Implementing partisan symmetry: Problems and paradoxes

Daryl DeFord, Natasha Dhamankar, Moon Duchin, Varun Gupta, Mackenzie McPike, Gabe Schoenbach, and Ki Wan Sim

Fall 2020

Abstract

We consider the measures of partisan symmetry proposed for practical use in the political science literature, as clarified and developed in [10]. Elementary mathematical manipulation shows the symmetry metrics derived from uniform partisan swing to have surprising properties. 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 using 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 approaches.

1 Introduction

In the political science literature, there is a long legacy of work on gerrymandering, or the act of drawing political boundary lines with an ulterior motive. 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, which has proved elusive. The family of fairness metrics with perhaps the longest pedigree is called partisan symmetry scores [5, 11, 12, 7, 4, 6], which got a conceptual and empirical overview and a timely renewed endorsement in [10]. 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 because, as we write, states are racing to adopt redistricting reform measures. 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. In Utah, partisan symmetry has now been adopted as a criterion to be considered by the new independent redistricting commission before plans can be approved.

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 and partisan bias) make systematic sign errors in terms of partisan advantage.

1“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. Cf. Utah Code, Chapter 7 Title 20A, Chapter 19 Part 1, Para. 103, https://le.utah.gov/xcode/Title20A/Chapter19/C20A-19-2018110620181201.pdf
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—arguably a sign error.

1.1 Literature review

1.1.1 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 the political science literature 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 a second construction of seats-votes curves that is emphasized in [10]: 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 partisan swing, so that the district vote shares $(v_1,\ldots,v_k)$ will shift to $(v_1+\alpha,\ldots,v_k+\alpha)$. This generates a step function spanning from $(0,0)$ to $(1,1)$ in the $V$-$S$ space, 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 evaluation. Katz–King–Rosenblatt explicitly make it Assumption 3 in their symmetry survey, noting that the curve-fitting alternative is more suited “for academic study... than for practical use” in evaluation of plans. Grofman noted in 1983 that linear swing is preferred in practice to more sophisticated models [5, n.14], it is touted as the standard technique in [3], and it has been invoked as recently as 2019 in expert reports and testimony [13].

1.1.2 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 [5] 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 [10, 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 obtain. The literature invoking MM and PB as measures of bias is too large to survey comprehensively. We note that the particular 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)” [14]. 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%” [16, p351].

1.1.3 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) [12] to Whitford v. Gill (2018) [4] to Rucho v. Common Cause (2019) [6]. The scores have been claimed to be “reliable

2Katz et al. also offer Assumption 4, a stochastic generalization of uniform swing, as the only alternative to linear UPS mentioned for general use. This would add many additional modeling decisions, so would be difficult to carry out in a practical setting. If used with recent election data, its impact would chiefly be to noise the seats-votes curve $\gamma$, which changes the precision of our findings but not the basic structure. In particular, this does not impact the prevalence of “paradoxes,” as described below.

3There is even more work centered on PB (notably [11]), but it is more rarely used in conjunction with linear swing, since that assumption makes its values move in large jumps.
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 [17, 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 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” [12, p11].

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

## 2 A mathematical characterization of the Partisan Symmetry Standard

We begin with definitions and notation needed to state the results in this paper, and particularly Theorem 3 which characterizes when $PG = 0$ (the Partisan Symmetry Standard from [10]). 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:

$$v = (v_1, \ldots, v_k)$$

where $0 \leq v_1 \leq \ldots \leq v_k \leq 1$. 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, $v_{\text{med}}$, 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 $\gamma$ generated by the vote share vector $v = (.21, .51, .61, .85, .87)$ under uniform partisan swing. This gives $\bar{v} = .61$ as the average vote share across the five districts. The jump points, which are values where an additional seat changes hands, are marked on the $V$ axis. Blue: the reflection of $\gamma$ about the center point $\star$. Since $MM$ is the horizontal displacement from $\star$ to a point on $\gamma$, this hypothetical election has a perfect $MM = 0$ score, but it is not very symmetric overall, with $PG = .112$, seen as the area of the shaded region between $\gamma$ 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 $\gamma(v+t) = \#\{i : v_i + t > \frac{1}{2}\}/k$, 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 $(v_1 + t, \ldots, v_k + t)$ is known as (linear) uniform partisan swing. The curve $\gamma$, 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 $\gamma$. 

