Three Practical Tests for Gerrymandering: Application to Maryland and Wisconsin

Samuel S.-H. Wang

ABSTRACT

Partisan gerrymandering arises when many single-district gerrymanders are combined to obtain an overall advantage. The Supreme Court has held that partisan gerrymandering is recognizable by its asymmetry: for a given distribution of popular votes, if the parties switch places in popular vote, the numbers of seats would change in an unequal fashion. However, the asymmetry standard is only a broad statement of principle, and no analytical method for assessing asymmetry has yet been held to be manageable. Recently I proposed (68 *Stanford Law Review* 1263) three statistical tests to reliably assess asymmetry in state-level districting schemes: (a) a discrepancy in winning vote margins between the two parties' seats; (b) undue reliable wins for the party in charge of redistricting, as measured by the mean-median difference in vote share, or by an unusually even distribution of votes across districts; and (c) unrepresentative distortion in the number of seats won based on expectations from nationwide district characteristics. These tests use district-level election outcomes, do not require the drawing of maps, and are accessible via nearly any desktop computer. Each test probes a facet of partisan asymmetry. The first two tests analyze intent using well-established, century-old statistical tests. Once intents are established, the effects of gerrymandering can be analyzed using the third test, which is calculated rapidly by computer simulation. The three tests show that two current cases, the Wisconsin State Assembly (*Whitford v. Nichol*) and the Maryland congressional delegation (*Shapiro v. McManus*), meet criteria for a partisan gerrymander. I propose that an intents-and-effects standard based on one or more of these tests is robust enough to mitigate the need to demonstrate predominant partisan intent. The three statistical standards offered here add to the judge's toolkit for rapidly and rigorously identifying the consequences of partisan redistricting.

Keywords: gerrymander, redistricting, Common Cause, First Amendment, *Vieth v. Jubelirer, LULAC v. Perry*

INTRODUCTION

THE TERM "GERRYMANDERING" DESCRIBES the act of drawing district lines to make an individual legislator's victory overwhelmingly likely, by virtue of creating a district with predictable voting patterns. Such a pattern contradicts the saying that "voters should choose their representatives, and not the other way around."[1] One special case of gerrymandering has attracted particular attention from the Supreme Court: that of a partisan gerrymander. In this sophisticated form of gerrymander, individual legislators of both political parties may benefit by gaining safe seats, but the overall effect is to give specific net advantage to one party. Partisan gerrymandering has been deemed justiciable since the 1986 ruling in *Davis v. Bandemer*,[2] in which Indiana Democrats asserted that they were

---

Samuel Wang is a faculty associate in the Program on Law and Public Affairs and a professor of molecular biology and neuroscience at the Princeton Neuroscience Institute at Princeton University in Princeton, New Jersey.

[1]Mitchell N. Berman, *Managing Gerrymandering*, 83 TEXAS L. REV. 781, 781 (2005).
[2]*Davis v. Bandemer*, 478 U.S. 109 (1986).

systematically disadvantaged by their state's legislative map.

In the thirty years since *Bandemer*, no manageable standard has been identified by the Supreme Court. The closest approach came with the *LULAC v. Perry*[3] case on mid-decade redistricting in Texas, in which a majority of the Court mentioned partisan asymmetry as a potentially applicable principle. In this guiding principle, suggested by political scientists Bernard Grofman and Gary King,[4] partisan symmetry is defined as a situation in which reversed positions in the popular vote lead to a reversed seat outcome. The absence of such symmetry would therefore define a partisan gerrymander. A remaining challenge is to translate this concept to a concrete standard for practical use.

Commonly, gerrymanders are diagnosed by analyzing specific districts. However, partisan gerrymandering emerges from patterns of districting, and examination of a single district does not clearly identify partisan asymmetry. Indeed, any given district may give an advantage to its own winner's party, to the opposing party, or to neither party, depending on the overall redistricting scheme. A partisan gerrymander can only be reliably diagnosed when considering a state's whole districting plan at once.

I recently developed a method for using patterns of election outcomes to detect partisan asymmetry.[5] I developed two analyses: one that detects intents, as evidenced by a pattern of district-level partisan outcomes that is unlikely to have arisen by chance, and therefore imply deliberate actions by those who drew the lines; and one that measures the effects of those actions, defined as the number of seats that exceed an appropriate range that would arise under partisan-symmetric principles.

My analysis of intents is based on mathematical tests that have been known for nearly a hundred years. Such tests are well established in the scientific community as a way of testing for differences between two groups of observations (in this case, groups of districts), or overall asymmetry (in this case, the pattern of advantages gained by two political parties). The tests are taught to undergraduates and are accessible to anyone with an introductory statistics textbook and a spreadsheet program. Judges may rapidly use these tests to analyze whether a pattern of election outcomes is likely to have arisen from partisan intent. This "analysis of intents" has the potential to place the initial diagnosis of partisan

gerrymandering under the control of judges, with expert testimony playing a role only after an initial determination has been made.

Once intent has been established, the question arises of effects: how many seats were gained by partisan gerrymandering? In my analysis of effects, I estimate the extent to which a party's elected number of seats exceeds an appropriate range that would arise under symmetric principles of districting. This measure overcomes the central difficulty that representation is not necessarily proportional to public support. The idea that representation should be proportional is intuitive but wrong and is violated in a system in which individual elections are winner-take-all.[6] A more sophisticated approach to quantifying the number of excess seats has relied on the detailed preparation of hypothetical maps[7] according to explicitly stated rules for how districts are drawn. However, such an approach may be criticized because it implicitly relies on the notion that specific standards for hypothetical districting represent an acceptable baseline for comparison. My calculation of effects takes the simplifying step of constructing a range of possibilities using national election results, without reference to specific geographic boundaries or districting rules.

