parties, Party A and Party B. A standard construction in political science is the "seats–votes curve," a plot representing the relationship of the vote share for Party A to the seat share for the same party. Observed outcomes generate single points in V - S space—for instance, (.3, .4) represents an election where Party A got 30% of the votes and 40% of the seats—but various methods have been used to extend from a scatterplot to a curve, such as by fitting a curve from a given class (linear or cubic, for instance) to observed data points. We will focus on the construction of seats-votes curves that is emphasized

1. At the time of writing, those measures stand in CO, MI, and OH. In Missouri, voters overwrote and partly erased the reform in the subsequent vote in 2020. In Utah, the state legislature repealed parts of the reform.

2. From the voter initiative: "The Legislature and the Commission shall use judicial standards and the best available data and scientific and statistical methods, including measures of partisan symmetry, to assess whether a proposed redistricting plan abides by and conforms to the redistricting standards" that bar party favoritism. The legislature later blocked this from taking effect. Cf. Utah Code, Chapter 7 Title 20A, Chapter 19 Part 1, Para. 103, https://le.utah.gov/xcode/Title20A/Chapter19/C20A-19_2018110620181201.pdf.

3. This is also the problem with axioms for partisan fairness metrics such as McGhee's "efficiency principle," which assumes that more votes will lead to more seats, absent manipulation. In reality, the spatial distribution of a party's votes, and not just the number, controls the possibilities for representation.

4. In the end, our conclusion is that methods centered on varying districting lines rather than varying votes should be preferred in the study of gerrymandering. This recommendation is consistent with a practice of perturbing recent vote patterns to get a view of the robustness of a gerrymander. A robustness check or a trend projection can be a valuable element of a redistricting analysis while playing a secondary rather than primary role.

in Katz, King, and Rosenblatt 2020: linear uniform partisan swing. Beginning with a single data point derived from a districting plan and a vote pattern, the vote share is varied by a uniform shift, so that the district vote shares (v_1, \dots, v_k) will shift to $(v_1 + t, \dots, v_k + t)$. This generates a step function spanning from $(0, 0)$ to $(1, 1)$ in the V - S plot, with a jump in seat share each time a district is pushed past 50% vote share for Party A. (See Figures 1-2 below for examples.)

Linear uniform partisan swing is the leading method proposed for use in implementation. Katz-King-Rosenblatt explicitly make it Assumption 3 in their symmetry survey—and use it throughout the paper—noting that the curve-fitting alternative is more suited "for academic study... than for practical use" in evaluation of plans. Katz et al. also mention Assumption 4, a stochastic generalization of uniform swing, as a preferred alternative to linear swing "whenever it makes a difference," and cite Brunell 1999; Jackman 1994 as examples that employ that model. This would add many additional modeling decisions, so would be difficult to carry out authoritatively in a practical setting. Nevertheless, we will identify some differences in shifting to a stochastic UPS approach in notes below. In short, neither including low levels of noise nor employing a vote vector obtained by regression from several elections will compromise the main findings here.

In any event, it is simple uniform partisan swing that is prevalent in practical applications. Grofman noted in 1983 that linear swing is preferred in practice to more sophisticated models (Grofman 1983, n.14); it is touted as the standard technique in (Garand and Parent 1991); and it has been invoked as recently as 2019 in expert reports and testimony (Mattingly 2019). Finally, Missouri voters actually wrote linear uniform partisan swing into their state constitution in 2018, requiring that an election index prepared by the state demographer be subjected to a swing of $t = -.05, -.04, \dots, +.05$ to test a plan's responsiveness.⁵ Because the present analysis is designed to address the prospects for implementation, we therefore focus on the linear swing model.

Deriving symmetry scores from the seats-votes curve

Given a seats-votes curve, many symmetry scores have been proposed; here, we focus on the mean-median score MM, the partisan bias score PB, and the partisan Gini score PG, which have all been considered for at least 35 years. (Definitions are found in the next section.) Grofman's 1983 survey paper (Grofman 1983) lays out eight possible scores of asymmetry once a seats-votes curve has been set. His Measure 3 is vote-denominated bias, which would equal MM under linear uniform partisan swing; similarly his Measure 4 corresponds to PB, and Measure 7 introduces PG.

Because the partisan Gini is defined as the area between the seats-votes curve and its reflection over the center (seen in the shaded regions in Figs 1-2), it is easily seen to "control" all the other possible symmetry scores: when $PG = 0$, its ideal value, all partisan symmetry metrics also take their ideal values, including MM and PB. This agrees with Katz-King-Rosenblatt (Katz, King, and Rosenblatt 2020, Def 1), where the coincidence of the curve and its reflection, i.e., $PG = 0$, is called the "partisan symmetry standard." In the current work, our Theorem 3 gives precise necessary and sufficient conditions for the partisan symmetry standard to be satisfied.

The literature invoking MM and PB as measures of bias is too large to survey comprehensively. We note that the interpretation of median-minus-mean as quantifying (signed) Party A advantage is fairly standard in the journal literature, such as: "The median is 53 and the mean is 55; thus, the bias runs two points against Party A (i.e., $53 - 55 = -2$)" (McDonald and Best 2015). The connection to the seats-votes curve is also standard: MM "essentially slices the S/V graph horizontally at the $S = 50\%$ level and obtains the deviation of the vote from 50%" (Nagle 2015, p351).⁶

5. "[T]he electoral performance index shall be used to simulate elections in which the hypothetical statewide vote shifts by one percent, two percent, three percent, four percent, and five percent in favor of each party. The vote in each individual district shall be assumed to shift by the same amount as the statewide vote." Cf. Missouri State Constitution, Article III Section 3.

6. There is even more work centered on PB (notably King and Browning 1987), but it is more rarely used in conjunction with linear swing, since that assumption makes its values move in large jumps.

We briefly note the impact of introducing stochasticity on the analysis below, we note that modifying the seats–votes curve γ by adding noising terms with mean zero will change the precision of our findings but not the basic structure, replacing exact equalities in the Characterization Theorem with approximate equalities. In particular, this does not impact the prevalence of "paradoxes," for two reasons. First, when the curve γ passes far from $(V, S) = (.5, .5)$, perturbations to γ will not move it past the center point, which would be needed to change the sign of MM and PB (as explained below in §2). Second, the standard mean-median score is simply calculated as the difference of the mean vote share by district and the median (Katz, King, and Rosenblatt 2020, p173), and thus relies on no swing assumption at all! Abandoning linear swing therefore does not fix the problems with the mean-median score, but only breaks its relationship to the seats–votes curve.

Applying symmetry scores in practice

The current work is designed to evaluate the techniques proposed by leading practitioners for practical use. Political scientists and their collaborators have advanced these scores in *amicus* briefs spanning from *LULAC v. Perry* (2006) (King *et al.* 2005) to *Whitford v. Gill* (2018) (Gerken *et al.* 2018) to *Rucho v. Common Cause* (2019) (Grofman and Gaddie 2019). The scores have been claimed to be "reliable and difficult to manipulate" and authors have argued that while "Symmetry tests should deploy actual election outcomes" (as we do here), they will nonetheless "measure opportunity," i.e., give information about future performance (Gerken *et al.* 2018, p17,24). That assertion is drawn from an *amicus* brief in the *Whitford* case explicitly proposing mean-median as a concrete choice of score for this task. As laid out earlier in the influential *LULAC* brief,

"Models applying the symmetry standard are by their nature predictive, just as the legislators themselves are predicting the potential impact of the map on partisan representation. The symmetry standard and the resulting measures of partisan bias, however, do not require forecasts of a particular voting outcome. Rather, by examining all the relevant data and the potential seat divisions that would occur for particular vote divisions, it compares the potential scenarios and determines the partisan bias of a map, separating out other potentially confounding factors. Importantly, those drawing the map have access to the same data used to evaluate it, and no data is required other than what is in the public domain" (King *et al.* 2005, p11).

This paper takes up precisely this modeling task in the manner explicitly proposed by its authors.

1.2 Premises and caveats

What is partisan gerrymandering?

To assess the success of partisan symmetry metrics at their task of identifying partisan gerrymandering, we should be clear about first principles. First, we agree on the definition from Katz, King, and Rosenblatt 2020: "Partisan gerrymanderers use their knowledge of voter preferences and their ability to draw favorable redistricting plans to maximize their party's seat share." That is, the express intent of a partisan gerrymander is to secure for their party as many seats as possible under the constraints of voter geography and the other rules of redistricting.