3
The locations (i.e., the $V_i = j_{MM}$ Figure 1 but should not happen with real-world data.) Below, we will focus on well-defined point unless there is a jump precisely at $1/\delta$. These scores can all be related to the shape of the seats-votes curve $\gamma$ (see Figures 1, 2). Partisan Gini measures the failure of $\gamma$ 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 $\gamma$ equals its reflection. Mean-median score is the horizontal displacement from $\star$ to a point on $\gamma$ 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 $\gamma$, and is therefore seats-denominated. (We note that $\left(\frac{1}{2}, \gamma(\frac{1}{2})\right)$ is a well-defined point unless there is a jump precisely at $1/2$, which occurs if some $v_i = \overline{\gamma}$ 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 $\text{MM} > 0 \iff \text{PB} \geq 0$ because of the geometric interpretation: if $\gamma$ passes to the left of $\star$ and is nondecreasing, then it must pass through or above $\star$.

We can see that the curve $\gamma$, 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 1.** The partisan Gini score $\text{PG}(v)$ is the area between the seats-votes curve $\gamma_v$ and its reflection over the center point $\star = (\frac{1}{2}, \frac{1}{2})$.

\[ \text{PG}(v) = \int_0^1 |\gamma(x) - \gamma(1-x) + 1| \, dx. \]

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

These scores can all be related to the shape of the seats-votes curve $\gamma$ (see Figures 1, 2). Partisan Gini measures the failure of $\gamma$ 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 $\gamma$ equals its reflection. Mean-median score is the horizontal displacement from $\star$ to a point on $\gamma$ 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 $\gamma$, and is therefore seats-denominated. (We note that $\left(\frac{1}{2}, \gamma(\frac{1}{2})\right)$ is a well-defined point unless there is a jump precisely at $1/2$, which occurs if some $v_i = \overline{\gamma}$ 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 $\text{MM} > 0 \iff \text{PB} \geq 0$ because of the geometric interpretation: if $\gamma$ passes to the left of $\star$ and is nondecreasing, then it must pass through or above $\star$.

We can see that the curve $\gamma$, 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.** The gaps in a vote share vector $v$ can be written in a gap vector

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

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

\[ t_1 := \frac{1}{2} - v_k, \quad t_2 := \frac{1}{2} - v_{k-1}, \quad \ldots, \quad t_k := \frac{1}{2} - v_1 \]

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

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

---

To see this, plug in $t = 1/2 - \text{MM} - \overline{\gamma} = 1/2 - v_{\text{med}}$ to deduce that $(1/2 - \text{MM}, 1/2)$ is on $\gamma$.

5Warning to the reader: if you try to draw your own examples to test some of these results, note 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 $j_i = \frac{1}{2} + \overline{\gamma} - v_{k+1-i}$. 

---

Figure 2: The seats-votes curve $\gamma$ 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 $\overline{\gamma} = 0.4372$, and earned Republicans 1 seat out of 5. The blue curve is the reflection of $\gamma$ 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 $\text{PG} = .05248$, $\text{MM} = -.0202$, and $\text{PB} = -.1$. The sign of the latter two scores is thought to indicate a Democratic advantage.
With this notation, we can manipulate the various partisan symmetry scores. For instance, the center-most rectangle(s) formed between \( v \) and its reflection have height \( 2PB \) and width \( 2MM \), allowing the derivation of inequalities relating these scores. For small \( k \), these reduce to extremely simple expressions: \( PG = \frac{1}{3}MM \) when \( k = 3 \), and \( PG = 2MM \) when \( k = 4 \). (Proved in the Supplement.)

For any number of districts, we obtain a very clean characterization of precisely which distribution of vote shares to districts satisfy the Partisan Symmetry Standard [10, Def 1].

**Theorem 3** (Partisan Symmetry Characterization). Given \( k \) districts with vote shares \( v \), jump vector \( j \), and gap vector \( \delta \), the following are equivalent under uniform partisan swing:

\[
\begin{align*}
PG(v) &= 0 & \text{(Partisan Symmetry Standard)} \\
\frac{1}{2}(v_i + v_{k+1-i}) &= \bar{v} & \forall i \\
\frac{1}{2}(v_i + v_{k+1-i}) &= v_{med} & \forall i \\
\delta_i &= \delta_{k-i} & \forall i
\end{align*}
\]

That is, this partisan symmetry standard is nothing but the requirement that the vote shares by district are distributed on the number line in a symmetric way. In particular, this tells you at a glance that an election with vote shares \((.75, .47, .57, .67)\) in its districts rates as perfectly partisan-symmetric, while one with vote shares \((.37, .47, .57, .60)\) falls short. 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.

### 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 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 a very low number of seats—the fewest seats it can possibly earn with its vote share, say. In other words, a paradox means that the score makes an apparent sign error.

