THE MEASURE OF A METRIC:
THE DEBATE OVER QUANTIFYING PARTISAN GERRYMANDERING

Nicholas O. Stephanopoulos[*]
Eric M. McGhee[**]

Over the last few years, there has been an unprecedented outpouring of scholarship on partisan gerrymandering. Much of this work has sought either to introduce new measures of gerrymandering or to analyze a metric—the efficiency gap—that we previously developed. In this Article, we reframe this debate by presenting a series of criteria that can be used to evaluate gerrymandering metrics: (1) consistency with the efficiency principle; (2) distinctness from other electoral values; (3) breadth of scope; and (4) correspondence with electoral history. We then apply these criteria to both the efficiency gap and other measures. The efficiency gap complies with the criteria under all circumstances. Other metrics, in contrast, often violate the efficiency principle and cannot be used in certain electoral settings.

[*] Professor of Law and Herbert and Marjorie Fried Research Scholar, University of Chicago Law School.
[**] Research Fellow, Public Policy Institute of California.

TABLE OF CONTENTS

INTRODUCTION

For several decades, there was "virtually a consensus [in] the scholarly community" about how to measure partisan gerrymandering.[1] An analyst would estimate the seat shares that the major parties would win in a state if (hypothetically) they each received the same vote share. The greater the divergence between the parties' seat shares for the same (counterfactual) vote share, the larger a district plan's *partisan bias*, and the more gerrymandered the plan.[2]

Despite the metric's wide acceptance among academics, the Supreme Court's pivotal member, Justice Kennedy, expressed misgivings about partisan bias. In a 2006 case, he did not "altogether discount[] its utility in redistricting planning and litigation," but he did worry that "[t]he existence or degree of [bias] may in large part depend on conjecture about where possible vote-switchers will reside."[3] He continued: "we are wary of adopting a constitutional standard that invalidates a map based on unfair results that would occur in a hypothetical state of affairs."[4]

The two of us agree with Justice Kennedy that it is odd to measure partisan gerrymandering based on how the major parties *would have* performed in counterfactual elections that did not actually materialize. It is more intuitive, in our view, to assess

[1] Bernard Grofman & Gary King, *The Future of Partisan Symmetry as a Judicial Test for Partisan Gerrymandering After* LULAC v. Perry, 6 ELECTION L.J. 2, 6 (2007).

[2] *See id.* at 6-13 (defining partisan bias).

[3] League of United Latin Am. Citizens v. Perry, 548 U.S. 399, 420 (2006) (opinion of Kennedy, J.).

[4] *Id.*

gerrymandering based on how the parties *did* perform in elections that, in fact, took place. This is why, a few years ago, we sought to unsettle the scholarly consensus in favor of partisan bias by introducing a new metric—the *efficiency gap*—that does not rely on predictions about what would occur in hypothetical electoral scenarios.[5]

The efficiency gap is rooted in the insight that partisan gerrymandering is always carried out in one of two ways: the *cracking* of a party's supporters among many districts, in which their preferred candidates lose by relatively narrow margins; or the *packing* of a party's backers in a few districts, in which their preferred candidates win by overwhelming margins.[6] Both cracking and packing produce what are known as *wasted votes* because they do not contribute to a candidate's election. In the case of cracking, all votes cast for the losing candidate are wasted; in the case of packing, votes cast for the winning candidate, above the 50% (plus one) threshold needed for victory, are wasted. The efficiency gap is simply one party's total wasted votes in an election, minus the other party's total wasted votes, divided by the total number of votes cast. It captures in a single figure the extent to which district lines crack and pack one party's voters more than the other's.[7]

It is fair to say that since we introduced the efficiency gap, there has been an explosion of judicial and academic interest in the measurement of partisan gerrymandering. In the courts, one of us has helped to litigate a pair of lawsuits based in part on the efficiency gap, one against Wisconsin's state house plan, the other against North Carolina's congressional map. The district court in the Wisconsin case struck down the plan in November 2016—the first such victory in more than thirty years—and the state's appeal is currently pending before the Supreme Court.[8] In the North Carolina case, the district court denied the state's motion to dismiss and a trial took place in October 2017.[9]

---

[5] McGhee first presented the efficiency gap (then known as relative wasted votes) in Eric McGhee, *Measuring Partisan Bias in Single-Member District Electoral Systems*, 39 LEGIS. STUD. Q. 55 (2014). The two of us then incorporated the measure into a potential legal test in Nicholas O. Stephanopoulos & Eric M. McGhee, *Partisan Gerrymandering and the Efficiency Gap*, 82 U. CHI. L. REV. 831 (2015).

We note that we recommend sensitivity testing (which requires the use of counterfactual elections) to establish the durability of a plan's efficiency gap. *See id.* at 889-90. This testing, however, is distinct from the measure itself. It also involves the use of counterfactual elections that are generally more plausible than those used to calculate partisan bias.

[6] This insight has not escaped the courts. *See, e.g.*, Vieth v. Jubelirer, 541 U.S. 267, 286 n.7 (2004) (plurality opinion) ("'Packing' refers to the practice of filling a district with a supermajority of a given group or party. 'Cracking' involves the splitting of a group or party among several districts to deny that group or party a majority in any of those districts.").

[7] *See* McGhee, *supra* note 5, at 68-70; Stephanopoulos & McGhee, *supra* note 5, at 850-55.

[8] *See Whitford v. Gill*, 218 F. Supp. 3d 837 (W.D. Wis. 2016) (*Whitford II*). Importantly, the efficiency gap is not the entirety of the legal test proposed in *Whitford*. Rather, the test requires (1) discriminatory intent; (2) a large and durable discriminatory effect; and (3) no legitimate justification for this effect. The efficiency gap is simply one way to establish the size of a discriminatory effect. *See id.* at 883-927.

[9] *See Common Cause v. Rucho*, ___ F. Supp. 3d ___, 2017 WL 876307 (M.D.N.C. Mar. 3, 2017). Beyond the efficiency gap, the mean-median difference has recently been presented to courts in two amicus briefs. *See* Br. of Amicus Curiae Samuel S. Wang, Ph.D. in Support of Appellees, *Harris v. Ariz. Indep. Redist. Comm'n*, 136 S. Ct. 1301 (2016) (No. 14-232); Mem. of Amici Curiae Common Cause and Virginia Majority Regarding Proposed Remedial Plans, *Personhubbalah v. Alcorn*, 155 F. Supp. 3d 552 (E.D. Va. 2016) (No. 3:13-cv-678).

In the academy, scholars have developed several new metrics and commented extensively on the efficiency gap. One of these new metrics is the *mean-median difference*: the difference between a party's mean vote share and median vote share across all of the districts in a plan. When a party's median vote share is lower than its mean vote share, the party is arguably the victim of gerrymandering.[10] Another new metric is the *difference between the parties' average margins of victory*. If one party's wins are more lopsided than the other party's, this may indicate that its supporters have been cracked and packed by the district lines.[11] Scholars have also proposed variants of both partisan bias (by averaging it across all electoral outcomes rather than for a single hypothetical election)[12] and the efficiency gap (by varying the definitions and weights of the two types of wasted votes).[13]

The academic discussion of the efficiency gap includes a number of criticisms of the measure. Writing in this journal's pages, Benjamin Cover contends that it is in tension with important democratic values. In his view, it favors uncompetitive elections, discourages proportional representation, and incentivizes voter suppression.[14] Cover and John Nagle also object to some of the methodological choices underpinning the efficiency gap: how wasted votes are defined and weighted, how imputations are made for uncontested races, and how variations in district-level turnout are addressed.[15] Wendy Tam Cho[16] and Jonathan Krasno and his coauthors[17] further complain about the metric's variability from election to election. This variability, they argue, is evident both in toy examples of one or two districts and in actual

---

[10] *See* Robin E. Best et al., *Considering the Prospects of Establishing a Packing Gerrymandering Standard*, 16 ELECTION L.J. (forthcoming 2017); Jonathan S. Krasno, et al., *Can Gerrymanders Be Measured? An Examination of Wisconsin's State Assembly* (May 22, 2016); Michael D. McDonald & Robin E. Best, *Unfair Partisan Gerrymanders in Politics and Law: A Diagnostic Applied to Six Cases*, 14 ELECTION L.J. 312 (2015); Samuel S.H. Wang, *Three Practical Tests for Gerrymandering: Application to Maryland and Wisconsin*, 15 ELECTION L.J. 367, 372 (2016) [hereinafter Wang, *Practical Tests*]; Samuel S.H. Wang, *Three Tests for Practical Evaluation of Partisan Gerrymandering*, 68 STAN. L. REV. 1263, 1304 (2016) [hereinafter Wang, *Three Tests*].

[11] *See* Wang, *Practical Tests*, *supra* note 10, at 371-72; Wang, *Three Tests*, *supra* note 10, at 1304.

[12] *See* John F. Nagle, *Measures of Partisan Bias for Legislating Fair Elections*, 14 ELECTION L.J. 346, 352 (2015). Other scholars have recently defended partisan bias as the measure was originally formulated. *See* Theodore S. Arrington, *A Practical Procedure for Detecting a Partisan Gerrymander*, 15 ELECTION L.J. 385, 390 (2016); Anthony J. McGann et al., *A Discernable and Manageable Standard for Partisan Gerrymandering*, 14 ELECTION L.J. 295, 308 (2015).

[13] *See* Christopher Chambers et al., *Flaws in the Efficiency Gap*, 33 J.L. & POL. 1, 15 (2017); Benjamin Cover, *Quantifying Political Gerrymandering: An Evaluation of the Efficiency Gap Proposal*, 70 STAN. L. REV. (forthcoming 2018) (manuscript at 40-51); John F. Nagle, *How Competitive Should a Fair Single Member Districting Plan Be?*, 16 ELECTION L.J. 196, 199-201 (2017). Still other scholars have recently sought to measure gerrymandering by determining a plan's deviation from the historical seat-vote curve, *see* Nicholas Goedert, *The Case of the Disappearing Bias: A 2014 Update to the "Gerrymandering or Geography" Debate*, RES. & POL., Oct.-Dec. 2015 [hereinafter Goedert, *Disappearing Bias*]; Nicholas Goedert, *Gerrymandering or Geography? How Democrats Won the Popular Vote but Lost the Congress in 2012*, RES. & POL., Apr.-June 2014 [hereinafter Goedert, *Gerrymandering or Geography*], or by sorting districts by vote share and then calculating the angle between lines connecting each party's median district to the 50% point, *see* Gregory S. Warrington, *Quantifying Gerrymandering Using the Vote Distribution* (May 15, 2017).

[14] *See* Cover, *supra* note 13 (manuscript at 26-40, 60-63)

[15] *See id.* (manuscript at 40-45, 51-60); Nagle, *supra* note 13, at 199-201.

[16] *See* Wendy K. Tam Cho, *Measuring Partisan Fairness: Guarding Against Sophisticated as well as Simple-Minded Modes of Partisan Discrimination*, 165 U PA. L. REV. ONLINE (forthcoming 2017) (manuscript at 3-13).

[17] *See* Best et al., *supra* note 10 (manuscript at 20-21, 26); Krasno et al., *supra* note 10, at 10-17.

district plans. And Christopher Chambers and his coauthors observe that the efficiency gap does not distinguish between moderate and extreme legislators. This oversight may allegedly lead to odd conclusions about certain maps.[18]

We find these criticisms unpersuasive, and we explain why later in this Article. We are also skeptical of the measures that other scholars have recently advanced. But before diving into the metrics' pros and cons, we think it is worthwhile to step back and ask what it is that we want from a quantitative measure of partisan gerrymandering. What properties would such a metric ideally exhibit?

One attribute is consistency with what one of us has labeled the *efficiency principle*. This is simply the idea that when a party's seat share increases while its vote share remains constant, a measure should reflect that party's growing advantage.[19] The essence of partisan gerrymandering is winning more seats without appealing to more voters, and a valid metric should capture that conceptual core. A second feature is *distinctness* from other democratic values. District plans implicate not just partisan fairness but also electoral competitiveness, voter participation, legislative polarization, and so on. But it is only partisan *un*fairness that lies at the heart of partisan gerrymandering, and thus should be revealed by a gerrymandering measure.

A third criterion is *breadth of scope*. In other words, a metric should be usable under a range of electoral conditions: when the parties are evenly matched statewide or when one party predominates, when turnout is or is not roughly equal in each district, and when there are two or more than two parties competing for office. Without such flexibility, a measure would be inapplicable to many common scenarios. And a fourth parameter is *empirical correspondence*. That is, the electoral ideal implied by a metric should not be too different from the American historical norm. Otherwise the measure would imply that most American plans have been gerrymanders—and its adoption would be so disruptive as to be infeasible.

These criteria are why we endorse the efficiency gap. First, as suggested by its name, it is indeed consistent with the efficiency principle. When a party wins a larger (smaller) seat share for the same vote share, the metric always shifts in favor of (against) that party. Second, the efficiency gap does not conflate partisan fairness with other democratic values. In particular, there is no connection between the measure and electoral competitiveness. Conceptually, any efficiency gap is compatible with any level of competitiveness, since it is only the *difference* between the parties' average margins of victory that affects the metric. Empirically too, American elections have exhibited essentially a zero correlation between the efficiency gap and competitiveness over the last half-century.

Third, as recently generalized by one of us,[20] the efficiency gap is usable in almost every electoral environment: competitive or uncompetitive, with or without equal district turnout, and whether two or more than two parties are running. The measure's only limitation is one that applies equally to all gerrymandering metrics: namely, that they are prone to large values (and

[18] *See* Chambers et al., *supra* note 13, at 23-29.

[19] *See* Eric McGhee, *Measuring Efficiency in Redistricting*, 16 ELECTION L.J. (forthcoming 2017) (manuscript at 5).

[20] *See id.* (manuscript at 22-36).

large swings) when the number of districts is small. And fourth, the efficiency gap corresponds closely to the empirical realities of American elections. Over the last fifty years, congressional and state house plans have both had mean and median efficiency gaps near zero. As the parties' vote shares have varied from year to year, their seat shares have also typically shifted to the precise extent necessary to maintain a low efficiency gap.

These criteria are also why we have reservations about other measures of partisan gerrymandering. Partisan bias and the mean-median difference, for instance, often violate the efficiency principle. In many circumstances, if a party's seat share changes while its vote share stays constant, the metrics either fail to move or move in the wrong direction. Additionally, partisan bias and the mean-median difference cannot be used in less competitive jurisdictions, where one party wins more than about 55% of the statewide vote. In these settings, the hypothetical election that is explicit in partisan bias, and implicit in the mean-median difference, becomes too unrealistic. And the variants of the efficiency gap proposed by Cover and Nagle, while theoretically defensible, are empirically inconsistent with American electoral history. They imply an ideal seat-to-vote relationship of either one (i.e., proportional representation) or three (i.e., hyper-responsiveness).

Despite these reservations, we wish to avoid the narcissism of small differences. One important point about the array of gerrymandering metrics that now exist is that no winner need be chosen among them. In other areas of election law, numerous measures of population inequality, racial polarization, and geographic compactness happily coexist. The same should be possible in the gerrymandering domain.

Indeed, détente should be *more* achievable here, since the various metrics are all linked mathematically to one another. The efficiency gap is equal to partisan bias whenever an election is tied statewide. The efficiency gap is also the same as the difference between the parties' average margins of victory, as long as each margin is weighted by the number of seats won by the party. And partisan bias can be transformed into the mean-median difference by dividing it by the slope of the seat-vote curve.

Furthermore, all of the measures produce similar results when jurisdictions are competitive statewide. These are the places where partisan gerrymandering can have the greatest impact, by turning the votes of an evenly split electorate into a state legislature or congressional delegation dominated by a single party. And in these places, it is largely immaterial which metric is consulted. Any of them will yield more or less the same conclusion.

The Article proceeds as follows. In Part I, we describe and justify the criteria we think should be used to evaluate measures of partisan gerrymandering. In Part II, we apply these criteria to the efficiency gap and show that it complies with them. In Part III, we respond to criticisms of the efficiency gap not already addressed: in particular, Cover's and Nagle's methodological objections, Cho's and Krasno et al.'s claims about variability, and Chambers et al.'s argument about ideology. And in Part IV, we discuss other gerrymandering metrics, both assessing them under our criteria and noting the extent of their convergence with the efficiency gap.

I. EVALUATIVE CRITERIA

During the several decades in which partisan bias was the only measure of partisan gerrymandering used by scholars, there was no need to develop criteria for judging such metrics. With only one option on the table, there was nothing to choose between. Times have changed, though, as measures of gerrymandering have proliferated in recent years. Now there is some urgency to the project of determining how to gauge them properly. Otherwise scholars, litigants, and courts risk being inundated by metrics whose strengths and weaknesses are only hazily grasped.

In this Part, we present a set of evaluative criteria that, in our view, reflect widely held intuitions about the character and quantification of partisan gerrymandering. These criteria seek: (1) to capture the essence of the activity; (2) to distinguish it from other electoral concepts; (3) to promote ease of use; and (4) to avoid unnecessary disruption. We freely admit that other academics might tinker with this list based on their opinions about gerrymandering. But as long as the list is seen as suggestive rather than definitive, we hope it will not provoke too much controversy.[21]

A. *Efficiency*

Our first criterion is consistency with the efficiency principle. This principle states that a measure of partisan gerrymandering "must indicate a greater advantage for (against) a party when the seat share for that party increases (decreases) without any corresponding increase (decrease) in its vote share."[22] The principle would be violated, for example, if a party received 55% of the vote and 55% of the seats in one election, and 55% of the vote and 60% of the seats in another election, but a metric did not shift in the party's favor. The principle would also be violated if a party's vote share increased from 55% to 60%, its seat share stayed constant at 55%, and a metric did not register a worsening in the party's position.