In this article I consider two current federal gerrymandering cases: the Maryland congressional delegation (*Shapiro v. McManus*[8]) and Wisconsin State Assembly districts (*Whitford v. Nichol*[9]). The tests show that gerrymandering has created partisan distortions that are statistically highly significant. I will end by suggesting ways in which these tests

---

[3]*League of United Latin American Citizens v. Perry*, 548 U.S. 399 (2004).

[4]Bernard Grofman and Gary King, *The Future of Partisan Symmetry as a Judicial Test for Partisan Gerrymandering after LULAC v. Perry*, 6 ELECTION LAW JOURNAL 2 (2007).

[5]Samuel S.-H. Wang, *Three Tests for Practical Evaluation of Partisan Gerrymandering*, 68 STANFORD LAW REVIEW 1263 (2016).

[6]Edward R. Tufte, *The Relationship Between Votes and Seats in Two-Party Systems*, 67 AMERICAN POLITICAL SCIENCE REVIEW 540 (1973).

[7]Jowei Chen and Jonathan Rodden, *Cutting Through the Thicket: Redistricting Simulations and the Detection of Partisan Gerrymanders*, 14 ELECTION LAW JOURNAL 331 (2015); Jowei Chen and Jonathan Rodden, *Unintentional Gerrymandering: Political Geography and Electoral Bias in Legislatures*, 8 QUARTERLY JOURNAL OF POLITICAL SCIENCE 239, 248 (2013).

[8]*Shapiro v. McManus*, 136 S. Ct. 450 (2015).

[9]*Whitford v. Nichol*, No. 3:15-cv-00421 (W.D. Wis. 2015).

can be used to construct a manageable standard for use by courts and legislatures.

This article was awarded second place in Common Cause's First Amendment Gerrymandering Standard competition of 2016. Parts of this article are modified from a previous publication.[10]

CONSTITUTIONAL INJURIES IN A PARTISAN GERRYMANDER

When districting plans are challenged for partisan gerrymandering, litigants assert that voters have lost the ability to elect representatives that fairly reflect their views. Redistricting efforts are also said to confer specific advantage on one political party at the expense of another. In most partisan gerrymanders, the districting scheme results in the election of delegations that do not naturally reflect the overall preferences of the state's voters.

Partisan gerrymandering's unconstitutionality rests on two rationales: the Fourteenth Amendment's Equal Protection Clause and "one person, one vote" principle, and the First Amendment-based protection of speech and association.[11] The justiciability of partisan gerrymandering arises from a series of Supreme Court cases starting with *Davis v. Bandemer* and continuing with *Vieth v. Jubelirer*[12] and *LULAC v. Perry*. In 1986, the Supreme Court established justiciability in *Davis v. Bandemer*.[13] The Court did not find a partisan gerrymander in *Bandemer*, but they did lay out a cause for action based on a two-prong test: 1) intent—an established purpose to create a legislative districting map to disempower the voters for one party; and 2) effect—proof that an election based on the contested districting scheme led to a distorted outcome.[14]

An equal protection-based approach might suggest the possibility of taking a disparate-impact approach to partisan gerrymandering. The *Arlington Heights v. Metropolitan Housing*[15] housing discrimination case established a framework in which courts evaluate a number of factors to identify housing discrimination in the form of disparate impact and/or disparate treatment of groups of differing socioeconomic or racial characteristics. However, the Supreme Court has thus far not adopted standards resembling *Arlington Heights* criteria in the context of partisan gerrymandering. Indeed, the Court has developed an explicit distinction between racial and partisan gerrymandering, as seen in *Vieth v. Jubelirer*.

The *Vieth* case concerned whether Pennsylvania's congressional districts constituted a partisan gerrymander. In that case, five justices voted to dismiss the claim. Justice Antonin Scalia wrote a plurality opinion for four justices. He wrote that "to the extent that our racial gerrymandering cases represent a model of discernible and manageable standards, they provide no comfort here [in the partisan context]."[16] Justice Kennedy wrote a separate concurrence, and also declined to join Justice Stevens's opinion stating that Stevens "would apply the standard set forth in the *Shaw* [race] cases" in "evaluating a challenge to a specific district" on partisanship grounds.[17]

Instead of the *Shaw* standard, Justice Kennedy suggested a basis for determining partisan gerrymandering under the First Amendment. Unlike ethnicity or socioeconomic status, identification with a political party can be changed with little effort. In this respect, partisan identification can be regarded as an act of speech or free association, both of which are protected by the First Amendment. In *Vieth*, Justice Anthony Kennedy has noted that the First Amendment can be interpreted as a mandate for "not burdening or penalizing citizens because of their participation in the electoral process, their voting history, their association with a political party, or their expression of political views.[18] Under general First Amendment principles those burdens in other contexts are unconstitutional absent a compelling government interest."[19]

Partisan gerrymandering can chill a voter's freedom to choose her or his favored political party. In gerrymandered districts, the noncompetitive nature

---

[10]Wang, *supra* note 5.
[11]*Bandemer*, 478 U.S. at 122–123; *Vieth v. Jubelirer*, 541 U.S. 267, 314 (2003) (J. Kennedy, concurring in judgment; "penalizing citizens because of their participation in the electoral process, … their association with a political party, or their expression of political views," citing *Elrod v. Burns*, 427 U.S. 347 (1976) (plurality opinion)).
[12]*Vieth*, 541 U.S. 267 (2003).
[13]*Bandemer*, 478 U.S. at 110.
[14]*Bandemer*, 478 U.S. at 128 (upholding the district court's finding that the Bandemer plaintiffs were required to prove discriminatory intent and effect).
[15]*Village of Arlington Heights v. Metropolitan Housing Development Corp.*, 429 U.S. 252 (1977).
[16]*Vieth*, 541 U.S. at 286.
[17]*Id.* at 321.
[18]*Vieth*, 541 U.S. at 314 (J. Kennedy, concurring); *Elrod v. Burns*, 427 U.S. 347, 362 (1976).
[19]*Vieth*, 541 U.S. at 314 (J. Kennedy, concurring).

of the general election leaves the primary election as the only avenue for voters to affect their representation. Such a situation creates a powerful incentive to compel voters to join the dominant political party, even if that party's issue positions do not encompass his or her political views. Since a partisan gerrymander creates noncompetitive districts for both major parties, voters on both sides may potentially feel the chill.

