ELECTION LAW JOURNAL
Volume 17, Number 1, 2018

DOI: 10.1089/elj.2016.0392

Original Articles

Considering the Prospects for Establishing a Packing Gerrymandering Standard

Robin E. Best, Shawn J. Donahue, Jonathan Krasno, Daniel B. Magleby, and Michael D. McDonald

ABSTRACT

Courts have found it difficult to evaluate whether redistricting authorities have engaged in constitutionally impermissible partisan gerrymandering. The knotty problem is that no proposed standard has found acceptance as a convincing means for identifying whether a districting plan is a partisan gerrymander with knowable unconstitutional effects. We review five proposed standards for curbing gerrymandering. We take as our perspective how easily manageable and effective each would be to apply at the time a redistricting authority decides where to draw the lines or, post hoc, when a court is asked to decide whether an unconstitutional gerrymander has been enacted. We conclude that, among the five proposals, an equal vote weight standard offers the best prospects for identifying the form of unconstitutional gerrymanders that all but ensure one party is relegated to perpetual minority status.

Keywords: gerrymander, vote dilution, efficiency gap, partisan symmetry

Partisan gerrymandering has become such a dark art that retired Justice John Paul Stevens proposed a constitutional amendment to curb it (Stevens 2014). After the 2000 round of redistricting, David Mayhew pointed to five cases of deft gerrymandering—Florida, Michigan, Ohio, Pennsylvania, and Texas (Mayhew 2011, 24; see also Toobin 2003), to which three others could have been added—California, Illinois, and South Carolina (McDonald and Best 2015, 321). After the 2012 round of redistricting, credible gerrymandering allegations have been leveled at no fewer than ten states: Florida, Georgia, Illinois, Louisiana, Maryland, North Carolina, Ohio, Pennsylvania, Tennessee, and Texas (Fang 2014). One could likely add Michigan and Wisconsin without any stretch of credibility. In all these cases the party in power is suspected of designing districts to perpetuate their majority control of a congressional delegation or state legislative chamber almost regardless of what a majority of voters would decide were they not preorganized in clusters favoring the party in power. The artistry, of this sordid sort, is accomplished through so-called packing gerrymanders. Very many partisans of one stripe are crammed into a small number of districts while partisans of the other stripe are given strong but not overwhelming majorities in the larger number of remaining districts.

Justice Stevens' call for a constitutional amendment comes in the face of two frustrations. Only a few states have shown a willingness to police partisan gerrymandering on their own, and courts have been unable to craft a diagnostic standard that identifies whether a districting plan produces constitutional harm. Needless to say, the wait for a constitutional amendment requires as much patience as the wait for states to adopt rules themselves. Instead

Robin E. Best is an associate professor of Political Science at Binghamton University in Binghamton, New York. Shawn J. Donahue is a JD and PhD candidate in Political Science at Binghamton University. Jonathan Krasno is an associate professor of Political Science at Binghamton University. Daniel B. Magleby is an assistant professor of Political Science at Binghamton University. Michael D. McDonald is a professor of Political Science at Binghamton University.

of waiting, we ask whether any of five recent proposals to assess partisan gerrymandering might be able to supply redistricting authorities in the first instance or courts, if needed later, with a manageable and effective diagnostic tool.

The five proposals are

(1) an *efficiency gap* test (Stephanopoulos and McGhee 2014);
(2) a test *comparing seats won to neutral expectations* (Chen and Rodden 2013a);
(3) an *equal vote weight* test (McDonald and Best 2015);
(4) a *partisan symmetry* test (Grofman and King 2007); and
(5) a *three-prong* test (Wang 2016).

Manageability refers to the clarity and ease with which an analyst can observe a standard's proposed showing of effect. Why? Absent a clear and easily observed effect, debatable aspects of the principal facts leave a conclusion in doubt. *Effectiveness* refers to the accuracy by which a standard's proposed showing of effect identifies gerrymandering as the cause of violating a constitutionally protected right. Why? Absent an accurate assessment of gerrymandering as the cause, doubts about the possibility of false negative or false positive inferences overtake a conclusion.

The next section lays a conceptual foundation by using the language of the Supreme Court to identify the constitutional harm packing gerrymanders can inflict. The third section, first, details the principles of manageability and effectiveness we use to evaluate each proposed standard and, next, describes the types of vote dilution the different standards are designed to uncover. The fourth section describes the reasoning associated with each of the five standards and, through a series of hypotheticals, offers preliminary evaluations of their manageability and effectiveness. Because hypotheticals are useful for illustrating general principles but are prone to doubts about how they operate in actual applications, the fifth section extends the evaluations by applying each standard to state senate districting plans in North Carolina and Iowa. North Carolina is a case where the intention to gain partisan advantage is acknowledged; Iowa is the poster child for a districting process that has neither the intent nor the effect of producing a partisan gerrymander. Thus, reliance on these two cases provides opportunities to check for false negative (North Carolina) and false positive (Iowa) readings.

While arguably manageable, we find that counting wasted votes (aka, the efficiency gap test) relies on a dubious definition of wasted votes and is decidedly ineffective because wasted votes occur for reasons other than gerrymandering. Comparing seats won to neutral expectations requires a set of neutrally drawn districts, a process that can encounter manageability problems due the black-box computer algorithms they require, and they can suffer effectiveness problems because a disadvantaged party hamstrung by a cracking gerrymander can win seats at or even above expectations when its votes amount to less than a majority. The equal vote weight test is manageable and mostly effective but not as aggressive as might be preferred. Testing for partisan symmetry is mostly effective but not entirely manageable because its reading of gerrymanders requires reliance on nonfactual hypotheticals. Finally, the three-prong approach fails on its own terms because the prongs do not fit together as a coherent whole and, worse, the prongs can operate at cross-purposes. All in all, the reviews lead to this conclusion: the equal vote weight standard is the most easily manageable and effective at identifying packing gerrymandering as the cause of a constitutional harm: diluting the votes of one set of partisans.

PARTISAN GERRYMANDERS OF THE PACKING VARIETY

All five proposed standards have been aimed at identifying packing gerrymanders.[1] As remarked, packing gerrymanders concentrate a large number of the disadvantaged party's voters in a small number of districts. When one party's voters are packed

[1]Wasted votes were the primary evidence of effect in a Wisconsin State Senate challenge (*Whitford v. Gill* 2016). Comparing wins was used in a challenge to Florida's congressional districts (*Romo v. Detzner* 2014). The equal vote weight standard was proposed by amici (Hebert and Lang 2015) at the remedy stage of the Virginia litigation that earlier found the State's congressional districts to be an unconstitutional racial gerrymander (*Page v. Virginia State Board of Elections* 2014). Seat-denominated symmetry was proposed to the Supreme Court by amici (King et al. 2005) for consideration in *LULAC v. Perry* (2006). One of the three prongs was proposed by amici (Wang 2015) in *Harris v. Arizona Redistricting Commission* (2016).

into a few districts, the packed partisans hold overwhelming majorities in those districts. Packing gerrymanders also serve to spread the packed party's remaining voters over a large number of districts where they constitute sizable but ineffective minorities.[2] By way of example, a competitive jurisdiction with 10 districts and a vote typically expected to split 52 percent Democrat and 48 percent Republican might enact a packing gerrymander by granting Republicans two districts that are 100 percent Republican and next set up the remaining eight so that they split 35 versus 65, Republican versus Democrat. The result is two safe Republican seats and eight safe Democratic seats, a seat split that would likely hold even if votes shifted substantially in the Republicans' favor. Notice that packing uses cracking at a second step. In the example, two districts are packed with Republicans; this recasts the system-wide percentages among the other eight, which are then cracked, safely for Democrats, so they all divide 35–65.

In theory an optimal partisan gerrymander can be shown to involve pure cracking (Freidman and Holden 2008), but as Owen and Grofman have shown, for reasons both of a party's desire for legislative majority control and of it and its individual candidate's risk aversion, an optimal gerrymander under competitive circumstances relies on packing (Owen and Grofman 1988; see also Gul and Pesendorfer 2010).[3] In any case, as we have noted (fn. 1), the five proposed standards have been aimed at packing gerrymanders and so, too, has the Supreme Court's attention in three major partisan gerrymandering decisions, *Davis v. Bandermer* (1986), *Veith v. Jubelirer* (2004), and *LULAC v. Perry* (2006).[4]

Justice Scalia, announcing the Court's judgment in *Veith*, defined gerrymandering as "[t]he practice of dividing a geographical area into electoral districts, often of highly irregular shape, to give a political party an unfair advantage by diluting the opposition's voting strength" (*Vieth v. Jubelirer*, 2004, 271 n. 1, quoting *Black's Law Dictionary* 1999, 696). Finding intention and observing weirdly shaped districts are seldom difficult (as in *Davis v. Bandemer* 1986; *Veith v. Jubilier* 2004, *LULAC v. Perry* 2006), but finding a standard that identifies a party's unfair advantage because the opposition party's votes have been diluted has proved elusive.

In *Bandemer*, Justice White explained the Court majority's holding of justiciability of partisan gerrymandering in response to a caution from Justice O'Connor. She worried that judicial attempts to police partisan gerrymandering would have courts give preference to proportionality. Justice White and the majority disagreed; justiciability of packing forms of partisan gerrymandering rests on the Court's preference not for proportionality but, rather, for ensuring that popular "majorities are not consigned to minority status" (*Davis v. Bandemer* 125, n. 9).[5] Such majority-to-minority consignment would signal vote dilution because turning a majority into a minority occurs only if the votes of those in the vote majority count less than those in the vote minority.

The Court's disagreement with Justice O'Connor came in a context of whether its approach to racial gerrymandering could also apply to partisan gerrymandering. It can, but with an important

[2]Gerrymandering is a term used to cover a large range of electoral manipulations. Aside from the packing gerrymander focus under review here, pure cracking gerrymanders spread one party's votes evenly across districts so that they constitute sizable but losing minorities in all districts. These are most effective, least risky, in jurisdictions with lopsided competition. At-large and multi-member district plurality elections with their super-majoritarian effects are referred to as institutional gerrymandering (Dixon 1971, 54). Creating under-populated districts for one versus the other partisan group is a form of malapportionment gerrymandering (Brunell 2012; *see also Harris v. Arizona Redistricting Commission* 2016). Creating a district adverse to or favorable to particular candidates are "personalized" gerrymanders or, when the candidates in question are incumbents, "incumbent-displacement" gerrymanders (Owen and Grofman 1988, 14–16). Each has its own means and methods for accomplishing its manipulation and thus is best approached with its own form of precisely aimed standard for detection.