What is the basis for the efficiency principle? It is our understanding of partisan gerrymandering as a practice aimed above all at enabling a party to convert its votes into seats more efficiently than its adversary. Gerrymandering, on this account, is fundamentally about the relationship between popular support and legislative representation—and *manipulating* this relationship to benefit one party and handicap its rival.[23]

This is not an idiosyncratic stance. It seems to be shared, rather, by many of the scholars who have devised gerrymandering metrics or criticized the efficiency gap.[24] Andrew Gelman and Gary King, the best-known proponents of partisan bias, write that they seek to measure "the degree to which an electoral system unfairly favors one political party in the translation of

---

[21] For a more sweeping attempt to define generally the qualities of good social scientific concepts, see JOHN GERRING, SOCIAL SCIENCE METHODOLOGY: A UNIFIED FRAMEWORK 117 (2012). For an argument that partisan symmetry (the concept underlying all measures of partisan gerrymandering) satisfies Gerring's criteria, see Tom Ginsburg & Nicholas Stephanopoulos, *The Concepts of Law*, 84 U. CHI. L. REV. 147, 164-66 (2017).

[22] McGhee, *supra* note 19 (manuscript at 5); *see also* McGhee, *supra* note 5, at 61.

[23] *See* Stephanopoulos & McGhee, *supra* note 5, at 858-59.

[24] *See* McGhee, *supra* note 5, at 57 ("Some version of efficiency is typically the core concept of interest in the literature on redistricting.").

statewide (or nationwide) votes into the partisan division of the legislature."[25] Michael McDonald and Robin Best, two of the most prominent advocates of the mean-median difference, state that "[t]he fact of a gerrymander is evident when . . . one set of voters cannot achieve majority status [in the legislature] with anything close to the same efficiency as the opposing set of voters."[26] And Cover argues that the key to gerrymandering is to "distort the way political parties translate popular support (votes) into governmental power (seats)."[27]

To be clear, we do not assert that the efficiency principle is universally endorsed. Our more limited claim is that it is *commonly* considered to be the concept at the core of partisan gerrymandering.[28] This, of course, is the view we hold as well.

B. *Distinctness*

Our second criterion is that a gerrymandering metric should capture efficiency and *only* efficiency. It should not try to gauge other electoral values, nor should it be (in part or in full) a function of those values. Here the values we have in mind are goals that district plans may be able to promote or impede: electoral competitiveness (or how close races tend to be), proportional representation (or whether parties' vote shares equal their seat shares), voter participation (or turnout), and so on. We consider a metric to be flawed to the extent it reflects these values in addition to (or instead of) efficiency.

This distinctness requirement has theoretical and doctrinal roots. Theoretically, social scientists who study concept formation stress "[t]he importance of differentiation."[29] "A poorly bounded concept has definitional borders which overlap neighboring concepts."[30] "Such a concept is (ceteris paribus) less useful" because it cannot be reliably distinguished from adjacent ideas.[31]

Doctrinally, the Supreme Court has made clear on a number of occasions that a measure of partisan gerrymandering must not take into account lack of competitiveness or deviation from proportional representation. In the 1973 case of *Gaffney v. Cummings*, the Court reviewed a plan that protected both parties' incumbents from viable challenges. The Court unanimously upheld the plan, declaring that "judicial interest should be at its lowest ebb when a State purports fairly

---

[25] Andrew Gelman & Gary King, *Enhancing Democracy Through Legislative Redistricting*, 88 AM. POL. SCI. REV. 541, 543 (1994); *see also* Grofman & King, *supra* note 1, at 8 ("The key idea is that candidates of each political party should have equal opportunity in translating voter support into the division of legislative seats between the parties.").

[26] McDonald & Best, *supra* note 10, at 318; *see also* Wang, *Practical Tests*, *supra* note 10, at 370 (statement by another proponent of the mean-median difference that "[p]artisan gerrymandering creates a situation in which the same overall statewide vote share would lead to a very different level of representation for the redistricting party and its opposing target").

[27] Cover, *supra* note 13 (manuscript at 9).

[28] As Justice Souter once wrote, "the usual point of gerrymandering . . . is to control the greatest number of seats." Vieth v. Jubelirer, 541 U.S. 267, 346 (2004) (Souter, J., dissenting).

[29] John Gerring, *What Makes a Concept Good? A Criterial Framework for Understanding Concept Formation in the Social Sciences*, 31 POLITY 357, 376 (1999).

[30] *Id.*

[31] *Id.* at 378.

to allocate political power to the parties."[32] In all three of its direct confrontations with partisan gerrymandering claims, the Court also emphatically rejected disproportionality as a proper metric. "[T]he Constitution contains no such principle [of proportional representation]," the plurality held in the 2004 case of *Vieth v. Jubelirer*.[33] "It nowhere says that farmers or urban dwellers, Christian fundamentalists or Jews, Republicans or Democrats, must be accorded political strength proportionate to their numbers."[34]

In the academy, scholars appear to be split as to whether a measure should incorporate values other than efficiency. For instance, John Nagle writes with respect to efficiency and competitiveness that "it has been appropriately stressed that these are separate concepts."[35] Cho and Cover, on the other hand, argue for conflation. In Cover's words, a gerrymandering metric should not "privilege[] one democratic norm—symmetric partisan efficiency—over other relevant, and equally important, democratic norms, including electoral competitiveness, seats-votes proportionality, and voter participation."[36]

We take Nagle's side in this dispute. We do so in part for the reasons alluded to above: As a matter of concept formation, differentiation is preferable to amalgamation, and as a doctrinal matter, a metric loses its appeal to the degree it is based on lack of competitiveness or disproportionality. We also agree with Nagle for a more practical reason. We simply do not see how one measure could possibly encompass efficiency, competitiveness, proportionality, participation—and maybe even more "relevant, and equally important, democratic norms." These values are all conceptually and empirically distinct, so there is no way to combine them into a single uber-metric.

C. *Scope*

Our third criterion is that a measure of partisan gerrymandering should be applicable to a wide range of electoral conditions. It should be possible to calculate a metric (and for the results of the computation to be meaningful) if elections are competitive statewide or if one party dominates the statewide vote; if turnout is roughly equal or if it varies sharply from district to district; and if two parties are competing or if a multiparty environment exists.

Our rationale for this requirement is usefulness. The norm in most states may be for the two major parties to be relatively evenly matched, for no other parties to be electorally relevant,

---

[32] Gaffney v. Cummings, 412 U.S. 735, 754 (1973); *see also Vieth*, 541 U.S. at 350 (Souter, J., dissenting) (rejecting an approach to gerrymandering based on lack of competition "out of caution about a wholesale conceptual transfer from economics to politics"); Davis v. Bandemer, 478 U.S. 109, 130-31 (1986) (plurality opinion) (voicing no objection to "creating as many safe seats as the . . . political characteristics of the State would permit" and thus "leav[ing] the minority in each safe district without a representative of its choice").

[33] *Vieth*, 541 U.S. at 288 (plurality opinion).

[34] *Id.*; *see also* LULAC v. Perry, 548 U.S. 399, 419 (2006) (opinion of Kennedy, J.) ("[T]here is no constitutional requirement of proportional representation . . . ."); *Bandemer*, 478 U.S. at 130 (plurality opinion) (" Our cases . . . clearly foreclose any claim that the Constitution requires proportional representation . . . .").

[35] Nagle, *supra* note 12, at 347; *see also* McGann et al., *supra* note 12, at 296 n.1 (noting that partisan bias "was a major advance" because "it separates symmetry/bias . . . from responsiveness").

[36] Cover, *supra* note 13 (manuscript at 6); *see also* Cho, *supra* note 16 (manuscript at 13) ("[A] measure of partisan fairness should be able to distinguish maps with non-competitive districts that are not responsive to voters from a map that is comprised of competitive districts that are responsive to the voters.").

and for turnout to differ only modestly from district to district (at least since the advent of the one person, one vote rule). But this norm is far from universally satisfied. At the congressional level, for example, one party won more than 55% of the statewide vote in close to half of the elections from 1972 to the present.[37] In these elections, the average difference between the highest- and the lowest-turnout district in each state was also more than 100,000 voters.[38] "Atypical" electoral scenarios are thus not all that uncommon, and a gerrymandering metric should not be foiled by them.

We are unaware of any dissent from the general proposition that breadth of scope is a desirable attribute for a gerrymandering metric.[39] A few scholars argue, however, that a measure need only be applicable to competitive jurisdictions, where the major parties enjoy similar levels of statewide support.[40] We disagree. It is only in these jurisdictions that *control* of the legislature is realistically at stake. But gerrymandering is harmful even when it does not produce a legislative majority for a party that receives a minority of the statewide vote. If the practice converts a strong statewide performance into an overwhelming legislative supermajority—or if it does the reverse, barely allowing a very popular party to retain a legislative majority—it still inflicts democratic injuries we think a metric should be able to detect. Legislative control does not exhaust our notion of legislative clout.[41]

At the congressional level, moreover, it is obvious that gerrymandering can affect the *national* balance of power whether or not a state is competitive. An extra seat is an extra seat in the House of Representatives, whether it comes from a red, blue, or purple state. A measure that is valid only in swing states would thus miss a large part of the picture—and, indeed, might distort it by ignoring redistricting dynamics in safe states.[42]

D. *Correspondence*

Our fourth and final criterion is that a measure of partisan gerrymandering should be broadly consistent with the empirical realities of American elections. Every metric implies a certain electoral ideal: a perfect score (usually zero) indicating that neither major party is benefited or handicapped by a district plan. This ideal should be one that maps have actually

[37] This figure is based on the database assembled by Simon Jackman in the North Carolina partisan gerrymandering litigation. This database includes 512 observations from 136 district plans and 25 states (all with at least 7 congressional seats), spanning elections from 1972 to 2016. *See* Simon Jackman, *Assessing the Current North Carolina Congressional Districting Plan* 20 (Mar. 1, 2017); *see also* Stephanopoulos & McGhee, *supra* note 5, at 861 (reporting a similar figure for state legislative elections).

[38] The average standard deviation of district turnout in these elections was about 30,000 voters. Both this figure and that in the main text take into account imputations for uncontested seats.

[39] *See, e.g.*, Cover, *supra* note 13 (manuscript at 21) (arguing that a metric should not "ignor[e] significant real-world electoral phenomena" such as "third-party, independent and write-in candidates" and "less competitive states with less competitive plans").

[40] *See, e.g.*, Grofman & King, *supra* note 1, at 31 ("[T]he concept of partisan symmetry is only appropriate for competitive situations where there is a potential for a change in partisan outcomes (majority control, in particular) as a result of shifting electoral tides . . . ."); McDonald & Best, *supra* note 10, at 319 ("For a . . . gerrymander to rise to the level of an ascertainable constitutional offense a jurisdiction has to be politically competitive.").

[41] For a longer version of this argument, see Stephanopoulos & McGhee, *supra* note 5, at 861.

[42] *See id.*

achieved with some regularity in prior elections. It should not be an aspiration that plans have almost always failed to realize.

We include this requirement primarily for the sake of practicality. Courts are unlikely to embrace a measure that suggests that most challenged maps are (and most maps historically have been) gerrymanders.[43] Likewise, line-drawers probably have little use for a metric that condemns most of their efforts—including plans designed without partisan intent, or even with the aim of partisan fairness.[44] Conceptually as well, a measure that is inconsistent with American electoral history may be inapplicable to the single-member-district plans employed by most U.S. jurisdictions. The inconsistency may arise because the metric is only valid for other electoral systems.[45]

Nagle is the lone scholar who seems to have thought about empirical correspondence. He observes that different measures of gerrymandering imply different ideal relationships between parties' vote shares and seat shares. To choose between the metrics, "it appears that the best one can do is to appeal to the empirical result" of how seats and votes have, in fact, been linked in previous elections.[46] We concur, though we would not put the point quite so strongly. In our view, the empirical record should be *a* (not the *dispositive*) factor in selecting a measure.

II. EVALUATING THE EFFICIENCY GAP

How does the efficiency gap fare under these criteria? Quite well, as we show in this Part. It is always consistent with the efficiency principle; it is distinct from other electoral values; it can be meaningfully calculated in almost any electoral environment; and it is highly congruent with American electoral history.

A few notes before beginning our evaluation: First, in the course of assessing the efficiency gap, we also respond to a number of critiques of the metric. We address remaining objections in the following Part. Second, where possible, we support our arguments with data. It is our strongly held stance that measures of partisan gerrymandering cannot be gauged productively in the abstract. And third, we refrain from commenting here on other gerrymandering metrics. In the next two Parts, though, we apply our criteria to partisan bias, the mean-median difference, the difference between the parties' average margins of victory, and other measures.

---

[43] *See* Vieth v. Jubelirer, 541 U.S. 267, 306 (2004) (Kennedy, J., concurring in the judgment) (warning against legal theories that "would commit federal and state courts to unprecedented intervention in the American political process").

[44] *See* Evenwel v. Abbott, 136 S. Ct. 1120, 1132 (2016) (rejecting an approach to apportionment that "would upset a well-functioning approach to districting that all 50 States and countless local jurisdictions have followed for decades").

[45] It is true, of course, that previous Court interventions have embraced electoral ideals (like one person, one vote and the absence of racial vote dilution) that historically were violated by many district plans. But the current Court does not seem to have any appetite for another reapportionment revolution: that is, for a partisan gerrymandering standard that would "commit federal and state courts to unprecedented intervention in the American political process." Vieth v. Jubelirer, 541 U.S. 267, 306 (2004) (Kennedy, J., concurring in the judgment).

[46] Nagle, *supra* note 13, at 204.

A. *Efficiency*

Starting with the efficiency principle, it is not terribly surprising that it is satisfied by a metric whose name is the efficiency gap. Still, it is worth unpacking the logic of *why* the efficiency gap always moves in the correct direction when a party wins a larger seat share for the same vote share. For a party's seat share to increase while its vote share stays constant, the distribution of the party's votes across the districts in the plan must change. Specifically, the party must receive more votes than the opposing party in at least one district where the opposing party previously received more votes. These additional votes for the first party must be deducted from this party's votes in other districts (since the party's overall vote share is fixed). This can happen by reducing the party's margin of victory in districts it was already winning, or by increasing its margin of defeat in districts it was already losing. Either way, the party wastes fewer votes in these districts than it did before, while winning at least one new district. The efficiency gap, which simply compares the parties' respective wasted votes, necessarily shifts in the party's favor.

Figure 1 makes this point with a simple ten-district example (to which we return in Part IV). In both elections, Party A receives 55% of the vote. But in the first election Party A wins six out of ten seats, while in the second election it wins seven out of ten seats. Consistent with the efficiency principle, the efficiency gap goes from 0% in the first election to 10% in Party A's direction in the second election.

What accounts for this shift? In the district that flips from Party B to Party A (District 7), the allocation of wasted votes swings dramatically: from 44 wasted A votes and 6 wasted B votes in the first election, to 3 wasted A votes and 47 wasted B votes in the second election. In three other districts, the allocation of wasted votes also differs modestly. Districts 2 and 3 each have 23 wasted A votes and 27 wasted B votes in the first election, and 20 wasted A votes and 30 wasted B votes in the second election. And District 9 has 44 wasted A votes and 6 wasted B votes in the first election, and 41 wasted A votes and 9 wasted B votes in the second election. All of these changes benefit Party A, and lead to it wasting just 200 votes in the second election compared to 250 in the first one. Party B, in contrast, sees its wasted votes rise from 250 to 300.[47]

[47] For ease of calculation, we assume that 50 votes are needed to win a district, not 51. Using 51 votes as the threshold instead, the efficiency gap is 0.2% in the first election and 10.4% in the second election (both times in favor of Party A).

FIGURE 1: EFFICIENCY GAP EXAMPLES

	ELECTION 1				**ELECTION 2**			
District	**Votes**		**Wasted Votes**		**Votes**		**Wasted Votes**	
	Party A	*Party B*	*Party A*	*Party B*	*Party A*	*Party B*	*Party A*	*Party B*
1	73	27	23	27	73	27	23	27
2	73	27	23	27	*70*	*30*	*20*	*30*
3	73	27	23	27	*70*	*30*	*20*	*30*
4	53	47	3	47	53	47	3	47
5	53	47	3	47	53	47	3	47
6	53	47	3	47	53	47	3	47
7	44	56	44	6	*53*	*47*	*3*	*47*
8	44	56	44	6	44	56	44	6
9	44	56	44	6	*41*	*59*	*41*	*9*
10	40	60	40	10	40	60	40	10
Total	**550**	**450**	**250**	**250**	**550**	**450**	**200**	**300**
EG	**(250 – 250) / 1000 = 0%**				**(300 – 200) / 1000 = 10%**			

Votes and wasted votes by district in two elections under the same ten-district plan. The efficiency gap is calculated using the full method. Votes and wasted votes are placed in italics in Election 2 where they are different from in Election 1.

Any single example, of course, may be misleading. To allay this concern, one of us used a computer program to simulate 5100 district plans, 100 for each one-point increment of statewide vote share between 25% and 75%.[48] Each of these plans had 100 districts whose turnout was allowed to vary by up to a factor of 15.[49] Each plan also included a third party that received anywhere from 0% to 20% of the statewide vote.[50]

Figure 2 displays the results of the simulations. The x-axis in each chart represents the vote share of the "benchmark" party (simply one of the two major parties). The y-axis indicates the proportion of the plans (paired in every possible permutation) that violate the efficiency principle. The black points assume equal turnout in each district while the gray points permit differential turnout. The five rows correspond to different vote shares for the third party. And of the four columns, only the first one, for the efficiency gap, is of interest here.[51] (The other three denote variants of the efficiency gap that we discuss in Part III.)

It is immediately evident that the efficiency gap *never* violates the efficiency principle in *any* of the simulations. The plot for the efficiency gap remains flat at zero no matter what vote share the benchmark party or the third party receives and whether turnout is equal or varies from district to district. This is powerful confirmation that there is nothing idiosyncratic about the ten-district example we presented above. As far as we can tell, the efficiency gap always complies with the efficiency principle, in every electoral setting.