The harms I have delineated above suggest two possibilities. First, packing voters into districts based on their partisan affiliation may constitute an infringement of public self-expression, or freedom of speech. Second, chilling of partisan choice may constitute an infringement of freedom of association. Together, these harms constitute a form of viewpoint discrimination. In this way, the purposeful creation of lopsided districts can be linked to First Amendment principles.

Justice Kennedy did not articulate an exact standard to evaluate partisanship under the First Amendment. Since *Bandemer*, a central difficulty has been establishing a manageable standard, i.e., one that provides a reliable and usable determination of whether an offense has occurred. In *Bandemer*, the justices described the effects prong in general terms. Justice White advocated an analysis of an entire districting plan: "A statewide challenge, by contrast, would involve an analysis of the voters' direct or indirect influence on the elections of the state legislature as a whole," while also acknowledging that this was "of necessity a difficult inquiry."[20] But eighteen years later in *Vieth*, the plurality opinion stated that no acceptable standard had been established in the intervening time, and therefore it was time to abandon the search.[21] The Court in *Vieth* was notably divided, culminating in five separate opinions.[22] In a separate concurrence, Justice Kennedy provided a fifth vote against invalidating the districts in Pennsylvania, but left the door open for future remedies in other cases if a clear standard could be established.[23] The dissenting four justices voted in favor of a finding of partisan gerrymandering and offered several possible standards, but none was backed by a majority of Justices.[24] *LULAC v. Perry* left this judicial stalemate unaltered, but it did contain various endorsements of the symmetry standard, spread across multiple opinions.[25]

In this article, I present three tests that address concerns expressed in the *Vieth* opinions of Justices Scalia and Kennedy, and which are rooted in the symmetry principle. My method has advantages offered by mathematical rigor previously absent from the Court's opinions on partisan gerrymandering. By translating principles that have emerged from constitutional jurisprudence into the language of classical statistics, these tests may plug a hole that has been left unfilled by the Court.

MATHEMATICAL METHODS CAN IDENTIFY STATE-LEVEL IMBALANCES

The most obvious harm from partisan gerrymandering is representational. Partisan gerrymandering creates a situation in which the same overall statewide vote share would lead to a very different level of representation for the redistricting party and its opposing target. For example, in the Pennsylvania congressional election of 2012, Democrats won only 5 out of 18 congressional House seats, despite winning slightly more than half of the statewide vote. Democratic winners were packed into districts where they won an average of 76 percent of the vote, while Republican winners won an average of 59 percent.[26] In other words, partisan gerrymandering creates representational asymmetry between the two major political parties.

---

[20]*Bandemer*, 478 U.S. at 143.
[21]*Vieth*, 541 U.S. at 279.
[22]*Vieth*, 541 U.S. 267, 271 (opinion. of J. Scalia, joined by C.J. Rehnquist, and O'Connor and Thomas, JJ.); *id.* at 306 (opinion of J. Kennedy, concurring in judgment); *id.* at 317 (opinion of J. Stevens, dissenting); *id.* at 343 (opinion of J. Souter, dissenting, joined by Ginsburg, J.); *id.* at 355 (opinion of J. Breyer, dissenting).
[23]*Vieth*, 541 U.S. at 306 ("I would not foreclose all possibility of judicial relief").
[24]*Id.*
[25]*LULAC*, 548 U.S. at 468 (n.9) (opn. of Stevens, J. P., joined by Breyer, S) ("a helpful [though certainly not talismanic] tool"). *LULAC*, 548 U.S. at 473 (n. 11) (opn. of Stevens, J.P.; asymmetry as one of eight criteria he would use for determining effects-based violations). *LULAC*, 548 U.S. at 466 (opn. of Stevens, J.) ("Plan 1374C [the challenged plan] is inconsistent with the symmetry standard"). *LULAC*, 548 U.S. at 483 (opn. of Souter, J.) ("do not rule out the utility of a criterion of symmetry"; "interest in exploring this notion is evident [on the Court]"). *LULAC*, 548 U.S. at 420 (opn. of Kennedy, J. joined by Justices Souter and Ginsburg) (indicating use as a standard based on election results, but not hypothetical future results).
[26]Sam Wang, *Let Math Save Our Democracy*, N.Y. TIMES, Dec. 5, 2015, at SR6, <http://www.nytimes.com/2015/12/06/opinion/sunday/let-math-save-our-democracy.html> (last visited Jan. 27, 2016).

However, anti-majoritarian outcomes do not by themselves constitute proof of deliberate distortion of electoral processes. Even if some imagined ideal of districting could maximize the likelihood of a majoritarian outcome, lack of congruence with this standard could still arise by chance and small variations in opinion. In 2012, if a few thousand voters in Arizona had cast their ballots for a Republican instead of a Democrat in the 1st or 2nd District, the delegation would have been, like the state's popular vote, majority Republican.[27] Thus anti-majoritarian outcomes are not always accurate indicators of partisan maneuvering. Furthermore, a simple majoritarian standard is incomplete because it only addresses the issue of whether seats or votes fall above or below a 50% threshold. For example, if a party receives 51% of the vote, receiving either 55% or 80% of the seats are both majoritarian outcomes, but the latter case might be viewed as an offense.