[3]Freidman and Holden's terminology can be misleading in that their title advises never cracking. Notice, however, they have in mind an uncommon meaning of cracking. They come at the issue from an approach that assigns individuals to districts and from there advises placing (packing in their meaning) the most staunch opposition partisans in districts with one's own staunch supporters. "Intuitively, extreme Democrats can be neutralized by matching them with a slightly larger mass of extreme Republicans" (Freidman and Holden 2008, 115). Discussions of gerrymandering normally refer to this as cracking or dispersal gerrymanders—spreading opposition partisans over many districts to deny them majority control in as many as possible (*see, e.g.*, Owen and Grofman 1988, 6).

[4]The Court considered allegations of a different form of partisan manipulation in *Harris v. Arizona Redistricting Commission* (2016). There, as remarked on in note 2, *supra*, the issue was neither packing nor cracking, as such, but malapportionment partisan manipulation by systematically underpopulating districts favoring Democrats (*see* Brunell 2012 for a general discussion of this form of manipulation).

[5]In relation to purely cracking forms of gerrymander, Justice White refers to the Court's concern for ensuring "significant minority voices are heard" (*Davis v. Bandemer* 1986, n. 9).

qualifying complication. In the same term that *Bandemer* was decided, the Court spelled out a three-prong test for racial gerrymandering (*Thornburg v. Gingles* 1986). While the allegation of racial vote dilution involved several of North Carolina's multimember districts, the *Gingles* standard could be (and later was) extended to strictly single-member district plans (*Growe v. Emison* 1993; *Voinovich v. Quilter* 1993; *Johnson v. DeGrandy* 1994). It calls for comparing the actual number of majority-minority districts to the number that could reasonably be expected to exist when a fair set of single-member districts is drawn.[6]

On its face, it would appear simple to transfer that diagnostic to partisan gerrymandering. One could ask whether Democrats and Republicans have won a number of districts compared to what could be expected under a fair set of compact and contiguous single-member districts. The resemblance is not quite as straightforward as it appears, however. Unlike counting people based on race or language minority status, where the relevant number is determined and essentially fixed by census count, vote counts vary from one election to another. In a packing gerrymander, an unfair allocation of seats of, say, 40 percent when a party wins 50 percent of the vote is readily apparent. However, when the same party receives only 40 percent of the vote and wins the same 40 percent of the seats, the plan would appear eminently fair. This sort of variable result could occur in a packing gerrymander precisely because a packing gerrymander is designed to grant the disadvantaged party some minority percentage of seats over a wide range of vote percentages. As we shall demonstrate, taking account of this understanding of how packing gerrymanders operate in differential ways when votes vary between low and high is a difficult problem that the five standards propose to but sometimes fail to resolve.

EVALUATIVE FRAMEWORK

We are looking for an easily manageable and effective standard for identifying packing gerrymanders that dilute the voting weights of one party's voters. Easy manageability refers to a diagnostic method that calls for a clear and self-evident observation of the facts as the basis upon which the ultimate inference is to rest. The more directly observable the facts, the more indisputable are the foundation stones of what everyone observes. Indubitably, such transparency fades to ambiguity the more the prescribed method requires leveraging assumptions. The fourth section identifies assumptions each standard relies on to establish the factual underpinning it calls for.

Effectiveness refers to a diagnostic method that avoids errors. A false negative error occurs when a method fails to identify a gerrymander even though the choice of where to place the district lines actually caused vote dilution. A false positive error occurs in either of two ways: a proposed standard identifies vote dilution when there is none, or it identifies gerrymandering as the cause of vote dilution when the cause is attributable to something else. In addition to highlighting assumptions relevant to manageability, the fourth section identifies possible reasons to be concerned about inferential errors. Because possible reasons for doubt are potentially more hypothetical than real, the fifth section evaluates effectiveness in two applications. If we accept that North Carolina's senate districts are a partisan gerrymander, which the state acknowledges, and Iowa's senate districts are not a partisan gerrymander, which most observers acknowledge, then a standard that fails to identify North Carolina's gerrymander or misidentifies Iowa's districts as a gerrymander is committing error. Moreover, if a standard sometimes identifies the same set of districts as a gerrymander with respect to some elections and a non-gerrymander with respect to other elections, we know with assurance it is committing errors.

As for the concept of vote dilution, it must be said that four of the five standards have in mind their own particular meaning. The discussions and analyses accept each standard's definition, and thus we evaluate manageability and effectiveness on each standard's own terms of what it means to dilute votes.

Comparing parties' wasted votes considers dilution to occur when one party's voters cast more

[6]Justice Brennan explained the Court's rationale this way. "The reason that a minority group making such a challenge must show, as a threshold matter, that it is sufficiently large and geographically compact to constitute a majority in a single-member district is this: Unless minority voters possess the potential to elect representatives in the absence of the challenged structure or practice, they cannot claim to have been injured by that structure or practice" (*Thornburg v. Gingles* 1986, 50 n. 17).

unneeded votes in the senses that they go to loosing candidates or exceed what is necessary to win a seat. If votes for one party are more likely to count for nothing, that party has more votes with zero weight and thus more votes that are diluted to a maximum extent. The comparison of wins standard sees dilution as existing to the extent that one set of partisan votes do not count as much as they should because they elect fewer of their party's candidates than would be expected under neutrally drawn districting procedures. This is the direct analogue to the approach taken by the Court in racial gerrymandering. The equal vote weight standard is a vote-denominated symmetry idea that says vote dilution is foretold by comparing the median district to mean district vote percentage. If all votes count the same, the median and mean have the same numerical value; if the median and mean differ, votes for the two major parties count differently as a consequence of being divided into districts. The partisan symmetry standard aims at non-dilution in the sense that whatever seat percentage one party wins with a given vote percentage, the other party is expected to win that same percentage of seats with that same percentage of votes. The idea here is that the same resources, votes, reap the same rewards, seats; otherwise, the two sets of voters are not counting equally. The three-prong test has more expansive interests that include vote dilution but carry concerns beyond just that concept. Its focus includes (1) seat-vote outcomes that hue towards proportional representation; (2) seat shifts that are responsive to vote shifts; and, (3) depending on competitiveness, a non-gerrymandered plan that either preserves symmetry or ensures the predominant party's district vote percentages are not too similar.

FIVE STANDARDS

Efficiency gap

Counting and comparing wasted votes is the basis for the efficiency gap standard proposed by Stephanopoulos and McGhee (2015; see McGhee 2014 for the underlying social science thinking). The approach proceeds from the insight that both winners and losers "waste" votes by inefficient allocation in an election. That is, any votes above the 50% +1 for the winner plus all votes for the loser are wasted in that they contribute nothing of determinative importance to deciding who wins. In a single-district election decided by a 60–40 margin, the winner wastes 10 percentage points above 50% (setting aside ties for the sake of simplicity), while the loser wastes all 40 percentage points. Comparing the magnitude of the waste on both sides, 10 versus 40, shows an efficiency gap (of 30 points) favoring the winner. McGhee and Stephanopoulos argue that in a non-gerrymandered system both sides waste the same number of votes, so ideally the efficiency gap should equal zero.

Their claim has an appealing label along with a seemingly simple, straightforward, and intuitive procedure for calculating a numerical indicator. Nevertheless, it runs into manageability difficulties in two regards: (1) it assumes wasted votes are to be counted in an odd way, and (2) it has no secure baseline for establishing the degree of wasted votes that indicates a gerrymander. Effectiveness difficulties arise for three reasons: (1) votes are wasted for reasons other than gerrymandering; (2) the wasted vote gap co-varies with a party's vote percentage; and (3) the method seeks to cover both cracking and packing gerrymanders in one calculation and thereby can allow some amount of cracking to disguise an undue amount of packing.

Even though the arithmetic required is simple, and in that sense would seem to clear the manageability bar, the efficiency gap's definition of votes wasted by the winning candidate is disputable.[7] In particular, decades ago Andrew Hacker, who refers to the winner's wasted votes as *excess* votes, defines them as one more than the votes received by the losing candidate (Hacker 1964, 55–7). McGhee (2014) and Stephanopoulos and McGhee (2015) define a winner's excess/surplus/wasted votes as votes beyond 50% +1. It runs into a second manageability problem when deciding how many wasted votes signal a gerrymander. Because no democratic or legal principle answers the question of how many wasted votes are needed to say a plan is a gerrymander, the approach calls for comparisons to the historical record in the same jurisdiction and contemporaneous results in other jurisdictions. Such relative baselines beg the question of whether what occurred previously in the same jurisdiction or

[7]Judge Greisbach, dissenting in *Whitford*, goes so far as to call the efficiency gap's method of counting excess wasted votes "absurd" (*Whitford v. Gill* 2016, 150).

is occurring contemporaneously in other jurisdictions are results contaminated by gerrymandering.[8]

The efficiency gap runs into three problems related to its effectiveness. First, and simply, under single-member district rules votes are wasted for reasons other than gerrymandering. One needs to look no further than a simple example of a congressional district in a one-district state such as Delaware to see this. Unless the vote splits 75–25, one party wastes more votes than the other, this despite the fact that a gerrymander is impossible in a one-district state. Maybe the efficiency gap is useful only in multi-district situations, but that can't be true either. Therein resides the efficiency gap's second effectiveness problem. In a three-district state, a symmetrical distribution of 48–52–56 has a gap of +8.3 in favor of the majority party and is, by the eight-point criterion, a gerrymander. Of course, if the vote shifts uniformly to 46–50–54, there is no gerrymander, even though it is the same districting plan. Then, if votes shift another two points to 44–48–52, the gerrymander would be said to run in the direction opposite of what was inferred from the original 48–52–56 distribution. In this scenario, the relative distribution of partisan voters did not change—neither party became relatively more (or less) packed—and yet the efficiency gap registered a substantial shift in partisan advantage. In fewer words, reading a gerrymander from the efficiency gap can and often will vary depending on the underlying percentage level of the votes a party receives.