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 \( MM \geq \bar{v} - \frac{1}{3} > 0 \). 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 \( \frac{1}{2}(v_2 + v_3) = v_{med} \), we have

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

Since \( v_1 \leq \frac{1}{2} \) (for B to win a seat) and \( v_4 \leq 1 \), we get \( MM = v_{med} - \bar{v} \geq \bar{v} - \frac{1}{2} \), 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 [2] 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 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 \( v_1 < 2\bar{v} - \frac{1}{2} \) 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% ≤ \( \bar{v} \) ≤ 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 of paradox.** 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

\[
MM = \frac{v_2 + v_3}{2} - \frac{v_1 + v_2 + v_3 + v_4}{4} = \frac{v_1 + v_2 + v_3 + v_4}{4} - \frac{v_1 + v_4}{2} = \bar{v} - \frac{v_1 + v_4}{2}.
\]

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

4 Investigations with observed vote data

4.1 Methods

In this section we illustrate the theoretical issues from above, using naturalistically observed election data together with a Markov chain technique that produces large ensembles of districting plans. All data and code are public and freely available for inspection and replication [9].

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. 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.

We have run trials on multiple elections in our dataset and all results are available for comparison [9]. Below, we have highlighted the most recent available Senate race from a Presidential election year in each state, to match conditions as closely as possible. The data bottleneck is access to a precinct shapefile matching geography to voting patterns, which is surprisingly difficult. A database of our research group’s vetted shapefiles is available at [15]. For the results we present here, we prefer to use the statewide U.S. Senate election to the 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 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.

The algorithmically generated plans are not offered as a statistical experiment and come with no probabilistic claims about plan selection, but mainly provide existence proofs to illustrate how readily gameable partisan symmetry standards will be for those engaged in redistricting. The methods also produce many thousands of examples of plans that are paradoxical in the senses described above, where partisan symmetry metrics identify the wrong party as the gerrymanderers, relative to the common understanding of that term.

---

6 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 [1].

7 This is exactly the use of ensemble methods that is endorsed in [10] p176 as productive and compelling: a demonstration of possibility and impossibility.
4.2 Utah and the “Utah Paradox”

We begin with Utah, where the elections that were available in our dataset all come from 2016. 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. Figure 3 shows outcomes from our 100,000-step ensemble.

We note in passing that the amount of linear partisan swing needed to reverse the partisan advantage could 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.

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, 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 algorithmic ensemble) would deterministically impose a partisan outcome: the one in which Democrats are locked out of representation.

- Furthermore, the signed scores make an apparent 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 upper-bound the possible Republican share of a district 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 \( MM > 0 \), by the derivation in the last section. 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 finding. 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 stands to lock Democrats out of representation, and all plans with D representation are reported as significant R-favoring gerrymanders.

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.

---

8 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 McMullen, giving that race an extremely different pattern. The Governor’s race shows similar results to the Senate, as the reader can verify in [9].

9 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.
Figure 3: 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. Recall that the mean-median score reports a Republican advantage when \( \text{MM} > 0 \). The PG score is unsigned, but larger magnitude indicates greater asymmetry.
Figure 4: 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 have an outlying number of seats 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. Recall that the mean-median score reports a Republican advantage when MM > 0. The PG score is unsigned, but larger magnitude indicates greater asymmetry.
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 4 shows the partisan properties in the ensemble of plans, allowing us to compare extreme symmetry scores (the ostensible indicator of partisan gerrymandering) to extreme seat shares (the explicit goal of partisan gerrymandering). We find no evidence of correlation or any kind of correspondence.

Over 98\% of all 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 outlying plans with the most Democratic seats (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 the partisan outlier plans are not proscribed by strict symmetry thresholds. Plans with the extreme outcome of 28 or more Republican seats occur with higher frequency among the plans with $MM \approx 0$ than among the full set of sampled plans—more than twice as often, in fact. This shows that restricting to “good” symmetry scores will not prevent extreme seat counts.

Next we present the reverse perspective, considering how the plans with the most extreme seat counts score on partisan symmetry. The last row in Figure 4 shows histograms made only from the seat outliers: blue for plans with $\leq 21$ Republican seats and red for plans with $\geq 28$ Republican seats. We find that a significant number of extreme D-favoring plans paradoxically register as major Republican gerrymanders under the MM score. 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 seat counts. 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 ($\sim 53\%$ Republican share).