This means that a (successful) gerrymander in favor of Party A is a districting plan that obtains an extreme Party A seat share. In the public perception, that will usually be assessed by comparing the seat share to the vote share, undergirded by an intuition that equates fairness with *proportionality*. But proportionality is not the neutral tendency of redistricting and in some cases it may be literally impossible to secure (Duchin *et al.* 2019). For the better part of a century, political scientists have investigated this neutral tendency by appealing to constructions like seats–votes curves and cube laws. This literature has been severely limited by its inattention (with a few notable exceptions) to

spatial factors, i.e., to the geography of the vote distribution.⁷ A powerful alternative has recently emerged through so-called *ensemble methods*: Markov chain algorithms (for example) can now build samples of alternative districting plans, holding a vote distribution fixed. Though we make use of ensembles of alternative plans below, the crux of this paper does not require the reader to commit to this or any particular choice of non-gerrymandered baseline. On any common view of the baseline, from proportionality to cube law to outlier analysis from an ensemble, the standard definition of partisan gerrymandering entails views like these:

- Circa 2016, the voter preferences in North Carolina were fairly even between Democratic and Republican candidates for statewide office. Both algorithmic techniques and human mapmakers can easily draw plans ranging from 7 to 10 districts with a Republican majority in the 13-member Congressional plan. In this context, a successful Republican gerrymander would secure a 10R-3D outcome, or an even more extreme outcome if possible.
- In Utah, voting patterns in this period tend to favor Republican over Democratic candidates by a roughly 70-30 split. This is tilted enough that a 4R-0D Congressional plan is in some sense typical or expected, and need not be viewed as a gerrymander. However, it would be an error to label a 3R-1D outcome as a Republican-favoring gerrymander.

We will treat these as premises in the treatment below.

Ensembles, not estimators

In this paper, the algorithmically generated plans are not offered as a statistical experiment and come with no probabilistic claims, but mainly provide *existence proofs* to illustrate how readily gameable partisan symmetry standards will be for those engaged in redistricting.⁸ The methods also produce many thousands of examples of plans that are paradoxical in the sense developed in this paper, where signed partisan symmetry metrics identify the wrong party as the gerrymanderers.

The algorithm used here (described in §4.1) builds a sample of plans that are plausible by the lights of traditional districting principles: they are population-balanced, contiguous, and relatively compact, using whole precincts as the building blocks. There are techniques to layer in other criteria in addition to these in a state-specific way to set up a more thorough outlier analysis, but that is not needed for this application. (See for instance DeFord and Duchin 2019; DeFord, Duchin, and Solomon 2020; Becker *et al.* 2020.) Nothing here, or in the broader logic of ensemble analysis, assumes that line-drawers are random agents.

The competitiveness caveat

Writing after the *LULAC v. Perry* decision in 2007, Grofman and King offer this key caveat: "[W]e are not proposing to apply this methodology in every situation, but only in *potentially* competitive jurisdictions, where the consequences of gerrymandering might be especially onerous in thwarting the will of the majority" (Grofman and King 2007, their emphasis). In the following paragraph they suggest that "reasonably competitive" settings could be those where each party receives 40-60% of vote share. This is a sizeable limitation on the scope of the symmetry approach: only about half of states have a recent U.S. Senate voting pattern in this range, for example.⁹ Two of the three cases presented below (North Carolina and Texas) are in this reasonably competitive zone; the third is Utah, where a partisan symmetry standard was recently enacted in law. We could not find any record of political scientists speaking out against Utah's adoption of this measure while it was on the ballot in 2018.

7. The absence of vote geography in the bulk of redistricting research is noted, for instance, by Johnston 2002; Rodden 2010; Calvo and Rodden 2015.

8. This is exactly the use of ensemble methods that is endorsed in (Katz, King, and Rosenblatt 2020, p176) as productive and compelling: a demonstration of possibility and impossibility.

9. See for instance <https://github.com/gerrymandr/party-tilt>.

The more recent Katz, King, and Rosenblatt 2020 mainly places its competitiveness caveat in the article supplement, though it is obliquely referenced in Assumption 2, which requires that there is a sufficiently large range of "possible values" for vote outcomes. Even there, the caveat is hedged: "Although Assumption 2 is defined in terms of possible electoral outcomes, those that are exceedingly unlikely, such as Washington DC voting overwhelming[ly] Republican, do not violate this assumption but may generate model dependence in estimation." On our reading, the authors do not rule out the use of partisan symmetry metrics even on states with an extreme partisan lean.

In any case, the analysis presented here, which shows that the partisan symmetry standard devolves to a simple numerical test, extends to competitive as well as uncompetitive situations.

2 A mathematical characterization of the Partisan Symmetry Standard

We begin with definitions and notation needed to state Theorem 3, which characterizes when $\text{PG} = 0$ (the Partisan Symmetry Standard from Katz, King, and Rosenblatt 2020). We describe the vote outcome in the election using an ordered tuple (i.e., a vector) whose coordinates record the Party A share of the two-party vote in each of the k districts as follows: $v = (v_1, \dots, v_k)$, where $0 \leq v_1 \leq \dots \leq v_k \leq 1$. Let the mean district vote share for Party A be denoted $\bar{v} = \frac{1}{k} \sum v_i$ and the median district vote share, v_{med} , be the median of the $\{v_i\}$, which equals $v_{(k+1)/2}$ if k is odd and $\frac{1}{2}(v_{k/2} + v_{(k/2)+1})$ if k is even because of the convention that coordinates are in non-decreasing order. We note that \bar{v} is not necessarily the same as the statewide share for Party A except in the idealized scenario that the districts have equal numbers of votes cast (i.e., equal turnout).

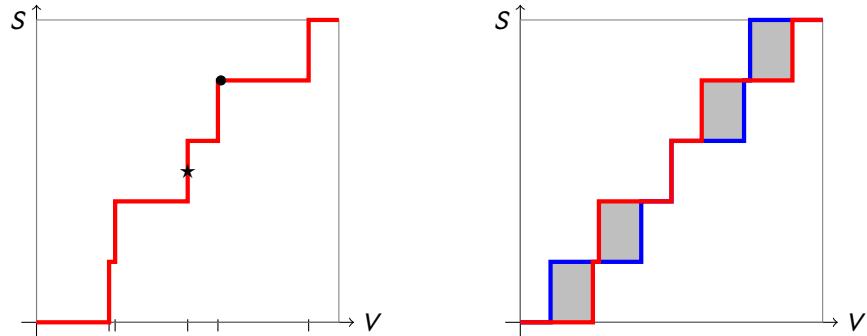


Figure 1. Red: The seats-votes curve γ generated by vote shares $v = (.21, .51, .61, .85, .87)$ under uniform partisan swing. This gives $\bar{v} = .61$ as the average vote share across the districts. The *jump points*, where an additional seat changes hands, are marked on the V axis. Blue: the reflection of γ about the center point \star . Since MM is the horizontal displacement from \star to a point on γ , this hypothetical election has a perfect $\text{MM} = 0$ score, but it is not very symmetric overall, with $\text{PG} = .112$, seen as the area of the shaded region between γ and its reflection. Because the step function jumps at $V = .5$, it is not clear how PB is defined in this case.

The number of districts in which Party A has more votes than Party B in an election with vote shares v is the seat outcome, $\#\{i : v_i > \frac{1}{2}\}$. This induces a seats-votes function $\gamma = \gamma_v$ defined as $\gamma(\bar{v} + t) = \#\{i : v_i + t > \frac{1}{2}\}/k$, which we can interpret as the share of districts won by Party A in the counterfactual that an amount t was added to A's observed vote share in every district. Varying t to range over the one-parameter family of vote vectors $(v_1 + t, \dots, v_k + t)$ is known as (linear) *uniform partisan swing*. The curve γ , treated as a function $[0, 1] \rightarrow [0, 1]$, has been regarded as measuring how a fixed districting plan would behave if the level of vote for Party A were to swing up or down. Below, we will refer to the function and its graph interchangeably, and we will call it the *seats-votes curve* associated to the vote share vector v . We begin with several scores based on v and γ .

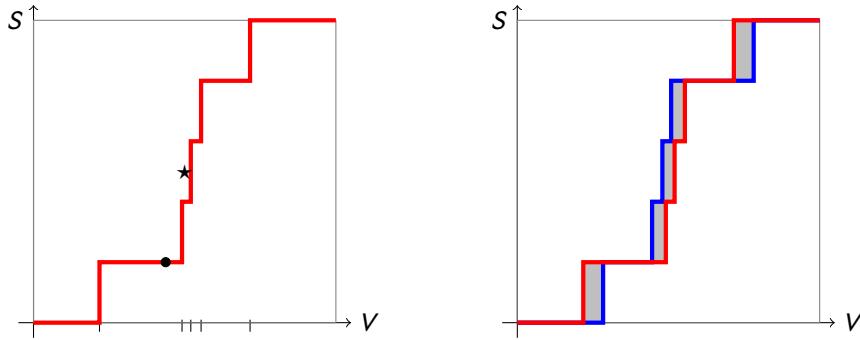


Figure 2. The seats-votes curve γ generated by the vote share vector $v = (.221, .383, .417, .446, .719)$, which was the observed outcome in the 2016 Congressional races in Oregon from the Republican point of view. This gives a mean of $\bar{v} = 0.4372$, and earned Republicans 1 seat out of 5. The blue curve is the reflection of γ about the center, so it shows seats at each vote share from the Democratic point of view. This could be regarded as a situation with reasonably good symmetry, since the red and blue curves are close. Its scores are $PG = .05248$, $MM = -.0202$, and $PB = -.1$. The sign of the latter two scores is thought to indicate a Democratic advantage.

