American Political Science Review (2020) 114, 1, 164–178
doi:10.1017/S000305541900056X
© American Political Science Association 2019

Theoretical Foundations and Empirical Evaluations of Partisan Fairness in District-Based Democracies

JONATHAN N. KATZ *California Institute of Technology*
GARY KING *Harvard University*
ELIZABETH ROSENBLATT *Harvard University*

We clarify the theoretical foundations of partisan fairness standards for district-based democratic electoral systems, including essential assumptions and definitions not previously recognized, formalized, or in some cases even discussed. We also offer extensive empirical evidence for assumptions with observable implications. We cover partisan symmetry, the most commonly accepted fairness standard, and other perspectives. Throughout, we follow a fundamental principle of statistical inference too often ignored in this literature—defining the quantity of interest separately so its measures can be proven wrong, evaluated, and improved. This enables us to prove which of the many newly proposed fairness measures are statistically appropriate and which are biased, limited, or not measures of the theoretical quantity they seek to estimate at all. Because real-world redistricting and gerrymandering involve complicated politics with numerous participants and conflicting goals, measures biased for partisan fairness sometimes still provide useful descriptions of other aspects of electoral systems.

INTRODUCTION

Partisan fairness in modern democracies is defined at the intersection of two grand representative institutions—political parties and district-based electoral systems. Whereas parties are mostly shaped by voters and candidates, the contiguous geographic districts that collectively tile a political system's landmass constitute the playing field on which the parties compete. This intersection is most visible during US redistricting processes, which are partisan and often explosively conflictual, but it is also crucial for the fairness of relatively fixed districts, such as the US Senate and electoral college, and many other district-based representative systems around the world.

In this article, we clarify the theoretical foundations of "partisan symmetry," the most widely accepted standard of partisan fairness, along with the standards not based on symmetry. We reveal definitions that have not been formalized, essential assumptions not discussed, and quantities of interest at best only implicitly defined. We pare down assumptions to their essential components, and then, to shore up or choose among assumptions with observable implications, we offer empirical evidence from 105,001 district-level state legislative elections in the US. (These data, which we arranged to make public, may be the largest collection of election data ever analyzed at once; see Klarner 2018.)

Although the literature dates back more than a century, the public outrage over what appear to be increasingly effective (and visually obvious) gerrymanders, court challenges in numerous jurisdictions, the proliferation of ballot initiatives, active anti-gerrymandering lobby groups, and new legislation in many states have contributed to a resurgence of scholarly interest in this field. In addition, the many scholars of redistricting who have served as expert witnesses, combined with the US Supreme Court's inability to speak with a unified voice on this topic, its highly predictable partisan divisions, and its unpredictable decisions that lack respect for precedent (most recently with partisan gerrymandering declared justiciable and then not justiciable), have also motivated scholars to publicly offer much judicial, legislative, and political advice.

Political scientists—as well as scholars from economics, statistics, computer science, mathematics, neuroscience, genetics, psychology, systems biology, physics, sociology, public policy, electrical engineering, law, computational biology, and other fields—have contributed many new approaches and generated much progress in recent years, although their differing perspectives, notation, methods, and theories have led to a scholarly chaos that may even rival some of the world we study. We thus use the theory and evidence introduced here to begin to unify this literature, to make sense of contrasting perspectives, and to build a more

Jonathan N. Katz, Kay Sugahara Professor of Social Sciences and Statistics, California Institute of Technology, jkatz@caltech.edu.
Gary King, Albert J. Weatherhead III University Professor, Institute for Quantitative Social Science, Harvard University, King@Harvard.edu.
Elizabeth Rosenblatt, Post-BA Affiliate, Institute for Quantitative Social Science, Harvard University, ERosenblatt@alumni.harvard.edu.

All appendices are available at GaryKing.org/symmetry; all data and information necessary to replicate our results are available in Katz, King, and Rosenblatt (2019). Our thanks to Steve Ansolabehere, Peter Aronow, Robin Best, Jowei Chen, Shawn Donahue, Moon Duchin, Jon Eguia, Andrew Gelman, Kosuke Imai, Jonathan Krasno, Eric Lander, Daniel Magleby, Michael D. McDonald, Eric McGhee, John Nagle, Nate Persily, Jameson Quinn, Nick Stephanopoulos, Tyler VanderWeele, Sam Wang, Greg Warrington, and Xiang Zhou for their helpful comments.

Received: December 5, 2018; revised: May 29, 2019; accepted: August 29, 2019; First published online: October 23, 2019.

https://doi.org/10.1017/S000305541900056X Published online by Cambridge University Press

solid foundation for future progress. To do this, we deploy a crucial principle of statistics that is often ignored in this literature—defining the quantities of interest rigorously and separately from the measures used to estimate them. This then enables us to apply standard statistical approaches to evaluate the many existing measures of partisan fairness. We distinguish between measures that are statistically appropriate and those that are in fact biased, limited, or not measures of the quantity they seek to estimate at all. We also show how some measures biased for a coherent standard of partisan fairness can still reveal other interesting features of complicated electoral systems unrelated to partisan fairness.

In the sections to follow, we define the partisan symmetry standard, consider alternative non-symmetry–based standards, clarify estimation assumptions, evaluate existing measures, and conclude. Online Appendices include extensive supporting evidence.

THE PARTISAN SYMMETRY STANDARD

In this section, we describe the partisan symmetry standard for a single member district, where it is easier to understand and afterward generalize it to an entire legislature. The concept of fairness-through-symmetry can be traced to "The Golden Rule" (part of almost every ethical tradition; Blackburn 2003) and the Bible (*Genesis* 13:8–9, *Matthew* 7:12). Even experts on opposing sides of the same court cases typically support partisan symmetry (Grofman and King 2007, 15).

Partisan symmetry, as an outcome-based measure, is of course distinct from process-oriented measures, such as nonpartisan redistricting commissions or bipartisan agreements (Cox 2006; McDonald 2007; Rave 2012). However, the two are related: Fair redistricting processes have consistently been shown to lead to symmetric outcomes, and processes controlled by partisan gerrymanderers predictably lead to asymmetric outcomes (Gelman and King 1994b; King and Gelman 1991). Obviously, a process designed to be "fair" that turns out to be substantially asymmetric, if there is a fairer alternative, would not normally be regarded as a legitimate outcome, absent other considerations.

Symmetry in a Single-Member District

Although our results generalize to any number of political parties (as in Katz and King 1999; King 1990), we use two parties throughout to simplify exposition. We also assume an odd number of voters to eliminate the possibility of a tie (or assume a coin flip in that instance) and then denote the Democratic proportion of the (two party) vote in district d as v_d (for $d = 1,\ldots, L$). In one single-member district, denote the *plurality voting rule* as $s(v) = \mathbf{1}(v > 0.5)$, which takes on the value one if $v > 0.5$ (meaning the Democratic candidate wins) and 0 otherwise (the Republican wins). In other words, when a political party receives more votes than any other party, it wins the seat. The reason this rule is universally judged as fair is because it is symmetric: It applies the same way to *any* party, regardless of its name or identity.

We formally express *district-level partisan symmetry* (like "neutrality" in formal theory; May 1952, 681–82) as $s(v) = 1 - s(1 - v)$, for all v. In other words, if we swapped the labels on the parties, nothing would change other than who wins the seat. For example, if the Democratic party received 0.55 of the vote in a district, it would win the seat because $s(0.55) = 1$, and if (instead) the Republican party received 0.55 of the vote, it would receive the seat because $1 - s(1 - 0.55) = 1$, the symmetric outcome.

Deviations from partisan symmetry in a single-member district, first-past-the-post electoral system can stem from fraud. For example, if a criminal surreptitiously stuffs the ballot box with an extra 0.1 Democratic proportion of the vote, then the Democratic party will win the seat if it receives more than 0.4 (rather than 0.5) of the votes—that is, $s'(v) = \mathbf{1}(v > 0.4)$, for all v—which is obviously not symmetric. To see this asymmetry formally, consider that a Democratic candidate receiving 0.45 of the vote would win the seat, $s'(0.45) = 1$, but a Republican candidate who (instead) receives the same proportion of the vote would lose: $1 - s'(1 - 0.45) = 0$.

Symmetry in a Legislature

We now show how partisan symmetry applies to fairness for an entire legislature.

The Seats–Votes Curve

We define here the seats–votes curve from its component parts. Denote the *populace* $\mathbb{P}$, the set of all individuals living in a state, including systematic patterns in their electoral behavior (or nonbehavior); an *electoral system*, $\mathbb{E}$, all factors that turn the populace's votes into seats, including district boundary lines, district-level voting rules (such as plurality voting), and whether the rules are followed (Cox 1997, 38); and other *measured exogenous influences* on voter behavior, X, such as demographic variables (e.g., percent African American or immigrant), candidate quality (e.g., incumbency status or uncontestedness), voter behavior (such as lagged vote), and campaign events. Together $\{\mathbb{P}, \mathbb{E}, X\}$ determine a "permutation invariant" joint probability density from which district-level vote proportions are drawn, $p(v_1,\ldots, v_L|X)$ (King 1989).[1]

Next, we aggregate the district vote proportions into the statewide average district vote $V = V(v_1, \ldots, v_L) = \text{mean}_d(v_d)$ and the statewide seat proportion $S = S(v_1, \ldots, v_L) = \text{mean}_d[s(v_d)]$, with $s(v_d)$

[1] Because all measures discussed are invariant to permutations of district labels, we only require densities specified up to a permutation of its arguments; e.g., $p(v_1, v_2, v_3|X) = p(v_1, v_3, v_2|X)$ (Wimmer 2010, 114). Permutation invariance is considerably less restrictive than exchangeability.

https://doi.org/10.1017/S000305541900056X Published online by Cambridge University Press

defined in the previous section.[2] Electoral systems $\mathbb{E}$, including changes such as redistricting, are important because sets of district votes that differ, $\{v_1,\ldots, v_L\} \neq \{v_1',\ldots, v_L'\}$, but which aggregate into the same average district vote $V(v_1,\ldots, v_L) = V(v_1',\ldots, v_L')$, can yield different statewide seat proportions $S(v_1,\ldots, v_L) \neq S(v_1',\ldots, v_L')$.[3]

We then define the *seats–votes function* by taking the expected value of the statewide seat proportion $S(v_1,\ldots, v_L)$ over the density $p(v_1,\ldots, v_L|\, X)$, constrained so that $V = \text{mean}(v_d)$:

$$E_p\left[S(v_1,\ldots,\, v_L)|X, \underset{d}{\text{mean}}(v_d) = V\right] = S(V|\mathbb{P},\mathbb{E},X) \equiv S(V). \quad \textbf{(1)}$$

The seats–votes function is a scalar property of the electoral system computed from random variables $\{v_1,\ldots, v_L\}$ and V, along with fixed characteristics X (King 1989). A coherent seats–votes function is defined independently of the observed realizations $\{v_1^{\text{O}},\ldots,\, v_L^{\text{O}}\}$ (and in turn independently of the observed realization of the average district vote V^{O}). We call this the *Stable Electoral System Assumption*:

Assumption 1 [SESA: Stable Electoral System]. *The probability density of district vote proportions is defined independently of any one set of realized district vote proportions:*

$$p(v_1,\ldots,\, v_L|X, v_1^{\text{O}},\ldots,\, v_L^{\text{O}}) = p(v_1,\ldots,\, v_L|X).$$

Assumption 1 can be thought of as Markov independence such that an election does not change the electoral system that generates vote proportions (after conditioning on X). However, the assumption will usually be applied to data from one election in isolation, at that one time point, with independence applying over hypothetical replications from the same (stable) electoral system. Violations of this assumption occur when certain (hypothetical) election outcomes prompt a new redistricting controlled by a different party or group, or if an electoral realignment changes the coalitions making up the parties (unless encoded in X). This seats–votes curve would then be incoherent because the electoral system it describes is not stable as it is defined differently depending on the observed vote. A simple numerical example of a violation of SESA, and no single seats–votes curve, is if $S(0.6) = 0.7$ for an election with $V^{\text{O}} = 0.6$ but $S(0.6) = 0.8$ following an election with $V^{\text{O}} = 0.5$. (Online Appendix A shows why the seats–votes curve is a function of the average district vote rather than the statewide proportion of the total vote.)