In this case, mean-median does much better than in Texas in terms of distinguishing the seat extremes: Figure 5 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. 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 6.
Figure 5: 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 have an outlying number of seats 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. Recall that the mean-median score reports a Republican advantage when $\text{MM} > 0$. The PG score is unsigned, but larger magnitude indicates greater asymmetry.
Figure 6: Each of these 13-district plans has 11 Republican-majority seats with respect to the SEN16 voting data, while having nearly perfect partisan symmetry: the PG score that describes the difference between the seats-votes curve and its reflection is near zero and in the best 2% of scores in the ensemble. These maps have PG scores of 0.0096, 0.0099, 0.0107, and 0.0115, respectively. This figure also illustrates the diversity of districting plans achieved by this Markov chain method.

5 Conclusion

In this note, we have characterized the partisan symmetry standard from [10] mathematically: under uniform partisan swing, it turns out to amount simply to a prescription for the arrangement of vote totals across districts (Theorem 2, Partisan Symmetry Characterization). 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.

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 reasoning would tell us not to 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.

If symmetry metrics measured something that was obviously of inherent value in the proper functioning of democratic representation, 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. Stated this way, it is more difficult to argue that symmetry captures an essential ingredient of civic fairness.

---

10 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. However, we know of no other detailed method that has been proposed for this task in practical or legal applications. See footnote.

11 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 or exogenous? Senate race or attorney general? Most single-score indicators have this problem.
Acknowledgments

MD is deeply grateful to Gary King for illuminating conversations about the partisan symmetry standard. All authors thank the other participants of the 2019 Voting Rights Data Institute at Tufts University, especially Cleveland Waddell, Brian Morris, Michelle Jones, and Thomas Weighill, for stimulating discussions. We are grateful to the Prof. Amar J. Bose Research Foundation at MIT and the Jonathan M. Tisch College of Civic Life at Tufts for their generous support of the Voting Rights Data Institute.

References

[1] Daryl DeFord, Moon Duchin, and Justin Solomon. Recombination: A family of Markov chains for redistricting. https://arxiv.org/abs/1911.05725, 2019.

[2] Moon Duchin, Taissa Gladkova, Eugene Henninger-Voss, Ben Klingensmith, Heather Newman, and Hannah Wheelen. Locating the representational baseline: Republicans in Massachusetts. Election Law Journal, 18(4):388–401, December 2019.

[3] James C. Garand and T. Wayne Parent. Representation, swing, and bias in U.S. presidential elections, 1872-1988. American Journal of Political Science, 35(4):1011–1031, November 1991.

[4] Heather K. Gerken, Jonathan N. Katz, Gary King, Larry J. Sabato, and Samuel S.-H. Wang. Amicus brief in Whitford v. Gill, 2018. No. 16-1161.

[5] Bernard Grofman. Measures of bias and proportionality in seats-votes relationships. Political Methodology, 9(3):295–327, 1983.

[6] Bernard Grofman and Ronald Keith Gaddie. Amicus brief in Rucho v. Common Cause, 2019. Nos. 18-422, 18-726.

[7] Bernard Grofman and Gary King. The future of partisan symmetry as a judicial test for partisan gerrymandering after LULAC v. Perry. 6 Election L.J. 2, 62–35, 2007.

[8] Voting Rights Data Institute. GerryChain. GitHub repository, 2018. https://github.com/mggg/gerrychain.

[9] Voting Rights Data Institute. Partisan Symmetry. GitHub repository, 2020. https://github.com/mggg/Partisan-Symmetry.

[10] Jonathan N. Katz, Gary King, and Elizabeth Rosenblatt. Theoretical foundations and empirical evaluations of partisan fairness in district-based democracies. American Political Science Review, 114(1):164–178, 2020.

[11] Gary King and Robert X. Browning. Democratic representation and partisan bias in congressional elections. American Political Science Review, 81(4):1251–1273, December 1987.

[12] Gary King, Bernard Grofman, Andrew Gelman, and Jonathan N. Katz. Amicus brief in LULAC v. Perry, 2005. Nos. 05-204, 05-254, 05-276, 05-439.

[13] Jonathan Mattingly. Expert report in Common Cause v. Lewis, 2019. https://sites.duke.edu/quantifyinggerrymandering/files/2019/09/Report.pdf see also https://sites.duke.edu/quantifyinggerrymandering/2019/06/26/the-animated-firewall/.

[14] Michael B. McDonald and Robin E. Best. Unfair partisan gerrymanders in politics and law: A diagnostic applied to six cases. Election Law Journal, 14(4):312–330, 2015.

[15] MGGG. MGGG States. GitHub repository, 2020. https://github.com/mggg-states.