A statistical approach is needed to distinguish what degree of inequity is allowable. I will use natural variation and basic concepts of statistics to build three tests for state-level partisan gerrymandering. My approach allows the user to consider conceptual subtleties and at the same time obtain unambiguous judgments without need for elaborate computation using methods whose details have either not been widely adopted by political science researchers and/or found by courts not to be persuasive in the outcome. I hope that a more straightforward approach may meet with wide approval and serve as a universal tool to assess claims of partisan gerrymandering objectively. In this way, the approach described here may eventually serve as a core part of a court's analysis of partisan gerrymandering. This approach recalls Justice Kennedy's statement that "new technologies may produce new methods of analysis that make more evident the precise nature of the burdens gerrymanders impose on the representational rights of voters and parties. That would facilitate court efforts to identify and remedy the burdens, with judicial intervention limited by the derived standards."[28]

Analysis of intents: Voter packing by intentional gerrymandering and self-association

Here I present an analysis of intents, which provides a way to identify characteristic patterns of voting results that are highly unlikely to have arisen by nonpartisan means. Partisan redistricting procedure creates a characteristic lopsided pattern of election results that can be used to identify when packing is likely to have occurred.

State-level gerrymandering is more elaborate than single-district gerrymandering and relies on an elaborate strategy. First, map drawers cram voters likely to favor their opponents so that they are "packed" into a few throwaway districts where the other side will win lopsided victories.[29] Second, state-level gerrymanders have a distinctive feature: the remaining, more numerous districts are drawn with boundaries to yield more-narrowly won victories. For example, voters can be "cracked" so that a bloc of votes is split across districts to dilute their impact and prevent them from contributing to a majority in any one district.[30] In this process, the critical requirement is asymmetry: the opposing party's voters must be more tightly packed than one's own voters.[31] The net result is an increased likelihood of unrepresentative outcomes.

A "lopsided-margins test" to detect when the targeted party wins with unusually large margins. The success of a gerrymandering scheme depends on the ability of the redistricting party to create safe margins of victory for both parties, with larger margins for their opponents. This pattern of outcomes can be quantified by sorting the districts into two groups, by winning party. Each party's winning vote shares can then be compared by what is said to be "the most widely used statistical test of all time":[32] the t-test for comparing the averages of two groups of observations. In

[27]State of Arizona, canvass of election results, <http://apps.azsos .gov/election/2012/General/Canvass2012GE.pdf>, at 4–6.

[28]*Vieth*, 541 U.S. at 312–313 (J. Kennedy, concurring).

[29]Justin Levitt, *A Citizen's Guide to Redistricting*, BRENNAN CENTER FOR JUSTICE, 57 (2010) <http://www.brennancenter .org/sites/default/files/legacy/CGR%20Reprint%20Single%20 Page.pdf> (last visited Feb. 17, 2016), at 12–13.

[30]Levitt, *supra* note 29.

[31]Because members of both major parties get packed into districts in a partisan gerrymander, individual members of the opposing party may acquiesce or even be complicit in the process. *See*, *e.g.*, *LULAC*, 548 U.S. at 418 (noting "a number of line-drawing requests by Democratic state legislators were honored"). In other words, a single-district gerrymander can favor one party even as a partisan gerrymander favors the other party. For this reason, the use of intent as a standard for gerrymandering should distinguish between district-level and party-level motivations.

[32]Richard Lowry, *Chapter 11: t-Test for the Significance of the Difference between the Means of Two Independent Samples*, VASSARSTATS, <http://vassarstats.net/textbook/ch11pt1.html> (last visited Dec. 5, 2015).

this way, the difference between each party's winning margins is used to test for intensive packing of the opposing party's voters.

The mean-median difference as a measure of skewness.   In a partisan gerrymander, district outcomes are distributed to favor the redistricter's party, even though the average vote may not favor that party. This discrepancy can be tested using a simple statistic: the difference between the mean (i.e., average) and the median vote share[33] for contested[34] districts. The median serves as a measure of the overall behavior of a state's district-level elections. The goal of a gerrymander is to maximize the number of districts won, which occurs when the median outcome is more unfavorable to the opposing party than that party's share of the vote. The mean-median difference is therefore a simple measure of asymmetry or skewness, and when it is allowed to develop without partisan acts, it has well-defined mathematical properties.[35]

As an example of the calculation, consider the 2012 Pennsylvania congressional election. The Democratic two-party share of the total vote in all 18 districts was, in terms of percentages and sorted in ascending order:

34.4, 36, 37.1, 38.3, 40.3, 40.6, 41.5, 42.9, <u>*43.2, 43.4,*</u> *45.2, 45.2, 48.3,* 60.3, 69.1, 76.9, 84.9, 90.6.[36]

Races won by Republicans are indicated in *italics* and the two middle values are underlined. The median percentage is defined as the midpoint of the two middle values, 43.3%. The mean Democratic vote share is 51.0%. The difference between the median and the mean is 7.7%. This difference reflects the fact that counterintuitively, Republican vote shares were above average in considerably more than half of the districts: 72% (13 out of 18), to be exact.

In other words, Pennsylvania's Democratic voters were empowered as if they comprised 43.3% of voters, even though they actually comprised 51.0%. The difference, 7.7%, is the number of voters who were effectively disenfranchised. Since approximately 5,400,000 Pennsylvanians cast votes in the 2012 congressional election, redistricting achieved an effect equivalent to over 400,000 Democratic voters casting their ballots for Republicans. The probability is less than 1% that this difference arose by chance in a nonpartisan process.[37]

Analysis of effects: What is an appropriate range of seats for a given share of votes?