SESA implies that the seats–votes function is *single-valued*, and not dependent on the election outcome, so that a complete representation of all values of $S(V)$ for populace $\mathbb{P}$, electoral system $\mathbb{E}$, and covariates X is the set $\mathcal{S} = \{S(V) : V \in [0,1]\}$, which we call the *seats–votes curve*. If SESA does not hold, then the seats–votes function is not single-valued and the seats–votes curve is not coherently defined and, as such, concepts like partisan symmetry cannot even be evaluated. Including sufficiently informative variables in X can correct for a violation of this assumption. If SESA holds, then we still need to consider how to estimate it, a subject we address in the section on estimation, below. (SESA is related to the "Stable Unit Treatment Value Assumption," SUTVA, commonly made in the causal inference literature; see Iacus, King, and Porro 2018; Rubin 1991; VanderWeele and Hernan 2012.)

Defining Symmetry

Here we (a) define the concept of partisan symmetry in a legislature, (b) offer point summaries of it, and then (c) discuss types of symmetric and asymmetric electoral systems. The most commonly accepted standard for fairness of voting in a legislature is statewide *partisan symmetry* (as proposed in King and Browning 1987), which we write as follows:

Definition 1 (Partisan Symmetry). *An electoral system satisfies the partisan symmetry standard if* $S(V) = 1 - S(1 - V)$, *for all* $V \in [0, 1]$.

Because of the impact of districting, even if $s(v) = 1 - s(1 - v)$ holds for every individual district, statewide partisan symmetry may not hold.

A deviation from partisan symmetry is known as *partisan bias*, which we define formally as follows:

Definition 2 (Partisan Bias). *Partisan bias is the deviation from partisan symmetry:* $\beta(V) = \{S(V) - [1 - S(1 - V)]\}/2$, *for any* $V \in [0, 1]$.

The quantity $\beta(V)$ is the (perhaps negative) proportion of seats that would need to be taken from the Democrats (and thus given to the Republicans) to make the system fair. (The division by two makes $\beta(V)$ the distance from each party to symmetry, as desired, rather than to each other.) Thus, special cases of partisan bias include (a) partisan symmetry, where $\beta(V) = 0$; (a) Democratic bias, where $\beta(V) > 0$; and (c) Republican bias, $\beta(V) < 0$. Although $\beta(V)$ is defined for any $V \in [0, 1]$, only half this range is needed, say $V \in [0.5, 1]$, because $\beta(V) = \beta(1 - V)$. (Partisan bias is unrelated to statistical bias, where the expected value of an estimator is not equal to the population quantity of interest.)

The chosen value of V in a seats–votes function must be a *possible* result of the electoral system so that there is a defined value of $S(V) \in [0, 1]$. For example, if one party would not tolerate the other party winning, so that war would break out and end the democracy if say $V > 0.5$, then $S(V)$ would be undefined for $V > 0.5$. Similarly, a party system defined based on fixed ethnic or racial divisions would mean that only slight variations in V from V^{O} would be possible (due to changes in turnout or demographic change).

[2] For set A with cardinality $\#A$, define the mean over i of function $g(i)$ as $\text{mean}_{i\in A}[g(i)] = \frac{1}{\#A}\sum_{i=1}^{\#A} g(i)$. When there is no ambiguity, we simplify notation by letting $\sum_d \equiv \sum_{d=1}^{D}$ and $\text{mean}_d \equiv \text{mean}_{d\in A}$.

[3] Redistricters often make calculations like these by assuming that individual votes are fixed, at least with respect to redistricting. Although convenient and often empirically accurate, this assumption is unnecessary. See Online Appendix D.

https://doi.org/10.1017/S0003055419000056X Published online by Cambridge University Press

The rotation in office principle (see Petracca 1996) says that it is conceivable for both parties to win office, if enough elections are run under the same electoral system:

Assumption 2 [Rotation in Office]. *For a given electoral system and "average district vote victory size" parameter* $\eta \in [0, 0.5]$ *chosen by the researcher, the range of possible values for the average district vote is no smaller than* $V \in [0.5 - \eta, 0.5 + \eta]$.

This assumption allows the range of possible vote proportions to be asymmetric, so long as it has as a subset a smaller symmetric range (e.g., [0.4, 0.8] includes [0.4, 0.6], so that $\eta = 0.1$). With the possible victory size parameter set to its maximum, $\eta = 0.5$, any value of $V \in [0, 1]$ may be used with $S(V)$ so that, e.g., the full version of partisan bias in Definition 2 can be used. We allow η to take smaller values so that special cases of the partisan symmetry standard can apply in electoral systems where certain lopsided outcome sizes are inconceivable as long as a symmetric range exists. For example, for $\beta(0.5)$, we can use $\eta = 0$. Generally, the range of conceivable values of V may be larger than $[0.5 - \eta, 0.5 + \eta]$. Although Assumption 2 is defined in terms of possible electoral outcomes, those that are exceedingly unlikely, such as Washington DC voting overwhelming Republican, do not violate this assumption but may generate model dependence in estimation (see the section on estimation below).

Online Appendix B discusses extending these ideas to noncompetitive electoral systems.

Summaries

Partisan bias may be summarized as (a) bias at 0.5, $\beta(0.5) = S(0.5) - 0.5$; (b) bias at another point such as $\beta(0.55) = \{S(0.55) - [1 - S(1 - 0.55)]\}/2 = \beta(0.45)$; (c) an indicator as in for whether $\mathbf{1}(V > 0.5) = \mathbf{1}[S(V) > 0.5]$ (Best et al. 2018); or (d) an average over a range of vote values, such as $E[\beta(V)] = \int_{0.5}^{0.55} \beta(V)p(V)\mathrm{d}V$, where $p(V)$ is the predictive density of likely votes or a uniform with range based on plausible average district vote values.

These summaries are easier to estimate than the entire curve, but if a summary differs from the value of partisan bias for other empirically reasonable values of V, then an electoral system judged to be fair by the summary can instead turn out to be biased in a real election. This pattern may even be intended by gerrymanderers who sometimes misjudge their likely average district vote and, instead of having an electoral system biased in their favor such as by winning a large number of districts by a small amount, they have one massively biased against them by losing all by a small amount.

For competitive electoral systems, (d) can be a reasonable summary if the values of V we are likely to observe are included in the specified range. In contrast, (a) is best used with another assumption because, even when $\beta(0.5) = 0$, $\beta(V)$ may be far from 0 for any other value of V. Summary (d) will normally be the most statistically stable of the three. These warnings do not mean that summaries should not be used, only that they come with an assumption that needs to be understood.

Types of Symmetry and Asymmetry

We order electoral systems meeting the partisan symmetry standard by the size of the bonus going to the statewide majority vote winner or, in other words, by the degree of *electoral responsiveness*, of $S(V)$ to changes in votes V, as follows:

Definition 3 (Electoral responsiveness). *Electoral responsiveness, which quantifies how much the statewide seat proportion is altered by a change in the average district vote, is* $\rho(V) = \partial S(V)/\partial V$.

Because the number of legislative seats is discrete, seats–votes curves are inherently discrete, and $\rho(V)$ is not uniformly continuous. Thus, in practice, the curve is summarized by smoothing via a discrete derivative $\rho(V, V') = [S(V) - S(V')]/(V' - V)$, given chosen values V and V'. We will use the shorthand $\rho(V)$ to refer to both the theoretical continuous and discrete quantities.[4]

Electoral responsiveness is commonly summarized at (a) $\rho(0.5)$; (b) an empirically reasonable value such as $\rho(V^{\mathrm{O}})$, where V^{O} is the observed average district vote for a real election; or (c) an empirically reasonable range, such as $\rho(0.45, 0.55)$.

We first use Definition 3 to define a minimal standard for a fair democratic electoral system, which we call *symmetric democracy*:

Definition 4 (Symmetric democracy). *An electoral system characterized by symmetric democracy satisfies (a)* partisan symmetry *(Definition 1), (b)* nonnegative responsiveness, $\rho(V) \geq 0$ *for all* V, *and (c)* unanimity, $S(0) = 0$.

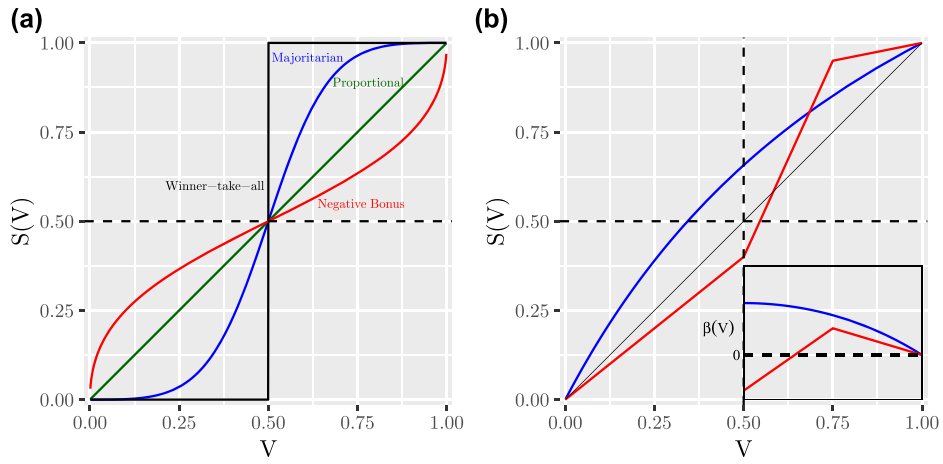
Conditions (a) and (c) imply also that $S(1) = 1$. Conditions (b) and (c) imply, for at least one point in $V \in [0, 1]$, that $\rho(V) > 0$. Condition (c) is referred to as "unanimity" or the "Pareto principle" in social choice theory (Sen 1976). (We suggest a modification of condition (c) in the section on non-symmetry–based standards and in Online Appendix B for when one party is unlikely to ever win a majority of votes.)