A third effectiveness problem has to do with the translation of votes to seats, the seat-vote ratio. Assuming equal turnout in all districts, a majoritarian seat-vote ratio of two to one is sufficient for equalizing wasted votes—i.e., having a seat percentage in excess of 50 equal to two times the vote percentage in excess of 50 produces an equal number of wasted votes (McGhee 2014, 79–80; Stephanopoulos and McGhee 2015, 853). For example, winning 60 percent of the seats (10 points above 50) in association with winning 55 percent of the votes (five points above 50) indicates there is no gerrymander. However, that is not necessarily so. A majoritarian seat-vote correspondence of two-to-one can occur even when a packing gerrymander is in place. Hence, a two-to-one seat-vote ratio is not a sufficient condition to conclude there is no gerrymander. For example, consider a 40–40–60–65–70 vote distribution. The distribution is asymmetrical (median 60 and mean 55), but the efficiency gap shows an equal number of wasted votes. Votes are five points above 50, and seats are ten points above 50; the majoritarian ratio is two-to-one even though the distribution is asymmetrical. Thus, despite its proponents' claims to the contrary, the efficiency gap standard does not comport with nor arise from the idea of partisan symmetry.[9]

The wasted vote approach has clear intuitive appeal. Nevertheless, it has several downsides. One, its computation poses a manageability problem because it relies on a shaky definition of what it means to waste a vote, given the alternative way of counting excess votes (as in Hacker 1964; *Whitford v. Gill*, 2016, 150–2, Greisbach dissenting). Two, it underachieves on the question of manageability because evaluation of the wasted vote computation requires using a relative comparison to the historical record of elections in the same jurisdiction or to elections in other jurisdictions. A historical comparison is liable to perpetuate gerrymanders in earlier years; comparison to other jurisdictions leaves one wondering whether the baseline involves a mix of fair and unfair outcomes elsewhere. What's more, it can under-reach and overreach on questions of effectiveness for three reasons, each functionally related to its implications that single-member district elections are fair if and only if they operate with a seat-vote majoritarian ratio of two to one. Underreaching occurs when it offers a false negative reading of gerrymandering because, despite substantial packing, the majoritarian ratio is two to one. It overreaches when it offers a false positive reading of gerrymandering by indicting a districting plan as a gerrymander because it has many competitive districts that slightly favor one party and thus a majoritarian ratio greater than two to one.

Comparing wins

This approach identifies diluted votes as winning fewer seats than expected in districting plans

[8]In some applications an efficiency gap beyond ± 8 indicates a gerrymander (Stephanopoulos and McGhee 2015, 831). In other applications, a gap beyond ± 7 is deemed indicative (Jackman 2015, 5). As applied to congressional districts, it is designed to be applied only to delegations of eight or more members; in this context a gerrymander is indicated, not by any particular magnitude of the gap, but when one party would have been expected to win two or more seats than it actually did win (Stephanopoulos and McGhee, 2015, 835–6).

[9]*See* Stephanopoulos and McGhee (2015, 834 and passim) for claims about the relationship between symmetry and the efficiency gap.

produced through partisan blind line-drawing. If an enacted plan is an outlier in a partisan-blind null set's expected seat distribution, one can infer that it was probably intended to hold a partisan advantage. This closely aligns with the Court's racial gerrymandering standard that asks for a comparison between how many districts a group actually wins and how many the group would win under a fairly drawn single-member district plan. Its manageability problem arises in association with the black-box nature of the computer algorithm needed to establish the factual baseline for comparison. Its effectiveness can be left wanting because the match of observed versus expected wins (or districts carried) depends on the vote percentage a party wins.

The basic idea behind generating the comparisons is to use a computer to draw a large number of districting plans. Using computers for this purpose is an idea that has been floated at least since William Vickrey made the point more than a half-century ago (Vickrey 1961). A few pioneers succeeded in advancing the idea in modest ways in the 1960s and 1970s (Nagel 1965; Engstrom and Wildgen 1977); then, with advances in processing speed, the approach was ready for a full-scale application years later (e.g., Cirincione, Darling, and O'Rouke, 2000; Altman and McDonald 2011; Chen and Rodden 2013a)—at least it seemed ready in the run up to the Florida proceedings involving the State's congressional districts. Both Thomas Darling along with Jowei Chen and Jonathan Rodden produced null sets in advance of the Florida trial (see Darling 2013; Chen and Rodden 2013b; 2014), and Rodden testified at length. In the end, however, neither the reports nor Rodden's testimony received any mention by the trial court or in subsequent court decisions (*Romo v. Detzner* 2014; *League of Women Voters of Florida v. Detzner* 2015).

For what it says about manageability, the Florida courts' silence is disquieting. It may have been benign. In the face of the smoking gun evidence of partisan maneuvering that violated Florida's newly operative state-constitution intent standard, the court might well have reasoned that nothing as sophisticated as a computer-generated null set was needed.[10] Perhaps, however, the court was dissuaded from crediting the method with probative value because one report identified a few contiguity problems (Hodge 2013) and another report, plus testimony, questioned whether the Chen-Rodden null set was randomly generated since no one can know the characteristics of the population of all possible plans (McCarty 2013; 2014). Or, perhaps and more simply, the black-box nature of the method left the court unsure what reliable conclusions could be drawn.

Because the null set approach has yet to be tried and tested in a full form application, questions about its effectiveness are open. Still, this much can be said. Not enough thought has gone into how the null set could be used to detect gerrymandering beyond forming a baseline to say whether an enacted plan is an outlier in the null set distribution and, on that basis, probably indicates a gerrymander. Engstrom and Wildgen (1977, 469–70) evaluate a plan in regard to how many competitive districts it contains. Cirincione et al. (2000), Darling (2013), along with Chen and Rodden (2013a, 2014), evaluate a plan in regard to the number of districts in which each racial group or political party holds a majority. We have to suppose that focusing solely on the central tendency is not enough. Why? Depending on the vote percentage won by a disadvantaged party, the expected number of competitive districts or of majority-held districts varies and might well include seat outcomes that square with the expectation—i.e., the central tendency—but involve packing.

As an example of the problem associated with a focus on seats won (more precisely, districts carried), consider Chen and Rodden's attempt to indicate a gerrymander by counting President Bush's 2000 or John McCain's 2008 district wins across Florida, in their academic and trial-related work, respectively (Chen and Rodden 2013a, 2013b, 2014). As noticed and noted by both Darling (2013) and McCarty (McCarty 2013, 2014), a match or mismatch between expected and observed number of districts carried is not a per se robust and structural feature of a districting plan. The match or mismatch varies depending on the vote percentage won. A packing gerrymander that all but guarantees that a party win, say, 40 percent of the districts whether it wins, say, 40, 50, or 60 percent of the vote—which is the type of result a packing gerrymander can and often does produce—will sometimes match the expected number of districts carried and

[10]The facts revealed such damning evidence as Republican legislators and their operatives enlisting mapmaking confederates to submit "citizen constructed plans" under fake names and writing scripts for "concerned citizens" to present the operatives' ideas at public meetings (*Romo v. Detzner* 2014, 20–31).

other times will not. In different words, the contours of a districting plan interact with a party's system-wide level of vote support to produce more, equal, or fewer than expected wins. As a consequence, the interaction produces variable readings of gerrymandering under the expected wins standard.[11]

Using computer-generated districts to form a null set holds promise. It removes all but inadvertent partisan effects in its construction of a null set and thus supplies a strong basis for probabilistic inferences about intentions. One problem it has to overcome is making the computer processing more intuitive and transparent. Another pressing matter is choosing a benchmark other than the expected number of competitive districts or the number of district wins. The approach supplies a useful tool, but we need to figure out how to make it transparent and how to use it effectively.

Equal vote weight

The *equal vote weight standard* relies on two observed facts: (1) compare the median district vote percentage to the mean district vote percentage received by the party, and (2) check whether majority rule is violated. When one group of partisans is relatively more packed than the other, a districting plan has the potential to violate the widely embraced principle of equal vote weights and, from the unequal weights, to entrench one party in majority status. Manageability of the equal vote weight standard is straightforward inasmuch as the essential facts are directly observable. Its effectiveness can be challenged, however, because its requirement to observe a violation of majority rule is not as assertive as some ideas about gerrymandering might require.

In all, the standard for a factual identification of a gerrymander rests on three manageable ideas.

(1) *Leading indicator:* Asymmetrical packing exists when the median district vote percentage for one party is persistently lower than its mean district vote percentage.
(2) *Objectionable harm:* A vote weight inequality is clearly identifiable when one set of partisan voters casts a majority of the votes but carries less than a majority of the districts, because violating majority rule occurs only when all votes do not count equally.[12]
(3) *Cause:* District line placements are the known cause of the unequal vote weights. Votes counted system-wide contribute equally to the count. Counting votes after division into districts changes only the manner of counting. To the extent the two forms of counting do not produce the same result, the difference must be caused by the line placements.

Manageable as it is with respect to the required facts, tying its focus to violating majority rule is an arguable shortcoming of its potential effectiveness. Equal median and mean district vote percentages indicate only average symmetry, not full-scale symmetry. Reaching for a full- or at least a fuller-scale approach would be more aggressive. For example, a five-district plan applied to two-party competition that has (expected) Republican district vote percentages of 44, 46, 51, 52, and 62 is symmetrical via the equal vote weight standard but asymmetrical under a full-scale symmetry requirement (i.e., as recorded by partisan symmetry considered next—see below). The median and mean are both 51. Thus, average symmetry is upheld inasmuch as deviations above and below the mean of 51 both average six. Majority rule is also preserved; the vote majority holds a three-to-two seat majority. Full-scale symmetry goes wanting, however, because something like uniform vote swings would result in Republicans winning only three seats with 52 percent of the vote—an upward shift of one point resulting in a 45, 47, 52, 53, 63 distribution—but Democrats win four seats when they have 52 percent of the vote—after a downward shift of three points resulting in a 41, 43, 48, 49, 59 distribution. While majority rule is maintained under both vote swings, the idea of equality is not as aggressive as it might be in the sense that different rewards (seats) can be acquired from the same resources (votes).

[11]Darling analyzed his 5,000-map null set for nine pre-2012 statewide Florida elections in addition to the McCain-Obama presidential contest. For the McCain-Obama contest he found, as did Chen and Rodden, the expected number of McCain wins under the 2012 lines was 14, whereas the enacted districting plan had McCain winning 17—a result observed in less than one percent of the null set plans. However, Darling's analysis of the nine other elections showed the actual versus expected wins either matched (three elections), differed by one in favor of Republicans (three elections), or differed by one or two in favor of Democrats (three elections)—*see* Darling (2013, 16).

[12]As McDonald and Best point out, violation of majority rule is evaluated against the two-party statewide vote percentage and not the district mean vote percentage, in order to ensure that the evaluation does not conflate a violation due to turnout bias with a violation due to gerrymandering bias (McDonald and Best 2015, 318).

The equal vote standard has pros and cons. Its required factual finding is easily observed: compare the median and mean district percentages and check for violations of majority rule. However, it is not as aggressively effective as some might demand. It can be charged with under-reaching by not accounting for situations when vote shifts produce different seat outcomes while winning the same vote percentage.

Partisan symmetry

A proposal for a partisan symmetry constructed on the basis of fair seat-vote translations at various levels of vote splits goes back decades (Gelman and King 1994). It has found favor among political scientists (e.g., Engstrom 2013; McGann at al. 2015, 2016). To some extent it has also found favor among members of the Supreme Court in *LULAC v. Perry* (2006; for a detailed discussion of the Justices' reactions see Grofman and King 2007, 1–6). Its effectiveness would not be much in doubt were it not for the assumptions required to establish baseline hypothetical seat results for making comparisons between the two parties.