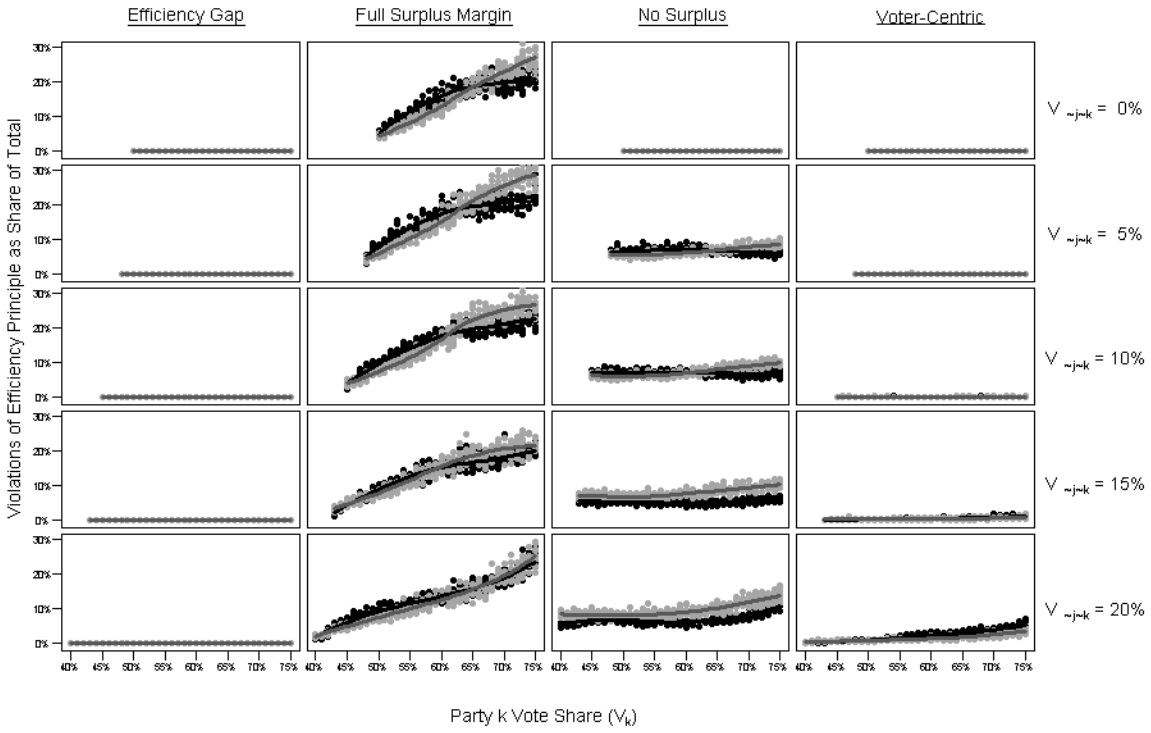
[48] *See* McGhee, *supra* note 19 (manuscript at 16-17).

[49] *See id.* (manuscript at 37).

[50] *See id.* (manuscript at 38).

[51] Both here and whenever district turnout is not assumed to be equal, we use the form of the efficiency gap presented in McGhee, *supra* note 19 (manuscript at 24-27).

FIGURE 2: EFFICIENCY GAP SIMULATIONS



The charts plot the proportion of simulated plans that violate the efficiency principle for the original efficiency gap and three of its variants. The x-axis represents the vote share of the benchmark party. The black points assume equal turnout in each district while the gray points permit differential turnout. The five rows correspond to different vote shares for the third party.

B. *Distinctness*

Turning to our distinctness criterion, Cover argues that the efficiency gap is *not* distinct from electoral competitiveness, proportional representation, or voter participation. In his view, the measure promotes uncompetitive elections, deters plans with seats in proportion to votes, and fosters voter suppression. We respond to these claims in turn.

With respect to competitiveness, Cover's reasoning is as follows: At the level of an individual district, the parties' wasted votes are equal when the vote is split 75% to 25%. (Each party wastes 25% of the vote in this case.) Therefore a district plan made up exclusively of 75%-to-25% districts would also feature equal wasted votes for each party, and hence an efficiency gap of zero. But 75%-to-25% districts are highly uncompetitive, with a whopping margin of victory of 50%.[52]

[52] *See* Cover, *supra* note 13 (manuscript at 27) ("Since a party can win a zero-disparity district with three quarters of ballots cast, a party can win every seat in a zero-gap plan with 75% vote share.").

Cover himself identifies the fatal flaw in this logic: While a district plan made up exclusively of 75%-to-25% districts has an efficiency gap of zero, *so do countless other plans, many of which are competitive*. As Cover puts it, the "simple minimizing plan" composed entirely of 75%-to-25% districts "is not the only minimizing plan" with a zero efficiency gap.[53] "A zero-gap plan can be highly competitive or highly uncompetitive, as long as it is not *differentially* competitive."[54] In other words, there is no necessary relationship between the efficiency gap and the average margin of victory under a plan. Any efficiency gap is compatible with any average margin of victory.

Cover has two rejoinders to his own rebuttal. One is that "each minimizing plan" with a zero efficiency gap "can be converted to . . . a [simple] minimizing plan" composed entirely of 75%-to-25% districts "by performing the appropriate series of voter swaps."[55] This observation does not establish a connection between the efficiency gap and competitiveness. That a competitive zero-efficiency-gap plan can, in principle, be redrawn into an uncompetitive zero-efficiency-gap plan does not render the former any less competitive. (Nor, in reverse, is an uncompetitive zero-efficiency-gap plan any more competitive just because, theoretically, it can be reshaped into a competitive zero-efficiency-gap plan.)

Cover's other point is that, of the universe of zero-efficiency-gap plans, only uncompetitive ones "can maintain a zero (or low) gap over a range of vote share."[56] This is incorrect. Take a "simple minimizing plan" made up exclusively of 75%-to-25% districts. This plan's efficiency gap is guaranteed *not* to remain low if the statewide vote shifts substantially in either party's favor. This is because none of the plan's (very safe) seats would flip, meaning that the parties' wasted votes would become increasingly unequal as the electoral environment evolved. Indeed, the only kind of plan whose efficiency gap *would* stay low in the face of significant input variation is a reasonably (though not maximally) competitive one.[57] In such a plan, enough districts *would* change hands as the statewide vote shifted to keep the parties' wasted votes in rough equilibrium.[58]

We mentioned earlier that, when evaluating measures of partisan gerrymandering, we are partial to data-driven analysis. In this spirit, Figure 3 shows how the efficiency gap and electoral competitiveness have been linked in state house and congressional elections from 1972 to the present.[59] The left chart is for state house elections (which we consider more informative due to their larger numbers of districts); the right chart is for congressional elections. The x-axis in both charts is the average margin of victory across all of a state's districts in a given year—a standard

[53] Cover, *supra* note 13 (manuscript at 33).

[54] *Id.* (manuscript at 35).

[55] *Id.* (manuscript at 33).

[56] *Id.* (manuscript at 36); *see also id.* (manuscript at 38) ("only two of ten districts are relatively competitive" in a durably zero-efficiency-gap plan).

[57] Specifically, as discussed below, a plan with enough competitive districts to yield a seat-to-vote responsiveness of two would maintain a given efficiency gap as the statewide vote shifted back and forth. This is the responsiveness that American elections have historically exhibited. *See infra* Part II.D.

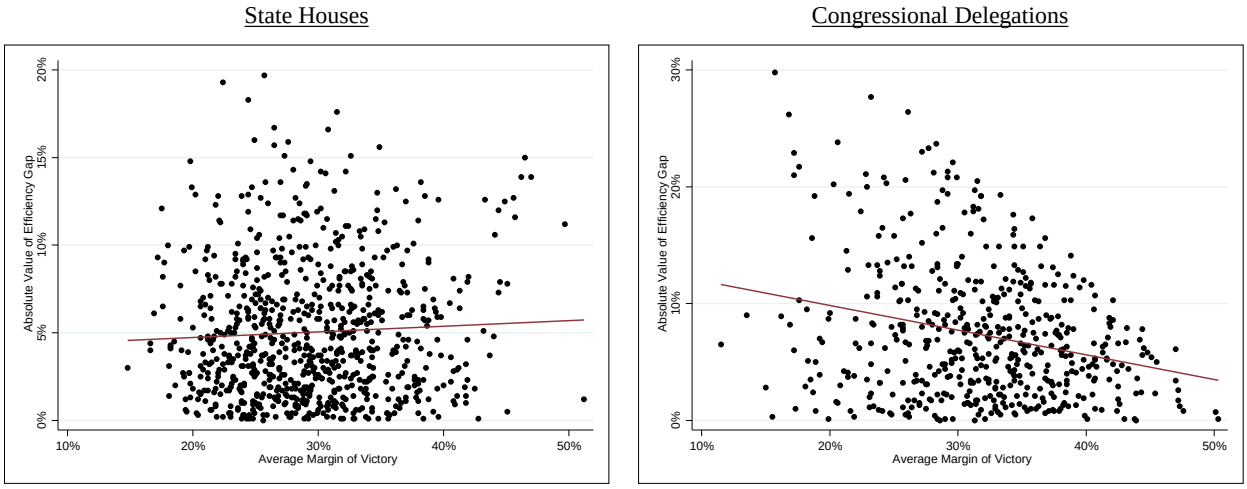
[58] Cover concedes the point. *See id.* (manuscript at 37-38).

[59] We described the congressional database above. *See supra* note 37. The state house database was assembled by Simon Jackman in the Wisconsin partisan gerrymandering litigation. It includes 786 observations from 207 district plans and 41 states spanning elections from 1972 to 2014. *See* Simon Jackman, *Assessing the Current Wisconsin State Legislative Districting Plan* 32 (July 7, 2015).

gauge of competitiveness.[60] The y-axis in both charts is the absolute value of the efficiency gap.[61]

At the state house level, there is plainly no relationship between electoral competitiveness and the magnitude of the efficiency gap. How close races tend to be tells us nothing about the size of a plan's partisan skew. At the congressional level, there *is* a negative correlation, but it is quite weak. In fact, the average margin of victory accounts for only about 6% of the variation in the absolute value of the efficiency gap.[62] The average margin of victory also fails to attain statistical significance when it is added to a causal model that one of us has constructed for the efficiency gap.[63] We think these findings dovetail nicely with the conceptual position we outlined above. In theory, any efficiency gap may coincide with any average margin of victory—and in practice too, these variables are virtually unrelated.[64]

FIGURE 3: ELECTORAL COMPETITIVENESS AND THE EFFICIENCY GAP



Scatter plots of the absolute value of the efficiency gap versus the average margin of victory in an election, for state houses and congressional delegations.

Next, with respect to proportional representation, Cover's objection is that it is *not* equivalent to the efficiency gap.[65] Because the concepts are distinct, "[a] plan may achieve the

[60] *See, e.g.*, Nicholas O. Stephanopoulos, *The Consequences of Consequentialist Criteria*, 3 UC IRVINE L. REV. 669, 685 (2013).

[61] We use the simplified form of the efficiency gap in the state house chart (because the full form is unavailable at this electoral level) and the full form in the congressional chart.

[62] The scatter plot suggests that the negative relationship is driven by a small number of competitive plans with large efficiency gaps.

[63] *See* Nicholas O. Stephanopoulos, *The Causes and Consequences of Gerrymandering*, 59 WM. & MARY L. REV. (forthcoming 2018) (manuscript at 24) (presenting the model's results at the congressional level).

[64] For a finding that the efficiency gap and competitiveness are distinct—but based on simulated rather than historical plans—see Bruce E. Cain et al., *A Reasonable Bias Approach to Gerrymandering: Using Automated Plan Generation to Evaluate Redistricting Proposals*, 59 WM. & MARY L. REV. (forthcoming 2018) (manuscript at 13).

[65] The reason why disproportionality is not equivalent to the efficiency gap is that the former measures a plan's deviation from a *1:1* seat-to-vote relationship. In contrast, a *2:1* seat-to-vote relationship is necessary to

ideal of equal wasted votes at the expense of . . . seats-votes proportionality."[66] Likewise, "[a] plan that produces rough proportionality" may also "produce a large efficiency gap."[67]

In our opinion, these possibilities are a feature of the efficiency gap, not a bug. Recall that the Supreme Court has disavowed disproportionality, in no uncertain terms, as an appropriate standard in partisan gerrymandering cases.[68] Thanks to this precedent, any metric that is synonymous with disproportionality is precluded from judicial consideration. It is thus the efficiency gap's lack of overlap with disproportionality that explains why judges have been willing to entertain it. If (as Cover wishes) the efficiency gap always led to the same results as disproportionality, it would not have made it through the courthouse door.[69]

There is another doctrinal problem with Cover's argument. If a jurisdiction intentionally enacted a district plan that achieved proportionality—but that also exhibited a large efficiency gap—this jurisdiction would be insulated from liability. After all, no one has proposed that maps be struck down whenever their efficiency gaps exceed a certain threshold. Rather, the only test on the table requires a large and durable efficiency gap *and* the presence of discriminatory intent *and* the absence of any legitimate justification before a plan is invalidated.[70] Under this test, the jurisdiction that aimed for proportionality would not have harbored any illicit intent.[71] There would also be a compelling justification for the map's large efficiency gap: namely, the jurisdiction's pursuit of proportionality.

Practically as well, Cover's argument is quite limited in its scope. American states rarely try to achieve proportionality when they draw district lines. In fact, not a single state (and as far as we know, not a single other American government) lists proportional representation as a statutory or constitutional redistricting criterion.[72] Additionally, as we discuss below, the single-member-district plans used by most American jurisdictions do not typically yield proportional outcomes.[73] This is because most of these plans include a substantial number of competitive districts. As the parties' fortunes ebb and flow, these districts tend to change hands too rapidly to maintain the 1:1 seat-vote ratio required for proportionality.

---

maintain an efficiency gap of zero (or any other value). Thus in the first election in Figure 1, where Party A wins 55% of the vote and 60% of the seats, disproportionality is 5% (60% – 55%) while the efficiency gap is zero.

[66] Cover, *supra* note 13 (manuscript at 39).

[67] *Id.*

[68] *See supra* notes 33-34 and accompanying text.

[69] *See* Whitford v. Nichol, 151 F. Supp. 3d 918, 929-30 (W.D. Wis. 2015) (*Whitford I*) (denying the state's motion to dismiss precisely because "an election's results may have a small efficiency gap without being proportional or they may be proportional and still have a large efficiency gap").

[70] *See Whitford II*, 218 F. Supp. 3d at 883-927 (describing and analyzing the three prongs of the test proposed by the plaintiffs). In this Article, we focus on the efficiency gap specifically rather than the entirety of the suggested legal standard for partisan gerrymandering.

[71] *See id.* at 910 ("[D]rafters who had the intent to create a proportional system hardly could be accused of harboring a discriminatory intent.").

[72] *See* NAT'L CONF. OF STATE LEGISLATURES, REDISTRICTING LAW 2010, at 172-217 (2009) (listing all fifty states' redistricting criteria). A handful of municipalities use different electoral systems that ensure proportional representation, but this is plainly not the same as trying to achieve proportionality in a single-member-district system.

[73] *See infra* Part II.D; *see also* Grofman & King, *supra* note 1, at 6 n.30 ("[I]t is widely recognized that a proportionality standard is simply not appropriate for use as a legal standard in plurality based elections, since plurality elections cannot be expected to yield proportional results.").

The last electoral value that Cover links to the efficiency gap is voter participation. He claims that if the efficiency gap were incorporated into a legal test for partisan gerrymandering, this would "incentivize and reward voter suppression."[74] Why? Because restrictive electoral rules, if they disproportionately affect the gerrymandered party's supporters, increase the vote share of the gerrymandering party. All else being equal, a larger vote share for a party produces a smaller efficiency gap in favor of this party, since it now wastes more votes in both the districts it wins (by wider margins) and those it loses (by narrower margins). A smaller efficiency gap, in turn, would reduce the party's exposure to gerrymandering liability.[75]

We highly doubt the plausibility of this causal chain. Over the last decade—before the efficiency gap was even invented, let alone invoked in lawsuits—many states adopted restrictive electoral rules.[76] They did so for the obvious reason that a larger vote share is valuable to parties *in and of itself*, even if it does not affect the odds that the parties will be found liable for gerrymandering. For Cover's contention to have any bite, there would have to exist parties that would not engage in voter suppression if the efficiency gap were excluded from the doctrine, but that would carry it out if the efficiency gap were included in the law. Such parties—so attuned to benefits in litigation, but so oblivious to ordinary political advantages—strike us as unicorns that the world has never seen.

Moreover, even if one of these odd creatures could be located, its efforts might very well backfire. Figure 4 plots the efficiency gap (on the y-axis) versus the Democratic share of the statewide vote (on the x-axis) for state house elections (left chart) and congressional elections (right chart).[77] At both electoral levels, there is a modest but unmistakable *positive* relationship between these variables. Thus if Democrats increase their vote share (through the suppression of Republican voters or other means), the efficiency gap tends to become more, not less, pro-Democratic. Similarly, if Republicans win a larger share of the vote, the efficiency gap again shifts toward rather than against them.

How can this be, given the statement above that, all else being equal, a larger vote share for a party produces a smaller efficiency gap in favor of this party? Because all else is *not* equal. Rather, when a party increases its vote share, it typically wins more seats too. And it typically wins *so many* more seats that the efficiency gap moves in its favor notwithstanding the extra votes it wastes in other districts. A party that curbed its opponents' voting in order to avoid gerrymandering liability, then, could easily find itself in a more precarious position. Its plan's efficiency gap could well become larger, placing the map on thinner legal ice.