Four ranges of electoral responsiveness that satisfy Definition 4 are often discussed, each of which we illustrate with a fair seats–votes curve in the left panel of Figure 1. First, *proportional representation* meets the partisan symmetric standard because $S(V) = V$ and $1 - S(1 - V) = V$, or in other words $\rho(V) = 1$ and $\beta(V) = 0$ for all V (green line in the figure). Legislatures with single member, plurality voting systems are not guaranteed to be proportional by law and tend to be *majoritarian* by empirical pattern, which means that they usually give a bonus to the party winning a majority of votes statewide, with $1 < \rho(V) < \infty$ (see blue line). For example, suppose the Democrats receive $V = 0.55$ proportion of the average district vote statewide and,

[4] Alternatively, let $V^{\mathrm{L}} = \arg\min_{V^{\mathrm{L}}} S(V^{\mathrm{L}}) = S(V)$ and $V^{\mathrm{H}} = \arg\min_{V^{\mathrm{H}}} S(V^{\mathrm{H}}) > S(V)$. Then $\rho(V) = [S(V^{\mathrm{H}}) - S(V^{\mathrm{L}})]/(V^{\mathrm{H}} - V^{\mathrm{L}})$, where $(V^{\mathrm{H}} - V^{\mathrm{L}})$ is the minimally sufficient change in the average district vote necessary to produce one additional legislative seat.

https://doi.org/10.1017/S000305541900056X Published online by Cambridge University Press

FIGURE 1. Types of Seats–Votes Curves



Notes: Left panel: Symmetric (fair) curves with differing levels of electoral responsiveness. Right panel: Asymmetric (biased) curves, including one consistently biased toward the Democrats (blue) and one with biases favoring different parties depending on V (red); the inset graph is for $\beta(V)$ for $V \in [0.5, 1]$ with the vertical axis scaled to be the same as the main plot, and lines color coded to the seats-votes curves.

because of how the district lines are drawn, receive $S(0.55) = 0.75$ proportion of the seats. This is not proportional, but it would be fair according to partisan symmetry because the Republicans, when receiving $1 - V = 0.55$ proportion of the vote, would also receive $1 - S(1 - 0.55) = 0.75$ proportion of the seats. Third, a more extreme type of electoral system still meeting partisan symmetry is *winner-take-all* (with $\rho \to \infty$), where the majority vote winner receives all of the seats (solid black line in the left panel of Figure 1). A final type of system that meets partisan symmetry is where the party winning a majority of votes receives a *negative bonus* ($0 < \rho < 1$); for example, if $S(0.65) = 0.55$ and $1 - S(1 - 0.65) = 0.55$ (red line).

Although partisan symmetry is widely viewed as a required standard of fair electoral systems (absent other considerations; see Online Appendix B), different levels of electoral responsiveness may reasonably be chosen as preferable or appropriate for and by different people and governments. Many would prefer that their electoral system meet partisan symmetry but not be proportional, winner-take-all, or negative bonus, and so would impose the restrictions of an unbiased ($\beta(V) = 0$) majoritarian ($1 < \rho < \infty$) electoral system. Similarly, although no US state constitution rejects partisan symmetry, the constitutions differ in their requirements regarding electoral responsiveness. Some state constitutions require their redistricters to draw highly responsive districts, to encourage competitive elections and party change in office, whereas others encourage their redistricters to draw minimally responsive districts, which protects their incumbents, perhaps to help them gain experience or seniority and thus power on congressional committees. Brunell (2010) even argues that less responsiveness (and thus less competitiveness) produces happier constituents (see also Gerber and Lewis 2004, 1378).

We also distinguish between two types of electoral systems that deviate from partisan symmetry—(a) those biased consistently in favor of one party and (b) those that switch from biased in favor of one party to the other as V changes. The right panel of Figure 1 gives one example of each of these seats–votes curves, along with an inset graph at the lower right with $\beta(V)$ plotted by V and color-coded to the corresponding seats–votes curve. The blue seats–votes curve indicates bias in favor of the Democratic party for every value of V, although by different amounts. We can see this by the corresponding blue line in the inset graph. For example, at $V = 0.5$, $S(V) = 0.66$, and so $\beta(0.5) = (0.66 - 0.5)/2 = 0.08$, which is also the height of the left end of the blue line in the inset graph (although numbers on the vertical axis of the inset graph have been removed to reduce clutter, distances from zero are the same as for the main graph). Whereas the blue $\beta(V)$ line in the inset graph is always above zero, indicating consistent bias toward the Democrats for all V, the red line indicates bias toward the Republicans for $V < 0.125$ and toward Democrats for larger average district vote values. Partisan bias that switches parties with V is important to consider when using summary measures of bias to represent the entire seats–votes relationship (Grofman and Brunell 2005). This type of seats–votes curve can also be the result of a gerrymandering strategy where the party in control draws district maps biased against it at values of V it sees as unlikely, so long as the same map has more bias in its favor at values of V in future elections it sees as likely.

NON-SYMMETRY–BASED STANDARDS

We consider here alternatives to and modifications of the partisan symmetry standard by studying

https://doi.org/10.1017/S000305541900056X Published online by Cambridge University Press

the effects of two variables that characterize every redistricting—partisan gerrymanderers' control over the redistricting process and the competitiveness of the party system. We first show how the goals of partisan gerrymandering affects electoral systems in terms of bias and responsiveness, and how these can differ, depending on competitiveness, from the often misleading "cracking and packing" stereotype used in the literature. We then show how a pure partisan gerrymandering perspective suggests alternative, but ultimately unsatisfactory, normative definitions of partisan fairness. (See also Online Appendix B, which discusses fairness standards for noncompetitive party systems and semi-permanent minority groups.)

Gerrymandering Goals

Consider an imaginary partisan gerrymanderer focused solely on advantaging the respective political party.[5] Partisan gerrymanderers use their knowledge of voter preferences and their ability to draw favorable redistricting plans to maximize their party's seat share. Gerrymanderers do not necessarily care about voter support, the efficiency of the translation of votes into seats, partisan bias, electoral responsiveness, or differential turnout—unless it helps them win more seats.

We show that these goals, when mapped into the concepts of partisan bias and electoral responsiveness, can be either consistent with or the opposite of those commonly described in the literature. Consider four situations, each of which leads to a different optimization function, effect on symmetry, and goal for bias and responsiveness (see Cox and Katz 1999, sec. 3.3, Friedman and Holden 2008; Puppe and Tasnadi 2009).

First is when the gerrymanderer is *running scared* (Mann 1978) and so is worried about what the party's statewide voting support may be in future elections. Here, optimizing means trying to win maximal seats with a safe margin, to insulate the party from potentially unfavorable future partisan swings. In this case, optimizing means seeking *high bias* and *low responsiveness*. Operationally, the gerrymanderer may do this by "packing" overwhelming numbers of opposition party votes into a few otherwise unwinnable districts and "cracking" the remaining opposition voting strength across a large number of districts to win each by a smaller but sufficiently safe vote margin. High bias helps the party in control of redistricting and low responsiveness protects their incumbents by locking in these gains for future elections.

Second is the opposite situation where the gerrymanderer is *confident of a statewide majority* of votes and so tries to make each district a microcosm of the entire state (i.e., $v_d = V$ for all d), producing a winner-take-all outcome overall (Cox and Katz 2002). In other words, the goal is an electoral system with *low bias and high responsiveness*. The "low bias" result is the consequence of optimizing primarily for high responsiveness, without preparing for the highly improbable situation where $V = 1 - V^{O}$. This situation involves neither packing nor cracking: If a Democratic gerrymanderer thinks his or her party can count on 55% of the total statewide vote, then packing, to give the Republicans a few seats, would be foolish and cracking, to win any seats by 50% plus a few votes, is irrelevant. Instead, the goal would be to win with $v_d = 0.55$ for all d. (Unless of course the gerrymanderer is overconfident about partisan swing and winds up losing all the districts; see Grofman and Brunell 2005.)

Third is where a partisan gerrymanderer reaches agreement with the other party (perhaps imposed by a redistricting commission). The result is a *bipartisan gerrymander*, which winds up optimizing for *low bias and low responsiveness*. Bias would be low because it is a zero-sum compromise between the parties, and low responsiveness reduces uncertainty in future elections by locking in the deal and protecting incumbents in both parties.

Finally, we consider the logical possibility of *high bias and high responsiveness* which, as it turns out, is rarely feasible in practice because of the dependence between the two quantities, as when additional packing to increase bias almost by definition reduces responsiveness. We can see this empirically in most real districtings (which we will see at scale when we get to Figure 3). Also, under a strategic behavior assumption, this is rarely an optimal strategy (Cox and Katz 2002).

Gerrymandering-Based Fairness Standards

We offer two ways of deriving a normative standard of partisan fairness from a purely partisan gerrymandering perspective.

First, consider as a thought experiment disallowing a redistricter to use knowledge of where its party's supporters live. This idea, which is equivalent to randomly permuting party labels on voters or on the gerrymanderer's voter forecasts, clearly removes *intent* to do harm. This step alone may be of value because human psychology and most judicial systems judge intentional harm more severely than accidental harm (Greene 2009). However, because plans drawn without knowledge of party support are drawn randomly with respect to party, any plan can be selected regardless of the degree of bias, responsiveness, or any other feature. In other words, gerrymandering without the knowledge of party removes intent but not harm. In fact, one possible districting plan that can occur is identical to that drawn by a partisan gerrymanderer with full knowledge

[5] This gerrymanderer is imaginary because those involved in redistricting balance numerous factors in addition to partisan gain (Gelman and King 1994a). These include incumbent protection or pairing, changing ideological polarization (McCarty, Poole, and Rosenthal 2009) or the legislature's median voter (Herron and Wiseman 2008), maintaining communities of interest, changing district compactness (Kaufman, King, and Komisarchik Forthcoming), not splitting local political subdivisions, keeping an incumbent's children's schools or parents houses in or out of their districts, keeping good challengers' homes out of certain districts, state legislators drawing congressional districts for them to run in, optimizing turnout differentials, swapping populations to hurt or encourage retirement of certain incumbents, and many others (Owen and Grofman 1988; Cox and Katz 2002, 39ff; Yoshinaka and Murphy 2009).

https://doi.org/10.1017/S000305541900056X Published online by Cambridge University Press

of where its party's voters live. In summary, the absence of intentional unfairness is not the same as fairness.