Distinguishing partisan distortion from Voting Rights Act Section 2 constraints.   Although partisan gerrymandering is considered justiciable, another practice that uses similar districting methods is permitted and even mandated under Section 2 of the Voting Rights Act: the establishment of districts in which an ethnic minority constitutes a majority of the district's inhabitants.[38] These "majority-minority" districts are constructed to ensure that the interests of identified subgroups are represented. When such minorities are much less than 50% of a state's population, they can end up on the losing side of every election. To counteract this risk, majority-minority districts are constructed to cluster groups with shared interests.[39]

---

[33]The mean-median difference has also been suggested by Robin E. Best and Michael D. McDonald, *Unfair Partisan Gerrymanders in Politics and Law: A Diagnostic Applied to Six Cases,* 14 ELECTION LAW JOURNAL 312 (2015). In the present paper I give mathematically rigorous confidence intervals on that statistic and describe the circumstances under which it is applicable.

[34]The presence of uncontested races reduces the value of the mean-minus-median statistic. In those cases, the partisan breakdown is not known with accuracy. Consider the example of a 20-district state where one district's election is uncontested. Assume that district's residents would have voted at a rate of 80% for their party, instead of the nominal 100%. If their district were drawn differently, the appropriate mean for comparison would be based on the 80% figure and shift the overall mean by 1%.

[35]The mean-median difference is a simple and old measure of "skewness," a statistical term for asymmetry. David P. Doane and Lori E. Seward, *Measuring Skewness: A Forgotten Statistic?*, 19 JOURNAL OF STATISTICS EDUCATION (2011), <http://www.amstat.org/publications/jse/v19n2/doane.pdf>; Karl Pearson, *Contributions to the Mathematical Theory of Evolution, II: Skew Variation in Homogeneous Material*, TRANSACTIONS OF THE ROYAL PHILOSOPHICAL SOCIETY, SERIES A, 186, 343–414 (1895); G. UDNY YULE AND MAURICE G. KENDALL, AN INTRODUCTION TO THE THEORY OF STATISTICS 162–163 (3d ed. 1950).

[36]KAREN L. HAAS, STATISTICS OF THE PRESIDENTIAL AND CONGRESSIONAL ELECTION OF NOVEMBER 6, 2012 (2013), <http://clerk.house.gov/member_info/electionInfo/2012election.pdf> (last downloaded Feb. 18, 2016).

[37]The level of statistical significance is calculated using Test 2 and Student's t-distribution. Richard Lowry, *Chapter 7: Tests of Statistical Significance: Three Overarching Concepts*, VASSARSTATS, <http://vassarstats.net/textbook/ch7pt1.html> (last visited Dec. 27, 2015).

[38]Stephen Ansolabehere and Maxwell Palmer, A Two Hundred-Year Statistical History of the Gerrymander, Presentation at the Congress and History Conference, Vanderbilt University (May 22–23, 2015).

[39]*How New York State's Approved Redistricting Lines Compare with Old District*s, REDISTRICTING AND YOU, <http://www.urbanresearchmaps.org/nyredistricting/map.html> (last visited Aug. 20, 2015).

This dual use of district-drawing methods opens the challenge of how to construct a criterion that identifies partisan gerrymandering as anomalous, but not single districts that are drawn to create ability-to-elect districts such as majority-minority districts. Such an analysis requires the evaluation of groups of districts at once. Existing doctrine may provide some guidance.

Among the standards for the proper establishment of majority-minority districts is the concept that majority-minority districts should comprise a fraction of all districts that does not exceed the proportion of the minority population.[40] In U.S. court precedent, the "no-more-than-proportional" concept contributes to "*Gingles* criteria" for evaluating districting schemes.[41] Where minority representation is concerned, *Gingles* criteria identify rough proportionality as a relevant factor in evaluating the fairness of a districting plan. Under that standard, the Court has held

> that no violation of § 2 can be found here, where, in spite of continuing discrimination and racial bloc voting, minority voters form effective voting majorities in a number of districts roughly proportional to the minority voters' respective shares in the voting-age population. While such proportionality is not dispositive in a challenge to single-member districting, it is a relevant fact in the totality of circumstances to be analyzed when determining whether members of a minority group have "less opportunity than other members of the electorate to participate in the political process and to elect representatives of their choice."[42]

For example, if a minority group with 20% of a state's eligible population is able to elect representatives in 20% of a state's districts, this argues against violation of *Gingles* criteria.[43]

The idea underlying the *Gingles* criteria can be used to address the question of appropriate representation by political parties. I suggest that a redistricting plan is acceptable if it moves the seats-to-votes outcome toward partisan proportionality (eu-proportionality) as measured by prevailing national standards and unacceptable if it moves the outcome away from proportionality (dysproportionality) beyond the zone of chance. This standard can be understood at a glance using a plot (Figure 1) that I term a "representation plot," or alter-
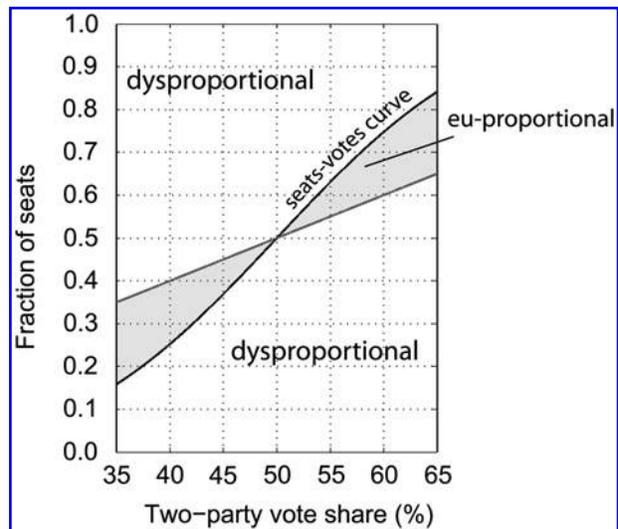