Definition 1 (Partisan symmetry scores). *The **partisan Gini** score $PG(v)$ is the area between the seats-votes curve γ_v and its reflection over the center point $\star = (\frac{1}{2}, \frac{1}{2})$.*

$$PG(v) = \int_0^1 |\gamma(x) + \gamma(1-x) - 1| dx.$$

The **mean-median** score is $MM(v) = v_{\text{med}} - \bar{v}$. The **partisan bias** score is $PB(v) = \gamma(\frac{1}{2}) - \frac{1}{2}$.

These scores can all be related to the shape of the seats-votes curve γ (see Figures 1,2). Partisan Gini measures the failure of γ to be symmetric about the center point $\star = (\frac{1}{2}, \frac{1}{2})$, in the sense that it is always non-negative, and it equals zero if and only if γ equals its reflection. Mean-median score is the horizontal displacement from \star to a point on γ ,¹⁰ which is why it is votes-denominated (vote-share being the variable on the x -axis). Similarly, partisan bias is the vertical displacement from \star to a point on γ , and is therefore seats-denominated. (We note that $(\frac{1}{2}, \gamma(\frac{1}{2}))$ is a well-defined point unless there is a jump precisely at $1/2$, which occurs if some $v_i = \bar{v}$ on the nose—this is shown in Figure 1 but should not happen with real-world data.) Below, we will focus on MM instead of PB , but we note that $MM > 0 \implies PB \geq 0$ because of the geometric interpretation: if γ passes to the left of \star and is nondecreasing, then it must pass through or above \star .

We can see that the curve γ , and consequently the partisan Gini score, is exactly characterized by the points at which Party A has added enough vote share to secure the majority in an additional district. For the following analysis, it will be useful to characterize this curve in terms of the v data.

Definition 2 (Gaps and jumps). *The **gaps** in a vote share vector v can be written in a gap vector*

$$\delta = (\delta_1, \delta_2, \dots, \delta_{k-1}) = (v_2 - v_1, v_3 - v_2, \dots, v_k - v_{k-1}).$$

The **jump points** for vote share vector v are the values of $\bar{v} + t$ such that some $v_i + t = \frac{1}{2}$. We have

$$t_1 := \frac{1}{2} - v_k, \quad t_2 := \frac{1}{2} - v_{k-1}, \quad \dots \quad t_k := \frac{1}{2} - v_1$$

as the times corresponding to these jumps, so we can denote the jumps as $j_i = \frac{1}{2} + \bar{v} - v_{k+1-i}$, and the jump vector as $j = (j_1, \dots, j_k)$.

¹⁰ To see this, plug in $t = 1/2 - MM - \bar{v} = 1/2 - v_{\text{med}}$ to deduce that $(1/2 - MM, 1/2)$ is on γ —note that this connection from arithmetic to geometry is only exact when γ has been constructed with linear swing, which is the standing assumption in this paper.

By definition of γ , these jump points j_i , marked in the figures, are exactly the x -axis locations (i.e., the V values) at which γ jumps from $(i-1)/k$ to i/k .¹¹

With this notation, we can rephrase and relate the various partisan symmetry scores. For instance, the center-most rectangle(s) formed between γ and its reflection have height $2PB$ and width $2MM$, which lets us relate the scores. For small k , these relationships reduce to extremely simple expressions: $PG = \frac{4}{3}|MM|$ when $k = 3$, and $PG = 2|MM|$ when $k = 4$. (Proved in the Supplement.)

For any number of districts, we obtain a clean characterization of precisely when vote shares by district satisfy the Partisan Symmetry Standard (Katz, King, and Rosenblatt 2020, Def 1).

Theorem 3 (Partisan Symmetry Characterization). *Given k districts with vote shares v , jump vector j , and gap vector δ , the following are equivalent:*

$PG(v) = 0$	(Partisan Symmetry Standard)
$j_i + j_{k+1-i} - 1 = 0 \quad \forall i$	(jumps)
$\frac{1}{2}(v_i + v_{k+1-i}) = \bar{v} \quad \forall i$	(mean vote)
$\frac{1}{2}(v_i + v_{k+1-i}) = v_{\text{med}} \quad \forall i$	(median vote)
$\delta_i = \delta_{k-i} \quad \forall i$	(gaps)

The proof is included in the Supplement. Note also that the theorem statement makes it clear that $PG = 0 \implies MM = 0$ by comparing the third equality to the fourth, which fits with the earlier observation that partisan Gini "controls" the other scores.

Theorem 3 asserts that the partisan symmetry standard under linear swing is nothing but the requirement that the vote shares by district are distributed on the number line in a symmetric way.¹² In particular, this tells you at a glance that an election with vote shares (.37, .47, .57, .67) in its districts rates as perfectly partisan-symmetric, while one with vote shares (.37, .47, .57, .60) falls short.

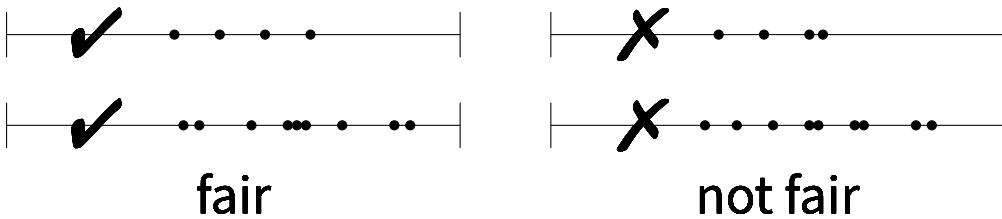


Figure 3. Four election outcomes, shown as vote shares by district. On the left-hand side, the v_i are symmetric about their center, so all partisan symmetry scores are perfect. On the right-hand side, non-symmetric outcomes. The partisan symmetry standard can be eyeballed by a glance at the vote shares in the districts.

3 Paradoxes with signed symmetry scores

Recall that mean-median and partisan bias are *signed* scores that are supposed to identify which party has an advantage and by what amount. A positive score is said to indicate an advantage for

11. Warning to the reader: if you try to draw your own examples to test some of these results, be aware that not just any step function can be generated by a vote vector. The jump points must satisfy $\sum j_i = k/2$, which follows directly from summing $j_i = \frac{1}{2} + \bar{v} - v_{k+1-i}$.

12. Consider replacing simple linear swing by the modeling approach frequently used in stochastic uniform partisan swing, such as in the JudgeIt software package. In this case, rather than using a single election's vote-share-by-district, a linear swing is applied to a different vote vector that is obtained by regression from several elections. In this case, the Characterization Theorem still holds, applied to this inferred vote vector instead of the observed vote vector.

Party A (the point-of-view party whose vote shares are reported in v). Let us say that a *paradox* occurs when the score indicates an advantage for one party even though it has the fewest seats it can possibly earn with its vote share, say.¹³ In other words, a paradox means that the score makes a sign error with respect to the standard definition of partisan gerrymandering.

When there is an extremely skewed outcome (with a vote share for one party exceeding 75%), we will show that paradoxes *always* occur, just as a matter of arithmetic. But even for less skewed elections with a vote share between 62.5 and 75% for the leading party—which frequently occur in practice!—mundane realities of political geography can force these sign paradoxes.

To illustrate these observations, we will begin with the case of $k = 4$ districts, where the algebra is simpler. The issues are not limited to small k , however: in the empirical section we will find paradoxes of this kind in $k = 13$ and $k = 36$ cases as well.

Example 4 (Paradoxes forced by arithmetic). *Suppose we have $k = 4$ districts and an extremely skewed election in favor of Party A, achieving $75\% < \bar{v} < 87.5\%$. With equal turnout, Party B can get at most one seat. However, every vote vector v achieving this outcome (one B seat) yields $MM \geq \bar{v} - \frac{3}{4} > 0$. In particular, such districting plans all have positive MM and PB, paradoxically indicating an advantage for Party A in every case where Party B gets representation.*

The demonstration is simple arithmetic. Since $\frac{1}{2}(v_2 + v_3) = v_{\text{med}}$, we have

$$\bar{v} = \frac{1}{4}(v_1 + v_2 + v_3 + v_4) = \frac{v_1 + v_4}{4} + \frac{v_2 + v_3}{4} \implies v_{\text{med}} - \bar{v} = \bar{v} - \frac{v_1 + v_4}{2}.$$

Since $v_1 \leq \frac{1}{2}$ (for B to win a seat) and $v_4 \leq 1$, we get $MM = v_{\text{med}} - \bar{v} \geq \bar{v} - \frac{3}{4}$, as needed.