Second, we compare the efficiency of each party's translation of votes into seats. In one observed election, the Democratic party receives $S(V^{\mathrm{O}})$ seats given V^{O} votes and the Republican party receives $1 - V^{\mathrm{O}}$ votes and $1 - S(V^{\mathrm{O}})$ seats. Which of these parties has a better or more efficient translation of seats into votes? Unless it happens that $V^{\mathrm{O}} = 1 - V^{\mathrm{O}} = 0.5$, this is an apples to oranges comparison because of the two different starting points. Making the vote comparison between the two parties comparable in any one election can be accomplished by imposing a counterfactual assumption. We consider two possibilities.

One possibility is to make an assumption enabling us to estimate what would happen if the parties switched their vote proportions, so that the election result was $1 - V^{\mathrm{O}}$ rather than V^{O}. This would allow us to estimate the unobserved seat proportion $1 - S(1 - V^{\mathrm{O}})$ and compare it with $S(V^{\mathrm{O}})$. This of course leads exactly to partisan symmetry.

Second, we could try assuming away the differential meaning of all, or some particular type of, votes cast for each party (e.g., "wasted votes," which are those cast for the losing party in a district, and possibly also those above 0.5 plus one vote in winning districts; see the section on the efficiency gap below). However, although all votes are observed, asserting that all or any subset has the same meaning for each party, when the parties have different expected vote proportions, requires an assumption with the same ontological status as imagining partisan swings that lead to partisan symmetry. For example, suppose the Democrats receive $V^{\mathrm{O}} = 0.6$ and are confident of a statewide majority in subsequent elections under the same redistricting plan. Then, the votes cast for each party in specific districts have markedly different meanings for Democrats than for Republicans now in the minority, with $1 - V^{\mathrm{O}} = 0.4$ votes. The Democrats in this scenario would benefit from districts that look like a microcosm of the state, whereas Republicans would benefit most by packing and cracking, and so assuming that these votes have the same meaning would be a stretch at best. Clearly, this does not seem like a promising direction for developing a new standard for partisan fairness.

ASSUMPTIONS FOR ESTIMATING SEATS–VOTES CURVES

We show, under three types of assumptions, how to estimate the full seats–votes curve and its features. Appendices C and D discuss assumptions for forecasting and modeling individual voters.

No Additional Assumptions

In what is usually the best case, where we have five elections occurring between the decennial censuses and thus which we could consider (close to being) under the same electoral system, we observe five data points $\left\{\hat{S}(V_t^{\mathrm{O}}), V_t^{\mathrm{O}} : t = 1, \ldots, 5\right\}$, where the observed statewide seat proportion $\hat{S}(V_t^{\mathrm{O}})$ is an estimate of the expected value $S(V_t^{\mathrm{O}})$ in election t.

From these data, two unusual circumstances may enable one to compute partisan bias with no modeling assumptions. In the first, if we happen to observe an election with a tied average district vote, $V^{\mathrm{O}} = 0.5$, then $\beta(0.5)$ is estimable simply by the observed seat proportion. In the second, an even luckier situation (encompassing the first), two elections happen to be observed under the same electoral system with average district vote proportions symmetric around 0.5. For example, in Wisconsin State House elections run under the same redistricting plan, $V^{\mathrm{O}} = 1 - 0.48$ in 2012 and $V^{\mathrm{O}} = 0.48$ in 2014 and where, as a result, statewide seat proportions were observed in each election. In this particular case, the results indicate severe bias favoring the Republicans because of the dramatic seat proportion differences: $1 - \hat{S}(1 - 0.48) = 0.6$ but $\hat{S}(0.48) = 0.36$ (approximately), and so $\hat{\beta}(0.48) = -0.12$. (This election was the subject of the Supreme Court case, *Gill v Whitford*, 585 US (2018).)

Functional Form Assumptions

One type of assumption is to specify a class of parametric functional forms for the seats–votes relationship and to estimate its parameters. Two examples of this form are *linear* (Tufte 1973),

$$S(V) = \alpha_0 + \alpha_1 V, \tag{2}$$

and (redefining parameters α_0 and α_1) *bilogit* (King and Browning 1987):

$$S(V) = \frac{1}{1 + \exp\left[-\alpha_0 - \alpha_1 \ln\left(\frac{V}{1-V}\right)\right]}. \tag{3}$$

In each equation, α_0 and α_1 are related in different ways to partisan bias and electoral responsiveness, respectively (and because $S(V)$ in each expression is an expected value, real data need not fit either form exactly). For example, we drew the fair seats–votes curves in Figure 1 with $\alpha_0 = 0$ for all four and $\alpha_1 = \{0.5, 1, 3, 10{,}000\}$ (10,000 being sufficiently close, for our figure, to winner-take-all, which is $\alpha_1 \to \infty$).

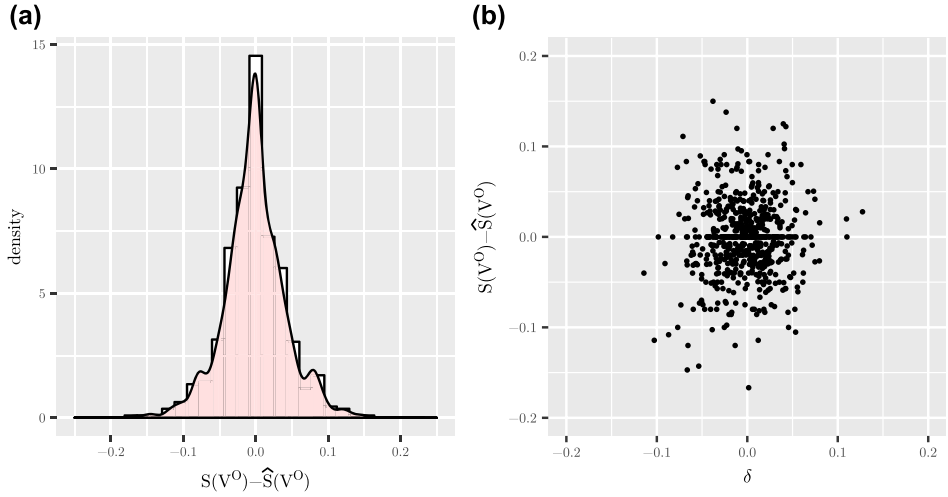
Once we estimate the seats–votes curve, we can then read off point estimates of $S(V)$ given any chosen V. This method enables one to compute partisan bias or any quantity of interest from the resulting estimated curve, along the appropriate level of uncertainty. Unfortunately, the few available observations from any one redistricting plan means that the result is usually uncertain and model dependent (and nonparametric approaches are not useful). As such, this strategy tends to be used more often for academic study of broad patterns across many electoral systems than for practical use evaluating individual redistrictings.

Partisan Swing Assumptions

An alternative approach is to use as inputs the district-level vote proportions. From this, we can estimate a single point on the seats–votes curve, $S(V^{\mathrm{O}})$, at the

https://doi.org/10.1017/S000305541900056X Published online by Cambridge University Press

FIGURE 2. Error Rates of Out-Of-Sample Statewide Seat Proportion Predictions Based on the Assumption of Uniform Partisan Swing



Notes: Histogram of errors (left panel) and errors by statewide partisan swing, δ, from one observed election to the next in predicting the seat proportion.

observed statewide vote V^O and then we can add an assumption to generate hypothetical elections from the same electoral system, for different V.

Fortunately, patterns in electoral data throughout the US and most parts of the world can usually be decomposed into (a) the average partisan swing from one election to the next affecting all districts and (b) the relative ordering of district votes within any one election. Whereas (b) is highly predictable, statewide swings are more volatile and harder to predict. Fortunately, (b) is more important for evaluating redistricting than (a).

A simple and, we show, remarkably accurate assumption that identifies $S(V)$ for any V is *uniform partisan swing*:

Assumption 3 [Uniform partisan swing (Butler 1951)]. *When the average district vote swings between elections under the same electoral system from V to V′, every district vote proportion moves uniformly by* $\delta \equiv V' - V$, *so that* $\{v_1,\ldots, v_L\}$ *from one election becomes* $\{v_1 + \delta,\ldots, v_L + \delta\}$ *in the next (with elements truncated to [0, 1] if necessary).*

Of course, uniform partisan swing is obviously violated when it generates district vote proportions outside [0, 1], but this is superfluous detail. We thus restate the assumption: *Given district-level vote proportions in one election,* $\{v_1,\ldots, v_L\}$, *and a chosen partisan swing parameter* δ, *the expected seat proportion in an election with average district vote* $V + \delta$ *under the same electoral system can be calculated via the (single-valued) function* $S(V + \delta) = \text{mean}_d[s(v_d + \delta)]$.

To study the empirical accuracy of this seats–votes function, we would need to compute the out-of-sample error rate for the statewide seat proportion in one election identical in all respects to the previous one—including candidates, the campaign, spending, weather on election day, and patterns of incumbency—except for the statewide partisan swing and the usual random uncertainties in vote outcomes. Finding such pairs of observed elections is obviously impossible, and so we instead use successive elections within the same redistricting regime. The consequence is that our error rate approximates an upper bound for the actual errors of uniform partisan swing–based predictions.

We begin with data from all regular elections in US state legislatures 1968–2016. We narrow these to the 646 elections for legislatures with all single-member districts, at least 20 total districts, with at least half the seats contested, and where no redistricting has occurred between this election and the one before.[6]

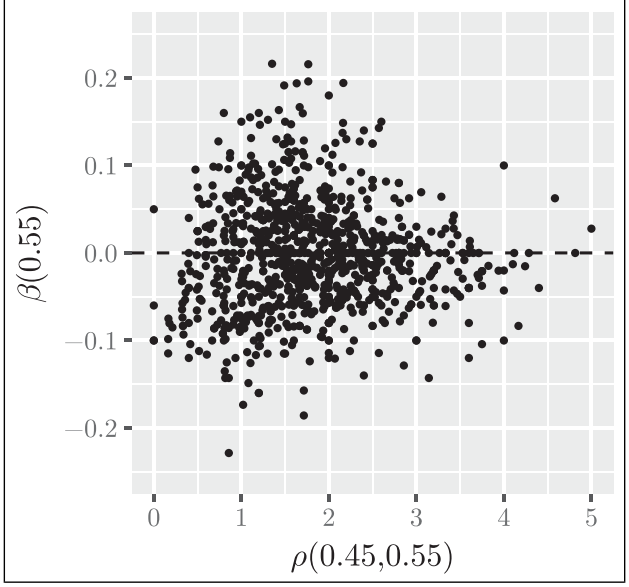
Thus, for each of 646 elections, we use the district-level vote proportions in election 1; the statewide swing to election 2, $\delta = V_2^O - V_1^O$; and the uniform partisan swing assumption to predict the expected statewide seat proportion for election 2, $S_2(V_2^O)$. We do not observe this expected value and so use the observed election 2 seat share $\hat{S}_2(V_2^O)$ (as a model-free estimate of the expected value) for validation. The error metric for the prediction $\hat{S}_1(V_2^O)$ is then simply $\hat{S}_2(V_2^O) - \hat{S}_1(V_2^O)$.