FIG. 1. A representation plot for classifying redistricting schemes. The seats/votes curve indicates the average relationship between seats won (*vertical* axis) and the popular vote share (*horizontal* axis), calculated by creating hypothetical delegations using 2012 House district election results. The *diagonal straight line* indicates proportional representation. Redistricting schemes that fall in the *shaded* zone between the *curve* and the *line* are termed eu-proportional; other outcomes are termed dysproportional. For clarity, the zone of chance (see text) is not shown.

natively a "bowtie plot," where eu-proportional outcomes are "inside the bowtie." Since dysproportional outcomes are a major result of partisan gerrymandering, a standard should distinguish between eu-proportionality and dys-proportionality.[44]

I note that the eu-proportionality concept specifically does *not* imply the establishment of proportional representation, a rule that is not to be found in the Constitution or in U.S. districting law and

---

[40]Section 2 of the Voting Rights Act of 1965, 79 Stat. 437, as amended, 42 U.S.C. § 1973.

[41]*Thornburg v. Gingles*, 478 U.S. 30 (1986); *Johnson v. De Grandy*, 512 U.S. 997 (1994).

[42]*Id.* at 1000 (finding no violation of § 2 of the Voting Rights Act of 1965, 79 Stat. 437, as amended, 42 U.S.C. § 1973).

[43]*Thornburg v. Gingles*, 478 U.S., 74–77 (describing near-proportional legislative representation of black voters as evidence of their ability to elect their preferred representatives).

[44]In this plot, the gray line indicates proportionality and is a straight line drawn from zero vote share and zero seat fraction to 100% vote share and 100% seat fraction. The seats/votes curve is calculated by resampling to build "fantasy delegations" (see the main text) and is approximated by the mathematical function that is the area under a bell-shaped curve whose average is 50% vote share and whose standard deviation is 14% vote share.

that does not arise in a single-member district system. Single-member districts usually generate outcomes in which a majority party's share of seats tends to exceed its proportion of popular support.[45] Instead, the eu-proportionality concept relies on the idea that some deviations from an average seats-to-votes relationship are beneficial for representation, whereas other deviations are detrimental. Good districting seeks to establish "fair and effective representation for all citizens."[46] The concept that deviations toward proportionality are good encompasses a wide range of concepts that includes (a) establishing appropriate levels of representation for minority groups (viz., *Gingles* criteria); (b) allowing the possibility that like a racial group, a political party with considerably less than 50% support might permissibly have enhanced representation relative to what would be predicted from national seats/votes relationships, but that reduced representation is impermissible; and (c) setting reasonable limits to how much enhancement from (b) is allowed. In this way, the Platonic ideal of proportionality does not set a specific goal but instead defines a direction of acceptable deviation. It is simple to state, it is flexible, and it contains many permissible outcomes.

Defining the zone of chance.   In addition to defining desirable and undesirable directions, a standard for partisan gerrymandering requires a method for determining whether a change could have arisen as part of normal variation in districting as practiced across the United States. In the three tests proposed here, I use the rules of probability to (a) describe that variation, (b) establish what the range of possible outcomes is, and (c) formulate a rule for identifying situations in which a state's new districting scheme has departed sufficiently from normal practice.

Faulty bright-line standards such as a majoritarian standard can be repaired by identifying a "zone of chance,"[47] which I define as the range of outcomes that could have arisen, without deliberate planning, from variations in how districts are drawn.[48] I will calculate zones of chance for (a) the pattern of voting outcomes across districts (Tests 1 and 2) and (b) the number of seats won in an election for any given statewide division of popular vote (Test 3).

The zone-of-chance approach recalls Justice Kennedy's statement that "new technologies may produce new methods of analysis that make more evident the precise nature of the burdens gerrymanders impose on the representational rights of voters and parties. That would facilitate court efforts to identify and remedy the burdens, with judicial intervention limited by the derived standards."[49] At the same time, I will also take advantage of longstanding statistical tests whose history assures their mathematical rigor. The use of statistical tests also allows judges to evaluate evidence more directly, with less need for assistance from external experts.

To understand the zone-of-chance concept, it is helpful to start by considering a case that is mathematically simple and does not require computer simulation: equally matched parties. I will focus on representation, i.e., the effects of redistricting.

As pointed out in the plurality opinion in *Vieth v. Jubelirer*, any districting scheme contains the possibility that a majority of votes will, by chance, lead to a minority of seats. To explore this concern, it is informative to calculate the exact probability that such a deviation could occur in the absence of intentional partisan districting. The calculation is simplest when the two-party popular-vote share (defined as the fraction of the top two parties' popular vote won by one party) is close to 50% for each party. In this circumstance, party A's seat-share for a random partitioning of N districts is on average N/2, and the probability of party A winning a particular district is 0.5. The actual number of districts won will vary, in the same way that a series of coin tosses are not guaranteed to yield equal numbers of heads and tails. The outcome will be within one standard deviation of the average about two-thirds of the time, and outcomes within this range would be fairly

---

[45]Proportional representation is achieved only in systems where it is enforced specifically and directly. For example, in Israel, members of the national legislative body, the Knesset, are assigned so that the number of a party's seats is proportional to the fraction of its popular vote. (Article 4 of the Basic Law: The Knesset.) Such a system embodies a legislature-centered form of the "one man, one vote" principle: each citizen's party preference is reflected proportionally at the national level.
[46]*Reynolds v. Sims*, 377 U. S. 533, 565 (1964).
[47]Sam Wang, *Let Math Save Our Democracy*, *supra* note 26.
[48]The zone of chance concept is a way to express the concept of significance testing in statistics. Statisticians calculate how far a measurement, such as the number of seats won by a party in a given election, is likely to stray from the expected average. In this article, I define the zone of chance as a region within which chance outcomes would fall 95% of the time, and outside the region 5% of the time. Statistics texts refer to this as a "p < 0.05" or "$\alpha < 0.05$" standard. *See* Lowry, *Chapter 7*, *supra* note 37. *See also* Wang, *supra* note 5.
[49]*Vieth*, 541 U.S. at 312–313 (J. Kennedy, concurring).

unsurprising.[50] And if the vote share is almost exactly 50%, then outcomes will give a majority to the other side close to half of the time.