A stronger statement can be made if one takes political geography into account. It was shown by Duchin–Gladkova–Henninger–Voss–Newman–Wheelen in a study of Massachusetts (Duchin *et al.* 2019) that, if the precincts are treated as atoms that are not to be split in redistricting, then several recent elections have the property that no choice of district lines can create even one district with Republican vote share over 1/2. This is because Republican votes are distributed extremely uniformly across the precincts of the state. While other states are not as uniform as Massachusetts, it is still true that there is some upper bound Q on the vote share that is possible for each party in any district. When this bound satisfies $Q < 2\bar{v} - \frac{1}{2}$, even moderately skewed elections are forced to exhibit paradoxical symmetry scores. As we will see below, having all $v_i < 2\bar{v} - \frac{1}{2}$ ensures both that one seat is the best outcome for Party B and that the median vote share is greater than the mean.

Example 5 (Paradoxes forced by geography). *Suppose we have $k = 4$ districts and a skewed election in favor of Party A, with $62.5\% \leq \bar{v} < 75\%$. Suppose the geography of the election has Party A support arranged uniformly enough that districts can not exceed a share Q of A votes, for some $Q < 2\bar{v} - \frac{1}{2}$. Then with equal turnout, Party B can get at most one seat. However, every vote vector v achieving this outcome (one B seat) has a positive MM and PB, paradoxically indicating an advantage for Party A in every case where Party B gets representation.*

Proof. First, it is easy to see that Party B can't achieve two seats: in that case, we would have $v_1, v_2 \leq \frac{1}{2}$. Since we also have $v_3, v_4 \leq Q < 2\bar{v} - \frac{1}{2}$, we can average the v_i to get the contradiction $\bar{v} < \bar{v}$.

To see that $MM > 0$, we write

$$MM = v_{\text{med}} - \bar{v} = \frac{v_2 + v_3}{2} - \frac{v_1 + v_2 + v_3 + v_4}{4} = \frac{v_1 + v_2 + v_3 + v_4}{4} - \frac{v_1 + v_4}{2} = \bar{v} - \frac{v_1 + v_4}{2}.$$

Since $v_1 < \frac{1}{2}$ and $v_4 \leq Q < 2\bar{v} - \frac{1}{2}$, we have $MM > \bar{v} - \frac{2Q+1}{4} > \bar{v} - \bar{v} = 0$. □

13. In practical settings, it may be hard to calculate the absolute minimum that is possible, so this may refer to an extremely low number of seats with respect to the known range of alternatives, as in the figures below.

4 Investigations with observed vote data

4.1 Methods

In this section we illustrate the theoretical issues from above, using naturally observed election data together with a Markov chain technique that produces large ensembles of districting plans.¹⁴ In each case, we have run a recombination ("ReCom") Markov chain for 100,000 steps—long enough to comfortably achieve heuristic convergence benchmarks in all scores that we measured—while enforcing population balance, contiguity, and compactness.¹⁵ Note that some Markov chain methods count every attempted move as a step, even though most proposals are rejected, so that each plan is counted with high multiplicity in the ensemble; in our setup, the proposal itself includes the criteria, and repeats are rare. 100,000 steps produces upwards of 99,600 distinct districting plans in each ensemble presented here.

We ran trials on multiple elections in our dataset and all results are available for comparison (VRDI 2020). Below, we highlight the most recent available Senate race from a Presidential election year in each state, to match conditions across cases as closely as possible. The data bottleneck is a precinct shapefile matching geography to voting patterns, which is surprisingly difficult to obtain. A database of shapefiles is available at (MGGG 2020). We use the statewide U.S. Senate election rather than an endogenous Congressional voting pattern because the latter is subject to uncontested races and variable incumbency effects. For instance, Utah's 2016 Congressional race had all four seats contested, but a Republican vote in District 3 went for hard-right Jason Chaffetz, while on the other side of the invisible line to District 4, the vote went to Mia Love, a Black Republican who is outspoken on racial inequities. When the district lines are moved, it's not clear (for example) that a Love voter stays Republican. The U.S. Senate race had a comparable number of total votes cast to the Congressional race (1,115,608 vs. 1,114,144) and offers a consistent choice of candidates around the state, making it better suited to methods that vary district lines.

4.2 Utah and the "Utah Paradox"

We begin with Utah, where the elections that were available in our dataset all come from 2016.¹⁶ Utah has only four congressional districts and has a heavily skewed partisan preference, with a statewide Republican vote share of 71.55% in the 2016 Senate race.¹⁷ Though it is far from clear that table-turning between the parties is conceivable in the near future,¹⁸ Utahns nonetheless recently enacted partisan symmetry consideration into state law.

Figure 4 shows outcomes from our 100,000-step ensemble. The vast majority (94.266%) of Utah plans found in our ensemble have all four R seats, with the remaining plans giving 3-1 splits. The chain found 5734 plans with 3 Republican and 1 Democratic seats, and we see that all of these have PG scores above 0.06. Below, we explore and explain these bounds on seats and scores.

When looking at the full PG histogram, we see a large bulk of plans with nearly-ideal PG scores,

14. All data and code are public and freely available for inspection and replication in GitHub (VRDI 2018, 2020) and Dataverse (DeFord *et al.* 2021).

15. The population balance imposed here is 1% deviation from ideal district size. Such plans are easily tuneable to 1-person deviation by refinement at the block level without significant impact to any other scores discussed here. Contiguity is enforced by recording adjacency of precincts. Compactness, at levels comparable to those observed in human-made plans, is an automatic consequence of the spanning-tree-based recombination step. For more information about the Markov chain used here, see (DeFord, Duchin, and Solomon 2019).

16. Out of GOV16, SEN16, and PRES16, none gives an especially pure partisanship signal, because the Democratic candidates for Governor and Senate were quite weak, while the partisanship in the Presidential race was complicated by the presence of a very strong third-party candidate in Evan McMullin, giving that race an extremely different pattern. The Governor's race shows similar results to the Senate, as the reader can verify in (VRDI 2020).

17. Recall that \bar{v} is the average of the district vote shares, which will not generally equal the statewide share except under an equal-turnout assumption.

18. The amount of linear partisan swing needed to reverse the partisan advantage should be viewed as unreasonably large under these conditions. With respect to SEN16, fully 199 out of 2123 precincts in the state have Republican vote share that reaches zero under this amount of swing. This is one of the reasons (though not the only reason) that this style of quantifying partisan advantage is poorly suited to Utah.