The approach, which could be called a seat-denominated symmetry standard, relies on an equal opportunity notion of fairness. Within practical and probabilistically knowable limits, each party is expected to win the same seat percentage for the same vote percentage. Suppose Democrats win 35 of 50 seats, 70 percent, with 55 percent of the vote. Seat-denominated symmetry requires that Republicans win 70 percent of the seats (35 of 50) when they win 55 percent of the vote. This notion of a partisan symmetry standard shares the same concern for asymmetry that violates majority rule as the equal vote weight approach, but it adds a requisite symmetrical operation of the swing ratio. At an even 50:50 vote split, seats should split 50:50, and in the competitive range of two-party vote splits, perhaps inside the 40 to 60 range, if Democrats win five more seats with 53 percent of the vote, then Republicans should be expected to add five seats when their vote is three points above 50. Its attention to the swing ratio bears a similarity to the wasted vote approach; however, it differs by being agnostic about the magnitude of the ratio, provided that the effect of the swing is symmetric.

One way to see the standard's manageability problem is from the example used to point to a shortcoming of the equal vote weight approach. There we had a five-district Democratic two-party vote percentage distribution of 44, 46, 51, 52, and 62. The median and mean are equal, and therefore a vote-denominated indicator of asymmetry is missing. However, as discussed, a three-point uniform shift in favor of the Republicans, moving the median and mean to 54, leaves them with three district wins, while a three-point swing in favor of Democrats leads to four district wins. That, of course, depends on the uniformity of the vote swing. If the swing is non-uniform—i.e., if it is mixed in the sense that some districts swing more than others—we need to know more, much more. Getting an assured handle on what else we need to know was the apparent stopping point for Justice Kennedy when he remarked favorably on the partisan symmetry approach but said courts are "wary of adopting a constitutional standard that invalidates a map based on unfair results that would occur in a hypothetical state of affairs" (*LULAC v. Perry* 2006, 420).

The partisan symmetry standard is more comprehensive than the equal vote weight standard. To realize the added value of it comprehensiveness, however, it can under reach in practice by requiring a supporting analysis that makes some decision makers wary of relying on it because it requires leveraging a variety of not easy to evaluate assumptions embedded in computationally intensive analysis of vote swings.

Three prongs

Because gerrymandering is a complex concept, it might seem to be a good idea to use multiple criteria to evaluate whether one has been enacted. Such is the apparent thought standing behind Samuel Wang's proposed three-prong test (Wang 2016). The three prongs are grounded in concerns for (a) a less than justifiable degree of seat-vote proportionality, (b) under-responsiveness of seat shifts to vote shifts, and (c) asymmetry in the vote distribution.

(1) *Excess seat test:* Seat-to-vote responsiveness is within a range between proportionality and what could be expected from the seat-vote relationship in other states (plus allowance for random variation).
(2) *Lopsided outcomes test:* Unequal average lopsidedness in the vote distribution is evaluated by comparing average values of each party's

winning margin above 50 (plus allowance for random variation).

(3) *Reliable wins test (two forms):* In a competitive jurisdiction a party's median district percentage equals its mean district percentage (plus allowance for random variation); in a non-competitive jurisdiction the dominant party's standard deviation of the vote percentages equals the standard deviation of the party's vote from simulations based on other jurisdictions (plus allowance for random variation).

Having three prongs gives the appearance of a more comprehensive set of concerns than the preceding four approaches. That much can be granted, but having three prongs creates at least two manageability problems. One is reliance on election results from other jurisdictions as a basis for comparison. As with the wasted vote approach, an external standard begs the question of whether what occurs in the jurisdiction in question is the consequence of something particular to the jurisdiction other than the manner in which the jurisdiction was divided into districts. Second, Wang advises that the three prongs can be used "separately or combined" (Wang 2016, 1308). Questions naturally follow: Is satisfying one of the prongs enough to say no gerrymander exists? Is violating one of the prongs enough to say a gerrymander has been enacted?

Wang's advice to use his three prongs independently or in combination also carries with it an effectiveness problem. The different prongs can provide indications running in opposite directions. For example, a five-district distribution of 40, 40, 60, 60, 60 satisfies both proportionality (prong 1) and equal average lopsidedness (prong 2) but fails the symmetry standard of prong 3 (median 60 and mean = 52). Likewise, a swing ratio could reside within the bounds of acceptable proportionality but fail on both lopsidedness and symmetry. And a districting plan could fail the lopsidedness test simply because an election-swing moves the vote percentage away from 50 percent even in the absence of gerrymandering. A second effectiveness problem also relates to a lack of clarity regarding which prongs apply. Requiring failure on all three prongs simultaneously leaves an opportunity for mapmakers to satisfy any one prong while enacting a gerrymander that would be indicated by either or both of the other two prongs. In all, and in other words, the three prongs lack a coherent framework that allows them to work together.

Evaluating gerrymanders through three different tests has an intuitive appeal. Nevertheless, it raises difficult questions for both manageability and effectiveness because, as it stands, no compelling coordinating principle supplies clarity about whether a gerrymander exists according to any or all three prongs.

TWO APPLICATIONS

Argument is instructive but not enough when evaluating standards to be applied not just in theory but also in fact. Below we put all five standards to the test in the contexts of North Carolina's and Iowa's post-2011 enacted state senate districts. We want to see whether any of the five produce false negative or false positive diagnoses.

We select North Carolina and Iowa because one case is rather assuredly a gerrymander (North Carolina) and the other is rather assuredly not (Iowa). That's because North Carolina's post-2011 districts are acknowledged by the state itself, assembly members, and, later, the courts to have been drawn with pro-Republican partisan advantage as one goal (*Dickson v. Rucho* 2014, 3). Iowa's redistricting process is often held up as an exemplar of neutral redistricting. Thus, we have opportunities to check on false negative (North Carolina) and false positive (Iowa) readings.

North Carolina

The North Carolina State Senate is a 50-member body elected every two years from 50 single-member districts. Following the 2010 elections, Republicans took control of the state senate and house for the first time since 1870. The 2010 census data were delivered in March 2011, and in July the legislature passed bills establishing state senate districts for the 2012 elections.[13] Those elections saw Republicans win 66 percent of the senate seats (33 of 50) with 52.8 percent of the vote. Two years

[13]While a Democrat, Beverly Perdue, occupied the governor's office, North Carolina's redistricting bills are not subject to gubernatorial veto.

later, 2014, Republicans won 70 percent of the seats with 54.9 percent of the vote.[14] Both are substantial seat victories, 16 to 20 points in seats beyond 50 percent for votes just three to five percentage points beyond 50. But important facts militate against reading too much into the senate results by themselves. Forty percent of the seats went uncontested by one or the other major parties: 19 of 50 in 2012 and 21 of 50 in 2014. This sort of noncompetitiveness, we have to think, reflects anticipated wins/losses as a consequence of the way the district lines were drawn in the first place, more so than a statement of accurate fact about the partisan disposition of the districts. More generally, prospective candidates in each of the various districts have to be thought to take account of their prospects of winning, in part—likely in substantial part—depending on a district's partisan leanings.

We can avoid the problem of district-by-district state senate election competition being endogenous to the enacted lines by turning to elections for statewide office (often referred to as *exogenous* elections) aggregated into separate counts within each of the 50 districts. The North Carolina General Assembly provides election returns for each of nine statewide offices elected in 2012 (the nine are identified in Table 1) aggregated to U.S. Census Defined Block Groups.[15] All nine elections resulted in vote percentage splits within a reasonably competitive range. We use these nine as the elections holding the most probative value for revealing whether the district lines are a pro-Republican gerrymander. In addition, with the state board supplying election returns for all nine election results disaggregated to the precinct level, we can run a large number of null set applications to generate expectations based on 50 districts drawn through a partisan-blind procedure.[16] This has a direct benefit for evaluating the observed versus expected district wins. In relation to two other proposed standards (not including the partisan symmetry and the three-prong tests) it has two additional benefits. The expectations provide a baseline for what partisan residential patterns alone could be expected to produce in regard to wasted votes and equal vote weights.

As a visual prelude, Figure 1 presents two histograms, one for the gubernatorial election, the least competitive of our nine elections, and the other for the lieutenant governor, the most competitive of our nine elections. Both distributions are bimodal. Just about two-thirds of the districts reside at percentages favorable to the Republicans regardless of whether Democrats won 44.2 or 49.9 percent of the vote. Indeed, when the vote percentage shifts in the Democrats' favor by 5.7 points, from 44.2 Democratic percent for governor to 49.9 percent Democratic for lieutenant governor, the gain in districts carried by the Democratic candidate is a mere one district. The electoral playing field is tilted substantially in favor of Republicans, leaving Democrats with a rather steep hill to climb before having any realistic prospect of winning a majority of districts.

Table 1 reports the Democratic two-party vote percentage for the nine statewide offices (column #1) and the relevant numbers for the five proposed standards (columns #2 through #6). The competitiveness noted above can be seen in the vote percentages; they range between 44.2–55.8 and 54.2–45.8, Democrat-Republican, two-party splits.

[14]Data from North Carolina State Board of Elections Nov 6, 2012 General Election Official Results and November 4, 2014 Official General Election Results are posted on the State Board of Elections (SBoE) website.

[15]We rely on the North Carolina General Assembly's (NCGA) 2016 Redistricting Base Data provided through the NCGA's website (NCGA.net). The state provides returns for statewide contests for the 2008 through 2014 general elections. These data are collected at the voter tabulation district (VTD) level (a Bureau of the Census term for a polling area such as a precinct) level; however, several VTDs in close proximity to military bases in North Carolina reported unusually high numbers of votes and contained unusually high numbers of residents. These extremely large VTDs caused problems for our development of a null set of neutral maps because districts containing extremely these large VTDs were liable to exceed reasonable levels of population parity. To circumvent this problem, we disaggregate the returns reported by the NCGA to census blocks. We achieve this by using the spatial join utility in the QGIS software package to determine into which VTD a census block falls (Quantum GIS Development Team 2016). We then assigned votes to a block according to the proportion of the VTD population that resides within the block. We then re-aggregate block level returns to the block groups.

[16]We use a neutral redistricting algorithm proposed by Daniel Magleby and Daniel Mosesson to draw a null set of maps of legislative districts for both North Carolina and Iowa (Magleby and Mosesson 2016). The null set we develop is partisan blind in that the maps that make up the distribution were drawn without reference to any factors besides geographic contiguity and population parity. The analysis uses a graph partitioning algorithm to randomly group geographic units (block groups in North Carolina and VTDs in Iowa). While maintaining district contiguity, it then uses a second algorithm to shift geographic units randomly between districts until all districts in a given plan have roughly equal populations. We repeat the process to draw 50,000 maps of North Carolina and Iowa's state senate districts. For the analysis presented here, we utilize the 25,000 maps with the lowest difference in population across districts. Among the maps included in our sample, the maximum population deviation is within ± 4.5%.