The left panel of Figure 2 gives a histogram of these out-of-sample prediction errors from uniform partisan swing. These results reveal highly accurate predictions, with a median error of 0.0000, a mean error of −0.001 (one-10th of one percentage point), and an interquartile

[6] Following standard practice (Gelman and King 1994b), we impute uncontested districts at 0.75 for Democratic wins and 0.25 for Republican wins, although this has no material impact on our results.

https://doi.org/10.1017/S000305541900056X Published online by Cambridge University Press

FIGURE 3. Estimates of Partisan Bias and Electoral Responsiveness, Both Evaluated in the Range $V \in [0.45, 0.55]$ in 963 Legislatures



range of only [−0.025, 0.021]. And these numbers are *upper* bounds.[7]

We also use these data to study how fast uniform partisan swing–based predictions degrade as we extrapolate farther from the original data, that is, when the statewide vote swing is larger. The right panel of Figure 2 plots the prediction error by the size of the statewide vote swing, δ. Remarkably, the graph shows that predictions do not degrade at all for larger swings (i.e., as δ deviates from zero). The implication is that uniform partisan swing is a relatively fixed feature of elections, with the more difficult-to-predict component of elections mostly relegated to statewide voter swings, which happen not to be important for studying partisan symmetry.

In our data, and in many elections all over the world, uniform partisan swing is a reasonable first approximation, especially for theoretical purposes like ours. What Assumption 3 ignores is that the world is stochastic and so is less useful for some empirical purposes. The simplicity can also generate inefficiency in part because of discreteness (Nagle 2015, 351). We thus generalize the deterministic uniform partisan swing assumption either directly via stochastic modeling (King 1989) or statistical modeling:

Assumption 4 [Stochastic uniform partisan swing (Gelman and King 1994a)]. *Hypothetical (denoted "(hyp)") district-level vote proportions, under the same electoral system, are generated as*

$$v_d^{(\text{hyp})} = X_d\theta + \gamma_d + \delta^{(\text{hyp})} + \epsilon_d^{(\text{hyp})}, \qquad (4)$$

where X_d is a vector of covariates describing the districts, candidates, voters, and lagged vote; θ is a vector of effect parameters; γ_d is an independent random normal district effect that is constant over hypothetical elections but varying over districts; $\delta^{(\text{hyp})}$ is the researcher-chosen uniform swing; and $\epsilon_d^{(\text{hyp})}$ is a stochastic normal error term, independent of γ and over d.

Assumption 4, rewritten as we did with its special case in Assumption 3, is thus less restrictive, more realistic, and more statistically efficient, and so should be used whenever it makes a difference. For our theoretical and methodological purposes, we will usually use the simpler Assumption 3 in this article to ease exposition, and because we analyze so many districts that inefficiency is a minor issue and because relevant empirical patterns of voter behavior are extremely regular across most elections in most countries (King et al. 2008, 952).

EVALUATING FAIRNESS MEASURES

We now evaluate measures of partisan fairness in district-based electoral systems. Where possible, we identify the corresponding estimate and implied notion of fairness. (Online Appendix E discusses uncertainty estimates.)

Estimation from Seats–Votes Curves

A straightforward way to estimate a feature of the seats–votes curve is with an estimate of the entire curve, using one of the assumptions in the section on estimation above. With the full curve, partisan bias $\beta(V)$, electoral responsiveness $\rho(V)$, and other quantities are easy to estimate for any relevant $V \in [0, 1]$, ensuring that Assumptions 1 and 2 hold. A few of the important articles computing bias and responsiveness in this way include Brunell (1999) and Jackman (1994), which use Assumption 4; Erikson (1972), using the functional form assumption in equation 2; Gilligan and Matsusaka (1999) and Niemi and Jackman (1991), using the functional form assumption in equation 3; and Brady and Grofman (1991) and Garand and Parent (1991), which use a combination of the functional form assumption in equation 3 along with Assumption 3.

As an illustration, we estimate partisan bias and electoral responsiveness using data from 963 legislatures (those from 1968 to 2016 with all single-member districts, at least 20 seats, and at least half of the seats are contested) via Assumption 3. Figure 3 plots bias vertically by responsiveness horizontally, both for $V \in [0.45, 0.55]$. The scatterplot shows that bulk of bias results is in $[-0.1, 0.1]$ and responsiveness in $[1, 3]$. The two quantities are uncorrelated in these data, but not independent in that as ρ increases, $|\beta|$ declines. This pattern is consistent with the scenario from the section on Gerrymandering Goals, where the redistricter is confident of a statewide majority

[7] Our replication dataset also includes an analysis of an alternative "proportional partisan swing" assumption, where each district moves the same proportion of its distance to 1 (or 0), so that when V moves by Δ, each district moves by $\Delta(1 - v_i)/(1 - V^{\text{O}})$. However, this alternative turns out to have about three times the error of uniform partisan swing.

https://doi.org/10.1017/S000305541900056X Published online by Cambridge University Press

and so seeks high responsiveness and is left with low bias.

Proportional Representation

We now discuss several individual measures that do not first estimate the entire seats–votes curve. For expository reasons, we begin with the simple *deviation from proportional representation* measure, $\text{PRD}(V^{\text{O}}) = S(V^{\text{O}}) - V^{\text{O}}$, which is easy to understand and turns out to fail as a measure of partisan fairness. We explain and then show how it is useful for other purposes.

Under this approach's standard of fairness, $\text{PRD}(V) = 0$, we solve for expected seats as a function of votes, $S(V^{\text{O}}) = V^{\text{O}}$, and then swap V^{O} with V, yielding $S(V) = V$. This is a coherent seats–votes curve because each value of V produces one value of $S(V)$, with Assumptions 1 and 2. See the green line in Figure 1, a symmetric electoral system with $\rho(V) = 1$, for all V. By writing partisan bias as $\beta(V) = \{S(V) - [1 - S(1 - V)]\}/2 = \{\text{PRD}(V) + \text{PRD}(1 - V)\}/2$, we can see that the proportional representation standard is a special case of partisan symmetry (because $\text{PRD}(V) = \text{PRD}(1 - V) = 0$ implies $\beta(V) = 0$ for all V), but partisan symmetry is not a special case of proportional representation (because $\beta(V) = 0$ whenever $\text{PRD}(V) = \text{PRD}(1 - V)$, even if $\text{PRD}(V) \neq 0$; see the other lines in Figure 1).

Although the proportional representation standard of $\text{PRD}(V^{\text{O}}) = 0$ is theoretically coherent, $\text{PRD}(V^{\text{O}})$ is inadequate as a measure of partisan symmetry. The problem is that V^{O} and $S(V^{\text{O}})$ produce only a single model-free estimate of a point on the seats–votes curve, which is insufficient for estimating the entire curve, because the second term in $\beta(V^{\text{O}}) = [\text{PRD}(V^{\text{O}}) + \text{PRD}(1 - V^{\text{O}})]/2$ is unobserved without further assumptions. For example, the election outcome $S(0.6) = 0.6$ is consistent with the standard because it falls on the line $S(V) = V$, which is proportional and symmetric. However, the same observed point is also consistent with the flat line $S(V) = 0.6$ (for all V) or with $1 - S(1 - 0.6) = 0$, neither of which is proportional, symmetric, or fair.

The deviation from proportional representation in one election is thus not a general measure of partisan symmetry, but it can be informative about its more specific standard ($\beta(V) = 0$ such that $\rho(V) = 1$): Although $\text{PRD}(V^{\text{O}}) = 0$ offers no information one way or the other, $\text{PRD}(V^{\text{O}}) \neq 0$ implies that this specific standard should be rejected. This may be useful on its own, but to go further requires estimating other points on the seats–votes curve, perhaps via the assumptions from the section above on estimation.

However, even when completely uninformative about partisan symmetry, $\text{PRD}(V^{\text{O}})$ is still an interesting and politically relevant summary of the outcome of an election. Certainly, small minority parties want to know whether they will receive at least some seats and so may compare it to their vote proportion as at least a benchmark. A forecast of $\text{PRD}(V^{\text{O}})$ would likely influence whether a small minority party would even compete in many districts or be likely to attract significant campaign contributions.[8]

Mean–Median

The *mean–median* measure is an easy-to-calculate difference: $\text{MM} = V^{\text{O}} - M$, where V^{O} is the average district vote and M is the median district vote, implicitly defined as $\frac{1}{L}\sum_{v_d > M} 1 = \frac{1}{2}$. Fairness according to this measure is when $\text{MM} = 0$. The measure is claimed "to reliably assess [partisan] asymmetry in state-level districting schemes" (Wang 2016a, 367). Essentially the same claim appears in Wang (2016b), Krasno et al. (2018), and McDonald and Best (2015), among others. Although no proof of this claim has appeared in the literature, we show that it is correct in different ways for two distinct theoretical quantities, vote- and seat-denominated partisan bias. In the first, we prove that MM provides limited information about $\beta(V)$; in the second, we show that MM is a more useful measure of a new theoretical quantity.

A Limited Measure of Partisan Bias

We begin by showing, under Assumptions 1, 2, and 3, that $\text{MM} = 0$ if and only if $\hat{\beta}(0.5) = 0$. Formally,

$$\hat{\beta}(0.5) = \hat{S}(0.5) - 0.5 \quad \text{(by definition)}, \tag{5}$$

$$= \frac{\sum_{v_d > V^{\text{O}}} 1}{L} - 0.5 \quad \text{(Assumption 3)},$$

$$= \frac{\sum_{v_d > M} 1}{L} - 0.5 \quad \text{(Assuming MM} = 0), \tag{6}$$

$$= 0.5 - 0.5 = 0. \qquad \square$$

As an estimate of $\beta(0.5)$, the mean–median measure has two limitations (in addition to the effects of discreteness; Nagle 2015). First, although our proof shows that MM is a useful indicator for whether $\beta(0.5)$ is zero, and so could be used for a hypothesis test, it is not a general measure of $\beta(0.5)$, as we have no proof that it is an unbiased or consistent estimator because the magnitude is not known to be correct when other than zero.

Second, if the electoral system is biased at a point other than $V = 0.5$, the mean–median measure will not necessarily reflect overall partisan symmetry (see the right panel, Figure 1). Consider an election with 10 districts and the following vote proportions:

[8] In practice, single-member district electoral systems are rarely even approximately proportional (see Figure 3). Paradoxically, even electoral systems that impose proportional representation at the statewide level wind up with considerable asymmetry, given how they are applied in multiparty contexts (see Grofman and King 2007, fn. 37). The Supreme Court has confirmed the empirically obvious: "the Constitution provides no right to proportional representation" (*Vieth v. Jubilerer*, 541 US 267 (2004)).

https://doi.org/10.1017/S000305541900056X Published online by Cambridge University Press

$$\{0.48, 0.49, 0.49, 0.49, 0.59, 0.61, 0.65, 0.65, 0.65, 0.90\}. \quad \textbf{(7)}$$

From these data, and the assumptions above, $\text{mm} = \hat{\beta}(0.5) = 0$, which would enable us to conclude that an aspect of the electoral system is fair. However, without any additional assumptions, we can show that in fact other aspects of the electoral system can be biased. For example, if the Democratic party receives an average district vote of $V^{\text{O}} = 0.6$ (the observed value of these district proportions), they would win a $\hat{S}(V^{\text{O}}) = 0.6$ seat proportion, but when the Republicans receive the same $1 - V^{\text{O}} = 0.6$, they would win a remarkable $1 - \hat{S}(1 - V^{\text{O}}) = 0.9$ of the seats, a 30 percentage point difference. This means that $\hat{\beta}(0.6) = -0.15$. In this example, the mean–median measure indicates that the electoral system represented is fair, but it is instead quite unfair.