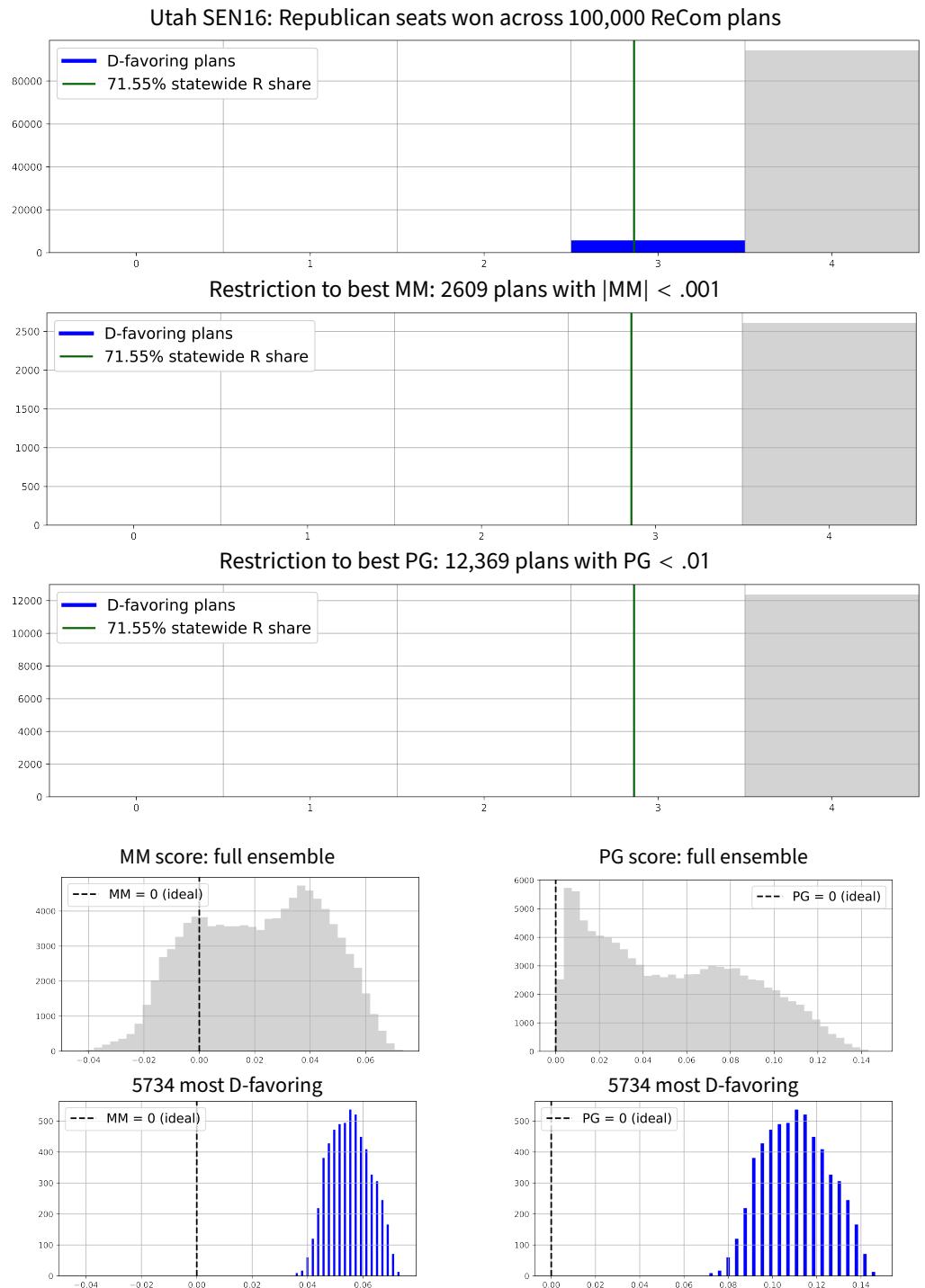


Figure 4. Ensemble outputs for 100,000 Utah Congressional plans with respect to SEN16 votes. Republicans received 71.55% of the two-way vote in this election, which is marked in the plots to show the corresponding seat share. There are 5734 plans in the ensemble in which Democrats get a seat; these are shown in blue in the top row, but they are absent from the next two rows because a D seat never occurs in plans with good symmetry scores. The last row of the figure shows the MM and PG histograms restricted to the plans with a D seat. The empirical data corroborates the prediction that good symmetry scores lock out Democratic representation, and illustrates the "Utah paradox" that a Democratic-won seat always receives the label of a Republican gerrymander.

all giving a Republican sweep (four out of four R seats). This is surprising enough to deserve a name.

The Utah Paradox

- Partisan symmetry scores near zero are supposed to indicate fairness, and signed symmetry scores are supposed to indicate which party is advantaged.
- There are many trillions of valid Congressional plans in Utah, and under reasonable geographical assumptions, every single one of them with PG close to zero is mathematically guaranteed to yield a Republican sweep of the seats. In particular, even constraining symmetry scores to better than the ensemble average (for any reasonably diverse neutral ensemble of alternatives) would deterministically impose a partisan outcome: the one in which Democrats are locked out of representation.
- Furthermore, the signed scores make a sign error: they report all plans with Democratic representation to be significant pro-Republican gerrymanders.

Geographic assumptions: The UT-SEN16 election has a statewide R share of .7155, so this is roughly equal to the district mean \bar{v} (or exactly equal in the equal-turnout case). If we can show that the possible Republican share of a district is bounded above by any $Q < .931$, then the arguments of the last section show that Democrats can secure at most one seat, and that every plan with Democratic representation has the sign error $MM > 0$. We consider the assumption that no district can exceed 93% Republican share to be very reasonable. Indeed, even a greedy assemblage of the 608 precincts with the highest Republican share in that race (which is the number needed to reach the ideal population of a Congressional district) only produces a district with R share .888. And this is even without imposing a requirement that districts be contiguous, which certainly limits the possibilities further and only strengthens the bound. As a further indication, our Markov chain run of contiguous plans never encounters a district with Republican share over .8595.

Example 6 (The Utah paradox, empirical). *The UT-SEN16 vote pattern can be divided into 4R-0D seats or 3R-1D seats. However, even though MM, PB, and PG scores can all get arbitrarily close to zero, there are no reasonably symmetric plans that secure a Democratic seat. In our algorithmic search, every plan with nonzero Democratic representation has $PG > .069$, $MM > .034$, and $PB \geq .25$, which is in the worst half of scores observed for each of those scores. Thus even a mild constraint on partisan symmetry has imposed an empirical Democratic lockout. As predicted by the analytic paradox, all plans with D representation are reported as significant R-favoring gerrymanders.*

The basic idea here is extremely simple and readers can try it for themselves. Choose any four numbers from 0 to 1 whose mean is .7155. If one of them is at or below .5 (a Democratic-won seat), you will find that the median of the four scores is greater than the mean, so the mean-median score is positive. It is simply false that a median higher than the mean is a flag of advantage for the point-of-view party.

As described in the introduction, Utah recently became the first state to encode partisan symmetry as a districting criterion in statute. This makes the Utah Paradox quite a striking example of the worries raised by using partisan symmetry scores in practice.

4.3 Texas

Next, we turn to Texas, creating a chain of 100,000 steps to explore the ways to divide up the 2012 Senate vote distribution. With 36 Congressional districts, Texas has one of the highest k values of any state (only California has more seats). The 2012 Senate race was won by a Republican with $\sim 58\%$ of the vote. Figure 5 shows the partisan properties in the ensemble of plans, allowing us to compare extreme symmetry scores (an ostensible indicator of partisan unfairness) to extreme seat shares (the explicit goal of partisan gerrymandering). We find no evidence of correlation or any kind of correspondence.

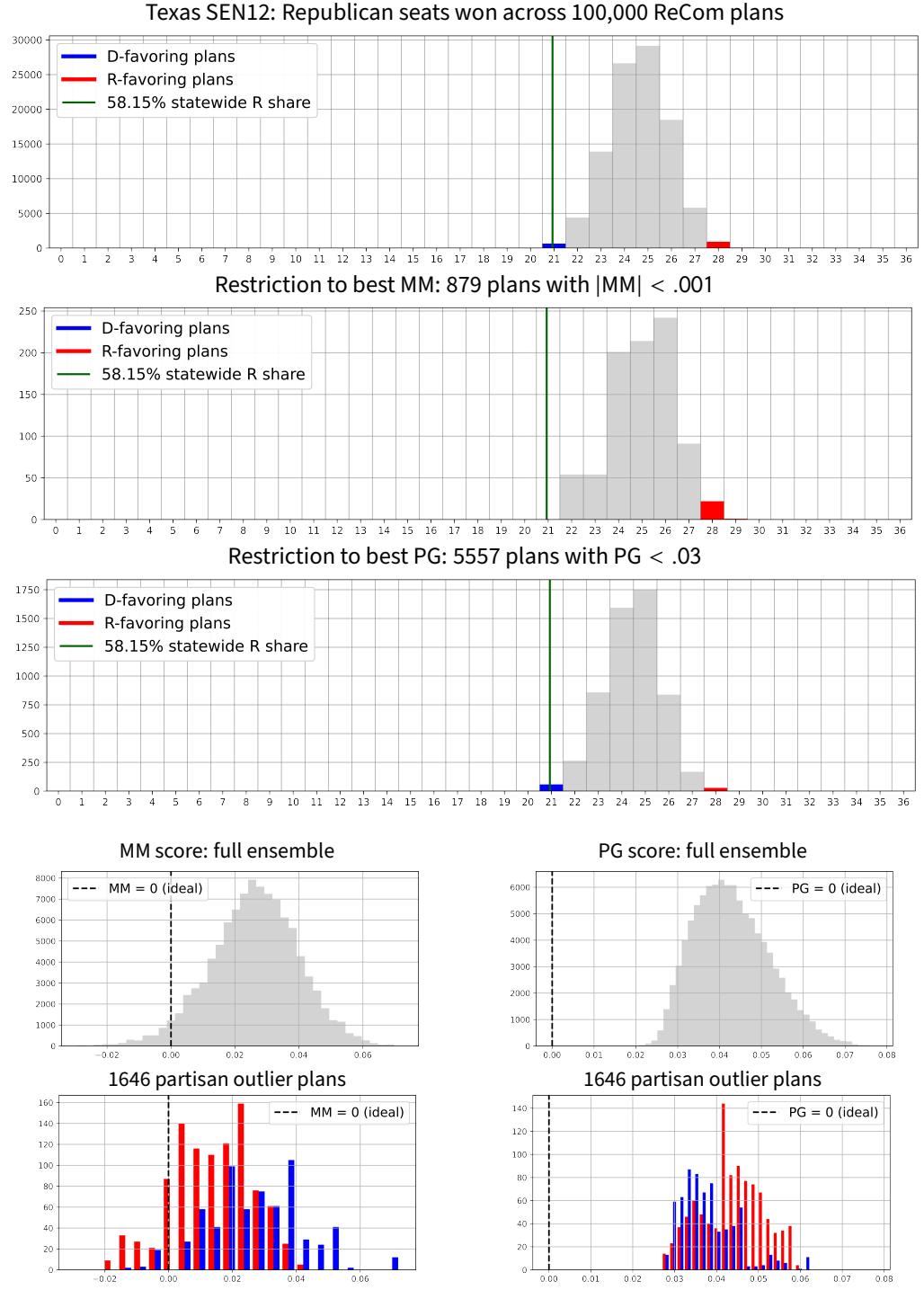


Figure 5. Ensemble outputs for Texas Congressional plans with respect to SEN12 votes. Republicans received 58.15% of the two-way vote in this election, which is marked in the plots to show the corresponding seat share. There are 1646 plans in the ensemble that are seats outliers for one party or the other; these are shown in red and blue in the top row and their relative frequency can be observed in the next two rows, which focus on plans with the best symmetry scores. The last row of the figure shows the MM and PG histograms restricted to the 1646 outlier plans flagged above. The scores are shown to be readily gamed: numerous extreme Democratic-favoring plans are found with near-optimal symmetry scores. In this sample, most extreme Democratic-favoring plans are labeled Republican gerrymanders by the mean-median score, and some extreme Republican-favoring plans are labeled Democratic gerrymanders.