Table 1. Results of Applying Five Standards for Evaluating Whether North Carolina's Senate Districts Are a Gerrymander

Office	#1 Obs Dem 2-pty vote %	#2 Wasted votes Obs	#2 Wasted votes Exp	#3 District wins Obs	#3 District wins Exp	#4 Equal vote weight Obs	#4 Equal vote weight Exp	#5 Partisan symmetry Dem Seat Advantage	#6 3-prong test Prong 1	#6 3-prong test Prong 2
Governor	44.2	6.8	13.2 (2.9)	16	15.3 (1.40)	−5.8	−1.6 (.91)	−8.5	2.02	−.44 (−0.22)
Lt Gov	49.9	16.5	5.8 (3.0)	17	21.5 (1.44)	−5.7	−2.0 (.95)	−9.5	1.80	9.22 (5.02)
Auditor	53.7	14.8	−1.6 (2.8)	21	26.9 (1.41)	−5.2	−1.8 (.99)	−8.2	1.72	11.36 (5.72)
Agri Comm	46.8	10.2	12.5 (2.8)	17	16.9 (1.35)	−7.1	−2.8 (.90)	−10.0	1.95	3.25 (1.74)
Insur Comm	51.9	16.2	2.3 (2.9)	19	24.1 (1.40)	−6.4	−2.2 (.98)	−9.5	1.81	10.11 (5.15)
Labor Comm	46.7	11.7	11.7 (2.9)	16	17.3 (1.39)	−6.1	−2.5 (.76)	−9.2	2.09	4.31 (2.33)
Sec of State	53.8	13.3	−3.1 (2.8)	22	27.7 (1.40)	−4.7	−1.8 (.82)	−8.5	1.97	10.49 (4.76)
Supt Pubic Ed	54.2	10.0	−3.9 (2.7)	24	28.3 (1.36)	−4.7	−1.7 (.88)	−8.1	1.91	9.38 (4.09)
Treasurer	53.8	15.1	−1.2 (2.9)	21	26.8 (1.45)	−5.3	−2.1 (.96)	−8.7	1.99	8.48 (3.86)

#1 = Percentages are for the statewide two-party vote.
#2 = Wasted votes are the difference in Dem vs Rep votes cast for a losing candidate plus votes above 50% +1 as a percentage of total two-party votes—i.e., {(Dem wasted – Rep wasted) / Total two-party votes} * 100. Positive numbers indicate more Dems wasted more votes.
#3 = District wins are the number of districts carried by the Dem candidate, observed and expected, with expectations based on 25,000 computer-generated results. Numbers in parentheses are the standard deviation of expectations among the 25,000 neutral plans.
#4 = Equal vote weights record the difference between the median district two-party Dem percentage and the mean two-party district Dem percentage. Negative numbers indicate Dem disadvantage, with the magnitude indicating approximately the percentage points above 50 Dems would need to carry a majority of districts. The column of expected results is the median-mean difference attributable to residential patterns, with standard deviations in parentheses.
#5 = Partisan symmetry is the average difference in Dem–Rep expected number of seats won in a competitive range of vote percentage (40 to 60) if each party won the same vote percentage. Negative numbers indicate Dems are expected to win fewer seats with the same vote percentage as Reps.
#6 = Prong 1 of the three-prong test is the estimated seat-vote swing ratio—e.g., a 2.02 value means a vote gain of one point brings a seat gain of 2.02 points. Prong 2 is the difference between Dem and Rep vote percentages above 50% in districts won by Dems vs Reps. Negative numbers indicate Dems have more extreme lopsided winning percentages. Numbers in parentheses are *t*-test values; values above 1.68 are statistically significant at $p < .05$, one-tail.

Efficiency gap. Applying the efficiency gap calculations produces mixed results for detecting a gerrymander. Eight of the nine elections show wasted vote percentage magnitudes exceeding the suggested demarcation line of 8.0, with the gubernatorial election falling below that line. What is one to say of these results? Sometimes the North Carolina senate districts appear to be a gerrymander, but once in a while they don't. The conclusion depends on which election one looks to as evidence. Notice,

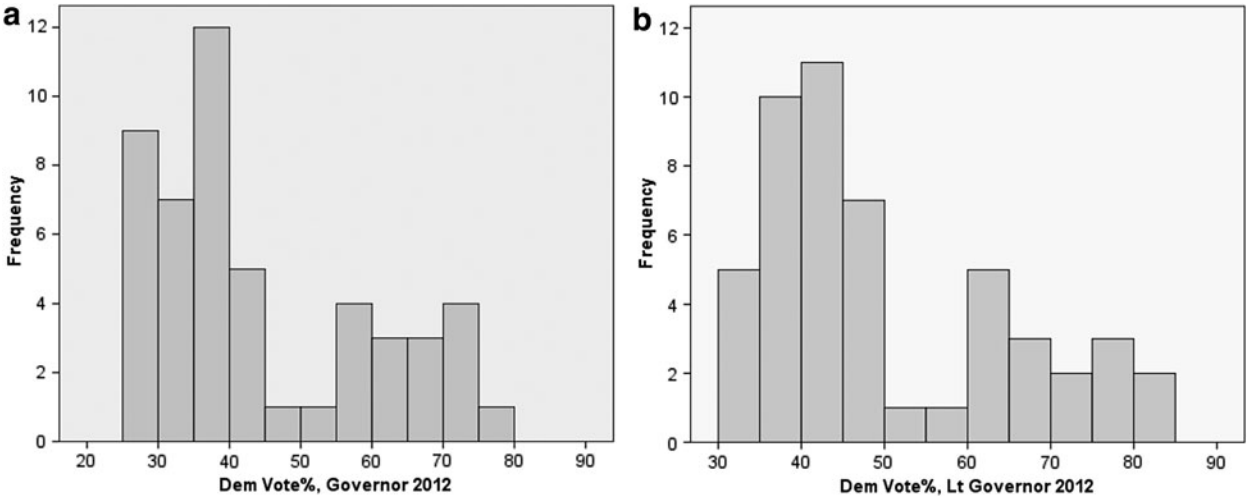

FIG. 1. Distribution of Democratic two-party vote percentages among North Carolina's state senate districts: 2012 governor and lieutenant governor elections. **(a)** *Left panel:* Dem Statewide % = 44.2; Dem Mean % = 44.4; Dem Median % = 38.6; Std. Dev. = 15.6; Dem Vote % > 50 = 16 of 50. **(b)** *Right panel:* Dem Statewide % = 49.9; Dem Mean % = 50.0; Dem Median % = 44.3; Std. Dev. = 15.0; Dem Vote % > 50 = 17 of 50.

also, the expected values rise and fall depending on the levels of the two-party vote. That is a serious problem because it tells us the magnitude of the wasted vote calculations depend on the vote percentage and not just whether the districts are gerrymandered. And notice that, despite being above the 8.0 threshold, two elections (commissioners of agriculture and labor) are not statistically distinguishable from expectations drawn for neutral plans.

What gives rise to the false negative reading from the gubernatorial election? The reason is directly related to the wasted vote requirement of a responsiveness ratio (aka, swing ratio) in the neighborhood of 2.0. When, as in North Carolina's gubernatorial election, Democrats win 44.2 percent of the vote, the wasted vote requirement for fairness is to have the Democrats winning 38.4 percent of the seats—i.e., the vote difference from 50 is 44.2 – 50 = −5.8. Two times that difference is −5.8 * 2 = −11.2, and an equal number of wasted votes would require that Democrats win 38.4 percent of the districts, since −11.6 + 50 = 38.4. Adding or subtracting the standard's requirement to be within eight points of the "fair" outcome implies that seat percentages in the range of 30.4 to 46.4 (38.4 ± 8) indicate no gerrymander effect. Given that a packing gerrymander might well be designed to grant Democrats some outcome in the vicinity of a third of the seats for a range of vote percentages, weak Democratic vote performances can fall within the safe-harbor range of the wasted vote standard. On the flip side, when Democrats receive something close to or exceeding 50 percent of the vote, a gerrymander effect becomes apparent, because seats are restricted to something such as 30 to 45 percent even when Democrats' votes approach or go above a majority. In short, the wasted vote standard can provide false negative readings in certain circumstances precisely because a gerrymander has been fashioned to allow one party to win a circumscribed minority number of districts unless and until it can win especially large vote majorities.

Comparing wins. The standard of counting the number of district wins suffers from the same shortcoming as the wasted vote standard. We see in Table 1 that in the three elections Democrats won with between 44 and 47 percent of the vote (governor, commissioner of agriculture, and commissioner of labor), they won close to the number of districts expected. When Democrats win votes in the vicinity of a majority or above, their shortfalls in seats are clear to see—just as when using the wasted vote standard. Put differently, when Democrats cast a minority of votes below 47, the safe seats granted to them by the gerrymander disguise the fact of the gerrymander. In short, comparing observed and expected district wins is subject to false negative readings under some circumstances.

Equal vote weights. This standard shows a consistent bias against Democrats. The median-mean differences run between 4.7 and 7.1 points adverse to Democrats, implying they would need something approaching 54.7 to 57.1 percent of the vote in order to carry a majority of districts—i.e., (50 + 4.7) to (50 + 7.1). Among the five elections when Democrats actually won a statewide vote majority, these various statewide candidates never carried a majority of the districts.[17] And, while the column of numbers on median-mean difference expectations shows Republicans have a natural 1.5- to 3.0-point advantage simply due to residential patterns, observed advantages attributable to gerrymandering fall far outside those expectations. Indeed, in none of the nine elections is the observed median-mean difference anywhere close to expectations. In the best-case circumstances, the secretary of state election, only 3 of 25,000 neutral maps (.012%, twelve-thousands of one percent) have a median-mean difference as large as the actual −4.7 value. In four elections, no expected value, among the 25,000 per election, is as large as the one observed. All indications from the equal vote weights standard indicate a rather harsh gerrymander favorable to Republicans, adverse to Democrats.