A Better Measure of "Vote-Denominated" Partisan Bias

More interestingly, we can prove that MM is a valid estimator of a new theoretical quantity, "vote-denominated" partisan bias VDB(0.5): how much more one party must earn in votes than the other party to win a given seat proportion (defined more formally in Online Appendix F). With Assumptions 1, 2, and 3, we have

$$\begin{aligned}\text{VDB}(0.5) &= -\{V(S) - [1 - V(1 - S)]\}/2 \\ &\quad \text{(by definition)}, \\ &= 0.5 - V(0.5), \\ &= V - M \quad \text{(Assumption 3)}, \\ &\equiv \text{MM}. \end{aligned}$$ □

If we use the more realistic Assumption 4 in place of Assumption 3, it is easy to show that MM is a statistically consistent estimator of VDB(0.5). Either way, this proof is important because, although the *magnitude* of MM has no clear relationship to $\beta(0.5)$, it is correct for the alternative quantity VDB(0.5), making the mean–median measure an easy-to-calculate and accurate estimator of this unusual but still coherent theoretical quantity.

Lopsided Outcomes

Wang introduces the *lopsided outcomes test* and claims it can be used "to reliably assess [partisan] asymmetry in state-level districting schemes," (Wang 2016b, 1263), or "to detect…" or "identify partisan asymmetry" (Wang 2016a, 368). We show here that this claim is false, and along the way describe other more productive uses of the measure.

Begin by denoting the average Democratic vote in Democratic-won districts as $D = \sum_d s(v_d)v_d / \sum_d s(v_d)$ and the average Democratic vote in Republican-won districts as $R = \sum_d [1 - s(v_d)]v_d / \sum_d [1 - s(v_d)]$, which implies $R < 0.5 < D$. Then we write an accounting identity with the average district vote as a weighted average of Democratic and Republican seat shares, $V = S(V^{\text{O}})D + [1 - S(V^{\text{O}})]R$, and solve for the generic seats–votes relationship, all without assumptions:

$$S(V^{\text{O}}) = \frac{V^{\text{O}} - R}{D - R}. \quad \textbf{(8)}$$

Equation 8 is true by definition, but swapping V^{O} for V is not sufficient to define a coherent seats–votes curve because $S(V)$ is not single-valued. The presence of R and D on the right side, which are functions of V, means that we need more constraints to ensure $S(V)$ has only one value. Only at that point can we add Assumption 2 and evaluate the claim that the resulting seats–votes curve meets the partisan symmetry standard. So the question for any measure is whether it imposes these sufficient constraints.

The lopsided outcomes measure is defined as the simple party difference in the average win size:

$$\text{lo} = D - (1 - R). \quad \textbf{(9)}$$

This measure seems intuitive because packed districts are sometimes a characteristic of successful partisan gerrymandering but, as we showed above, the intuition is often wrong because packing (and cracking) can be counterproductive. The measure deserves credit as a fine measure of the skewness of the vote proportion distribution because nonskewness implies that the center of mass on either side of $v_d = 0.5$ will be equidistant from this midpoint. A forecast of this measure may indeed be useful to partisans or others trying to understand the competitive playing field and what it takes on average to win a district for their party.

Unfortunately, lopsided outcomes is not necessarily related to partisan symmetry or any other measure of fairness. To show this, we now study how equation 9 is constrained by the measure's notion of fairness, $\text{LO} = 0$. Thus, by substituting $D = 1 - R$ into equation 8, we have

$$S(V^{\text{O}}) = \frac{V^{\text{O}} - R}{1 - 2R}. \quad \textbf{(10)}$$

Unfortunately, after substituting V^{O} with V, we are still left with multiple values of $S(V)$ for any one V because of the presence of R on the right side, which indicates the lack of a coherent seats–votes curve. This means that the lopsided outcomes test, and the implied *set* of multiple seats–votes curves it considers "fair," can be consistent with either symmetry or asymmetry. As a result, LO does not imply particular values of $\beta(V)$ and is not a measure of (the deviation from) partisan symmetry.

Thus, to construct an example, we add information the framework omits in the form of Assumption 3. We then construct examples of votes from three hypothetical legislatures:

$$\{0.25, 0.25, 0.25, 0.75, 0.75, 0.75, 0.75, 0.75, 0.75, 0.75\}, \quad \textbf{(11)}$$

$$\{0.30, 0.30, 0.40, 0.55, 0.55, 0.65, 0.65, 0.90, 0.90, 1\}, \quad \textbf{(12)}$$

$$\{0.40, 0.40, 0.40, 0.40, 0.40, 0.60, 0.60, 0.60, 0.60, 0.60\}, \quad \textbf{(13)}$$

and then compute partisan bias and LO for each. The inconsistency is apparent: Legislature (11) is judged fair

https://doi.org/10.1017/S000305541900056X Published online by Cambridge University Press

by the lopsided outcomes test but is in fact asymmetric (LO = 0, $\hat{\beta}(V^{\mathrm{O}}) = \hat{\beta}(0.5) = 0.2$). Legislature (12) is judged unfair by lopsided outcomes but is in fact symmetric (LO = 0.08, $\hat{\beta}(V^{\mathrm{O}}) = \hat{\beta}(0.5) = 0$). And Legislature (13) is also judged fair by lopsided outcomes and is in fact symmetric (LO = 0, $\hat{\beta}(V^{\mathrm{O}}) = 0$).

Declination

Warrington (2018a, 2) introduced a measure called *declination* and claims it "is a measure of partisan symmetry" (or "a new measure of partisan asymmetry"; Warrington 2018b). We prove that this claim is incorrect, but along the way convey the measure's intuition and potential descriptive uses.

Warrington found a clever geometric interpretation of his measure, intuitive from the perspective of his field of mathematics, by defining it as DECLINATION = $2(\theta_D - \theta_R)/\pi$, where $\theta_D = \arctan[(2D - 1)/S(V^{\mathrm{O}})]$ and $\theta_R = \arctan\{(1 - 2R)/[1 - S(V^{\mathrm{O}})]\}$. For our intended audiences, the measure is easier to understand without the arctan transformation or constant normalizations, which only adjust the scale. We thus define the unnormalized declination

$$\mathrm{DEC} = \frac{D - 0.5}{S(V^{\mathrm{O}})} - \frac{0.5 - R}{1 - S(V^{\mathrm{O}})}. \tag{14}$$

Equation 14 (which can also be thought of as a normalized version of lopsided outcomes; cf. equation 9) is similar to the difference in the magnitude of electoral responsiveness on each side of $V = 0.5$. This is important because under partisan symmetry, the difference in (actual) responsiveness is zero. The problem is that responsiveness is a change in votes divided by a change in seats (see Definition 3), whereas each of the two terms in DEC is a change in votes divided by an absolute seat proportion. That means that DEC is not an unbiased measure of partisan symmetry, but is related and serves as a measure of the skewness of the distribution of district vote proportions.

To formally prove the connection between declination and partisan symmetry, consider how its notion of fairness, DEC = 0, constrains the generic equation 8. The result is

$$S(V^{\mathrm{O}}) = \frac{D - 0.5}{2D - 0.5 - V^{\mathrm{O}}}. \tag{15}$$

As with lopsided outcomes, even after swapping V^{O} for V, $S(V)$ is not a single-valued function of V and so, even under its notion of "fairness," is not a coherent seats–votes curve. That is, we can try to adjust V to see how $S(V)$ changes, but how D changes with V is left unspecified, which leaves many possible values of $S(V)$. The proposed standard is thus consistent with both symmetric and asymmetric seats–votes curves and, as such, declination not a measure of partisan symmetry.

In parallel to the previous section, we now offer examples of these inconsistencies with three hypothetical 10-district legislatures:

$$\{0.45, 0.45, 0.55, 0.55, 0.55, 0.60, 0.80, 0.80, 0.80, 0.95\}, \tag{16}$$

$$\{0.40, 0.40, 0.55, 0.55, 0.55, 0.60, 0.60, 0.63, 0.70, 0.70\}, \tag{17}$$

$$\{0.48, 0.48, 0.52, 0.52, 0.55, 0.59, 0.60, 0.60, 0.63, 0.63\}. \tag{18}$$

As in the previous section, we add missing information in the form of Assumption 3 and compute partisan bias and DEC. We find that Legislature (16) is judged fair by declination but is in fact asymmetric (DEC = 0, $\hat{\beta}(V^{\mathrm{O}}) = \hat{\beta}(0.5) = -0.1$. Legislature (17) is judged unfair by declination but is in fact symmetric (DEC = −0.36, $\hat{\beta}(0.5) = \hat{\beta}(V^{\mathrm{O}}) = 0$). And Legislature (18) is also judged fair by declination and is symmetric (DEC = 0, $\hat{\beta}(0.5) = \hat{\beta}(V^{\mathrm{O}}) = 0$).

Efficiency Gap

Stephanopoulos and McGhee (2015) introduce the *efficiency gap* and claim it is "a new measure of partisan [a] symmetry" (quote repeated on pages 831, 834, 838, 849, and 899). We prove that this claim is false, and also convey the intuition and productive uses of the measure.

The efficiency gap redefines the classic definition of "wasted votes" of all votes cast for losing candidates (Campbell 1996) by adding those for winning candidates above the 50%-plus-one-vote threshold. The article then claims that partisan symmetry is satisfied when these wasted votes are equally divided between the parties. We show this claim is incorrect. Although the efficiency gap is controversial (Chambers, Miller, and Sobel 2017; Cho 2017; Tapp 2018), it comes with important intuition; the authors also deserve substantial credit for bringing many, including the US Supreme Court, back to this venerable field (see *Gill v. Whitford*, 585 US (2018); see Stephanopoulos and McGhee 2018).

The intuition works best in highly competitive situations, when one party is in control of redistricting and running scared. Here, redistricters try to pack and crack and thus reduce wasted votes. In other situations, such as when confident of a statewide vote majority, packing is against the redistricter's interests. Here, the efficiency gap becomes confused; for example, if a party receives 80% of the votes and *all* the seats, the measure indicates that the electoral system treats it unfairly (see also Veomett 2018).