North Carolina SEN16: Republican seats won across 100,000 ReCom plans

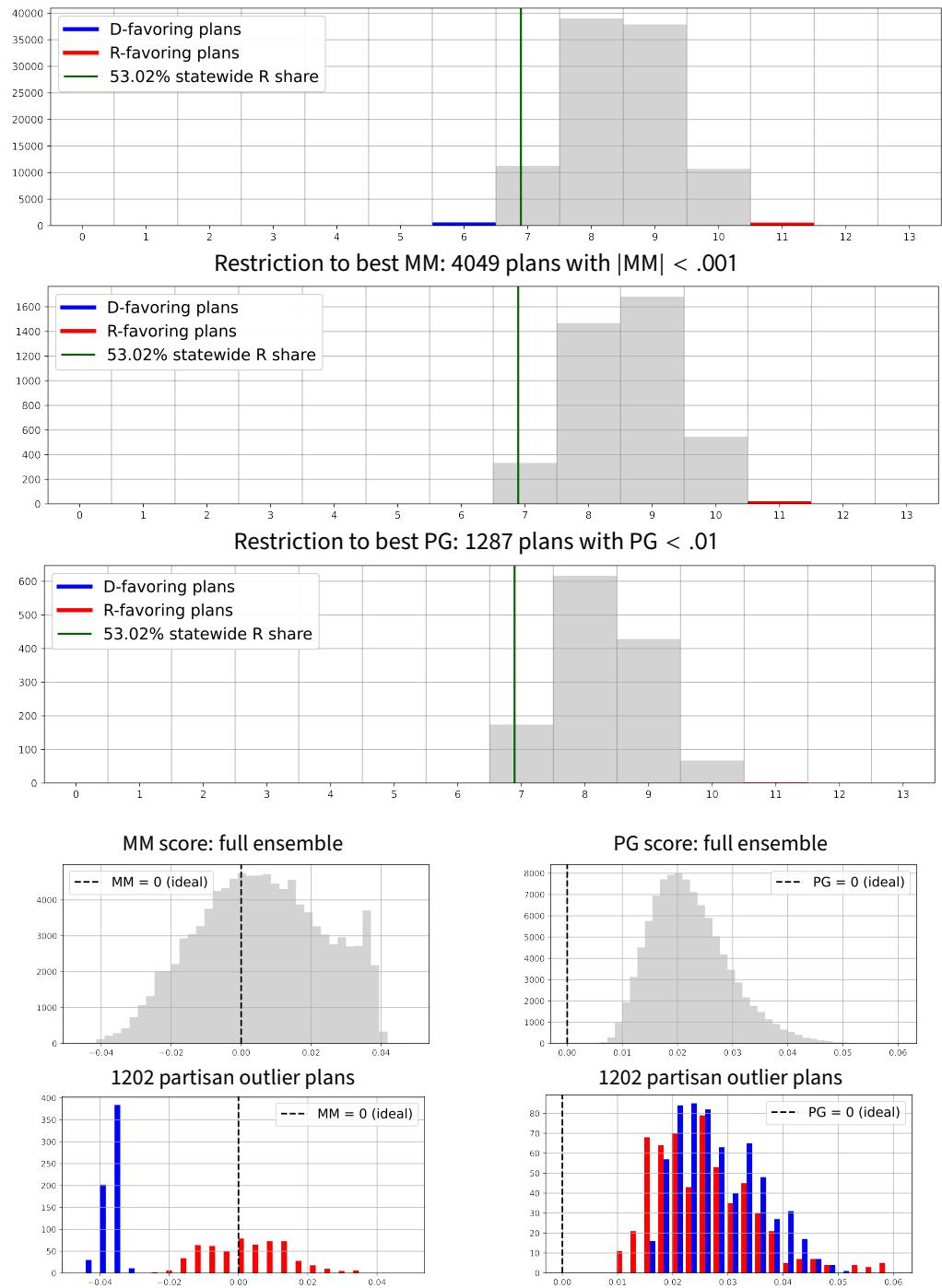


Figure 6. Ensemble outputs for North Carolina Congressional plans with respect to SEN16 votes. Republicans received 53.02% of the two-way vote in this election, which is marked in the plots to show the corresponding seat share. There are 1202 plans in the ensemble that are seats outliers for one party or the other; these are shown in red and blue in the top row and their relative frequency can be observed in the next two rows, which focus on plans with the best symmetry scores. The last row of the figure shows the MM and PG histograms restricted to the 1202 outlier plans flagged above. In this setting, symmetry can easily be gamed in favor of Republicans, with thousands of 11-2 plans receiving near-perfect mean-median scores.

Over 98% of the sampled plans give Republicans 22 to 27 seats out of 36, seen in gray in the histogram. The red bars mark the outlying plans with the most Republican seats (28 or more R seats), while the blue bars mark the most Democratic plans (21 or fewer R seats). We can then study the histograms formed by the winnowed subsets of the ensemble with the best PG and MM scores, which in each case fall in the top 6%. Note that these severely winnowed subsets not only have a shape similar to the full ensemble (indicating a lack of correlation), but that there are still many partisan outlier plans even with strict symmetry standards in place. Plans with the extreme outcome of ≥ 28 R seats actually occur with *higher* frequency among the $MM \approx 0$ plans than in the full sample—more than twice as often, in fact. This shows rather emphatically that restricting to "good" symmetry scores is no impediment to partisan gerrymandering.

For the reverse perspective, we consider how plans with extreme seat counts score on symmetry. The last row in Figure 5 shows only the seat outliers: blue for plans with ≤ 21 and red for plans with ≥ 28 Republican seats. A significant number of maximally D-favoring plans (which are also close to vote proportionality) paradoxically register as major Republican gerrymanders under the MM score, outpacing by a significant margin the most extreme R-favoring plans. The mean-median score utterly fails at identifying partisan advantage even in an election regarded as "reasonably competitive" by the proponents of partisan symmetry. In Texas, as in Utah, it is simply false that a median higher than the mean is a flag of advantage for the point-of-view party.

In terms of the overall symmetry measured by PG, extreme plans for both parties can be found with scores that are as good as nearly anything observed in the ensemble. So from this perspective as well, neither MM nor PG signals anything with respect to political outcomes. Even if the proponents of symmetry standards never intended to constrain extreme seat imbalances, this runs counter to the common expectations of anti-gerrymandering reforms in popular discourse, in legal settings, and even in much of the political science literature.

4.4 North Carolina

Finally, we move to a state with a much closer to even partisan split: North Carolina ($k = 13$ seats), with respect to the 2016 Senate vote (~53% Republican share). In this case, mean-median does much better than in Texas in terms of distinguishing the seat extremes: Figure 6 shows consistently higher scores for the maps with the most Republican seat share than the ones with the most Democratic outcomes. However, the extreme Republican maps still straddle the "ideal" score of $MM = 0$, and both sides can still find very extreme plans whose PG scores report that their symmetry is essentially as good as anything in the ensemble.

Overall it is fair to say that partisan symmetry imposes no constraint on partisan gerrymandering in North Carolina, at least for one side: this method easily produces hundreds of maps with 10-3 outcomes (which was clearly reported in the Rucho case to be the most extreme that the legislature thought was possible) while securing nearly perfect symmetry scores, and the gerrymanderer only needs *one*. Indeed, the ensemble even finds highly partisan-symmetric maps that return an 11-2 outcome for this particular vote pattern. Four of these are shown in Figure 7.

To sum up, we give a recipe for how to hide your partisan gerrymander from detection by symmetry scores, even in a 53-47 state. To begin, leverage differential turnout. A recent analysis of 2014-16 Congressional voting showed that several states have voter turnout that is 25 or even 40% higher in the districts won by one party than the other (Veomett 2018). This can easily push the average vote share in the districts two points higher than the statewide share. Now give your party a 55% majority in many districts, and let the others be landslides, being sure to arrange the vote shares symmetrically around 55%. This secures sterling symmetry scores and a windfall of seats for your side without even risking any close contests.¹⁹

19. Example: Suppose Party A has 53% of the statewide vote in a ten-district state, and suppose the average turnout in

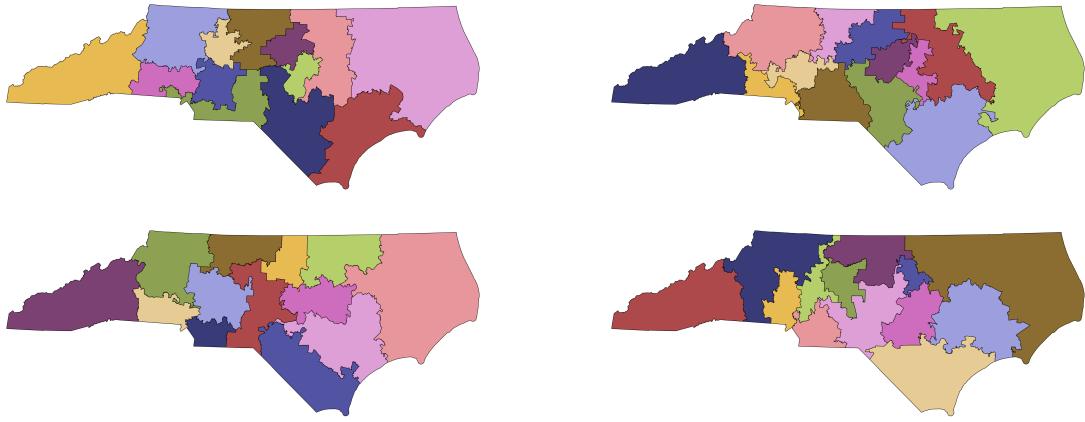


Figure 7. Algorithmic methods are mainly used here for example generation. Each of these 13-district plans comes out 11R-2D with respect to the SEN16 voting data, while having nearly perfect partisan symmetry. These maps have PG scores of 0.0096, 0.0099, 0.0107, and 0.0115, respectively, all in the best 2% of the ensemble. This figure also illustrates the diversity of districting plans achieved by this Markov chain method.