To generalize the zone-of-chance calculation, I will use computer simulation. I will use existing districts in the year under examination as a source of information about how vote totals in districts may vary. The inputs to the calculation are the congressional vote totals for the state under examination and national district-by-district congressional results from the same year. This process escapes the burden of drawing boundaries, which requires the researcher to apply her or his standards about "good districting." This calculation will yield both a general seats/votes relationship and a statistical confidence interval (a.k.a. zone of chance) for the range of outcomes that could be expected in the absence of directed partisan intent. The zone of chance provides an answer to the question of whether a set of election outcomes has deviated sharply from national standards.

National districting patterns can be used to identify a natural seats/votes relationship. Computer simulations can be used to ask a simple question: if a given state's popular House vote were split into differently composed districts carved from the same statewide voting population, what would its congressional delegation look like? The answer allows the definition of a range of seat outcomes that would arise naturally from districting standards that are extant at the time of the election in question.

It is possible to calculate each state's appropriate seat breakdown—in other words, how a congressional delegation would be constituted if its districts were not contorted to protect a political party or an incumbent. This is done by randomly selecting combinations of districts from around the United States that add up to the same statewide vote total for each party. Like a fantasy baseball team, a delegation put together this way is not constrained by the limits of geography. On a computer, it is possible to create millions of such unbiased delegations in short order. In this way, one can ask[51] what would happen if a state had districts whose distribution of voting populations was typical of the pattern found in rest of the nation. Because this approach uses existing districts, it uses as a baseline the asymmetries that are present nationwide.[52] Indeed, the average result of these simulations approximates a "natural" seats/votes relationship that can be defined with mathematical rigor and exactitude. In short, these simulations de-

tect distortions in representativeness in one state, relative to the rest of the nation.

Using a standard ThinkPad X1 Carbon laptop computer equipped with the mathematical program MATLAB, simulation code[53] can perform one million simulations for a state in less than 20 seconds. Figure 2 shows 1,000 such "simulated delegations" for the state of Pennsylvania, along with the actual outcome. The thick curve defines a mathematically expected average seats/votes relationship.

I will develop an analysis of intents test that uses the zone-of-chance concept. The standard deviation,

---

[50]For example, if all N races are perfect toss-ups, then they behave like coin tosses, and according to the laws of probability the standard deviation of the outcome, a measure of variation often referred to as "sigma," or σ, is $0.5 * \sqrt{N}$. Thus if political parties A and B compete in a state that is composed of 16 congressional districts, all of which are closely contested, then each party can expect to get eight seats on average. Sigma for the specific case of all-close-races is $0.5 * \sqrt{16} = 2$ seats, suggesting that each party would typically get 6 to 10 seats. It must be noted that the foregoing formula for sigma is a substantial overestimate of real-life situations, because districting generates a mixture of more and less closely contested districts, and only close contests contribute to uncertainty. To estimate the true value of sigma, which is typically smaller, a more sophisticated approach is required, as detailed in Wang, *supra* note 5, in the section titled "National districting patterns can be used to identify a natural seats/votes relationship."

[51]This can be done by using all 435 House race outcomes. For a state X with N districts, calculate the total popular vote across all N districts. Now pick N races from around the country at random and add up their vote totals. If their vote total matches X's actual popular vote within 0.5%, score it as a comparable simulation. *See* Sam Wang, *The Great Gerrymander of 2012*, N.Y. TIMES, Feb. 2, 2013, at SR1.

[52]It is possible to explore the properties of this simulation procedure by giving it a variety of hypothetical nationwide distributions of districts as starting data. These hypothetical scenarios reveal that the "fantasy delegation" procedure has important features that are required of a descriptor of partisan asymmetry. First, for a symmetric distribution of congressional districts, i.e., a scenario in which Democrat-dominated districts are no more packed than Republican-dominated districts, fantasy delegations are typically majoritarian, awarding more representatives to the party that receives more votes. Second, the fantasy delegations have the same natural variation in partisan composition as the nationwide distribution, as measured by standard deviation. Third, when the nationwide distribution of districts has asymmetry, for instance containing a number of districts that are very packed with one party (as is the case in real life for Democrats), the fantasy delegations show a bias toward the other party, a phenomenon that is well analyzed (reviewed in Jowei Chen and Jonathan Rodden, *Unintentional Gerrymandering: Political Geography and Electoral Bias in Legislatures*, 8 QUARTERLY JOURNAL OF POLITICAL SCIENCE 239, 248 [2013]).