To formalize, denote the proportion of wasted votes in district d for Democrats as $w_d = v_d - s(v_d)/2 \in [0, 0.5]$ and Republicans as $(0.5 - w_d) = (1 - v_d) - [1 - s(v_d)]/2$. Then define the efficiency gap as follows:

$$\mathrm{EG}(V^{\mathrm{O}}) = \frac{\sum_d n_d(0.5 - w_d) - \sum_d n_d w_d}{\sum_d n_d}, \tag{19}$$

$$= S(V^{\mathrm{O}}) - 2V^{\mathrm{O}} + 0.5 - C, \text{ where } C = 2\frac{\sum_d t_d w_d}{\sum_d n_d}, \tag{20}$$

https://doi.org/10.1017/S000305541900056X Published online by Cambridge University Press

where $t_d = n_d - \text{mean}_d(n_d)$ (see McGhee 2017). We can solve this expression as $S(V^{\text{O}}) = 2V^{\text{O}} - 0.5 + C$. However, because C is a function of V, a single-valued seats–votes function does not result, violating Assumption 1. Stephanopoulos and McGhee (2015, 853) tried to remove the problem by assuming the turnout is constant, implying $C = 0$, but because this claim is observable, making it an "assumption" does not make sense. A minimally necessary condition for which $C = 0$ is $\text{Cov}(t_d, w_d) = 0$, but this too does not solve the problem because this covariance is rarely zero.

This result means that the claims for the efficiency gap are mistaken: it is not a measure of partisan symmetry. The slope of the implied seats–votes curve is not two because it does not imply a coherent seats–votes curve. The claim that the efficiency gap and partisan bias "are mathematically identical in the special case in which both parties receive exactly 50% of the vote" (p. 856) is incorrect. The claim that "a party can win more than half the seats with half the votes only by exacerbating the efficiency gap in its favor" (p. 856) is also untrue.

Corrected Efficiency Gap

We give the efficiency gap idea the benefit of the doubt here with the same simplicity sought in Stephanopoulos and McGhee (2015) by computing a *corrected efficiency gap* (CEG). This measure involves moving C to the left side of equation 20, and defining:

$$\text{CEG}(V^{\text{O}}) \equiv \text{EG}(V^{\text{O}}) + C = S(V^{\text{O}}) - 2V^{\text{O}} + 0.5. \quad \textbf{(21)}$$

(cf. McGhee 2017, 427ff). We study this measure's standard of fairness $\text{CEG}(V^{\text{O}}) = 0$ by solving equation 21 for $S(V)$, adding Assumptions 1 and 2, and writing what turns out to be a coherent (single-valued) seats–votes curve:

$$S(V) = 2V - 0.5. \quad \textbf{(22)}$$

The *assumed fair* seats–votes curve in equation 22 meets the partisan symmetry standard in Definition 1 because $S(V) = 2V - 0.5 = 1 - [2(1 - V) - 0.5]$, but it is a special case because of the additional constraints of a slope of $\rho(V) = 2$ for $V \in [0.25, 0.75]$ and $\rho(V) = 0$ for $V \notin [0.25, 0.75]$ (the red line in the left panel of the figure in Online Appendix G); note that all four symmetric electoral systems in Figure 1 would be judged unfair according to this standard. Equation 22 is an unpopular normative standard (e.g., Chambers, Miller, and Sobel 2017, 16; McGann et al. 2015, fn. 1), but it is coherent and so meets Assumption 1.

We move now from the fair seats–votes curve assumed under the efficiency gap framework to estimation. Unfortunately, an estimated CEG in one election is insufficient to determine whether the electoral system is symmetric. In particular, $\beta(V^{\text{O}}) = [\text{CEG}(V^{\text{O}}) + \text{CEG}(1 - V^{\text{O}})]/2$ equals zero only when $\text{CEG}(V^{\text{O}}) = -\text{CEG}(1 - V^{\text{O}})$. However, an election with $1 - V^{\text{O}}$ is unobserved and so $\text{CEG}(1 - V^{\text{O}})$ is not identified, nor is $\beta(0.5)$ or $\beta(V)$.

Examples and empirical evidence about CEG are given in Online Appendix G.

Simulation Measures

Simulation measures seek to compare a redistricting plan with all possible plans for a state, or all that meet the chosen criteria (Chen and Rodden 2013; Chikina, Frieze, and Pegden 2017; Duchin 2018; Magleby and Mosesson 2018). Unfortunately, all possible plans are too large, proper random sampling remains an unsolved problem (Fifield et al. 2018; Tam Cho and Rubinstein-Salzedo 2019), and using actual plans from different states is unrealistic (Wang 2016b).

We thus describe two simulation-based measures (for when the random sampling problem is solved) without separately defined standards and then two with clear standards. First, purely relative measures are of little value: Is a plan fair if it is at the 50th percentile of possible plans but, when the parties split the vote equally, Republicans receive 85% of the seats? Is a plan unfair if a party receives more seats than 99.99% of all plans but, when the parties split the votes equally, they split the seats equally? This approach might work if achieving a fairer plan on an absolute scale were known to be impossible, but that is rare in real redistrictings and in any event is not addressed by relative measures.

Second, some assign uniform probability to each randomly drawn plan. The resulting distribution has been compared with estimating the "cone of uncertainty" in hurricane predictions (Lander 2018), but this analogy does not hold. The cone of uncertainty is a posterior distribution based on informative data, whereas a uniform distribution is a prior assumed without evidence (based on something like Laplace's discredited "Principle of Insufficient Reason"). Uniform hurricane predictions would put equal probability over the entire globe. In real redistricting cases, when (naive) judges or special masters propose drawing plans randomly, or arbitrarily like a checkerboard, experts on all sides strenuously object because of the well-known likelihood of unintended consequences.

We also consider two approaches that make good use of randomly drawn districts. First, sampling can be productively used for "producing a large set of legally viable maps with respect to multiple criteria" (Cain et al. 2017, 1538) to convey what is *possible*, such as plans with de minimis levels of partisan bias while also meeting other criteria. The approach is also useful for identifying the characteristics of plans that are *impossible* given the state's geography, which can be compelling (e.g., Chen and Rodden 2013, sec. 5 and Duchin et al. n.d.). These deterministic statements are also useful for studying what constrains gerrymanderers, such as compactness or the tendency of Democrats to live packed in cities; after all, there is no reason to think gerrymanderers are influenced by any number of possible plans except for the one they would choose as optimal.

Finally, suppose a redistricter claims that the *only* criteria used in selecting a plan were (say) compactness and equal population. Then choice among plans that meet these criteria is random (due to exchangeability). This leads to a coherent hypothesis test: For a criterion outside the original two, such as partisan bias, the null—that no information was used in drawing the plan other than the three criteria—can be computed directly

https://doi.org/10.1017/S000305541900056X Published online by Cambridge University Press

by computing the probability of observing partisan bias as or more extreme than the existing plan. One may also be able to develop a coherent hypothesis test by drawing plans from distributions like $f\big(\hat{\beta}(0.5) \,|\, \beta(0.5) = 0\big)$. Although this density is not identified from uniform draws, it would be useful if we can develop a way to draw from the null directly.

CONCLUDING REMARKS

The literature on partisan fairness in district-based electoral systems dates back more than a century, predating the invention of most of modern statistics. This time period even includes the invention of one of the most fundamental principles of statistical inference—separating the estimator and the quantity of interest being estimated—and all the ways of using this principle to evaluate and improve statistical estimators. We update the venerable partisan fairness literature, apply this statistical principle, reveal essential assumptions not discussed or formalized, and shore them up with extensive empirical evaluations when observable implications are available. We then prove which of the many new estimates claimed to be measures of partisan symmetry or other specific fairness quantities are appropriate and which are biased or otherwise limited.

We hope the theoretical foundations and empirical evidence we offer here will help sort out some of the conflicting arguments and proposals that have appeared in the literature in recent years and enable faster progress going forward. We especially look forward to work that pushes forward the Frontier of statistical estimation and helps formalize other concepts in this area, such as those which adjudicate trade-offs between partisan fairness and other goals such as racial fairness, representing communities of interest, district compactness, and others.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit https://doi.org/10.1017/S000305541900056X.

Replication materials can be found on Dataverse at: https://doi.org/10.7910/DVN/FTYHPJ.

REFERENCES

Best, Robin E., Shawn J. Donahue, Jonathan Krasno, Daniel B. Magleby, and Michael D. McDonald. 2018. "Considering the Prospects for Establishing a Packing Gerrymandering Standard." *Election Law Journal* 17 (1): 1–20.

Blackburn, Simon. 2003. *Ethics: A Very Short Introduction*. Vol. 80. New York: Oxford University Press.

Brady, David W., and Bernard Grofman. 1991. "Sectional Differences in Partisan Bias and Electoral Responsiveness in US House Elections, 1850–1980." *British Journal of Political Science* 21 (2): 247–56.

Brunell, Thomas. 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): 316–37.

Brunell, Thomas. 2010. *Redistricting and Representation: Why Competitive Elections Are Bad for America*. New York: Routledge.

Butler, David E. 1951. "Appendix." In *The British General Election of 1950*, ed. H.G. Nicholas. London: Macmillan, 306–33.

Cain, Bruce E., Wendy K. Tam Cho, Yan Y. Liu, and Emily R. Zhang. 2017. "A Reasonable Bias Approach to Gerrymandering: Using Automated Plan Generation to Evaluate Redistricting Proposals." *William & Mary Law Review* 59: 1521.

Campbell, James E. 1996. *Cheap Seats: The Democratic Party's Advantage in US House Elections*. Columbus, OH: The Ohio State University Press.

Chambers, Christopher P., Alan D. Miller, and Joel Sobel. 2017. "Flaws in the Efficiency gap." *Journal of Law & Politics* 33: 1.

Chen, Jowei, and Jonathan Rodden. 2013. "Unintentional Gerrymandering: Political Geography and Electoral Bias in Legislatures." *Quarterly Journal of Political Science* 8 (3): 239–69.

Chikina, Maria, Alan Frieze, and Wesley Pegden. 2017. "Assessing Significance in a Markov Chain without Mixing." *Proceedings of the National Academy of Sciences* 114 (11): 2860–4.

Cho, Wendy K. Tam. 2017. "Measuring Partisan Fairness: How Well Does the Efficiency gap Guard against Sophisticated as Well as Simple-Minded Modes of Partisan Discrimination." *University of Pennsylvania Law Review Online* 166: 17.

Cox, Adam B. 2006. "Designing Redistricting Institutions." *Election Law Journal* 5 (4): 412–24.

Cox, Gary W. 1997. *Making Votes Count: Strategic Coordination in the World's Electoral Systems*: Cambridge University Press.

Cox, Gary W., and Jonathan N. Katz. 1999. "The Reapportionment Revolution and Bias in US Congressional Elections." *American Journal of Political Science* 43 (3): 812–41.