---

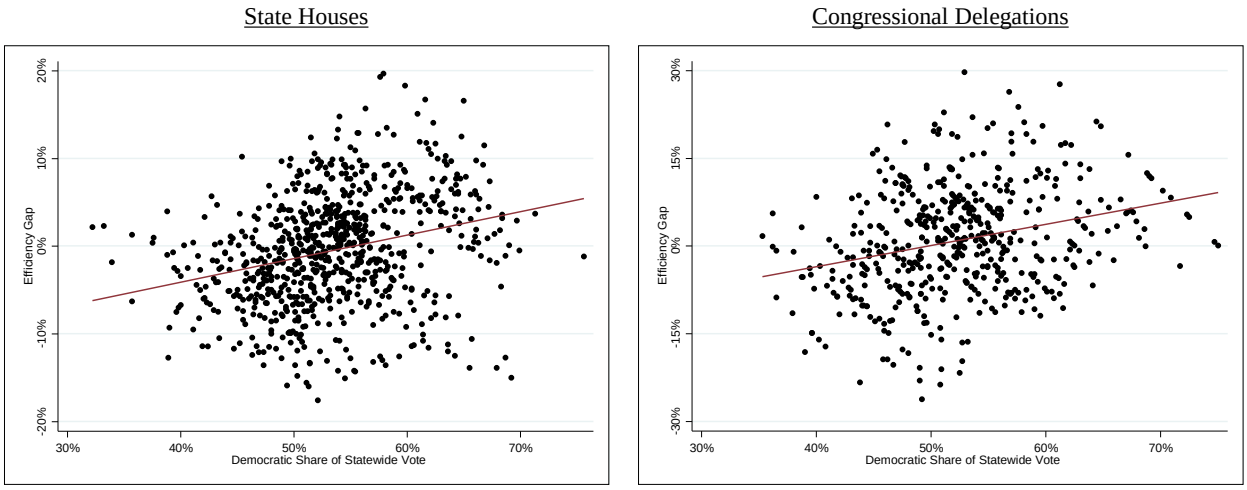
[74] Cover, *supra* note 13 (manuscript at 55).

[75] *See id.* (manuscript at 60-61) ("Suppressing one party's statewide vote total can have the effect of reducing the overall gap.").

[76] *See New Voting Restrictions in America*, BRENNAN CTR. FOR JUSTICE (Mar. 1, 2017), http://www.brennancenter.org/new-voting-restrictions-america (tracking the adoption of voting restrictions).

[77] Positive efficiency gap scores are pro-Democratic; negative scores are pro-Republican.

FIGURE 4: STATEWIDE PARTY PERFORMANCE AND THE EFFICIENCY GAP



Scatter plots of the efficiency gap versus the Democratic share of the statewide vote in an election, for state houses and congressional delegations.

C. *Scope*

Shifting gears from distinctness to breadth of scope, a nice feature of the efficiency gap is that it can be calculated—meaningfully—no matter how competitive or uncompetitive a jurisdiction happens to be. Whether the parties are evenly matched or one party predominates in a state, the parties' respective wasted votes can be tallied and compared.[78] Cho[79] and Cover[80] complain that when a party receives 75% or more of the statewide vote, the efficiency gap can be minimized only if this party wins all of the seats.[81] But this is not actually a troublesome result. In principle, we would expect it to be quite difficult for a small minority party to muster a districtwide majority given the electorate's overwhelming support for the opposing party. And empirically, in both of the elections in our databases where a party received at least 75% of the statewide vote, it indeed won every seat.[82] Alabama Democrats won 75.6% of the state house vote along with all 105 state house seats in 1974. Massachusetts Democrats also won 75.0% of the congressional vote along with all 10 congressional seats in 2006.

Another appealing attribute of the efficiency gap is that it can be computed if turnout is assumed to be equal or if it varies from district to district. If turnout is assumed to be equal, the "simplified form" of the efficiency gap can be used: $(S - 0.5) - 2 \times (V - 0.5)$, where S is a party's

[78] *See* Stephanopoulos & McGhee, *supra* note 5, at 855 (making this point).

[79] *See* Cho, *supra* note 16 (manuscript at 15) (discussing a scenario where "party A wins both districts with an overwhelming 75–25 margin").

[80] *See* Cover, *supra* note 13 (manuscript at 34) ("When one party earns 75% of the votes, the simple minimizing plan accords that party all the seats . . . .").

[81] We first noted this implication ourselves. *See* McGhee, *supra* note 5, at 70; Stephanopoulos & McGhee, *supra* note 5, at 863-64.

[82] It is revealing, of course, that out of the nearly 1300 cases in our databases, only *two* feature a statewide vote share above 75%. The scenario to which Cho and Cover object is thus vanishingly rare. For more information on our databases, see notes 37, 59, *supra*.

statewide seat share and *V* is the party's vote share averaged across all districts.[83] To take into account district-level variations in turnout, the "full form" of the efficiency gap must be used instead. Under this method, the parties' respective wasted votes are counted (district by district) and then totaled. Next, one of these sums is subtracted from the other, and the difference is divided by the total number of votes cast.[84]

Although he cites no evidence to justify his worry, Cover "raises concerns about the degree of divergence of results calculated using the long-form formula and those calculated using the simplified formula."[85] Figure 5 supplies the data that is missing from Cover's critique. It plots the full form of the efficiency gap (on the y-axis) versus the simplified form of the efficiency gap (on the x-axis) for congressional elections. (We do not include an analogous chart for state house elections because we do not currently have full-form estimates at this electoral level.) It is obvious from the scatter plot that the two forms of the efficiency gap are very closely linked. The points hug the best fit line, and the correlation between the two methods is an impressive *0.97*.[86] There is thus little chance that one's substantive conclusions would change based on the technique one used to calculate the efficiency gap. Turnout does vary from district to district, of course, but it does *not* vary much between districts won by one party and districts won by the other. As Cover observes, it is only such "*average differential voter turnout*" that would drive a wedge between the two forms of the efficiency gap.[87]

[83] *See* McGhee, *supra* note 5, at 68-70; Stephanopoulos & McGhee, *supra* note 5, at 850-53. The reason to use the simplified form of the efficiency gap is that sometimes only vote shares (not raw votes) are available by district. In this case, the full form cannot be employed.

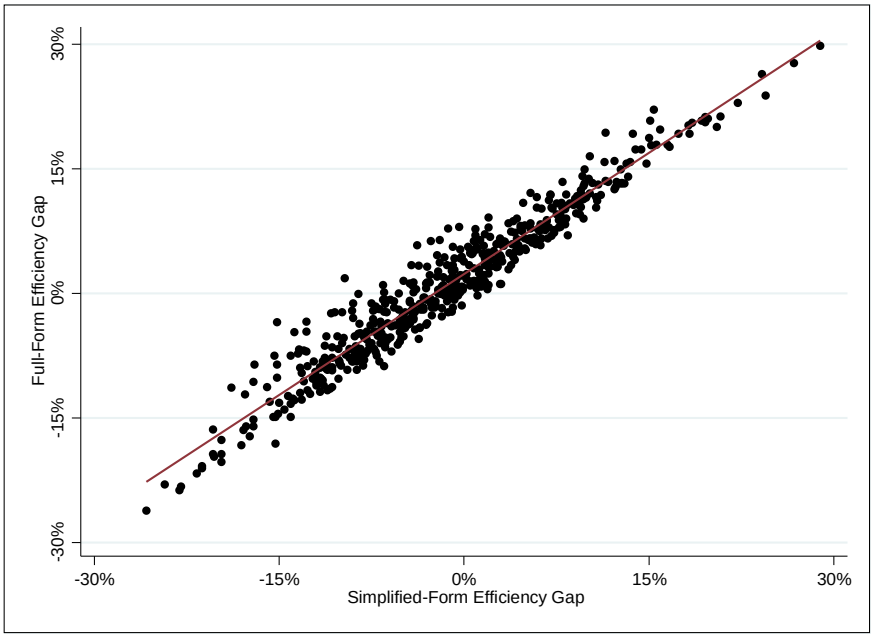
[84] *See id.* As noted earlier, whenever we use the full form of the efficiency gap in this Article, we include the adjustment described in McGhee, *supra* note 19 (manuscript at 24-27). *See supra* note 51.

[85] Cover, *supra* note 13 (manuscript at 55-56).

[86] The cases furthest from the best fit line are concentrated in New York, which has numerous low-turnout districts that are almost always won by Democrats.

[87] *Id.* (manuscript at 56).

FIGURE 5: THE FULL FORM AND THE SIMPLIFIED FORM OF THE EFFICIENCY GAP



Scatter plot of the efficiency gap calculated using the full form versus the efficiency gap calculated using the simplified form, for congressional elections.

The last electoral setting in which the efficiency gap is applicable is the multiparty context. One of us recently showed how the measure can be calculated when there are more than two parties competing for office.[88] We do not repeat the technical details here, but rather highlight three points. First, the efficiency gap is always computed in *pairwise* fashion—that is, for two focal parties while combining all of the seats and votes won by other parties. These two focal parties can be the two major parties, a major party and a minor party, or two minor parties.[89] Second, when using the full method for more than two parties, a party can have positive *or negative* wasted votes when it wins a district. As in the two-party case, a party's wasted votes are positive when it receives more than 50% of the district vote. But a party's wasted votes are negative (or beneficial to it) when it receives less than 50% of the district vote but nevertheless prevails because of how the remaining vote is split among its adversaries.[90]

And third, the simplified form of the efficiency gap in the multiparty context is: $(S_A - 0.5) - 2 \times (V_A - 0.5) - V_C + 0.5 \times S_C$, where S_A is one focal party's statewide seat share, V_A is this party's statewide vote share, V_C is the combined statewide vote share of all non-focal parties, and S_C is the combined statewide seat share of all non-focal parties.[91] Assume, for example, that Party A wins 45% of the vote and 55% of the seats, that Party B wins 40% of the vote and 40%

[88] *See* McGhee, *supra* note 19 (manuscript at 30-36).

[89] *See id.* (manuscript at 30).

[90] *See id.* (manuscript at 31-32). Assume, for example, that a party wins a district 40% to 30% to 30%. Then it has -10% wasted votes because it won the district with 10% less of the vote than was necessary to *guarantee* victory.

[91] *See id.* (manuscript at 33).

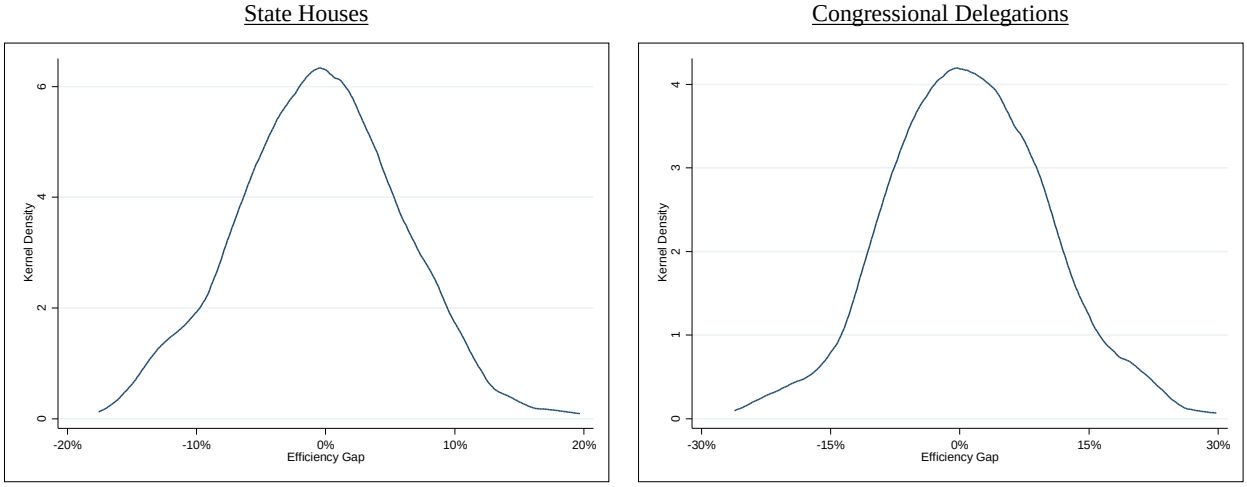
of the seats, and that all other parties win 15% of the vote and 5% of the seats. Then the efficiency gap for Party A relative to Party B is: (0.55 – 0.5) – 2 × (0.45 – 0.5) – 0.15 + 0.5 × 0.05 = 0.05 + 0.1 – 0.15 + 0.025 = 0.025 (or 2.5%) in favor of Party A.

D. *Correspondence*

Our final criterion is empirical correspondence with American electoral history. One way to assess the efficiency gap's correspondence is to examine its distribution over a long period of time. If the distribution is normal (or shaped like a bell curve) and centered on zero, this would indicate that many plans historically have managed to score well on the measure. Conversely, if the distribution is skewed and there are few values near zero, this would suggest that a low efficiency gap is an unrealistic aspiration for many American jurisdictions.

Figure 6 presents two density curves (or smoothed versions of histograms): on the left, for the efficiency gap in state house elections from 1972 to the present; and on the right, for the efficiency gap in congressional elections from 1972 to the present.[92] Both curves are textbook normal distributions with means and medians very close to zero. In other words, over nearly half a century, most plans have not significantly benefited (or handicapped) either party, and there has been no overall skew in either party's direction. We think this is strong evidence in favor of the efficiency gap's empirical correspondence. The metric does not expect the impossible of American jurisdictions; rather, it is highly consistent with their actual electoral outcomes over many years.

FIGURE 6: EFFICIENCY GAP DENSITY CURVES



Kernel density curves for the efficiency gap, for state houses and congressional delegations.

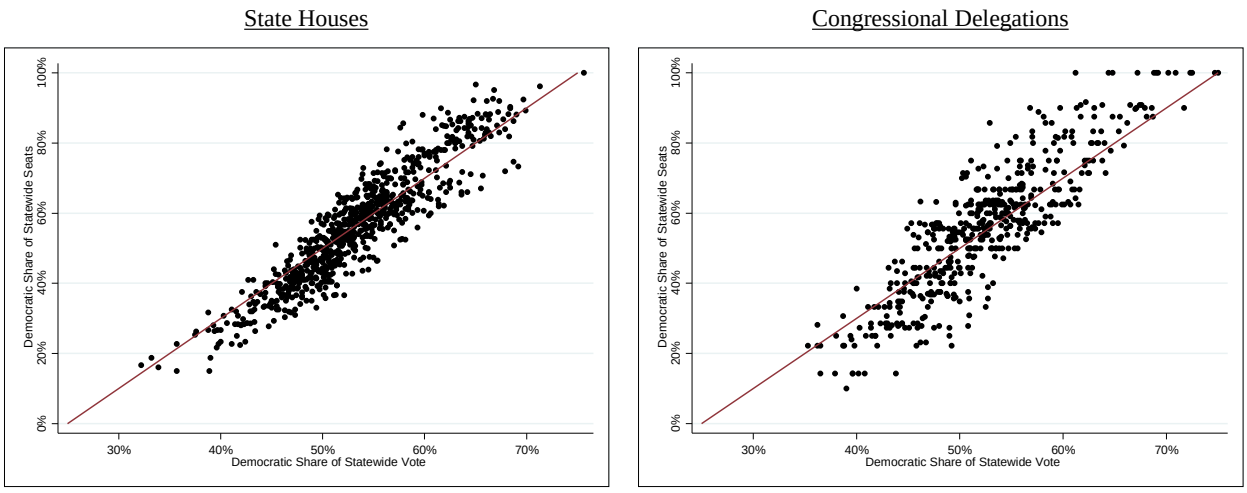
Another test for the efficiency gap is whether its implied seat-to-vote relationship resembles the one in fact exhibited by American elections. Using either the full or the simplified method, the efficiency gap stays constant (at zero or any other value) only if a party's seat share

[92] For similar density curves, see Stephanopoulos & McGhee, *supra* note 5, at 870.

changes at twice the rate of its vote share.[93] It would therefore be encouraging if seats and votes have been linked in roughly a 2:1 ratio in prior American elections. But it would be concerning if historical responsiveness has diverged markedly from two.

Figure 7 plots the Democratic share of statewide seats (on the y-axis) versus the Democratic share of the statewide vote (on the x-axis) for state house elections (left chart) and congressional elections (right chart). The charts also include a seat-vote line with a slope of precisely two. In both cases, the points cluster quite tightly around this line. The fit is not *perfect*—at both electoral levels, historical responsiveness has been slightly higher than two—but it is still very good.[94] In our view, this is further confirmation of the efficiency gap's empirical correspondence. Not only have most elections produced efficiency gaps close to zero, but when seats and votes have changed from year to year, they have done so at just about the right rates to maintain a low efficiency gap.

FIGURE 7: PARTIES' SEAT SHARES AND VOTE SHARES



Scatter plots of the Democratic share of statewide seats versus the Democratic share of the statewide vote, for state house and congressional elections. A seat-vote line with a slope of two is included in each chart.

III. CRITIQUES OF THE EFFICIENCY GAP

We responded to several critiques of the efficiency gap in the previous Part while discussing its compliance with our criteria for measures of partisan gerrymandering. But we did not address three further sets of objections. The first involves the methodological choices that underpin the efficiency gap: in particular, how wasted votes are defined and weighted, and how imputations are made for uncontested races. The second relates to the metric's variability, both in toy examples of one or two districts and in actual district plans. And the third stems from the

[93] *See* McGhee, *supra* note 5, at 68-69; Stephanopoulos & McGhee, *supra* note 5, at 854.

[94] For additional findings that responsiveness at the congressional level has been close to two, see Goedert, *Gerrymandering or Geography*, *supra* note 13, at 2-3; Edward R. Tufte, *The Relationship Between Seats and Votes in Two-Party Systems*, 67 AM. POL. SCI. REV. 540, 542-43 (1973).

efficiency gap's conflation of all legislators affiliated with a given party, regardless of their ideologies.

As to methodology, we agree that our approach is not the only reasonable way to treat wasted votes. The trouble with other positions, though, is that they result in violations of the efficiency principle or a lack of correspondence with American electoral history. As to variability, the toy examples are largely irrelevant and do not distinguish the efficiency gap from other measures. Real-world evidence is how we think the efficiency gap's stability should be evaluated—and in our opinion, this data tells quite a positive story. And as to ideology, no gerrymandering metric distinguishes between moderate and extreme legislators who belong to the same party. There is also little reason to make this distinction, given the extraordinary polarization of both state legislatures and Congress.