Partisan symmetry. As Justice Kennedy stated in *Veith*, the partisan symmetry standard runs into manageability problems because it relies on hypothetical estimates for the number of seats that would be won were one versus the other party to win the same vote percentage. We address the seat-denominated symmetry question in two ways, one more and one less factual. The facts from among our nine elections show that in the lieutenant governor's election the vote splits 49.9 to 50.1. Partisan symmetry would expect Democrats to win 24 or 25

[17]Turnout bias never exceeds 0.8 percent, and among the nine elections it averages 0.17 percent favoring Democrats.

seats for such an evenly split vote. They actually won only 17 districts. Furthermore, in three elections that Democrats won with 53.7 or 53.8 vote percentages (auditor, secretary of state, and treasurer), they won 21 or 22 seats. By way of contrast, in close to comparable circumstances, when Republicans won 53.2 or 53.3 percent of the vote (agriculture and labor commissioners), they won 33 or 34 seats. Clearly, large discrepancies in equal opportunities exist in the seat-vote relationship. Very similar resources (vote percentages) carry with them hugely different seat rewards. Through this more factual version of applying the seat-denominated symmetry standard we arrive at a clear indication of gerrymandering. Democrats win far fewer seats than Republicans when they win something close to the same vote percentages.

The less factual analysis takes a form more closely aligned with that described by Grofman and King (2007). We construct it through four steps: (1) accept as given the vote percentages and the number of districts won for each of our nine elections, (2) allow for hypothetical uniform vote swings so that they range from 40 and 60, (3) record the number of districts carried by Democrats at each of the 21 percentage points, and (4) compare the differences when both Democrats and Republicans won 40, 41, 42, … , 60 percent of the vote. The seat-denominated column in Table 1 records the results. On average, across the 21 percentage points, Democrats are at an eight- to nine-seat disadvantage despite, hypothetically, winning the same vote percentages as Republicans. Moreover, were we to restrict the comparisons to a vote range of 45 to 55, the Democrats' seat disadvantage runs, on average, between 13 and 15 districts. By this second form of analysis, too, the partisan standard indicates a substantial pro-Republican gerrymander.

Three prongs. Vote-denominated symmetry is the third prong in the proposed test. As discussed, by that prong we see an indication of a pro-Republican gerrymander.

Prong 1, the excess seats test, calls for calculating "whether the outcome … was disproportional relative to the seats/votes curve" by checking whether "the actual seats and the simulated number of seats" correspond beyond chance deviations (see Wang 2016, 1306). One method of checking is to revisit the district wins comparison in the null set test. That would tell us that in some elections district wins are in line with expectations but some are not. Another check is through a simulated seats/votes curve based on the simulation analysis we described for the less factual version of the partisan symmetry analysis but, here, by reporting the seat/vote slope value. Those results show seat/vote relationships between 1.7 and 2.1 (column 5 of Table 1). All results are within the range of one and three, which the standard supposes indicates no gerrymander (Wang 2016, 1286–89).

The reason for the sometime false negative readings from comparing actual and expected seat results is similar to the reasons we reported for the wasted votes and null set comparisons. The expectation ebbs and flows depending on the level of the vote, and when the disadvantaged party's votes are below 47, the districts the gerrymander grants to that party turn out to be about as expected in a non-gerrymandered plan. As the disadvantaged party votes rise to something approaching or beyond a majority, however, few additional districts are won. In fewer words, North Carolina created an effective packing gerrymander, and an associated consequence of packing gerrymanders is to reduce seat responsiveness toward proportional seat-to-vote results. The disadvantaged party wins its granted set of packed districts with relatively small statewide vote percentages, but as its vote percentages approach and go above 50, to say 54 or 55, the seats gains respond only modestly. All in all, therefore, we have to conclude the prong 1 test cannot be considered an effective standard by which to evaluate whether a packing gerrymander was enacted in North Carolina. It is prone to false negative readings because the standard it sets for a non-gerrymander is actually an outcome we expect a gerrymander to produce.

Prong 2 also runs into a problem, where again the problem is a failure to take account of how a gerrymander functions as vote percentages for the disadvantaged party vary between low versus high. It calls for a comparison of average vote percentages above 50 for districts won by Democrats compared to districts won by Republicans. To check whether the comparisons show systematic differences going beyond mere chance, prong 2 applies t-tests for the differences between two means. In contradiction of a pro-Republican gerrymander that North Carolina enacted, applying prong 2 to the Governor's election shows a difference slightly adverse to Republicans, not Democrats. The difference is not statistically

significant, and therefore the inference indicated from the gubernatorial election is that there is no gerrymander. Put differently, the prong 2 results tell us that sometimes the North Carolina senate districts appear to be a gerrymander, but sometimes they do not. The conclusion depends on which election is analyzed.

North Carolina Summary. North Carolina's senate districts were drawn for the purpose, in part, of providing Republicans with electoral advantage. Prong 1 of the three-prong standard misses that fact completely. The wasted vote, district wins, and prong 2 of the three-prong standard are not fully reliable indicators of that advantage. More often than not they indicate a Republican advantage, but depending on the size of statewide vote percentage they can, and in North Carolina do, give false negative readings. At the very least we have to conclude that indicators of gerrymandering that vary depending on how the vote splits are undesirable. More to the point, the false negatives exist because packing gerrymanders are intended to produce the seat outcome that the standards misidentify—i.e., packing gerrymanders grant the disadvantaged party some minority number of seats whether their vote percentage is small or substantial. The two symmetry standards, on the other hand, provide consistent indicators of North Carolina's designed partisan advantage. No false negatives appear. Thus, in application to North Carolina the symmetry standards are the dependable indicators, at least in the sense of avoiding false negatives.

Iowa

The Iowa Senate is a 50-member body elected to four-year terms from 50 single-member districts. Elections are staggered, with 25 members elected in presidential years and 25 elected in presidential midterms. Iowa's Legislative Service Agency (LSA) and its subordinate affiliated redistricting commission serve in an advisory capacity by presenting congressional and state legislative districts for the legislature's approval/disapproval, subject to veto by the governor.[18] The LSA is required to ignore partisan-related information of party registration, voting patterns, incumbency, candidate residences, and the like. The process has long drawn praise for its fair-mindedness (*Economist* 2002; Martin 2016).

Following the 2010 round of redistricting, the combined 2012 and 2014 senate elections saw the Democrats win 52 percent of the seats (26 of 50) with only 46.5 percent of the vote. As we noted in regard to North Carolina, however, the senate elections themselves do not offer especially probative evidence because the choices by candidates about whether and how to compete depend on where the lines are located. In Iowa, for instance, nearly one-third of all districts (16 of 50) went uncontested. Among the 34 districts contested by major-party candidates, Democrats cast 51.2 percent of the vote and won 20 districts. Thus, as with North Carolina, the more probative evidence is drawn from analyses of Iowa's statewide elections, here ten of them between 2008 and 2012.

As prelude, Figure 2 presents two vote percentage histograms: one for the secretary of state and the other for the treasurer, the two most competitive elections among our ten. The obvious fact apparent in both graphs is that Iowa has a large number of competitive districts. The numbers of districts in a competitive vote percentage range between 45 and 55 are 26 (secretary of state) and 27 (treasurer). Notice, also, a difference of just 4.4 vote points is associated with seat splits of 17 Democratic and 33 Republican versus 38 Democratic and 12 Republican. Small vote shifts apparently bring large district win rewards.

The numbers relevant to evaluating the five standards are reported in Table 2. Our various analyses track the same path as those reported and discussed for the North Carolina application.

Efficiency gap. The news about whether the wasted vote standard provides the correct reading of no gerrymander in Iowa is mixed. Nine of ten values exceed the suggested line of demarcation for distinguishing a gerrymander from a non-gerrymander, i.e., a value below −8 or above +8. If analysts rely on just one exogenous election to evaluate a gerrymandering allegation, they are likely to arrive at a false positive conclusion. If, however, two or more elections are investigated and each party wins a vote majority in at least one of the elections, it would be possible to see that the wasted votes rise and fall depending on whether a party receives a vote majority or minority. In Iowa, Democrats

[18]If disapproved, the Legislative Service Agency (LSA) is required to draw new maps. After three disapprovals, the legislature is allowed to draw new maps, but this has not occurred since implementation in the 1980 round of redistricting.

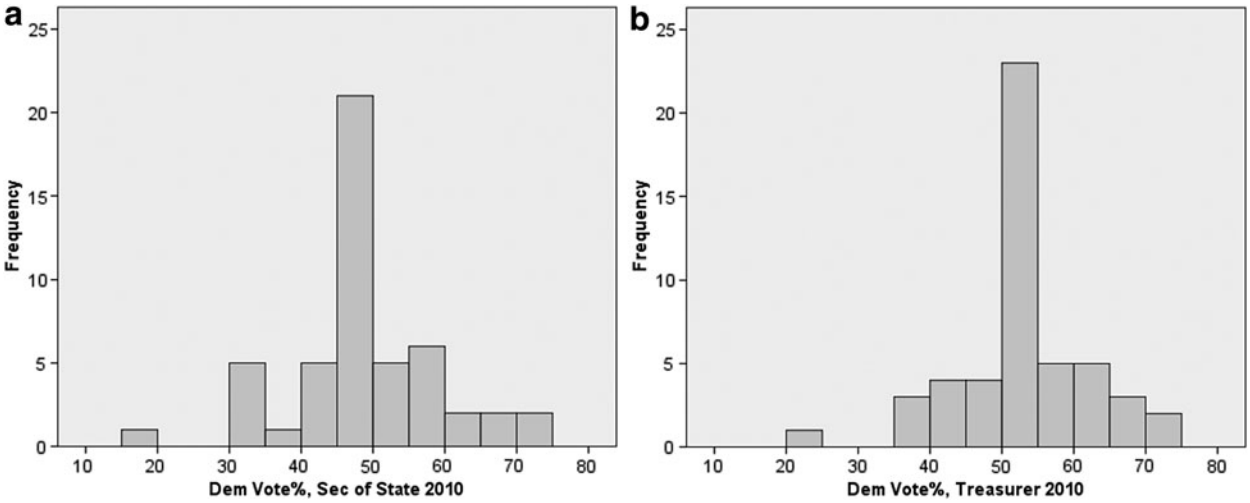

FIG. 2. Distribution of Democratic two-party vote percentages among Iowa's state senate districts: 2010 secretary of state and treasurer elections. **(a)** *Left panel:* Dem Statewide % = 48.5; Dem Mean % = 48.7; Dem Median % = 48.3; Std. Dev. = 10.0; Dem Vote % > 50 = 17 of 50. **(b)** *Right panel:* Dem Statewide % = 52.9; Dem Mean % = 53.0; Dem Median % = 52.8; Std. Dev. = 8.8; Dem Vote % > 50 = 38 of 50.

waste fewer votes than Republicans (indicated by the negative values in column 2) when they win a vote majority but waste more votes (positive values in column 2) when Republicans win a vote majority.