Cox, Gary W., and Jonathan N. Katz. 2002. *Elbridge Gerry's Salamander: The Electoral Consequences of the Reapportionment Revolution*. New York: Cambridge University Press.

Duchin, Moon. 2018. "Gerrymandering Metrics: How to Measure? What's the Baseline?" arXiv:1801.02064.

Duchin, Moon, Taissa Gladkova, Eugene Henninger-Voss, Ben Klingensmith, Heather Newman, and Hannah Wheelen. n.d. *Obstructions to Proportional Representation: Republicans in Massachusetts*. Tufts University.

Erikson, Robert S. 1972. "Malapportionment, Gerrymandering, and Party Fortunes in Congressional Elections." *American Political Science Review* 66 (4): 1234–45.

Fifield, Benjamin, Michael Higgins, Kosuke Imai, and Alexander Tarr. 2018. "A New Automated Redistricting Simulator Using Markov Chain Monte Carlo." Working Paper, Princeton, NJ: Princeton University.

Friedman, John N., and Richard T. Holden. 2008. "Optimal Gerrymandering: Sometimes Pack, but Never Crack." *The American Economic Review* 98 (1): 113–44.

Garand, James C., and T. Wayne Parent. 1991. "Representation, Swing, and Bias in US Presidential Elections, 1872–1988." *American Journal of Political Science* 35 (4): 1011–31.

Gelman, Andrew, and Gary King. May 1994a. "A Unified Method of Evaluating Electoral Systems and Redistricting Plans." *American Journal of Political Science* 38 (2): 514–54. URL: http://j.mp/unifiedEc.

Gelman, Andrew, and Gary King. September 1994b. "Enhancing Democracy through Legislative Redistricting." *American Political Science Review* 88 (3): 541–59. URL: http://j.mp/redenh.

Gerber, Elisabeth R., and Jeffrey B. Lewis. 2004. "Beyond the Median: Voter Preferences, District Heterogeneity, and Political Representation." *Journal of Political Economy* 112 (6): 1364–83.

Gilligan, Thomas W., and John G. Matsusaka. 1999. "Structural Constraints on Partisan Bias under the Efficient Gerrymander." *Public Choice* 100 (1–2): 65–84.

Greene, Joshua D. 2009. "The Cognitive Neuroscience of Moral Judgment." *The cognitive neurosciences* 4: 1–48.

Grofman, Bernard, and Thomas L. Brunell. 2005. "The Art of the Dummymander: The Impact of Recent Redistrictings on the Partisan Makeup of Southern House Seats." In *Redistricting in the New Millennium*, ed. Peter F. Galderisi. New York: Lexington Books, 183–99.

https://doi.org/10.1017/S000305541900056X Published online by Cambridge University Press

Grofman, Bernard, and Gary King. January 2007. "The Future of Partisan Symmetry as a Judicial Test for Partisan Gerrymandering after LULAC V. Perry." *Election Law Journal* 6 (1): 2–35. http://gking.harvard.edu/files/abs/jp-abs.shtml.
Herron, Michael C., and Alan E. Wiseman. 2008. "Gerrymanders and Theories of Law Making: A Study of Legislative Redistricting in Illinois." *The Journal of Politics* 70 (1): 151–67.
Iacus, Stefano M., Gary King, and Giuseppe Porro. 2018. "A Theory of Statistical Inference for Matching Methods in Causal Research." *Political Analysis*: 1–23. URL: http://j.mp/Nt9TkZ.
Jackman, Simon. 1994. "Measuring Electoral Bias: Australia, 1949–93." *British Journal of Political Science* 24 (3): 319–57.
Katz, Jonathan N., and Gary King. March 1999. "A Statistical Model for Multiparty Electoral Data." *American Political Science Review* 93 (1): 15–32. URL: http://bit.ly/mtypty.
Katz, Jonathan N., Gary King, and Elizabeth Rosenblatt. 2019. *Replication Data for: Theoretical Foundations and Empirical Evaluations of Partisan Fairness in District-Based Democracies*. doi: 10.7910/DVN/FTYHPJ.
Kaufman, Aaron, Gary King, and Mayya Komisarchik. Forthcoming. "How to Measure Legislative District Compactness if You Only Know it when You See it." *American Journal of Political Science*. URL: http://j.mp/2u9OWrG.
King, Gary. November 1989. "Representation through Legislative Redistricting: A Stochastic Model." *American Journal of Political Science* 33 (4): 787–824. URL: http://j.mp/2o46Gkk.
King, Gary. May 1990. "Electoral Responsiveness and Partisan Bias in Multiparty Democracies." *Legislative Studies Quarterly* XV (2): 159–81. URL: http://bit.ly/ErespMP.
King, Gary, and Robert X. Browning. December 1987. "Democratic Representation and Partisan Bias in Congressional Elections." *American Political Science Review* 81 (4): 1252–73. URL: http://j.mp/parSym.
King, Gary, and Andrew Gelman. February 1991. "Systemic Consequences of Incumbency Advantage in the US House." *American Journal of Political Science* 35 (1): 110–38. URL: http://bit.ly/SystCs.
King, Gary, Ori Rosen, Martin Tanner, and Alexander F. Wagner. 2008. "Ordinary Economic Voting Behavior in the Extraordinary Election of Adolf Hitler." *The Journal of Economic History* 68 (4): 951–96. URL: http://bit.ly/nazivote.
Klarner, Carl. 2018. *State Legislative Election Returns, 1967–2016: Restructured for Use*. doi: 10.7910/DVN/DRSACA.
Krasno, Jonathan, Daniel B. Magleby, Michael D. McDonald, Shawn J. Donahue, and Robin E. Best. 2018. "Can Gerrymanders Be Detected? An Examination of Wisconsin's State Assembly." *American Politics Research*: 1–40. Published Online May 14, 2018.
Lander, Eric S. 2018. "Brief for Amicus Curiae Eric S. Lander in Support of Appellees." *Gill v. Whitford* 585 U.S. (2018).
Magleby, Daniel B., and Daniel B. Mosesson. 2018. "A New Approach for Developing Neutral Redistricting Plans." *Political Analysis* 26 (2): 147–67.
Mann, Thomas E. 1978. *Unsafe at Any Margin: Interpreting Congressional Elections*. Vol. 220. Washington DC: American Enterprise Institute for Public Policy Research.
May, Kenneth O. 1952. "A Set of Independent Necessary and Sufficient Conditions for Simple Majority Decision." *Econometrica* 20 (4): 680–4.
McCarty, Nolan, Keith T. Poole, and Howard Rosenthal. 2009. "Does Gerrymandering Cause Polarization?" *American Journal of Political Science* 53 (3): 666–80.
McDonald, Michael D., and Robin E. Best. 2015. "Unfair Partisan Gerrymanders in Politics and Law: A Diagnostic Applied to Six Cases." *Election Law Journal* 14 (4): 312–30.
McDonald, Michael P. 2007. "Regulating Redistricting." *PS: Political Science & Politics* 40 (4): 675–9.
McGann, Anthony J., Charles Anthony Smith, Michael Latner, and J. Alex Keena. 2015. "A Discernable and Manageable Standard for Partisan Gerrymandering." *Election Law Journal* 14 (4): 295–311.
McGhee, Eric M. 2017. "Measuring Efficiency in Redistricting." *Election Law Journal: Rules, Politics, and Policy* 16 (4): 417–42.
Nagle, John F. 2015. "Measures of Partisan Bias for Legislating Fair Elections." *Election Law Journal* 14 (4): 346–60.
Niemi, Richard G., and Simon Jackman. 1991. "Bias and Responsiveness in State Legislative Districting." *Legislative Studies Quarterly* 16 (2): 183–202.
Owen, Guillermo, and Bernard Grofman. 1988. "Optimal Partisan Gerrymandering." *Political Geography Quarterly* 7 (1): 5–22.
Petracca, Mark P. 1996. "A History of Rotation in Office." In *Legislative Term Limits: Public Choice Perspectives*, ed. Bernard Groffman. Boston: Kluwer Academic Publishers, 247–77.
Puppe, Clemens, and Attila Tasnadi. 2009. "Optimal Redistricting under Geographical Constraints: Why 'pack and Crack' Does Not Work." *Economics Letters* 105 (1): 93–6.
Rave, D. Theodore. 2012. "Politicians as Fiduciaries." *Harvard Law Review* 126: 671.
Rubin, Donald B. 1991. "Practical Implications of Modes of Statistical Inference for Causal Effects and the Critical Role of the Assignment Mechanism." *Biometrics* 47: 1213–34.
Sen, Amartya. 1976. "Liberty, Unanimity and Rights." *Economica* 43 (171): 217–45.
Stephanopoulos, Nicholas O., and Eric M. McGhee. 2015. *Partisan Gerrymandering and the Efficiency Gap*. The University of Chicago Law Review 82: 831–900.
Stephanopoulos, Nicholas O., and Eric M. McGhee. 2018. "The Measure of a Metric: The Debate over Quantifying Partisan Gerrymandering." *Stanford Law Review* 70: 1503.
Tam Cho, Wendy K., and Simon Rubinstein-Salzedo. 2019. "Understanding Significance Tests from a Non-mixing Markov Chain for Partisan Gerrymandering Claims." *Statistics and Public Policy* 6 (1): 1–10.
Tapp, Kristopher. 2018. "Measuring Political Gerrymandering." arXiv:1801.02541.
Tufte, Edward R. 1973. "The Relationship between Seats and Votes in Two-Party Systems." *American Political Science Review* 67 (2): 540–54.
VanderWeele, Tyler J., and Miguel A. Hernan. 2012. "Causal Inference under Multiple Versions of Treatment." *Journal of Causal Inference* 1: 1–20.
Veomett, Ellen. 2018. "Efficiency Gap, Voter Turnout, and the Efficiency Principle." *Election Law Journal: Rules, Politics, and Policy* 17 (4): 249–63.
Wang, Samuel. 2016a. "Three Practical Tests for Gerrymandering: Application to Maryland and Wisconsin." *Election Law Journal* 15 (4): 367–84.
Wang, Samuel. 2016b. "Three Tests for Practical Evaluation of Partisan Gerrymandering." *Stanford Law Review* 68: 1263.
Warrington, Gregory S. 2018a. "Introduction to the Declination Function for Gerrymanders." arXiv preprint arXiv:1803.04799.
Warrington, Gregory S. 2018b. "Quantifying Gerrymandering Using the Vote Distribution." *Election Law Journal* 17 (1): 39–57.
Wimmer, Karl. 2010. "Agnostically Learning under Permutation Invariant Distributions." In *51st Annual IEEE Symposium on Foundations of Computer Science*: IEEE, 113–22.
Yoshinaka, Antoine, and Chad Murphy. 2009. "Partisan Gerrymandering and Population Instability: Completing the Redistricting Puzzle." *Political Geography* 28 (8): 451–62.

https://doi.org/10.1017/S000305541900056X Published online by Cambridge University Press