A. *Methodology*

When we developed the full form of the efficiency gap, we defined *surplus votes* as votes cast for the winning candidate in a race, above the 50% (plus one) threshold needed for victory.[95] We also weighted surplus votes and *lost votes* (votes cast for the losing candidate) equally. At the time, both of these methodological choices seemed so intuitive to us that we barely bothered defending them. In a two-candidate race, 50% (plus one) is the line between victory and defeat. It therefore struck us as the obvious benchmark for defining surplus votes. Similarly, surplus votes and lost votes are equally wasted in that neither one contributes to a candidate's election. We thus saw no reason to weight one kind of wasted vote more or less heavily than the other.

These choices, it turns out, are not as incontrovertible as we thought. Cover[96] and a dissenting federal judge[97] both suggest defining surplus votes relative to the losing candidate's performance, not 50% (plus one). Under this approach, *20%* of the vote would be wasted when a candidate wins 60% to 40% (60% – 40%), not *10%* (60% – 50%). Cho, Cover, and Nagle also propose different weighting schemes for surplus and lost votes. Cho argues that only surplus votes should be taken into account.[98] Cover adjusts the weight of surplus votes (relative to lost votes) from zero to one.[99] And Nagle varies this weight from zero (meaning only lost votes would be counted) to infinity (counting only surplus votes).[100]

We regard all of these variants of the original efficiency gap as unwise. Defining surplus votes relative to the losing candidate's performance overlooks the fact that, in a two-candidate race, all votes not cast for one candidate are cast for her opponent. If a candidate wins 60% to 40%, for example, this candidate would *lose* if 20% of the vote was deemed surplus, deducted from her total, and then added to her adversary's tally. The maximum vote the candidate could

---

[95] *See* McGhee, *supra* note 5, at 68-70; Stephanopoulos & McGhee, *supra* note 5, at 850-53.

[96] *See* Cover, *supra* note 13 (manuscript at 41-46); *see also* Best et al., *supra* note 10 (manuscript at 8) (also mentioning this point).

[97] *See Whitford II*, 218 F. Supp. 3d at 958 (Griesbach, J., dissenting) ("[A]ll you need to win an election in a two-candidate race is one more vote than the other candidate, not 50%-plus-one of the total votes.").

[98] *See* Cho, *supra* note 16 (manuscript at 6).

[99] *See* Cover, *supra* note 13 (manuscript at 43-46).

[100] *See* Nagle, *supra* note 13, at 200-01; *see also* Warrington, *supra* note 13, at 12-14 (also noting several different ways to treat wasted votes).

afford to cede to her rival while still prevailing is 10% (minus one)—or exactly our definition of surplus votes.

Analogously, we disagree with Cho that it is "more intuitive" to "count[] only excess *winning* votes."[101] Doing so implies that lost votes are immaterial and that the cracking that produces them need not be incorporated into the efficiency gap. But cracking is a fundamental gerrymandering technique that no valid measure should ignore. By the same token, we disagree with Cover and Nagle that surplus votes are less problematic than lost votes. True, "the voter who casts a surplus vote gets to be represented by the candidate of her choice."[102] But gerrymandering is inherently a *statewide* activity whose essence is the more efficient conversion of votes into seats *in the aggregate*.[103] From this perspective, district-specific representation is beside the point, and surplus votes are every bit as ineffective as lost votes.

While we find our rejoinders to be persuasive, they rely on understandings of wasted votes that may not be universally shared. There are two more problems with the variants of the efficiency gap, though, that are less conceptual and more rooted in the criteria we presented earlier. First, one of us used district plan simulations to assess the variants' consistency with the efficiency principle.[104] (See the last three columns of Figure 2.) *None* of the variants satisfies the principle. The worst offender is the efficiency gap calculated with the alternative definition of surplus votes, which violates the principle in almost all electoral environments. The efficiency gap calculated with only lost votes fares better, but still increasingly violates the principle as more of the statewide vote goes to third parties. The same is true for a "voter-centric" variant of the efficiency gap introduced by Nagle[105] (and reprised by Cover[106]), which compares the parties' respective *shares* (rather than *numbers*) of wasted votes. It too increasingly violates the efficiency principle as third-party performance improves.[107]

The second problem with the efficiency gap variants is that they are at odds with American electoral history. Recall that the original efficiency gap stays constant if seat share changes at twice the rate of vote share—and that responsiveness has, in fact, been very close to two in state house and congressional elections over the last half-century.[108] In contrast, the efficiency gap calculated with the alternative definition of surplus votes implies an ideal responsiveness of *three*.[109] This is a higher level of responsiveness than American elections have exhibited in the modern era. Likewise, the efficiency gap calculated with only lost votes,[110] and

---

[101] Cho, *supra* note 16 (manuscript at 6).

[102] Cover, *supra* note 13 (manuscript at 42).

[103] *See supra* notes 23-27 and accompanying text.

[104] *See* McGhee, *supra* note 19 (manuscript at 37-42).

[105] *See* Nagle, *supra* note 13, at 201-03.

[106] *See* Cover, *supra* note 13 (manuscript at 46-52); *see also* McGhee, *supra* note 5, at 69 (noting that the efficiency gap could also be defined in this way).

[107] As for Cho's surplus-vote-only variant, Nagle has shown that surplus votes can *only* be equalized when each party receives exactly 50% of the statewide vote. Cho's variant is also exclusively a function of vote share, and thus completely ignores seat share. *See* Nagle, *supra* note 13, at 201 & n.18.

[108] *See supra* notes 93-94 and accompanying text.

[109] *See* Cover, *supra* note 13 (manuscript at 45) (option 1 in the table); McGhee, *supra* note 19 (manuscript at 28).

[110] *See* Cover, *supra* note 13 (manuscript at 45) (option 5 in the table); McGhee, *supra* note 19 (manuscript at 29-30).

Nagle and Cover's voter-centric efficiency gap,[111] each imply an ideal responsiveness of *one*. This is both a lower level of responsiveness than American elections have historically exhibited and a measure of deviation from proportional representation. Whatever its normative appeal, disproportionality, of course, is precluded as a metric by Supreme Court precedent.[112]

While Cover focuses on devising variants of the efficiency gap, he makes one more methodological critique: that the measure is highly sensitive to how votes are imputed for uncontested races.[113] Votes have to be imputed when races do not feature a candidate from each major party because data is required for each district and voter opinion is not actually unanimous when a candidate wins 100% of the two-party vote.[114] These imputations are typically based on a regression model for *contested* races where the legislative vote is the dependent variable, the presidential vote is the key independent variable, and indicators are also included for incumbency status. This model is then deployed to estimate how each party would have performed in *uncontested* races if the races had, in fact, featured two major-party candidates.[115]

To begin with, Cover's objection applies to *all* measures of partisan gerrymandering—partisan bias, the mean-median difference, the difference between the parties' average margins of victory, and so on—not just the efficiency gap. All of these metrics require data for each district to be calculated. All of the metrics are also more accurate when uncontested races are neither omitted from the analysis nor treated as if they were won unanimously. There is simply nothing about imputation that is specific to the efficiency gap. Rather, it is a generally applicable technique that social scientists use whenever data is necessary but unavailable.[116]

Additionally, as far as we can tell, it makes little substantive difference how exactly imputations are carried out. In our original article on the efficiency gap, we used the presidential vote and incumbency status to estimate the parties' vote shares in uncontested congressional districts.[117] We then relied on these estimates to calculate the simplified form of the efficiency gap. Using similar data—but without any methodological input from us—Simon Jackman imputed the parties' votes and vote shares in uncontested congressional districts in the ongoing North Carolina partisan gerrymandering litigation.[118] He then computed both the full form and the simplified form of the efficiency gap.

Figure 8 plots the simplified form of the efficiency gap as calculated by Jackman (on the y-axis) versus the simplified form as computed by us (on the x-axis). The correlation between the

---

[111] *See* Nagle, *supra* note 13, at 202. Deviation from proportional representation is also endorsed by Chambers et al., *supra* note 13, at 15.

[112] *See supra* notes 33-34 and accompanying text.

[113] *See* Cover, *supra* note 13 (manuscript at 52-55).

[114] *See* Stephanopoulos & McGhee, *supra* note 5, at 865-67; *see also* Andrew Gelman & Gary King, *A Unified Method of Evaluating Electoral Systems and Redistricting Plans*, 38 AM. J. POL. SCI. 514, 549-50 (1994).

[115] *See* Stephanopoulos & McGhee, *supra* note 5, at 865-67; *see also, e.g.*, Jackman, *supra* note 37, at 20-26; Jackman, *supra* note 59, at 24-31.

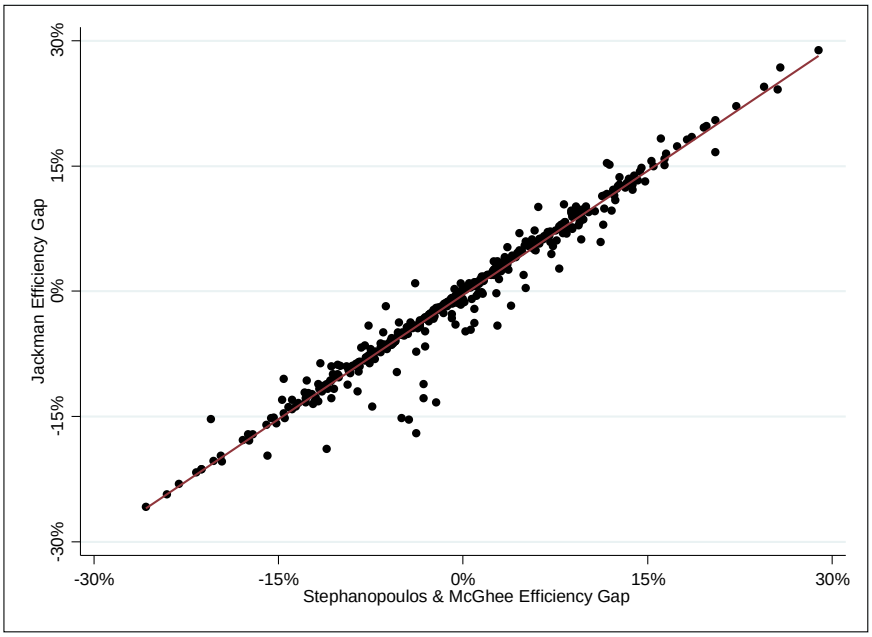
[116] *See* Gelman & King, *supra* note 114, at 550 ("For many applications, we believe some imputation for uncontested districts is essential.").

[117] *See* Stephanopoulos & McGhee, *supra* note 5, at 866.

[118] *See* Jackman, *supra* note 37, at 20-26. We do not carry out this exercise for state house elections because we did not have access to the presidential vote aggregated by state house district while Jackman generally did. *Compare* Stephanopoulos & McGhee, *supra* note 5, at 866-67, *with* Jackman, *supra* note 59, at 26-27.

two variables is nearly perfect—a whopping *0.98*—and almost all of the points fall on or very near the best fit line.[119] Based on this evidence, we think Cover is simply wrong when he claims that "the imputation method chosen" would often "have a significant, and even outcome-determinative, impact on the measure of a plan's efficiency gap."[120] On rare occasions, it might make a difference which technique was used. But imputation cannot be nearly as influential as Cover supposes if separate scholars, proceeding independently with their analyses, arrived at essentially the same results for hundreds of congressional elections over dozens of states and years. To the contrary, the better conclusion is the one reached by Grofman and King with respect to partisan bias: namely, that while "[e]xperts can disagree about which set of input data is relevant for a given case," "the resulting measures . . . normally will not differ to any significant degree because the best current methods are quite robust to changes in model specification."[121]

FIGURE 8: IMPUTATIONS AND THE EFFICIENCY GAP



Scatter plot of the efficiency gap calculated by Jackman versus the efficiency gap calculated by Stephanopoulos & McGhee, using the simplified method for congressional elections.

B. *Variability*

Turning from methodology to variability, Cho and Krasno et al. present toy examples of one or two districts, and argue based on them that the efficiency gap is either not changeable enough (Cho) or *too* changeable (Krasno et al.). Cho emphasizes the fact that if the parties' statewide vote shares are fixed, the efficiency gap can exhibit at most as many values as there are

[119] The cases furthest from the best fit line are mostly from the 1970s and 1980s. From 1992 to the present, the correlation between the two sets of efficiency gap scores is 0.99.

[120] Cover, *supra* note 13 (manuscript at 40).

[121] Grofman & King, *supra* note 1, at 17.

seats in a plan. This property makes the measure too granular for her taste.[122] Krasno et al, on the other hand, point out that as the parties' vote shares vary within a particular district, the efficiency gap flips signs at 75%-to-25%, at 50%-to-50%, and at 25%-to-75%. They assert that a metric should not register so many shifts in the identity of the advantaged party.[123]

Krasno et al., first, improperly apply the efficiency gap to a single district. The measure is only valid for a district map *as a whole*, for which it indicates which party more efficiently converts its *aggregate* votes into *aggregate* seats. The efficiency gap simply has no intelligible meaning for one district taken in isolation. Whatever edge a party happens to enjoy in that district may be offset—or magnified—by outcomes in a plan's other districts.

Indeed, *no* gerrymandering metric yields meaningful results for a single district. Consider partisan bias: It cannot even be calculated for one district, because in a hypothetical tied election, neither party can be deemed to have won the lone available seat. Or take Krasno et al.'s preferred measure, the mean-median difference: A party's mean and median vote share are always identical for a single district, meaning that the difference between them is always zero. Or think of the difference between the parties' average margins of victory: It too cannot be computed for one district, since only one party records a victory in the first place.

Next, the granularity to which Cho objects is a function not of the efficiency gap but rather of seats themselves in a single-member-district system. Such seats are exceedingly granular. They transform a continuous variable (a party's vote share) into a binary variable (whether a party wins or loses a district). Accordingly, any metric that takes seats into account must exhibit a level of granularity similar to that of the efficiency gap. The granularity follows unavoidably from seats' winner-take-all nature.

Figure 9 makes this point by displaying a revised version of a table created by Cho.[124] Cho outlines seven scenarios for a two-district plan, in which both districts are won by the same margin and the winning party's statewide vote share ranges from 51% to 99%. Cho calculates the efficiency gap for each scenario, finding that it goes from 48% against Party A (in Scenario 1) to 48% in favor of Party A (in Scenario 7). Cho also shows (and this is the nub of her critique) that in each scenario, no matter how votes are allocated between Party A and Party B, the efficiency gap can take on either one value (if Party B lacks enough votes to win a district) or two (if Party B is able to win a district).

Figure 9 adds four columns to Cho's table: one computing partisan bias and another the mean-median difference for each scenario, and two more listing the possible partisan biases and

---

[122] *See* Cho, *supra* note 16 (manuscript at 3-9). The efficiency gap's performance in a two-district setting is also criticized by Chambers et al., *supra* note 13, at 32.

[123] *See* Krasno et al., *supra* note 10, at 13-14).

[124] *See* Cho, *supra* note 16 (manuscript at 4). Cho includes columns that we omit for Party A's and Party B's wasted votes and for the efficiency gap calculated with only surplus votes. *See id.* As noted earlier, her surplus-vote-only variant is strictly a function of vote share. *See supra* note 107. It thus varies dramatically as Party A's vote share rises from 51% to 99%—and yields the odd conclusion that Party A's advantage grows as its voters become more packed in each district.

mean-median differences if the parties' votes were distributed in different ways.[125] Partisan bias cannot be calculated for any of the scenarios because, in all of them, both districts would be split 50%-to-50% in a hypothetical tied election and so could not be assigned to either party. Aside from this null result, the only possible partisan bias if the parties' votes were reallocated is zero, indicating that each party would win one district in a counterfactual tied election. Similarly, the mean-median difference is zero both in every scenario and in every possible reallocation of the parties' votes. This is because, in a two-district plan as in a one-district plan, a party's mean and median vote share are always identical.

The conclusion we draw from this exercise is that, even in a highly artificial two-district setting, the efficiency gap performs *better* in terms of granularity than other measures of partisan gerrymandering. The efficiency gap varies from one scenario to another; partisan bias cannot be computed for any of them while the mean-median difference stays constant at zero. The efficiency gap could also exhibit up to two values if the parties' votes were distributed in different ways; partisan bias and the mean-median difference could only exhibit one. Cho's complaint about the efficiency gap thus seems misplaced: If anything, it is the alternatives to the efficiency gap that are more affected by seats' binary nature.[126]

[125] Figure 9 also tweaks the possible efficiency gap values for Scenario 4. If the parties' votes were reallocated so that one district was split 50%-to-50% and the other was won 100%-to-0%, then the efficiency gap could not be calculated because no winning or losing party could be identified for the 50%-to-50% district.

[126] Cho also argues that because of the efficiency gap's granularity, "a comparison of efficiency gap values across states, over different years, or whenever the underlying voter population changes, is problematic." *Id.* (manuscript at 5). But again, this point is *less* applicable to the efficiency gap than to other, more granular measures. Moreover, both the efficiency gap and other metrics become far less granular—indeed, essentially continuous—once real-world conditions of many districts and shifting seat and vote shares are introduced. *See, e.g.*, Nagle, *supra* note 13, at 205 n.37 ("[W]hen there are many legislative districts . . . such jumps become statistically small, and the EG method is likely to be appropriate."). And as we have previously explained, gerrymandering metrics may be presented in *seats* rather than *seat shares* when there exist concerns about granularity. *See* Stephanopoulos & McGhee, *supra* note 5, at 869. There is nothing unusual about this move; to the contrary, analysts frequently switch between absolute values and percentages depending on the issue being addressed.