Comparing wins. Comparing actual district wins to expected wins from maps drawn using a neutral process comes close to getting to the right conclusion that Iowa's senate districts are not a gerrymander. The observed results are never too far

Table 2. Results of Applying 5 Standards for Evaluating Whether Iowa's Senate Districts Are a Gerrymander

	#1	#2 Wasted votes		#3 District wins		#4 Equal vote weight		#5 Partisan symmetry	#6 3-prong test	
Office	*Obs Dem 2-pty vote%*	*Obs*	*Exp*	*Obs*	*Exp*	*Obs*	*Exp*	*Dem Seat Disdvantage*	*#1*	*#2*
Pres 2012	53.0	–9.6	–8.6 (2.8)	33	32.4 (1.37)	.47	0.1 (.48)	.2	4.60	1.26 (0.71)
Pres 2008	54.8	–7.8	–12.7 (2.7)	34	36.4 (1.32)	.40	–0.3 (.50)	–.2	4.98	4.87 (2.75)
U.S. Senate 10	34.1	14.4	9.0 (1.2)	2	2.3 (0.63)	–.88	–1.14 (.55)	.2	4.82	–11.20 (–1.99)
U.S. Senate 08	62.7	–22.6	–24.6 (1.2)	49	48.6 (0.59)	.47	0.4 (.46)	0	5.91	2.00 (0.29)
Governor	45.0	17.1	15.9 (2.1)	12	12.6 (1.04)	.42	–0.5 (.44)	.2	4.63	–0.60 (–.29)
Sec of State	48.5	13.1	8.7 (3.2)	17	19.3 (1.60)	–.38	–0.3 (.43)	–.2	5.15	2.20 (1.07)
Treasurer	52.9	–20.8	–17.4 (3.1)	38	35.0 (1.53)	–.25	0.1 (.39)	–.9	5.50	–1.42 (–0.67)
Auditor	43.5	22.7	25.0 (2.5)	11	11.1 (1.14)	–.11	–0.1 (.61)	.7	4.36	–3.41 (–1.55)
Sec of Agri	37.1	15.7	15.0 (1.8)	5	5.00 (1.01)	–1.93	–1.6 (.63)	1.1	3.90	–9.39 (–2.57)
Atty Gen	55.6	–21.7	–18.7 (2.6)	41	39.5 (1.28)	–.11	0.2 (.42)	–.6	5.20	0.78 (0.33)

#1 = Percentages are for the statewide two-party vote.
#2 = Wasted votes are the difference in Dem vs Rep votes cast for a losing candidate plus votes above 50% +1 as a percentage of total two-party votes—i.e., {(Dem wasted – Rep wasted) / Total two-party votes} * 100. Positive/negative numbers indicate more Dems/Reps wasted more votes.
#3 = District wins are the number of districts carried by the Dem candidate, observed and expected, with expectations based on 25,000 computer generated results. Numbers in parentheses are the standard deviation of expectations among the 25,000 neutral plans.
#4 = Equal vote weights record the difference between the median district two-party Dem percentage and the mean two-party district Dem percentage. Negative numbers indicate Dem disadvantage, with the magnitude indicating approximately the percentage points above 50 Dems would need to carry a majority of districts. The column of expected results is the median-mean difference attributable to residential patterns, with standard deviations in parentheses.
#5 = Partisan symmetry is the average difference in Dem–Rep expected number of seats won in a competitive range of vote percentage (40 to 60) if each party won the same vote percentage. Negative numbers indicate Dems are expected to win fewer seats with the same vote percentage as Reps.
#6 = Prong 1 of the three-prong test is the estimated seat-vote swing ratio—e.g., a 4.60 value means a vote gain of one point brings a seat gain of 4.60 points. Prong 2 is the difference between Dem and Rep vote percentages above 50% in districts won by Dems vs Reps. Negative numbers indicate Dems have more extreme lopsided winning percentages. Numbers in parentheses are *t*-test values; values above 2.02 are statistically significant at $p < .05$, two-tails.

off expectations. For six of ten elections, the difference is just a fraction of one seat. The one hitch is that two elections are statistically significantly different from expectations (i.e., more than 1.65 standard deviations removed from expectations). Because the differences run in both partisan directions—once with Democrats carrying fewer than expected (treasurer) and once with Republicans carrying fewer (president 2008)—an evaluation of several elections could be used to demonstrate no systematic favoritism serving to advantage one but not the other party. So, even though the comparison of wins standard generally avoids false positives more often than not, the statistical significance consideration is a reminder that it is worthwhile to apply the standard to more than one exogenous election.

Equal vote weight. The equal vote weight standard (aka vote-denominated symmetry) reaches the correct conclusion of no Iowa gerrymander. The median-mean differences are small; they run in different directions (six negative versus four positive); and never is majority rule violated.[19] All this leaves the no gerrymander conclusion on secure footing.

Partisan symmetry. Seat-denominated symmetry involves a degree of ambiguity but essentially reaches the right conclusion. By the method that pairs comparable situations where Democrats and Republicans win the same vote percentage, four comparisons come close to filling the bill: (1) President 2008 vs Governor, (2) Attorney General vs Governor, (3) Treasurer vs Secretary of State, and (4) U.S. Senator vs Secretary of Agriculture. In order, respectively,

(1) D vote % 54.8 and R vote % 55.0 → D seats = 34 vs R seats = 38
(2) D vote % 55.6 and R vote % 55.0 → D seats = 41 vs R seats = 38
(3) D vote % 52.9 and R vote % 51.5 → D seats = 38 vs R seats = 33
(4) D vote % 62.7 and R vote % 62.9 → D seats = 49 vs R seats = 45

The results in any one election are three, four, or five seats off—hence the ambiguity—but one election shows a Republican advantage and the other three a Democratic advantage. In other words, there is no indication of a persistent partisan advantage running in one direction. Alternatively, applying the less factual, simulation analysis reported in Table 2's column 6 (see the details of how this approach works in our discussion of the North Carolina analysis, above), we see mostly fractional seat differences with none amounting to as many as two seats. On this evidence, seat-denominated symmetry indicates about as little of a gerrymandering seat effect as one might imagine in a fair set of districts, but with a touch of ambiguity.

Three prongs. The third prong of the three-prong test has already been covered as it repeats the calculation of the equal vote weight test. On that score, the test indicates no gerrymandering. One version of evaluating the first prong, from the standpoint of a party winning more or fewer seats than expected, also indicates there is no gerrymander inasmuch as that is what the district wins test indicates (i.e., from column 3). That follows, however, when the expectation is based on the null set. Compared to outcomes in other elections nationwide (Wang 2016, 1289–92), the rather large seat swings in response to vote shifts might very well lead to a different conclusion. As can be seen in the prong 1 column of the three-prong test, simulated seat-vote relationships have values above 3.90. All ten simulated slopes are beyond the test's zone of acceptability (Wang 2016, 1286). Taking all of these considerations on board makes it difficult to say what conclusion should be drawn from the prong 1 test.

Finally, prong 2 offers mixed readings. Two of ten differences in the lopsidedness of district-win percentages are statistically significant—viz., president 2008 and secretary of agriculture. On the one hand, because one significant result shows a Democratic win is too lopsided and the other shows a Republican win is too lopsided, one could conclude the lopsidedness shows no partisan favoritism and thus no gerrymandering. On the other hand, the results more generally show that comparing lopsidedness is not a reliable indicator of gerrymandering in any case. Large vote percentage outcomes for a party, as in Iowa's 2010 U.S. Senate and secretary of agriculture elections, can produce disparities in lopsidedness as the result of the vote percentages, not as a result of gerrymandering.

[19]As is true for North Carolina (fn. 17), turnout bias in Iowa does not amount to much. It favors Democrats in all ten elections but never exceeds 0.6 percent and averages just 0.22 percent.

Iowa summary. Iowa's senate districts are widely viewed as fair. All five standards could be made to confirm that they are. Three of the five arrive at that conclusion only as contingencies, however. By way of counting wasted votes in any one election, the results actually look like a gerrymander. The important fact revealed by this contingency is that counting wasted votes and checking whether they exceed the proposed threshold of ±8 is not anything close to a standard for identifying a gerrymander because wasted votes exceed the threshold for reasons other than gerrymandering. In Iowa they occur in nine of ten elections because many senate districts are highly competitive, something that is neither an ill in and of itself nor something that operates to the detriment of only one party. That same high degree district competitiveness hampers prong 1 of the three-prong approach, and prong 2 is subject to false positives simply when one party wins considerably more votes than the other. Comparing observed to expected wins fares better. It usually arrives at the right conclusion, though it is subject to possible false positive reading as in two of ten elections when the differences are not large but nevertheless statistically significant. Both the equal vote weight and partisan symmetry standards offer credible readings of Iowa's non-gerrymander. One finds no indication of a gerrymander from the equal vote weight standard and, at most, not so much a false positive reading as a degree of ambiguity from the partisan symmetry standard. In all, on questions of avoiding false positives, just as with avoiding false negatives, the two symmetry standards are the dependable indicators, one slightly more so (equal vote weight) and the other slightly less so (partisan symmetry).

DISCUSSION

What have we learned? The two symmetry standards hold the best prospects for identifying a packing gerrymander that dilutes the votes of one party's voters relative to the vote weight enjoyed by the other party's voters. Between the two, the equal vote weight standard is the more convincing as it more readily meets manageability and effectiveness considerations. Considered as matters of principle and checked against hypotheticals, the equal vote weight standard is faulted only for not being aggressive enough to cover the contingency that, while a districting plan is fair in the sense of not violating majority rule, it could miss the fact that one party can expect more seats when it wins a vote majority with X percent of the vote compared to when the other party wins the same X percent of the vote. This lack of aggression has to be balanced against the less manageable partisan symmetry standard, which relies on observed outcomes where the votes are mirror images—e.g., 45–55 and 55–45—or engages in hypothetical projections of what reasonably could be expected to result were votes to shift in some particular way. Also, as the Iowa application illustrates, the equal vote weight standard avoids a few of the modest ambiguities that arise when the partisan symmetry standard is applied.[20]

The three other standards leave much to be desired. Each suffers manageability problems: wasted votes for both its arguable counting procedure and its need to look externally to create a relative metric by which to say whether a gerrymander exists; comparing observed versus expected wins for its black box computer algorithms; and the three-prong test for its possible internal contradictions. All three also suffer effectiveness problems, each and all, in essence, because their results vary depending on the level of the vote each party receives. Their missing effectiveness is especially damning because it means these three approaches misapprehend a key feature of how packing gerrymanders work. Packing gerrymanders grant the disadvantaged party some number of seats that can look fair when that party wins a modest vote percentage but is clearly unfair when the same or similar limited number of seats is all it wins with vote totals approaching or exceeding a majority. The series of false negative readings in the North Carolina applications make this shortcoming ever so clear. To be sure, each of the three can be saved from full-scale rejection. When applied to the "right" mix of elections each can be argued to come to the right conclusion. At that juncture, however, there is nothing to be gained over applying the symmetry standards and

[20]In application, the choice does not need to be treated as a stark either/or. The equal vote choice is easier to manage and, in most cases, is highly likely to reach the same conclusion were one, instead, to apply the partisan symmetry standard. When and where circumstances warrant, a need for the greater aggressiveness of the partisan symmetry approach can be explained and the case for its broader notion of vote dilutions can be pressed.

something to be lost by doubts and arguments about just what is the "right" mix of elections.