[16] John F. Nagle. Measures of partisan bias for legislating fair elections. Election Law Journal, 14(4):346–360, 2015.
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 \( \delta \) with \( \delta_i = v_{i+1} - v_i \); and jump vector \( j \) with \( j_i = \frac{1}{k} + \overline{v} - v_{k+1-i} \), where \( \overline{v} \) is the mean of the \( v_i \). These expressions define \( j, \delta \) in terms of \( v \); neither \( j \) nor \( \delta \) completely determines \( v \) because they are invariant under translation of the entries of \( v \), but one additional datum (such as \( v_1 \) or \( \overline{v} \)) suffices, with \( j \) or \( \delta \), to fix the associated \( v \). In this section, we begin by expressing \( \text{PG} \) in terms of the jumps \( j \), then giving equivalent conditions for \( \text{PG} = 0 \) in terms of \( j, \delta, \) or \( v \).

As outlined above, \( \text{PG} \) measures the area between the seats-votes curve \( \gamma \) and its reflection. The shape of the region between those curves depends directly on the points \( j = (j_1, j_2, \ldots, j_k) \), 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 I makes it clear that it is complicated to decompose the integral into vertical rectangles in the style of a Riemann sum, because the jump in the curve and the \( 1 - j_i \) 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 \((j_i, \frac{i-1}{k})\), \((j_i, \frac{1}{k})\) 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 \((1 - j_{k+1-i}, \frac{1}{k})\), 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

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

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 \( \overline{v} \). The only condition on the gap vector \( \delta \) is that its entries sum to less than one. Putting these observations together we may define the set of achievable \( v, \delta, j \) respectively as

\[ \mathcal{V} = \{(v_1, \ldots, v_k) : 0 \leq v_1 \leq \cdots \leq v_k \leq 1 \}, \]

\[ \mathcal{D} = \{ (\delta_1, \ldots, \delta_{k-1}) : \delta_i \geq 0 \forall i, \sum \delta_i \leq 1 \}, \]

\[ \mathcal{J} = \{ (j_1, \ldots, j_k) : 0 \leq j_1 \leq \cdots \leq j_k \leq 1, \sum j_i = \frac{k}{2} \}. \]

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 \( \gamma \).

Now we can prove the Characterization theorem.

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

\[ \text{PG}(v) = 0 \quad \text{(1)} \]

\[ j_i + j_{k+1-i} - 1 = 0 \quad \forall i \quad \text{(2)} \]

\[ \frac{1}{2} (v_i + v_{k+1-i}) = \overline{v} \quad \forall i \quad \text{(3)} \]

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

\[ \delta_i = \delta_{k-i} \quad \forall i \quad \text{(5)} \]

**Proof.** The condition that \( \text{PG}(v) = 0 \) has been rewritten in terms of \( j \) above, and converting back to the \( v_i \) we get

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

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 \( i = \left[ \frac{k}{2} \right] \) in \( (3) \) to obtain \( v_{\text{med}} = \overline{v} \); in the other direction, average both sides over \( i \) in \( (4) \) to obtain \( \overline{v} = v_{\text{med}}. \) Finally, the symmetric gaps condition \( (5) \) is clearly equivalent to the symmetry of the values of \( v \) about the center \( v_{\text{med}} \), which is \( (4) \). \( \square \)
B  Supplement: Bounding partisan Gini in terms of mean-median

Recall that the mean-median score $\text{MM}$ 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 $\text{PG}$ is a non-negative score that simply quantifies the failure of symmetry, interpreted as a magnitude of unfairness. We easily see that $\text{PG} = 0 \iff \text{MM} = 0$ 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 $\text{discrep}(i) = \frac{v_i + v_{i+1}}{2} - \overline{v}$, measuring the difference between the average of a pair of vote shares from the average of all the vote shares. This gives

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

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

**Theorem 7.** The partisan Gini score satisfies

$$\begin{cases} 
\text{PG} \geq \frac{4}{k} |\text{MM}|, & k \text{ odd} \\
\text{PG} \geq \frac{8}{k} |\text{MM}|, & k \text{ even,}
\end{cases}$$

with equality when $k = 3, 4$.

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

$$\text{PG} = \frac{2}{k} \sum_{i=1}^{k} \text{discrep}(i) = \frac{2}{k} \left( |\text{discrep}(m)| + \sum_{i\neq m} |\text{discrep}(i)| \right) \geq \frac{2}{k} \left( |\text{discrep}(m)| + \left| \sum_{i\neq m} \text{discrep}(i) \right| \right)$$

$$= \frac{2}{k} \left( |\text{MM}| - |\text{MM}| \right) = \frac{4}{k} |\text{MM}|.$$

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

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

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

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

15