FIGURE 9: MEASURES OF PARTISAN GERRYMANDERING IN TWO-DISTRICT SCENARIOS

Scenario	Party A Votes	Party B Votes	*EG*	Possible *EG* Values	*Bias*	Possible *Bias* Values	*MMD*	Possible *MMD* Values
1	99	1	0.48	{0.48}	N/A	{N/A, 0.00}	0.00	{0.00}
	99	1						
2	90	10	0.30	{0.30}	N/A	{N/A, 0.00}	0.00	{0.00}
	90	10						
3	80	20	0.10	{0.10}	N/A	{N/A, 0.00}	0.00	{0.00}
	80	20						
4	75	25	0.00	{0.00, N/A}	N/A	{N/A, 0.00}	0.00	{0.00}
	75	25						
5	70	30	-0.10	{-0.10, 0.40}	N/A	{N/A, 0.00}	0.00	{0.00}
	70	30						
6	60	40	-0.30	{-0.30, 0.20}	N/A	{N/A, 0.00}	0.00	{0.00}
	60	40						
7	51	49	-0.48	{-0.48, 0.02}	N/A	{N/A, 0.00}	0.00	{0.00}
	51	49						

For seven different scenarios, all with two districts, the table shows the parties' votes as well as actual and possible efficiency gap, partisan bias, and mean-median difference values.

Krasno et al. criticize the efficiency gap for one more reason relating to its variability. They calculate the measure for Iowa's and North Carolina's current state senate plans[127] and Wisconsin's current state house plan[128] using exogenous election results. (Exogenous elections are for offices unrelated to the map in question, such as the presidency or the governorship.) They then note that the efficiency gap changes based on which exogenous election is used to compute it. This mutability is problematic, in their view, because it means that one's substantive conclusion about a map might vary along with the electoral environment.[129]

To start, it is bad methodological form to analyze plans with exogenous election results.[130] Voters may well behave differently when casting their ballots for the office actually at issue. The better approach is to use election results for the relevant office whenever races are contested, and to impute outcomes through the methods described above whenever races are uncontested.[131] As Grofman and King wrote more than a decade ago, one cannot "assum[e] that votes in statewide elections for statewide candidates have any particular ex ante relationship with votes for legislative candidates."[132] Instead, one should examine this relationship empirically

[127] *See* Best et al., *supra* note 10 (manuscript at 20-21, 26).

[128] *See* Krasno et al., *supra* note 10, at 9-12.

[129] *See id.*; Best et al., *supra* note 10 (manuscript at 20-21, 26). Krasno et al. also apply their exogenous election results to simulated district plans, again stressing the range of efficiency gaps that arise.

[130] This critique also applies to Cho's calculation of the efficiency gap using the presidential vote and then (even further removed from endogenous elections) party registration data. *See* Cho, *supra* note 16 (manuscript at 9-13).

[131] *See supra* note 115 and accompanying text.

[132] Grofman & King, *supra* note 1, at 12.

using "recent election results, the presence of an incumbent in the district, and whether the race is contested."[133]

Krasno et al. also evaluate the efficiency gap's variability rather simplistically. They merely point out that it changes when one or another exogenous election is used to calculate it. But all measures of partisan gerrymandering shift as the parties' seats and votes fluctuate.[134] The key questions are *how much* the metrics vary and whether their volatility is high enough to undermine their usefulness.

One way to answer these questions is by calculating the measures' standard deviations. From 1972 to the present, the standard deviation of the efficiency gap was 6.3% at the state house level and 9.2% at the congressional level. By comparison, the standard deviation of partisan bias (which is also denominated in units of seat share) was 8.2% at the state house level and 10.6% at the congressional level over this period—or slightly higher than that of the efficiency gap. Analogously, the standard deviation of the mean-median difference was 3.8% at the state house level and 5.6% at the congressional level. Because the mean-median difference is denominated in units of vote share, these figures must be roughly doubled to be made comparable.[135] When they are doubled, they are again slightly higher than the efficiency gap's standard deviation.

Another way to tackle these issues is to compute the proportion of a metric's total variation that occurs *between* (rather than *within*) district plans. (Plans are typically redrawn each decade and used for five elections.) The higher this proportion, the better a measure captures a durable, plan-specific characteristic. From 1972 to the present, 74% of the efficiency gap's variation was between plans at the state house level and 60% was between plans at the congressional level.[136] The analogous figures for partisan bias were 86% and 74%; for the mean-median difference they were 83% and 75%. Thus the bulk of all three metrics' variation was between plans over the last half-century, and this plan-specific component was modestly stronger for partisan bias and the mean-median difference than for the efficiency gap.

Why does the efficiency gap for a given plan vary somewhat more than partisan bias or the mean-median difference? The explanation is that only the efficiency gap takes into account the seats actually won or lost by the parties. Partisan bias is based on the seats won or lost in a hypothetical tied election—which often stay the same even as the seats won or lost in reality change. The mean-median difference ignores the parties' seat shares altogether; it is calculated using only the parties' vote shares. The efficiency gap therefore sacrifices some stability for the sake of better correspondence with the concept of partisan gerrymandering. At its core,

---

[133] *Id.*; *see also* Gelman & King, *supra* note 114, at 524 ("[I]t is clear that the variables [used to calculate a gerrymandering metric] should certainly include *past legislative election results* . . . ."); Wang, *Three Tests*, *supra* note 10, at 1280 ("For analyzing congressional or legislative districts, the results of congressional or legislative elections themselves have the advantage of being a direct measure of voter preference for the type of office under dispute . . . .").

[134] *See* Warrington, *supra* note 13, at 5 ("[A]ny measure of asymmetry in vote distributions will vary from election to election . . . .").

[135] Recall that the seat-to-vote relationship in American elections has been very close to two over the last half-century. *See supra* notes 93-94 and accompanying text.

[136] For similar calculations, see Jackman, *supra* note 37, at 31, and Jackman, *supra* note 59, at 48.

gerrymandering is about translating actual votes into actual seats more efficiently.[137] Actual votes and seats fluctuate more than either hypothetical seats or actual votes alone. The efficiency gap reflects this fluctuation, yet remains only moderately more volatile than the other measures.[138]

Informative as they are, standard deviations and proportions of variation are a bit abstract. A more practical test—one of particular interest to litigants and courts[139]—is whether a plan that exhibits a certain efficiency gap in its *first* election is likely to exhibit a similar efficiency gap *over its lifetime*. If so, then a plan's initial efficiency gap would be a reliable indicator of the plan's future performance, and decisions could be made on the basis of that first score. If not, then the efficiency gap's within-plan variation would prevent observers from reaching any robust conclusions after a single election.[140]

Figure 10 plots plans' lifetime average efficiency gaps (on the y-axis) versus their initial efficiency gaps (on the x-axis) for state house elections (left chart) and congressional elections (right chart). Only plans that are in place for at least three elections (for which averages are more meaningful) are included. It is clear from the charts that there is a strong relationship between these variables. The correlation is 0.85 at the state house level and 0.76 at the congressional level, and the points cluster around the best fit lines.[141] We think these findings further refute Krasno et al.'s critique. While the efficiency gap does change from year to year, plans' first scores are an excellent guide to their lifetime performances. Small initial scores tend to remain small, large initial scores typically stay large, and observers indeed learn something significant about plans from the first data points they produce.

---

[137] *See supra* Part I.A.

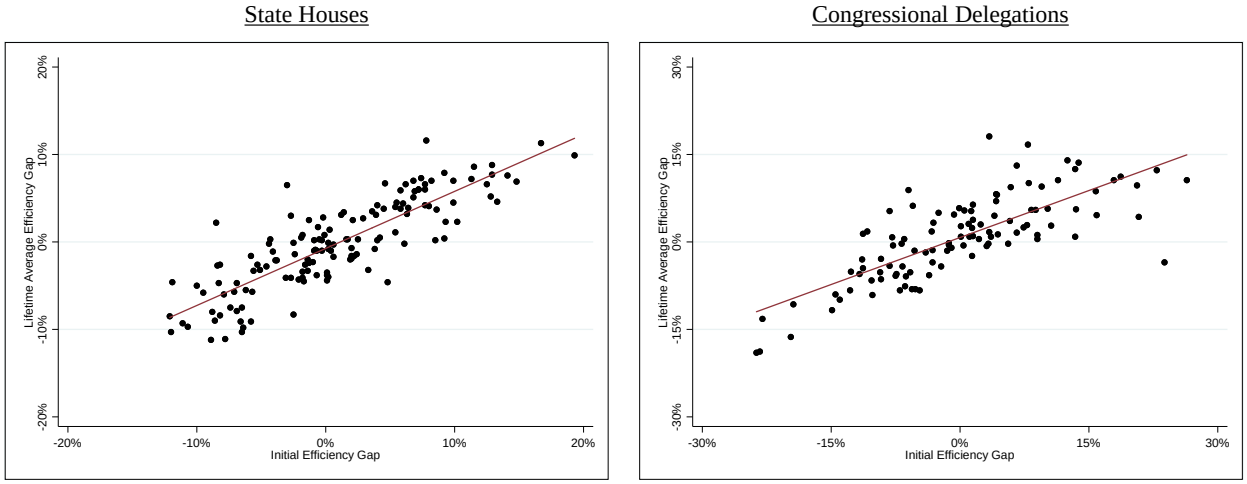
[138] For a related argument, see Stephanopoulos & McGhee, *supra* note 5, at 864. *See also* McGhee, *supra* note 5, at 76 ("A measure should not force stability . . . or it risks ignoring some of the very properties it is meant to test."); Warrington, *supra* note 13, at 18 ("[T]he mean-median difference doesn't keep track of the number of seats won by each party. This is probably the main reason it is slightly more stable from election to election than the efficiency gap . . . .").

[139] *See Whitford II*, 218 F. Supp. 3d at 905 (relying on evidence that "[t]he party that 'wastes' more votes in the first election year is likely to continue 'wasting' more votes in future elections").

[140] We note that it is just as easy to calculate a plan's efficiency gap before an election (using predicted election results) as afterwards (using actual ones). Ex ante calculation may even be preferable because it can take into account the full range of plausible outcomes, thus avoiding an excessive focus on a single, possibly idiosyncratic election.

[141] For a similar finding, see Jackman, *supra* note 37, at 47-51. Jackman further demonstrated the efficiency gap's stability by examining how often plans' scores flip signs, *see* Jackman, *supra* note 59, at 56-63, by carrying out a series of prognostic tests, *see* Jackman, *supra* note 37, at 41-47, and by subjecting current plans to electoral "perturbations" and then recalculating their efficiency gaps, *see id.* at 54-57.

FIGURE 10: INITIAL AND LIFETIME AVERAGE EFFICIENCY GAPS



Scatter plots of district plans' lifetime average efficiency gaps versus their initial efficiency gaps, for state houses and congressional delegations. Only plans used in at least three elections are included.

C. *Ideology*

Lastly, Chambers et al. point out that the efficiency gap does not distinguish between ideologically moderate and extreme legislators who belong to the same party.[142] As far as the metric is concerned, a Democratic member is a Democratic member, and a Republican is a Republican. But, claim Chambers et al., competitive districts tend to elect centrists while extremists typically win safe seats.[143] Maps that appear to benefit one party, considering partisanship alone, may thus actually favor the opposition. Compare, for instance, a plan with five moderate Democrats, all elected from competitive districts, to a plan with three extreme Democrats and two extreme Republicans, all winning safe seats.[144] A Republican voter might prefer the former plan, notwithstanding its more pro-Democratic efficiency gap, because it would yield a center-left rather than a far-left legislative median.[145]

It is perfectly true, of course, that the efficiency gap does not take into account legislators' (or voters') political philosophies. This is because it is a measure of *partisan* gerrymandering, not of *ideological* divergence between the legislature and the electorate. Such divergence is important, and we have studied it extensively in previous work.[146] But it is simply

[142] *See* Chambers et al., *supra* note 13, at 23 ("[T]he efficiency gap does not contemplate that political parties may be heterogeneous.").

[143] *See id.* at 26 ("[P]oliticians elected in politically lopsided districts may be more extreme than politicians from politically competitive districts . . . .").

[144] *See id.* at 24-29 (presenting these plans).

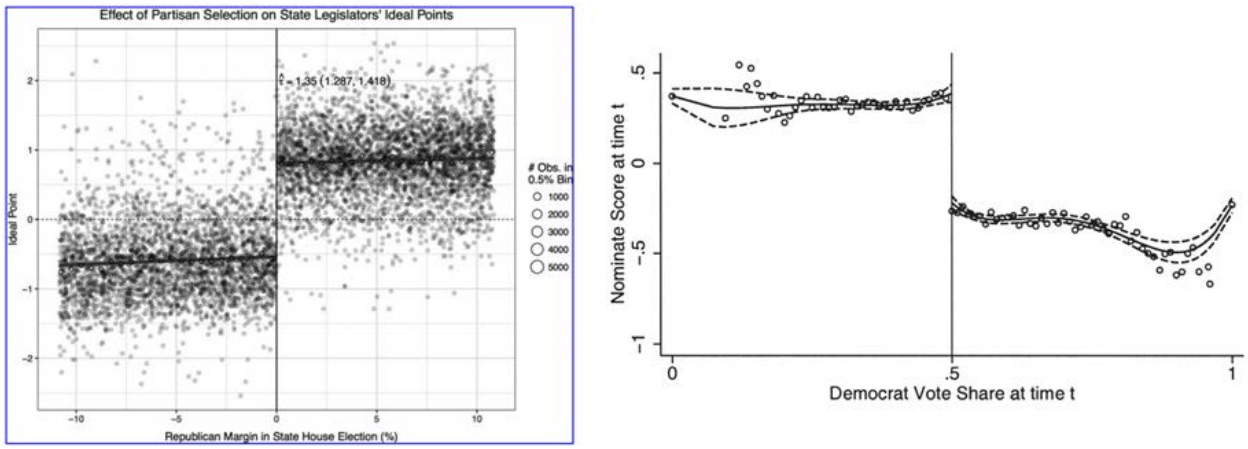
[145] *See id.* at 29 ("All Republicans prefer the first plan, even though every seat is captured by the Demcorats."). The dissenter on the *Whitford* trial court made a similar argument. *See Whitford II*, 218 F. Supp. 3d at 954 ("[A] Republican who has won with only 51% of the vote will very likely govern differently than one who has a safe seat . . . .").

[146] *See generally* Nicholas O. Stephanopoulos, *Elections and Alignment*, 114 COLUM. L. REV. 283 (2014); Nicholas O. Stephanopoulos et al., *The Realities of Electoral Reform*, 68 VAND. L. REV. 761 (2015).

not the same as gerrymandering, "the usual point of [which]," according to Justice Souter, "is to control the greatest number of *seats*."[147] Precisely because gerrymandering is about votes and seats (not voters' and legislators' policy preferences), no other metrics consider ideology either. Partisan bias, for example, asks how different the parties' seat shares would be in a hypothetical tied election, while the mean-median difference looks only at the parties' vote shares. Ideology is nowhere to be found in these (or any other) gauges of gerrymandering.

Additionally, the empirical premise of Chambers et al.'s argument—that centrists are elected from competitive districts while extremists win safe seats—is wrong. At least in recent years, there is virtually no connection between district composition and legislator ideology. Rather, Democratic members tend to be liberal, and Republicans conservative, no matter how precarious or secure their seats happen to be. The left chart in Figure 11, borrowed from Devin Caughey and his coauthors, makes this point at the state house level using data from 1995 to 2012.[148] Legislators' ideal points barely budge as their margins of victory vary. A Democrat (Republican) from a competitive district is nearly as liberal (conservative) as one from a safe seat. The right chart, taken from David Lee and his coauthors, presents similar results at the congressional level using data from 1946 to 1995.[149] Again, House Democrats are very ideologically different from House Republicans, and again, representatives' ideal points are almost unrelated to their electoral performances.

FIGURE 11: LEGISLATOR IDEOLOGY VERSUS DISTRICT COMPOSITION



Scatter plots of legislator ideal point versus electoral performance. The left chart is for state houses from 1995 to 2012 and considers only relatively close elections. The right chart is for the House of Representatives from 1946 to 1995 and considers all elections.

[147] Vieth v. Jubelirer, 541 U.S. 267, 346 (2004) (Souter, J., dissenting) (emphasis added).

[148] *See* Devin Caughey et al., *Partisan Gerrymandering and the Political Process: Effects on Roll-Call Voting and State Policies*, 16 ELECTION L.J. (forthcoming 2017) (manuscript at 6).

[149] *See* David S. Lee, *Do Voters Affect or Elect Policies? Evidence from the U.S. House*, 119 Q.J. ECON. 807, 840 (2004). For similar findings at the congressional level, see Stephen Ansolabehere et al., *Candidate Positioning in U.S. House Elections*, 45 AM. J. POL. SCI. 136, 142 (2001), and Nolan McCarty et al., *Does Gerrymandering Cause Polarization?*, 53 AM. J. POL. SCI. 666, 671 (2009).

One implication of these findings is that Chambers et al.'s first district plan is wildly implausible. Under modern political conditions, a map with five competitive but Democratic-leaning seats usually would *not* elect five center-left Democrats. In general, five liberal Democrats would prevail instead, their ideologies unaffected by their districts' makeups.

The findings' other implication is that the efficiency gap should be a potent driver of ideological divergence between the legislature and the electorate. Since the efficiency gap does not distinguish among legislators who belong to the same party—and since each party's legislators are, in fact, mostly indistinguishable in their voting records—the metric should be strongly linked to ideological noncongruence between representatives and voters. Caughey et al. test and confirm this hypothesis at the state house level. They show that the efficiency gap has a large and statistically significant effect on the chamber median, even controlling for a host of other factors.[150] One of us arrives at the same conclusion at the congressional level. There too, an efficiency gap in a party's favor leads the midpoint of a state's delegation to shift in this party's direction, even holding voter opinion constant.[151] Accordingly, Chambers et al. fail to undermine the appeal of the efficiency gap by changing the subject from partisan gerrymandering to ideological divergence. The efficiency gap is not a *measure* of such divergence, but it is a powerful *cause* of it.