CONCLUSION

The ballot box is the essential institution of any democracy, with more than a few thousand up through hundreds of millions of people coming together to exercise self-government. It is remarkable that centuries beyond the widespread recognition that gerrymandering can be and has been used to distort the self-governing process we are still struggling to find ways to identify and combat it. Our evaluation of five proposals for curbing packing gerrymanders reveals both the difficulties and possibilities.

Our focus has been on packing, as it is the most commonly alleged form. Its clear harm to democratic principles protected by the U.S. Constitution is unequal treatment of voters by implicitly assigning them different vote weights. Its contra-democratic systemic consequence is relegation of a popular majority to minority status. The three proposals of computing the efficiency gap, comparing wins, and applying a three-prong test encounter manageability problems. More damning, the three ask for evidence of gerrymandering that, when the specified evidence does not appear, can actually be absent because a gerrymander has been wrought—i.e., the false negative readings North Carolina's senate districts. Just as damning for two of the three proposals, not including comparing wins, is their asking for evidence that when it does appear it is for reasons other than gerrymandering—i.e., the false positive readings of Iowa's senate districts. The two symmetry-based standards, equal vote weights and partisan symmetry, are both more or less easily manageable—the equal vote weight test is the more manageable of the two. By argument and confrontation with evidence we have shown both to be effective at identifying when the placement of lines is the cause of diluting votes—here, again, with the equal vote weight standard providing more clarity—i.e., avoiding the arguable claims that could be focused on why a party did not win more seats at each and various level of its votes. On this review, it is clear that the equal vote weight symmetry standard offers the best prospects for redistricting authorities and courts to confront the perniciousness we know as packing partisan gerrymanders.

REFERENCES

Altman, Micah and Michael P. McDonald. 2011. "BARD: Better Automated Redistricting." *Journal of Statistical Software* 42: 2–36.

Brunell, Thomas L. 2012. "The One Person One Vote Standard in Redistricting: The Uses and Abuses of Population Deviations in Legislative Redistricting." *Case Western Reserve Law Review* 62: 1057–77.

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

Chen, Jowei and Jonathan Rodden. 2013b. "Report on Computer Simulation of Florida Congressional Districting Plans" (February 15). Available at Justin Levitt's *All about Redistricting* website, <http://redistricting.lls.edu/>.

Chen, Jowei and Jonathan Rodden. 2014. "Response to Expert Report of Stephen W. Hodge" (February 18). Available at Justin Levitt's *All about Redistricting* website, <http://redistricting.lls.edu/>.

Cirincione, Carmen, Thomas A. Darling, and Timothy O'Rourke. 2000. "Assessing South Carolina's 1990 Congressional Districting." *Political Geography* 19: 189–211.

Darling, Thomas A. 2013. "Report of Thomas A. Darling in *Romo v. Detzner*" (April 8). Available at Justin Levitt's *All about Redistricting* website, <http://redistricting.lls.edu/>.

Davis v. Bandemer. 1986. 478 U.S. 109.

Dickson v. Rucho. 2014. 766 S.E.2d 238, vacated and remanded, 575 U. S. ___, aff'd on remand, 781 S.E.2d 404 (2015), cert. pending, No. 16–24.

Dixon, Robert G. Jr. 1971. "The Court, the People, and 'One Man, One Vote.'" In Nelson W. Polsby (ed.). *Reapportionment in the 1970s*. Berkeley, CA: University of California Press, 7–55.

Economist. 2002. "In Praise of Iowa." *The Economist* (October 17).

Engstrom, Erik J. 2013. *Partisan Gerrymandering and the Construction of American Democracy*. Ann Arbor, MI: University of Michigan Press.

Engstrom, Richard L. and John K. Wildgen. 1977. "Pruning Thorns from the Thicket: An Empirical Test of the Existence of Racial Gerrymandering." *Legislative Studies Quarterly* 2: 465–79.

Fang, Lee. 2015. "Gerrymandering Rigged the 2014 Elections for Republican Advantage." *The Nation* (November 5) <www.the.nation.com/article/gerrymandering-rigged-2014-elections-republican-advantage/>.

Friedman, John N. and Richard T. Holden. "Optimal Gerrymandering: Sometimes Pack, but Never Crack." *American Economic Review* 98: 113–144.

Gelman, Andrew and Gary King. 1994. "A Unified Method of Evaluating Electoral Systems and Redistricting Plans." *American Journal of Political Science* 38: 514–54.

Grofman, Bernard N. and Gary King. 2007. "The Future of Partisan Symmetry as a Judicial Test for Partisan Gerrymandering after *LULAC v. Perry*." *Election Law Journal* 6: 2–35.

Growe v. Emison. 1993. 507 U.S. 25.

Gudgin, Graham and Peter J. Taylor. 1979. *Seat, Votes, and the Spatial Organisation of Elections*. London, UK: Pion.

Gul, Faruk and Wolfgang Pesendorfer. 2010. "Strategic Redistricting." *American Economic Review* 100: 1616–1641.

Hacker, Andrew. 1964. *Congressional Districting: The Issue of Equal Representation* (revised ed.). Washington, DC: Brookings.

Harris v. Arizona Independent Redistricting Commission. 2016. 578 U.S. ___.

Hebert, J. Gerald and Danielle Lang. 2015. "Memorandum of Amici Curiae Common Cause and Virginia New Majority Regarding Proposed Remedial Plans." U.S. District Court for the Eastern District of Virginia (3:13-cv-678).

Hodge, Stephen W. 2013. "Assessments of the Redistricting Simulations of Professors Chen and Rodden" (November). Available at Justin Levitt's *All about Redistricting* website, <http://redistricting.lls.edu/>.

King, Gary, Bernard Grofman, Andrew Gelman, and Jonathan Katz. 2005. "Amicus Brief in Jackson v. Perry Submitted on Behalf of Neither Party, in the Supreme Court." (No. 05–276).

Jackman, Simon. 2015. "Assessing the Current Wisconsin State Legislative Districting Plan" (July). Available at Justin Levitt's *All about Redistricting* website, <http://redistricting.lls.edu/>.

Johnson v. DeGrandy. 1994. 512 U.S. 997.

League of United Latin American Citizens v. Perry. 2006. 548 U.S. 399.

League of Women Voters of Florida v. Detzner. 2015. Supreme Court of Florida (SC14-1905).

Martin, David O. 2016. "Iowa's Redistricting Model Deserves Bipartisan Support." *Wisconsin State Journal* (May 7).

Magleby, Daniel B. and Daniel B. Mosesson. Forthcoming. "Neutral Redistricting Using a Multi-Level Weighted Graph Partitioning Algorithm." *Political Analysis*.

Mayhew, David 2011. *Partisan Balance: Why Political Parties Don't Kill the U.S. Constitutional System*. Princeton, NJ: Princeton University Press.

McCarty, Nolan. 2013. "Response to Chen and Rodden's 'Report on Computer Simulation of Florida Congressional Districting Plans'" (No date stated). Available at Justin Levitt's *All about Redistricting* website, <http://redistricting.lls.edu/>.

McCarty, Nolan. 2014. "Response to Jowei Chen and Jonathan Rodden's February, 2014 Report" (April 28). Available at Justin Levitt's *All about Redistricting* website, <http://redistricting.lls.edu/>.

McDonald, Michael D. and Robin E. Best. 2015. "Unfair Partisan Gerrymandering in Politics and Law: A Diagnostic Applied to Six Cases." *Election Law Journal* 14: 312–30.

McGann, Anthony J., Charles Anthony Smith, Michael Latner, and Alex J. Keena. 2015. "A Discernable and Manageable Standard for Partisan Gerrymandering." *Election Law Journal* 14: 295–311.

McGann, Anthony J., Charles Antony Smith, Michael Latner, and Alex J. Keena. 2016. *Gerrymandering in America: The House of Representatives, the Supreme Court, and the Future of Popular Sovereignty*. New York: Cambridge University Press.

McGhee, Eric. 2014. "Measuring Partisan Bias in Single-Member District Electoral Systems." *Legislative Studies Quarterly* 55: 68–69.

Nagel, Stuart. 1965. "Simplified Bipartisan Computer Redistricting." *Stanford Law Review* 17: 863–69.

Owen, Guillermo and Bernard Grofman. 1988. "Optimal Partisan Gerrymandering." *Political Geography Quarterly* 7(1): 5–22.

Page v. Virginia State Board of Elections. 2014. 3:13-cv-678 (E.D. Va).

Quantum GIS Development Team. 2016. *Quantum GIS Geographic Information System*. Open Source Geospatial Foundation Project. <http://qgis.osgeo.org>.

Romo v. Detzner 2014. Final Judgement, CA-412 (2nd Jud. Cir. Fla.).

Stephanopoulos, Nicholas and Eric McGhee. 2015. "Partisan Gerrymandering and the Efficiency Gap." *University of Chicago Law Review* 82: 831–900.

Stevens, John Paul. 2014. *Six Amendments: How and Why We Should Change the Constitution*. New York: NY: Little, Brown.

Thornburg v. Gingles. 1986. 478 U.S. 30.

Toobin, Jeffrey. 2003. "The Great Election Grab: When Does Gerrymandering Become a Threat to Democracy?" *New Yorker* (December 8): 63–80.

Vickrey, William. 1961. "On the Prevention of Gerrymandering." *Political Science Quarterly* 76: 105–10.

Voinovich v. Quilter. 1993. 507 U.S. 146.

Wang, Samuel S.-H. 2016. "Three Tests for Practical Evaluation of Partisan Gerrymandering." *Stanford Law Review* 68: 1263–1321.

Wang, Samuel S.-H. 2015. "Brief of Amicus Curiae Samuel S. Wang, Ph.D. in Support of Appellees." U.S. Supreme Court, *Harris v. Arizona Independent Redistricting Commission* (14-232).

Whitford v. Gill. 2016. 15-cv-421-bbc, 2016 U.S. Dist. LEXIS 160811 (W.D. Wis. November 21, 2016), jurisdiction postponed, 582 U.S. ___ (June 19, 2017) (No. 16–1161).

Veith v. Jubelirer. 2004. 541 U. S. 267.

Address correspondence to:
Michael D. McDonald
Department of Political Science
Binghamton University, SUNY
Bartle Library, G251
Binghamton, NY 13902-6000

E-mail: mdmcd@binghamton.edu