5 Conclusion

In this piece, we have characterized the partisan symmetry standard from Katz, King, and Rosenblatt 2020 mathematically: it turns out to amount simply to a prescription for the arrangement of vote totals across districts (Theorem 3, Partisan Symmetry Characterization).²⁰ We follow this with examples of realistic conditions under which the adoption of strict symmetry standards not only (a) fails to prevent extreme partisan outcomes but even (b) can lock in unforeseen consequences on these partisan outcomes. Finally, again under realistic conditions, signed partisan symmetry metrics (c) can plainly mis-identify which party is advantaged by a plan.²¹

None of these findings gives a theoretical reason for rejecting partisan symmetry as a definition of fairness. A believer in symmetry-as-fairness can certainly coherently hold that symmetry standards do not *aim* to constrain partisan outcomes, but merely to reinforce the legitimacy of district-based democracy by reassuring the voting public that the tables can yet turn in the future. This view casts aside, or holds irrelevant, the standard definition of a partisan gerrymander as a plan designed to maximize the seats for a party. With this reasoning, we should not worry that Democrats in Utah may for now be locked out of Congressional representation *by the symmetry standard itself*; this is still fair because Democrats would enjoy a similar advantage of their own if election patterns were to linearly swing by 40 percentage points in their favor.

For those who do want to constrain the most extreme partisan outcomes that line-drawing can secure, these investigations should serve as a strong caution regarding the use of partisan symmetry metrics, whether in the plan adoption stage or in plan evaluation after subsequent elections have been conducted.

A-won districts is 80% that of B-won districts. Then Party A can get a 7A-3B outcome with all safe seats by dividing up the votes as $v = (0.26, 0.26, 0.26, 0.55, 0.55, 0.55, 0.55, 0.84, 0.84, 0.84)$, while acing every symmetry test.

20. One possible response is to try to preserve the partisan symmetry standard but abandon linear uniform partisan swing in favor of a different way of drawing seats-votes curves. We have discussed alternatives above in several places. We reiterate that noising the seats-votes curve will only change the precision of the characterization. We also note that no change to the swing assumption impacts the findings on the mean-median score, which is defined as the difference between the mean district vote share and the median.

21. While this is beyond the scope of the current paper, there is also every reason to believe that partisan symmetry metrics can (d) give answers that depend unpredictably on which vote pattern is used to assess them: endogenous (imputing uncontested races) or exogenous? Senate race or attorney general? Most single-score indicators have this problem. This may give us a reason to prefer modeling approaches that incorporate multiple races, but those approaches come with a host of modeling decisions that make them impractical for real-world use.

If symmetry metrics measured something that was obviously of inherent value in the healthy functioning of representative democracy, then we might reasonably choose to live with the consequences of the definition, no matter the partisan outcomes. However, the Characterization Theorem shows that a putatively perfect symmetry score is nothing more and nothing less than a requirement that the vote shares v_i in the districts be arranged symmetrically on the number line (see Figure 3). Someone who wishes to assert that partisan symmetry is *really about* some principle—majority rule, responsiveness, equality of opportunity, etc—would have to explain why that principle is captured by the simple arithmetic of vote share spacing. With this framing, it is more difficult to argue that symmetry captures any essential ingredient of civic fairness.

Funding, Acknowledgments, Data Availability

The 2019 Voting Rights Data Institute was generously supported by the Prof. Amar J. Bose Research Grant at MIT and the Jonathan M. Tisch College of Civic Life at Tufts. MD was partially supported by NSF DMS-1255442.

MD is deeply grateful to Gary King for illuminating conversations about the partisan symmetry standard. All authors thank the other participants of the Voting Rights Data Institute, especially Brian Morris, Michelle Jones, and Cleveland Waddell, for stimulating discussions.

Replication code for this article is available in VIDI (2020) at <https://doi.org/10.7910/DVN/MVVLQC>. We also make use of the code and data in (VIDI 2018; MGGG 2020).

Finally, we offer our warmest thanks to Crimson Groh, an undergraduate math research student who caught a sign error in the PG expression in the accepted version of this paper!

Supplementary Material

For supplementary material accompanying this paper, see the appendices below the references.

References

- Becker, A., M. Duchin, D. Gold, and S. Hirsch. 2020. “Computational Redistricting and the Voting Rights Act.” Preprint.
- Brunell, T. 1999. “Partisan Bias in US Congressional Elections, 1952–1996: Why the Senate Is Usually More Republican Than the House of Representatives.” *American Politics Quarterly* 27 (3): 316–337.
- Calvo, E., and J. Rodden. 2015. “The Achilles Heel of Plurality Systems: Geography and Representation in Multiparty Democracies.” *American Journal of Political Science* 59 (4): 789–805.
- DeFord, D., N. Dhamankar, M. Duchin, V. Gupta, M. McPike, G. Schoenbach, and K. W. Sim. 2021. *Replication Data for: Implementing partisan symmetry: Problems and paradoxes*. V. 3. <https://doi.org/10.7910/DVN/MVVLQC>.
- DeFord, D., and M. Duchin. 2019. “Redistricting Reform in Virginia: Districting Criteria in Context.” *Virginia Policy Review* 12(2):120–146.
- DeFord, D., M. Duchin, and J. Solomon. 2019. “Recombination: A family of Markov chains for redistricting.” <https://arxiv.org/abs/1911.05725>.
- . 2020. “A Computational Approach to Measuring Vote Elasticity and Competitiveness.” *Statistics and Public Policy* 7 (1): 69–86.

- Duchin, M., T. Gladkova, E. Henninger-Voss, B. Klingensmith, H. Newman, and H. Wheelen. 2019. “Locating the Representational Baseline: Republicans in Massachusetts.” *Election Law Journal* 18, no. 4 (December): 388–401.
- Garand, J. C., and T. W. Parent. 1991. “Representation, Swing, and Bias in U.S. Presidential Elections, 1872–1988.” *American Journal of Political Science* 35, no. 4 (November): 1011–1031.
- Gerken, H. K., J. N. Katz, G. King, L. J. Sabato, and S. S.-H. Wang. 2018. *Amicus Brief in Whitford v. Gill*. No. 16-1161.
- Grofman, B. 1983. “Measures of Bias and Proportionality in Seats-Votes Relationships.” *Political Methodology* 9 (3): 295–327.
- Grofman, B., and R. K. Gaddie. 2019. *Amicus Brief in Rucho v. Common Cause*. Nos. 18-422, 18-726.
- Grofman, B., and G. King. 2007. “The Future of Partisan Symmetry as a Judicial Test for Partisan Gerrymandering after *LULAC v. Perry*.” *6 Election L.J.* 2 6:2–35.
- Jackman, S. 1994. “Measuring Electoral Bias: Australia, 1949–93.” *British Journal of Political Science* 24 (3): 319–357.
- Johnston, R. J. 2002. “Manipulating Maps and Winning Elections: Measuring the Impact of Malapportionment and Gerrymandering.” *Political Geography* 21 (1): 1–31.
- Katz, J. N., G. King, and E. Rosenblatt. 2020. “Theoretical Foundations and Empirical Evaluations of Partisan Fairness in District-Based Democracies.” *American Political Science Review* 114 (1): 164–178.
- King, G., and R. X. Browning. 1987. “Democratic Representation and Partisan Bias in Congressional Elections.” *American Political Science Review* 81, no. 4 (December): 1251–1273.
- King, G., B. Grofman, A. Gelman, and J. N. Katz. 2005. *Amicus Brief in LULAC v. Perry*. Nos. 05-204, 05-254, 05-276, 05-439.
- Mattingly, J. 2019. *Expert Report in Common Cause v. Lewis*. <https://sites.duke.edu/quantifyinggerrymandering/files/2019/09/Report.pdf> see also <https://sites.duke.edu/quantifyinggerrymandering/2019/06/26/the-animated-firewall/>.
- McDonald, M. B., and R. E. Best. 2015. “Unfair Partisan Gerrymanders in Politics and Law: A Diagnostic Applied to Six Cases.” *Election Law Journal* 14 (4): 312–330.
- MGGG. 2020. “MGGG States.” <https://github.com/mggg-states>, GitHub repository, accessed June 26, 2020.
- Nagle, J. F. 2015. “Measures of Partisan Bias for Legislating Fair Elections.” *Election Law Journal* 14 (4): 346–360.
- Rodden, J. 2010. “The Geographic Distribution of Political Preferences.” *Annual Review of Political Science* 13:321–340.
- Veomett, E. 2018. “Efficiency Gap, Voter Turnout, and the Efficiency Principle.” *Election Law Journal* 17 (4): 249–263.
- VRDI. 2018. “GerryChain.” <https://github.com/mggg/gerrychain>, GitHub repository, accessed May 17, 2019.
- . 2020. “Partisan Symmetry.” <https://github.com/mggg/Partisan-Symmetry>, GitHub repository, accessed August 6, 2020.