IV. OTHER GERRYMANDERING METRICS

We began to address gerrymandering metrics other than the efficiency gap in the previous Part. We explained that all measures—not just the efficiency gap—require imputations to be made when races are uncontested and do not consider voters' or legislators' ideologies. We also showed that partisan bias and the mean-median difference are even more granular than the efficiency gap in a two-district example. We found as well that partisan bias and the mean-median difference vary somewhat less than the efficiency gap, though at the cost of worse conceptual correspondence.

In this Part we make three further points about other gerrymandering metrics. First, there is no need for scholars or courts to embrace a single measure. In other redistricting areas, multiple metrics harmoniously coexist, and the same should be possible in this domain. Second, there exist deep mathematical relationships between the various measures of gerrymandering. Factor analysis also confirms that the metrics mostly reflect the same underlying concept. And third, despite our generally catholic stance toward other measures, we do harbor some reservations about them. In particular, they perform poorly on our criteria of compliance with the efficiency principle and breadth of scope.

A. Multiplicity

Partisan gerrymandering is far from the only redistricting field that relies heavily on statistics. In contexts including reapportionment, racial vote dilution, and racial gerrymandering, quantitative metrics also take center stage. In these areas, interestingly, there has been little interest in crowning a single "best" measure. Instead, academics and judges have been happy to

[150] *See* Caughey et al., *supra* note 148 (manuscript at 11).
[151] *See* Stephanopoulos, *supra* note 63 (manuscript at 27).

calculate and cite multiple metrics, using them in tandem to bolster their conclusions. We see no reason why the same inclusive approach could not be emulated in the partisan gerrymandering domain.

Consider reapportionment cases, where a key issue is the extent to which a plan's districts deviate from perfectly equal population. Courts most commonly analyze this issue using *total deviation*: the percentage gap in population between a plan's most populous and least populous districts.[152] But courts also often refer to *average deviation*: the mean difference between each district's population and the population required for perfect equality.[153] Courts further invoke the proportion of a state's population that could elect a majority of the legislature under a plan.[154] These measures are not equivalent; they sometimes point in different directions, and they have distinct strengths and weaknesses. Courts, however, have been uninterested in this theoretical debate. They have recognized that the metrics are usually consistent with one another, and they have cited them more or less interchangeably.[155]

Or take racial vote dilution doctrine, where a prerequisite for liability is a showing of racial polarization in voting.[156] Racial polarization can be calculated through both homogeneous precinct analysis (comparing outcomes in heavily white and heavily minority precincts) and ecological regression (with election results as the dependent variable and racial composition as the main independent variable). In a pivotal 1986 case, the Supreme Court endorsed *both* of these techniques, calling them "complementary methods of analysis" that are "standard in the literature."[157] Ever since, as James Greiner notes in a survey of the vote dilution case law, "courts [have] relied on expert testimony resting on both homogenous precinct analysis and regression."[158] Indeed, courts have also been receptive to a *third* method for computing racial polarization: King's ecological inference, which more effectively deploys the same data as conventional regression.[159]

Geographic compactness—how "normal" or "strange" districts look—offers one more example of metric multiplicity.[160] Compactness is legally significant in racial vote dilution cases (because plaintiffs must show that a compact remedial district exists)[161] and in racial

---

[152] *See, e.g.*, Karcher v. Daggett, 462 U.S. 725, 728 (1983); White v. Regester, 412 U.S. 755, 761 (1973); Kirkpatrick v. Preisler, 394 U.S. 526, 528-29 (1969).

[153] *See, e.g.*, *Karcher*, 462 U.S. at 728; Gaffney v. Cummings, 412 U.S. 735, 737 (1973) (also referring to the median deviation); Mahan v. Howell, 410 U.S. 315, 319 (1973).

[154] *See, e.g.*, *Mahan*, 410 U.S. at 319; Swann v. Adams, 385 U.S. 440, 442 (1967); Reynolds v. Sims, 377 U.S. 533, 545 (1964).

[155] Cases citing multiple measures of population deviation include *Karcher*, 462 U.S. at 728; *Gaffney*, 412 U.S. at 737; *Mahan*, 410 U.S. at 319; and *Swann*, 385 U.S. at 442.

[156] *See* Thornburg v. Gingles, 478 U.S. 30, 52-74 (1986).

[157] *See id.* at 52, 53 n.20.

[158] D. James Greiner, *Ecological Inference in Voting Rights Act Disputes: Where Are We Now, and Where Do We Want to Be?*, 47 JURIMETRICS 115, 155 (2007).

[159] *See id.* at 158 ("[J]udicial reaction to King's EI has generally been favorable."); *cf.* GARY KING, A SOLUTION TO THE ECOLOGICAL INFERENCE PROBLEM: RECONSTRUCTION INDIVIDUAL BEHAVIOR FROM AGGREGATE DATA (1997).

[160] Nagle has also analogized partisan gerrymandering to compactness. *See* Nagle, *supra* note 12, at 347 ("[H]aving different measures of partisan bias is somewhat similar to having different measures of compactness.").

[161] *See Gingles*, 478 U.S. at 50 ("[T]he minority group must be able to demonstrate that it is sufficiently large *and geographically compact* . . . ." (emphasis added)).

gerrymandering litigation (where noncompactness is probative evidence of a predominant racial purpose).[162] Compactness can also be quantified in literally *dozens* of ways. As a recent study observes, "[n]umerous specific compactness measures have been proposed in the academic literature, each one fitting different . . . conceptual definitions and intuitions."[163] Scholars have not selected a "gold standard" among these metrics, but rather "have managed to use [them] productively in research by combining multiple measures, adjusting or weighing each for specific purposes, or making careful qualitative decisions."[164] Courts too have no preferred compactness metric, instead relying on several without any evident unease.[165]

The lesson of other redistricting contexts is thus that no quest for a single holy grail has been—or has to be—attempted. Academics may develop a range of partisan gerrymandering measures, each capturing a separate aspect of the practice. It is less important for them to *rank* metrics, to endorse one while rejecting all other options. Judges, similarly, may benefit from experts who quantify plans' partisan skews in several ways, each adding to their understanding of the maps' properties. Judges may be more confident in their rulings when different measures tell the same story, and less certain when metrics diverge.[166]

B. Relationships

Another reason why we are sanguine about the use of multiple gerrymandering measures is that most of them are linked mathematically to one another. Because they have been published elsewhere, we do not include formal proofs here, but rather describe qualitatively the ties between the metrics. We do, however, present a factor analysis of gerrymandering measures. This analysis shows that they are connected arithmetically *and* empirically.

First, as we pointed out in our original article on the efficiency gap, the metric is equal to partisan bias if it is calculated using the simplified method and an election is tied statewide.[167] The simplified formula for the efficiency gap, again, is $(S - 0.5) - 2 \times (V - 0.5)$, where S is a party's statewide seat share and V is the party's average vote share across all districts.[168] If an election is tied statewide, V is 0.5, so this formula shortens further to $(S - 0.5)$. This, of course, is the very definition of partisan bias: a party's seat share minus 50% in a tied election.

---

[162] *See, e.g.*, Miller v. Johnson, 515 U.S. 900, 913 (1995) ("[B]izarreness . . . may be persuasive circumstantial evidence that race for its own sake . . . was the legislature's dominant and controlling rationale in drawing its district lines.").

[163] Aaron Kaufman et al., *How to Measure Legislative District Compactness if You Only Know It when You See It* 4 (July 10, 2017).

[164] *Id.* at 5.

[165] *See, e.g.*, Vieth v. Jubelirer, 541 U.S. 267, 348 (2004) (Souter, J., dissenting) (noting that compactness "can be measured quantitatively in terms of dispersion, perimeter, and population ratios"); Karcher v. Daggett, 462 U.S. 725, 756 n.19 (1983) (Stevens, J., concurring) ("The scholarly literature suggests a number of different mathematical measures of compactness, each focusing on different variables.").

[166] *See* Cain et al., *supra* note 64 (manuscript at 4) ("[A]ny systematic evaluation of partisan bias should utilize several political fairness measures."); Nagle, *supra* note 12, at 348 ("[T]here are several measures that essentially agree, so having different measures should not block the inclusion of partisan bias in election reform."). However, as we discuss below, certain metrics should not be used in certain electoral environments. Inclusiveness has its limits. *See infra* Part IV.C.

[167] *See* Stephanopoulos & McGhee, *supra* note 5, at 856; *see also* McGhee, *supra* note 5, at 70.

[168] *See supra* note 83 and accompanying text.

Second, as Cover elegantly establishes,[169] the simplified form of the efficiency gap is *always* equal to the difference between the parties' average margins of victory—at least if two small tweaks are made to the latter measure. One is that each party's margin in each district must be subtracted from 50%. The other is that the parties' average margins must be weighted by the shares of seats that they each win. If these two adjustments are performed, then a mathematical identity follows, as does an important substantive implication: The efficiency gap connotes not just differential cracking and packing, but also differential competitiveness. Indeed, differential cracking and packing *is* differential competitiveness.

And third, as Nagle and others demonstrate,[170] the mean-median difference is equal to partisan bias divided by the slope of the seat-vote curve. Partisan bias is a party's *seat* share minus 50% in an election where the party's *vote* share is 50%, while the mean-median difference is a party's *vote* share minus 50% in an election where the party's *seat* share is 50%. In charts like those displayed in Figure 7, partisan bias is the *vertical* distance between a seat-vote curve and the point where vote share and seat share are both 50%, while the mean-median difference is the *horizontal* distance. This close relationship explains a theme that is already apparent and that becomes even clearer below: namely, that partisan bias and the mean-median difference have nearly identical strengths and weaknesses, and are linked to the efficiency gap in almost the same ways.

To further illuminate the connections between measures of partisan gerrymandering, we carry out a factor analysis for the efficiency gap, partisan bias, the mean-median difference, and an even newer metric called the *declination*. Factor analysis is a data reduction technique that seeks to identify latent variables that cannot be measured directly but that influence the values of a set of observed variables. The latent variables yielded by the method are linear functions of the observed variables and explain as much as possible of the variance in the data.[171] The declination is calculated by plotting all of a plan's districts in order from lowest to highest vote share. Two lines are then drawn, one between the median Democrat-won district and the 50% point, and another from the median Republican-won district to the 50% point. The angle between these lines is the declination.[172] We include it in the factor analysis because, despite its novelty, it strikes us as a promising statistic.

Figure 12 tabulates the results of the analysis for state house and congressional elections. At both levels, two factors are retained. The first of these is by far the more important, accounting for 97% of the variance at the state house level and 95% at the congressional level. All of the gerrymandering metrics load heavily onto this factor. The second retained factor

---

[169] *See* Cover, *supra* note 13 (manuscript at 35-36).

[170] *See* Nagle, *supra* note 12, at 351; *see also* McDonald & Best, *supra* note 10, at 314; McGhee, *supra* note 19 (manuscript at 13).

[171] For a more detailed discussion, see Nicholas O. Stephanopoulos, *Spatial Diversity*, 125 HARV. L. REV. 1903, 1938 (2012).

[172] *See* Warrington, *supra* note 13, at 3-5. The intuition behind the declination is that "[p]artisan gerrymandering . . . modifies a natural distribution in a manner that treats the 0.5 threshold as special." *Id.* at 3. If the district distribution does not shift at the 50% point, then the two lines have the same slope and the declination is zero. But if there is a break at the 50% point, typically caused by differential cracking and packing, then the two lines' slopes diverge and the declination indicates the extent of each party's advantage or disadvantage.

explains a much more modest 7% of the variance at the state house level and 13% at the congressional level.[173] All of the metrics have smaller loadings onto this factor. The *signs* of the loadings also diverge here, with pairings emerging between the efficiency gap and the declination on the one hand, and partisan bias and the mean-median difference on the other.

These findings support two conclusions. One is that, to a significant extent, all of the measures of partisan gerrymandering reflect the same latent variable, which we call *generic party advantage*. This is why the first retained factor accounts for such an impressive share of the variance in the data, and why all of the metrics load so heavily onto this factor.[174] The second takeaway is that to the degree the measures diverge, the efficiency gap and the declination fit into one category, while partisan bias and the mean-median difference fit into another. This is why each group loads with different signs onto the second factor, which we dub *seat-specific party advantage*. This name stems from the fact that the efficiency gap and the declination consider the seats actually won by each party, while partisan bias and the mean-median difference do not. This distinction likely explains the behavior of the second factor—and why a second factor emerges in the first place.

[173] The retained factors in principal factor analysis account for more than 100% of the variance in the data when the non-retained factors account for negative proportions of the variance. All of the factors together explain exactly 100% of the variance.

[174] *See also* Cain et al., *supra* note 64 (manuscript at 13, 20, 23) (finding generally strong connections between partisan bias, the efficiency gap, and disproportionality in simulated plans).

FIGURE 12: FACTOR ANALYSIS RESULTS

	Factor 1: Generic Party Advantage	Factor 2: Seat-Specific Party Advantage
STATE HOUSES		
Efficiency gap	0.794	0.277
Partisan bias	0.933	-0.184
Mean-median difference	0.844	-0.282
Declination	0.909	0.209
Eigenvalue	3.039	0.234
Proportion of variance	0.972	0.075
CONGRESSIONAL DELEGATIONS		
Efficiency gap	0.825	-0.342
Partisan bias	0.817	0.293
Mean-median difference	0.735	0.337
Declination	0.904	-0.227
Eigenvalue	2.705	0.368
Proportion of variance	0.951	0.129

The table lists the loadings of each measure of partisan gerrymandering onto each of the two retained factors, for state houses and congressional delegations. It also lists each factor's eigenvalue and proportion of variance explained.

C. *Reservations*

To this point, our discussion has emphasized the virtues of multiplicity, especially given the relationships that exist between different gerrymandering metrics. We now turn to the drawbacks, under our evaluative criteria, of measures other than the efficiency gap. They do not always comply with the efficiency principle, and they cannot be used in certain electoral environments. We hasten to add that these shortcomings do *not* mean the metrics should be discarded. The implication is only that their properties should be fully understood before they are employed.

To see why partisan bias and the mean-median difference[175] sometimes violate the efficiency principle, it is helpful to return to the efficiency gap examples we presented in Figure 1.[176] Figure 13 reproduces these examples, and also computes partisan bias and the mean-median

[175] We do not address the declination in the same detail due to its novelty. It is clear that the metric sometimes violates the efficiency principle since a seat can cross the 50% point without changing the angle between the two lines used to calculate the measure. We also do not discuss the difference between the parties' average margins of victory since it is essentially identical to the efficiency gap.

[176] For additional examples along these lines, see McGhee, *supra* note 19 (manuscript at 9, 15-16); McGhee, *supra* note 5, at 61-63.

difference for each election. With respect to partisan bias, it is 20% in favor of Party B in the first election. Since Party A received 55% of the vote and Party B received 45%, 5% of the vote must be deducted from Party A and added to Party B in each district to simulate a tied election. When this adjustment is carried out, Party A wins three districts (1-3) and Party B wins seven districts (4-10). The difference between Party A's counterfactual seat share (30%) and vote share (50%) is the first election's partisan bias.[177]

Critically, this -20% score does not change in the second election even though Party A earns an additional seat. In a hypothetical tied election, Party A would still win three districts, so the difference between its counterfactual seat share (30%) and vote share (50%) is still -20%. Partisan bias therefore violates the efficiency principle in these examples because it does not shift in Party A's favor when Party A extracts a larger seat share for the same vote share. The metric fails to register Party A's stronger performance because this improvement is evident only in the election that actually occurred. It is *not* evident in the hypothetical tied election that is used to calculate partisan bias.

Similarly, Party A's mean vote share in the first election is 55% and its median vote share is 53%, generating a mean-median difference of -2%. All of these figures are unchanged in the second election. The swings in Districts 2, 3, 7, and 9 affect neither Party A's mean vote share (because they offset one another) nor its median vote share (because Districts 5 and 6, at the center of the distribution, stay the same). The mean-median difference thus also violates the efficiency principle in these examples. It does not move in Party A's direction when it captures an additional seat with the same popular support because this extra seat is not reflected in Party A's mean or median vote share.[178]

[177] This result is itself problematic. Partisan bias registers a large *disadvantage* for Party A, based on the hypothetical tied election, even though it won six of ten seats in the election that in fact occurred. *See* Stephanopoulos & McGhee, *supra* note 5, at 861-62 (also making this point).

[178] Warrington similarly notes that "it is quite possible for packing and cracking to occur without having any effect on the mean-median difference." Warrington, *supra* note 13, at 18.

FIGURE 13: EXAMPLES WITH MULTIPLE GERRYMANDERING METRICS

	ELECTION 1				**ELECTION 2**			
District	**Votes**		**Wasted Votes**		**Votes**		**Wasted Votes**	
	Party A	*Party B*	*Party A*	*Party B*	*Party A*	*Party B*	*Party A*	*Party B*
1	73	27	23	27	73	27	23	27
2	73	27	23	27	*70*	*30*	*20*	*30*
3	73	27	23	27	*70*	*30*	*20*	*30*
4	53	47	3	47	53	47	3	47
5	53	47	3	47	53	47	3	47
6	53	47	3	47	53	47	3	47
7	44	56	44	6	*53*	*47*	*3*	*47*
8	44	56	44	6	44	56	44	6
9	44	56	44	6	*41*	*59*	*41*	*9*
10	40	60	40	10	40	60	40	10
Total	**550**	**450**	**250**	**250**	**550**	**450**	**200**	**300**
EG	**(250-250) / 1000 = 0%**				**(300-200) / 1000 = 10%**			
Bias	**(3 / 10) – 50% = -20%**				**(3 / 10) – 50% = -20%**			
MMD	**53% – 55% = -2%**				**53% – 55% = -2%**			

Votes and wasted votes by district in two elections under the same ten-district plan. The efficiency gap is calculated using the full method. Partisan bias and the mean-median difference are also calculated. Votes and wasted votes are placed in italics in Election 2 where they are different from in Election 1.

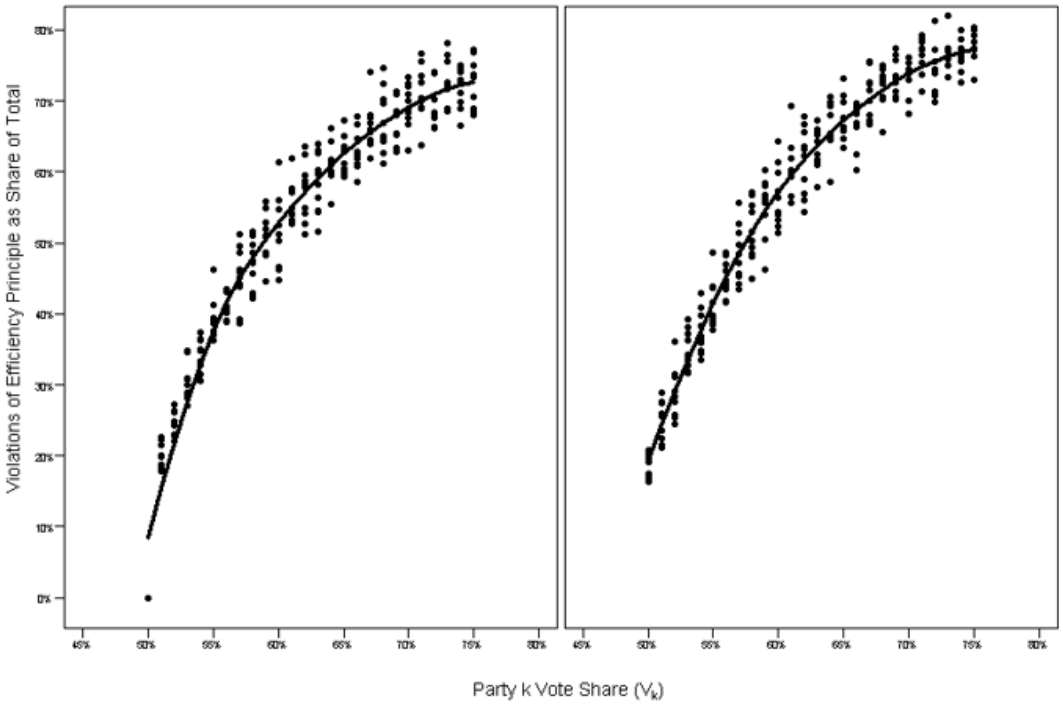
How typical are these examples? In other words, how often do partisan bias and the mean-median difference contravene the efficiency principle over a larger set of cases? To find out, one of us used the simulation technique we described earlier, generating 5100 district plans, 100 for each one-point increment of vote share between 25% and 75%.[179] For each increment, Figure 14 plots the proportion of the simulated plans (paired in every possible permutation) that are inconsistent with the efficiency principle. The chart on the left is for partisan bias; the one on the right is for the mean-median difference.

It is clear from Figure 14 that partisan bias and the mean-median difference violate the efficiency principle quite often indeed. The rate of violations is relatively low when both parties receive close to 50% of the statewide vote. (In fact, in a perfectly tied election, partisan bias never violates the efficiency principle.[180]) But as elections become less competitive statewide, the rate of violations rises dramatically. This rate is about 50% when one party receives three-fifths of the statewide vote, indicating that in this electoral environment, partisan bias and the mean-median difference violate the efficiency principle for every other pair of simulated plans. When one party receives three-fourths of the statewide vote, the violation rate surges above *70%*, meaning that in this setting, partisan bias and the mean-median difference breach the efficiency principle more frequently than they abide by it.

[179] *See supra* notes 48-51 and accompanying text; *see also* McGhee, *supra* note 19 (manuscript at 16-19).

[180] This, of course, is the one point at which partisan bias is equal to the efficiency gap. *See supra* notes 167-168 and accompanying text.

FIGURE 14: SIMULATIONS WITH MULTIPLE GERRYMANDERING METRICS



Scatter plots of the proportion of simulated plans that violate the efficiency principle versus the statewide vote share of the benchmark party, for partisan bias and the mean-median difference.

This analysis relates not just to our first criterion (compliance with the efficiency principle) but also to our third one (breadth of scope). Partisan bias and the mean-median difference become less usable as elections grow less competitive statewide because they produce more violations of the efficiency principle. What accounts for this pattern? In the case of partisan bias, the explanation is that the hypothetical tied election diverges more from the actual election as the actual election gets less competitive. When both parties receive close to 50% of the statewide vote, only small adjustments have to be made to the parties' votes in each district to simulate a tied election, and few seats change hands as a result. But when one party predominates in a state, larger adjustments must be made, and more seats consequently flip. These flipped seats are assigned to the party that actually won them for efficiency purposes, but to the opposing party to calculate partisan bias. Hence the more flipped seats there are, the more often the efficiency principle is violated.[181]

In the case of the mean-median difference, it is unsurprising that it performs like partisan bias since, as discussed above, it is equal to partisan bias divided by the slope of the seat-vote curve.[182] Additionally, the weak link that exists between the mean-median difference and the

[181] For similar arguments, see McGhee, *supra* note 5, at 62; and Stephanopoulos & McGhee, *supra* note 5, at 860-62. The proponents of partisan bias concede that the measure "is only appropriate for competitive situations where there is a potential for a change in partisan outcomes (majority control, in particular) as a result of shifting electoral tides." Grofman & King, *supra* note 1, at 31.

[182] *See supra* note 170 and accompanying text.

parties' seats when elections are competitive statewide disappears when they are more one-sided. In competitive settings, inferences can be drawn about the parties' seats from their mean and median vote shares. For example, if a party's mean vote share is a bit higher than 50% and its median vote share is a bit lower, it is likely (though not certain) that the party has a narrow legislative minority despite its slim popular majority.[183] Speculative as such inferences are when elections are competitive statewide, they become entirely infeasible when one party predominates. In these circumstances, the median vote share tells us which party controls more seats, but this is the only information we can glean from the components of the mean-median difference. Virtually unmoored from the parties' seats, the metric rarely satisfies the efficiency principle—a test whose centerpiece is the relationship between the parties' seats and votes.[184]

Since the efficiency gap always complies with the efficiency principle,[185] while partisan bias and the mean-median difference increasingly violate it as elections grow less competitive, we might expect the measures to behave similarly in competitive environments but to diverge in uncompetitive ones. The mathematical links between the metrics support this expectation. When an election is perfectly tied, the efficiency gap and partisan bias are identical, and the mean-median difference is (as ever) a function of partisan bias.[186] So the closer an election is to an outright tie, the more highly correlated all three measures should be. And the further an election is from an even split, the more attenuated the metrics' connections should become.

Figure 15 confirms this hypothesis. It plots the efficiency gap (on the y-axis) versus either partisan bias or the mean-median difference (on the x-axis) for state house elections (left side) and congressional elections (right side). The top row displays the measures' relationships in competitive settings, where the parties' statewide vote shares are closer than 55% to 45%. The bottom row shows their connections in uncompetitive settings, where one party receives more than 55% of the statewide vote. Plainly, the efficiency gap correlates strongly with both partisan bias and the mean-median difference in competitive environments, in both state house and congressional elections. In the top row of charts, the points cluster around the best fit lines. And equally plainly, the metrics' ties are weaker in uncompetitive environments, at both the state house and congressional levels. In the bottom row of charts, the points are widely dispersed and the best fit lines are closer to flat.[187]

[183] *See* McGhee, *supra* note 19 (manuscript at 16) ("If the median equals . . . a vote share of 0.5 . . . then when a seat changes hands, the median will also change and the median-mean difference will reflect that change.").

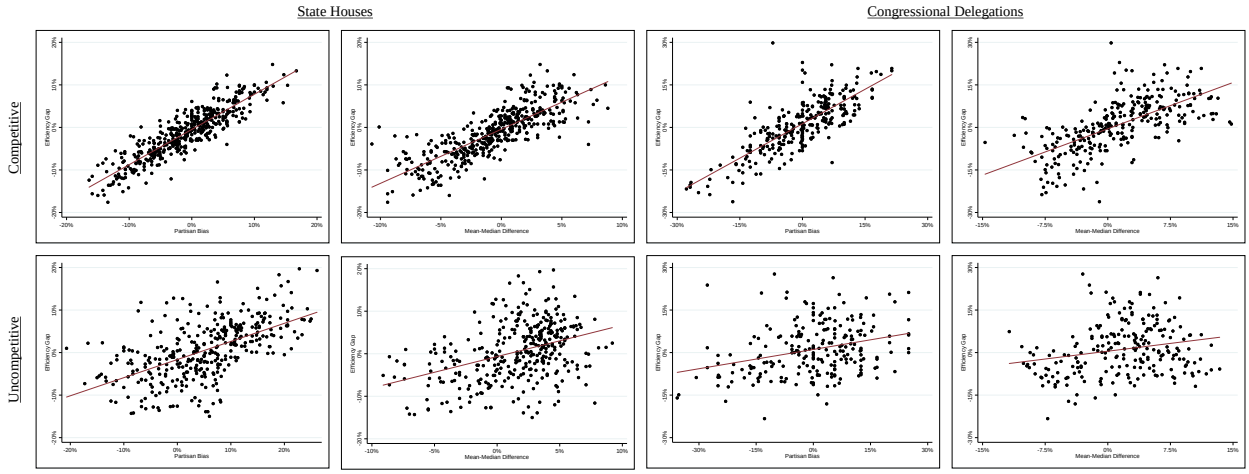
[184] The proponents of the mean-median difference agree that for the measure to be useful, "a jurisdiction has to be politically competitive" and "[b]oth parties have to be capable of winning a vote majority." McDonald & Best, *supra* note 10, at 319.

[185] *See supra* Part II.A.

[186] *See supra* Part IV.B.

[187] For a similar finding—albeit based on simulated rather than actual election results, and only with respect to partisan bias—see Stephanopoulos & McGhee, *supra* note 5, at 857-58.

FIGURE 15: RELATIONSHIPS BETWEEN MULTIPLE GERRYMANDERING METRICS



Scatter plots of the efficiency gap versus partisan bias or the mean-median difference, for state houses and congressional delegations. The top row is for competitive settings, where the parties' statewide vote shares are closer than 55% to 45%. The bottom row is for uncompetitive settings, where one party receives more than 55% of the statewide vote.

In combination, we think Figures 14 and 15 counsel against using partisan bias and the mean-median difference when elections are uncompetitive statewide. The measures frequently violate the efficiency principle in these settings, and are linked only loosely to the efficiency gap. We suppose, though, that one might disagree with this recommendation if one dismissed the relevance of the efficiency principle. The argument would then be that it is the *efficiency gap* that should not be used in uncompetitive environments since it diverges from partisan bias and the mean-median difference.[188]

To respond to this challenge, we invoke a value distinct from efficiency: impact on legislative representation.[189] Even if the point of a partisan gerrymander is not to win more seats for the same votes, it might be to shift the ideological makeup of the legislature in the direction of the gerrymandering party. That way the gerrymandering party is better positioned to pass its preferred laws—to convert mere seats into tangible policy outcomes. To assess the effects of the efficiency gap, partisan bias, and the mean-median difference on legislative representation, we run a series of regressions based on a model that one of us previously constructed.[190] All of these regressions are for congressional elections, because it is only at this level that representational data exists dating back to 1972.[191] The dependent variable in each case is the average DW-

[188] Notably, the proponents of partisan bias and the mean-median difference do not make this argument. *See supra* notes 181, 184.

[189] This, of course, is the same value that Chambers et al. bring up in the critique that we addressed in Part III.C., *supra*.

[190] *See* Stephanopoulos, *supra* note 63 (manuscript at 17-18, 27).

[191] For a similar analysis of state house representation, restricted to the period from 1995 to 2014, see Caughey et al., *supra* note 148.

Nominate score of a congressional delegation in a particular term.[192] (DW-Nominate scores are derived from House members' roll call votes and are the most widely used measure of congressional ideology.[193]) The independent variables are a given gerrymandering metric's value in a given election, the Democratic share of the statewide vote, and fixed effects for states and years. And we run separate regressions for each measure and for competitive and uncompetitive elections, clustering standard errors by state.

Figure 16 plots the impact on a congressional delegation's average DW-Nominate score when the efficiency gap, partisan bias, or the mean-median difference increases by one standard deviation, in competitive and uncompetitive settings.[194] The influence of the efficiency gap is unaffected by the electoral environment. In both competitive and uncompetitive elections, a one-standard-deviation rise in the efficiency gap results in an increase of about 0.06 (or one-third of a standard deviation) in a delegation's average DW-Nominate score. In competitive settings, one-standard-deviation rises in partisan bias and the mean-median difference have comparable effects: increases of roughly 0.05 and 0.03, respectively, in a delegation's average DW-Nominate score. But in uncompetitive environments, the impact of partisan bias and the mean-median difference disappears entirely. When these metrics rise by one standard deviation, a delegation's average DW-Nominate score does not change at all—or even moves slightly, though not statistically significantly, in the *opposite* direction.

The upshot of this analysis is that even if one rejects the efficiency principle, one still should not use partisan bias or the mean-median difference in uncompetitive settings. When one party predominates, an increase in one of these measures has no consistent effect on the ideological makeup of the legislature. An increase in one of these measures, that is, does not skew representation in favor of the gerrymandering party or help it enact its preferred policies.[195]

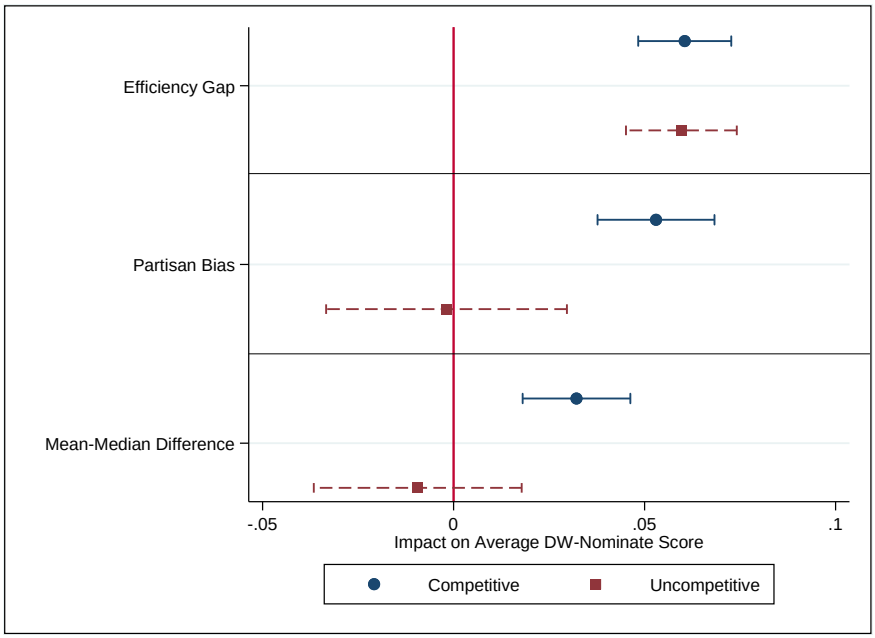
[192] We use the mean rather than the median DW-Nominate score because it is less prone to large jumps from year to year. We also re-sign DW-Nominate scores so that positive values are liberal and negative values are conservative.

[193] *See DW-NOMINATE Scores With Bootstrapped Standard Errors*, VOTEVIEW.ORG (Sept. 17, 2015), http://www.voteview.com/dwnomin.htm.

[194] The results of the regressions are reported in the Appendix, *infra*.

[195] For a similar finding, see McGhee, *supra* note 5, at 66-68 (finding that in uncompetitive elections, partisan bias has virtually no predictive power for a party's seat share).

FIGURE 16: IMPACT OF MULTIPLE GERRYMANDERING METRICS ON CONGRESSIONAL REPRESENTATION



The point estimates represent the impact on a congressional delegation's average DW-Nominate score of increasing the efficiency gap, partisan bias, or the mean-median difference by one standard deviation. The bars to the left and right of the point estimates denote 95% confidence intervals. Separate point estimates are provided for competitive and uncompetitive electoral environments.

CONCLUSION

As partisan gerrymandering enjoys a burst of academic attention, we find ourselves pleased but still dissatisfied by the burgeoning literature. Pleased because a vitally important topic is being studied more rigorously than ever by a range of talented scholars. But still dissatisfied because most recent work has sought only to devise new measures of gerrymandering or to criticize existing metrics. In this Article, we have tried to do something different: to take a step back and think about what it is that we want from a gerrymandering measure. In our view, a metric should be consistent with the efficiency principle, distinct from other democratic values, broad in scope, and empirically congruent with American electoral history. Also in our view, the efficiency gap achieves these goals while other measures often do not. The second point, though, is less important than the first. Our primary aim here is not to evangelize in favor of the efficiency gap. Rather, it is to push the academic debate in new and hopefully more productive directions. If we manage to do so, we will be satisfied even if readers remain unsure about the efficiency gap itself.

APPENDIX

VARIABLES	(1) EG Competitive Elections	(2) EG Uncompetitive Elections	(3) Bias Competitive Elections	(4) Bias Uncompetitive Elections	(5) MMD Competitive Elections	(6) MMD Uncompetitive Elections
Efficiency Gap	0.650***	0.641***				
	(0.0631)	(0.0755)				
Partisan Bias			0.498***	-0.0174		
			(0.0694)	(0.144)		
Mean-Median Difference					0.576***	-0.168
					(0.122)	(0.236)
Democratic Vote Share	1.153***	1.328***	0.891***	1.482***	1.049***	1.518***
	(0.169)	(0.253)	(0.186)	(0.396)	(0.227)	(0.374)
Constant	-0.563***	-0.719***	-0.456***	-0.819***	-0.540***	-0.828***
	(0.0725)	(0.133)	(0.0849)	(0.230)	(0.101)	(0.210)
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	261	203	261	203	261	203
R-squared	0.943	0.960	0.907	0.930	0.886	0.931

Dependent variable is a congressional delegation's average DW-Nominate score

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1