A Supplement: Proof of characterization theorem

We briefly recall the needed notation from above: vote share vector v with i th coordinate v_i ; gap vector δ with $\delta_i = v_{i+1} - v_i$; and jump vector j with $j_i = \frac{1}{2} + \bar{v} - v_{k+1-i}$, where \bar{v} is the mean of the v_i . These expressions define j , δ in terms of v ; neither j nor δ completely determines v because they are invariant under translation of the entries of v , but one additional datum (such as v_1 or \bar{v}) suffices, with j or δ , to fix the associated v . In this section, we begin by expressing PG in terms of the jumps j , then giving equivalent conditions for $PG = 0$ in terms of j , δ , or v .

As outlined above, PG measures the area between the seats-votes curve γ and its reflection. The shape of the region between those curves depends directly on the points $j = (j_1, j_2, \dots, j_k)$, since each j_i is the x value of a vertical jump in the curve and the $1 - j_i$ are the values of the jumps in the reflection. But looking at Figure 1 makes it clear that it is complicated to decompose the integral into vertical rectangles in the style of a Riemann sum, because the $\{j_i\}$ and the $\{1 - j_{k-i}\}$ do not always alternate. Fortunately, it is always easy to decompose the picture into horizontal rectangles (analogous to a Lebesgue integral), where it is now clear which red and blue corners to pair as the seat share changes from i/k to $(i+1)/k$. The curve contains the points $(j_i, \frac{i-1}{k})$, $(j_i, \frac{i}{k})$ as well as $(j_{k+1-i}, \frac{k-i}{k})$, $(j_{k+1-i}, \frac{k-i+1}{k})$. The rotated curve therefore contains the points $(1 - j_{k+1-i}, \frac{i-1}{k})$ and $(1 - j_{k+1-i}, \frac{i}{k})$, which means that the i th rectangle has height $1/k$ and width $|j_i + j_{k+1-i} - 1|$. Summing over these rectangles gives us the expression

$$PG = \frac{1}{k} \sum_{i=1}^k |j_i + j_{k+1-i} - 1|.$$

Recall that the set of vote share vectors \mathcal{V} is the cone in the vector space \mathbb{R}^k given by the condition that the v_i are in non-decreasing order in $[0, 1]$. The j vector is simply the v vector reversed and re-centered at $1/2$ rather than \bar{v} . The only condition on the gap vector δ is that its entries are non-negative and sum to at most one. Putting these observations together we may define the set of achievable v, δ, j respectively as

$$\begin{aligned} \mathcal{V} &= \{(v_1, \dots, v_k) : 0 \leq v_1 \leq \dots \leq v_k \leq 1\}, \\ \mathcal{D} &= \left\{(\delta_1, \dots, \delta_{k-1}) : \delta_i \geq 0 \forall i, \quad \sum \delta_i \leq 1\right\}, \\ \mathcal{J} &= \left\{(j_1, \dots, j_k) : 0 \leq j_1 \leq \dots \leq j_k \leq 1, \quad \sum j_i = \frac{k}{2}\right\}. \end{aligned}$$

The condition on j is of interest because it exactly identifies the possible seats–votes curves $\Gamma = \{\gamma_v : v \in \mathcal{V}\}$. (That is, not just any step function is realizable as a valid γ .)

Now we can prove the Characterization theorem.

Theorem 3. *Given k districts with vote shares v , jump vector j , and gap vector δ , the following are equivalent:*

$$PG(v) = 0 \tag{1}$$

$$j_i + j_{k+1-i} - 1 = 0 \quad \forall i \tag{2}$$

$$\frac{1}{2} (v_i + v_{k+1-i}) = \bar{v} \quad \forall i \tag{3}$$

$$\frac{1}{2} (v_i + v_{k+1-i}) = v_{\text{med}} \quad \forall i \tag{4}$$

$$\delta_i = \delta_{k-i} \quad \forall i \tag{5}$$

Proof. The condition that $PG(v) = 0$ has been rewritten in terms of j above, and converting back to

the v_i we get

$$\frac{1}{k} \sum_{i=1}^k |j_i + j_{k+1-i} - 1| = \frac{2}{k} \sum_{i=1}^k \left| \frac{v_i + v_{k+1-i}}{2} - \bar{v} \right| = 0,$$

which immediately gives (1) \iff (2) \iff (3) since a sum of nonnegative terms is zero if and only if each term is zero. To see (3) \iff (4), just consider $i = \lceil \frac{k}{2} \rceil$ in (3) to obtain $v_{\text{med}} = \bar{v}$; in the other direction, average both sides over i in (4) to obtain $\bar{v} = v_{\text{med}}$. Finally, the symmetric gaps condition (5) is clearly equivalent to the symmetry of the values of v about the center v_{med} , which is (4). \square

B Supplement: Bounding partisan Gini in terms of mean-median

Recall that the mean-median score MM is a signed score that is supposed to identify which party has a structural advantage, and by what amount. On the other hand, the partisan Gini PG is a non-negative score that simply quantifies the failure of symmetry, interpreted as a magnitude of unfairness. We easily see that $PG = 0 \implies MM = 0$ by comparing (3) and (4) in Theorem 3. In this section we strengthen that to a bound from below that is sharp in low dimension.

Let us define $\text{discrep}(i) = \frac{v_i + v_{k+1-i}}{2} - \bar{v}$, measuring the difference between the average of a pair of vote shares from the average of all the vote shares. This gives

$$PG = \frac{1}{k} \sum_{i=1}^k |2\bar{v} - v_i - v_{k+1-i}| = \frac{2}{k} \sum_{i=1}^k |\text{discrep}(i)|.$$

Note that $\text{discrep}(\lceil \frac{k}{2} \rceil) = MM$, as observed above, and that $\sum_{i=1}^k \text{discrep}(i) = 0$ by definition of \bar{v} .

Theorem 7. *The partisan Gini score satisfies* $\begin{cases} PG \geq \frac{4}{k} |MM|, & k \text{ odd} \\ PG \geq \frac{8}{k} |MM|, & k \text{ even,} \end{cases}$ *with equality when $k = 3, 4$.*

Proof. First suppose k is odd, say $k = 2m+1$. Then $\text{discrep}(m) = v_m - \bar{v} = MM$, so $\sum_{i \neq m} \text{discrep}(i) = -MM$. We have

$$\begin{aligned} PG &= \frac{2}{k} \sum_{i=1}^k |\text{discrep}(i)| = \frac{2}{k} \left(|\text{discrep}(m)| + \sum_{i \neq m} |\text{discrep}(i)| \right) \geq \frac{2}{k} \left(|\text{discrep}(m)| + \left| \sum_{i \neq m} \text{discrep}(i) \right| \right) \\ &= \frac{2}{k} (|MM| + |-MM|) = \frac{4}{k} |MM|. \end{aligned}$$

The argument for even $k = 2m$ is very similar, except that $\text{discrep}(m) = \text{discrep}(m+1) = \frac{v_m + v_{m+1}}{2} - \bar{v} = MM$. So now those terms contribute $2|MM|$ to the sum and the remaining terms contribute at least $2|-MM|$, for a bound of $PG \geq \frac{8}{k} |MM|$. That completes the proof of the inequalities.

For $k = 3$ or $k = 4$, the term $\sum_{i \neq m} |\text{discrep}(i)|$ is just $2|\text{discrep}(1)|$, making the inequality into an equality. \square

By a dimension count, it is easy to see that PG is not simply a function of MM for $k \geq 5$. A direct calculation confirms this, and shows that MM is not simply a function of PG either. Let

$$v = (.2, .3, .4, .5, .7), \quad v' = (.2, .31, .39, .5, .7), \quad v'' = (.19, .31, .4, .5, .7),$$

giving $PG(v) = PG(v')$ while $MM(v) \neq MM(v')$. On the other hand, $MM(v) = MM(v'')$ while $PG(v) \neq PG(v'')$.