[53]The MATLAB software is available at GITHUB, <https://www.nytimes.com/2015/12/06/opinion/sunday/let-math-save-our-democracy.html> (last visited Aug. 24, 2015).
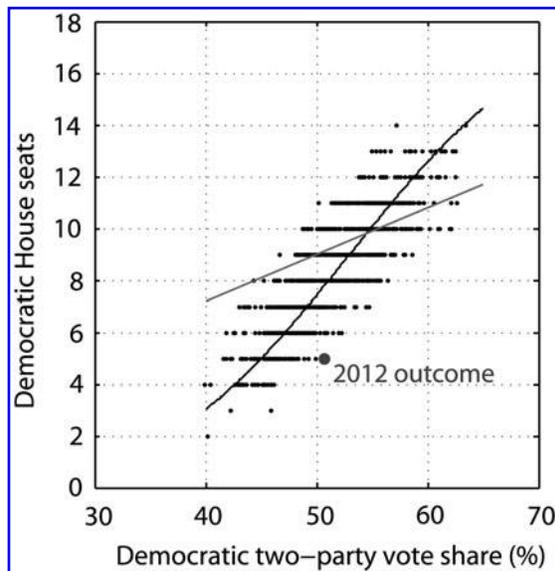
FIG. 2. Simulated Pennsylvania House delegations. Each *point* indicates one hypothetical delegation composed of 18 House districts drawn at random from the national House election of 2012. One thousand simulations are shown. The *thick curved line* indicates the average seats/votes curve and the *thinner diagonal line* indicates proportionality, both as defined in Fig. 1. The labeled larger *point* indicates the actual outcome, which falls in a zone of dys-proportionality, "outside the bowtie."

sigma, will be used as a yardstick of deviations from the average expected outcome. The general idea is that an average outcome only reflects one point in a range of outcomes, and the standard deviation (often referred to as sigma, or σ) is necessary to define a zone of chance. A difference would then be expressed as Delta, defined as the difference divided by sigma. Generally speaking, for a bell-like curve, which these simulations approximately follow, a difference of 1.6 standard deviations or more (Delta ≥ 1.6) occurs by chance in 5% of cases. Five percent is a common threshold for determining statistical significance.[54] In this way, the standard deviation is a handy and universal reference measure for detecting extreme outcomes, and it applies to all the analyses and tests in this article.

THREE QUANTITATIVE TESTS OF INTENTS AND EFFECTS IN PARTISAN GERRYMANDERING

Converting the analyses to practical tests

I will now use the analyses of intents and effects to propose three tests. I use the analysis of intents,

which identifies narrow-but-reliable wins as a hallmark of gerrymandering, to construct two tests: Test 1, the lopsided outcomes test; and Test 2, a reliable-wins test. I use the analysis of effects, which is based on numerical simulation of seat outcomes, to construct Test 3, the excess seats test.

Test 1 (the lopsided outcomes test). Compare the difference between the share of Democratic votes in the districts that Democrats win and the share of Republican votes in the districts that Republicans win. This test works because in a partisan gerrymander, the targeted party wins lopsided victories in a small number of districts, while the gerrymandering party's wins are engineered to be relatively narrow. To compare the winning vote shares for the two parties, use a grouped t-test, an extremely common statistical test.

Test 2 (the reliable-wins test). Systematic rigging of total statewide outcomes occurs by the construction of districts that offer secure wins for the party in control of the map. These wins would be wide enough to guarantee victory but not so wide as to waste votes that could be used to shore up other districts. How this intent is detected depends on whether the state's partisan vote is closely divided or whether one party is dominant. In a closely divided state, reliable wins occur when the average and median vote differ from one another. In a state that is dominated by one party, reliable wins occur when that party's strength is spread highly evenly across districts.

In a closely divided state. Calculate the difference between a party's statewide average district vote share on the one hand and the median vote share it receives on the other. In this situation a systematic gerrymander can be detected when a party's median vote share is substantially below its average vote share across districts.[55] For this test, calculate Delta by dividing the mean-median difference by σ, which is defined as 0.756 * (standard deviation

---

[54] A difference of Delta = 1 or more in a dysproportional direction occurs in approximately 16% of cases. A difference of Delta = 2 or more occurs in approximately 2.3% of cases. A difference of Delta = 3 or more occurs in approximately 0.13% of cases. These values are for Analysis #1, which uses a bell-shaped curve. Analyses #2 and #3 use the t-distribution, which gives slightly different values.

[55] This is the mean-median test described in Wang, *supra* 5 and Wang, *supra* 26, and by Best and McDonald, *supra* note 33.

of vote share across all N congressional districts in a state)/$\sqrt{N}$.[56]

In a state where the redistricting party is dominant. Calculate the standard deviation of the redistricting party's vote share in the districts that it wins. Calculate the standard deviation of the party's vote share in the districts that it wins nationwide. Compare these two standard deviations using a well-established testing tool, the chi-square test for comparison of variances,[57] to define zones of chance.

Test 3 (the excess seats test). Calculate whether the outcome of an election after redistricting was dysproportional relative to a simulated seats/votes curve and whether that outcome favors the redistricting party. For a state containing N districts, calculate the difference between the actual seats and the simulated expected number and divide by the standard deviation to obtain Delta.

Tests 1 and 2 determine whether the pattern of data could have arisen by chance; if not, this indicates an intent to gerrymander. A residual possibility exists of a false-positive result, i.e., identifying that a gerrymandering event occurred when in fact it did not. To reduce the possibility of such a false alarm, partisan gerrymandering could be assessed by evaluating both Test 1 and Test 2. Finally, Test 3 evaluates whether a party gained a significant advantage in terms of seats, and calculates the size of the effect.

Advantages and disadvantages of the three tests

The tests proposed here have several advantages. First, the tests do not require the detailed drawing of maps. Second, because they are derived from election results only, the tests can be applied independently from evaluating the details of the districting process. Third, because the results of the tests are highly correlated with one another, in situations where one test is unsuitable, another can be used instead. In this way the tests can be used separately or combined to reduce the risk of falsely identifying a gerrymander where none occurred. Conversely, the use of multiple tests also reduces the risk of failing to detect a gerrymander where one did occur. Finally, because the three tests do not use geography, they can easily be combined with other standards which may require circuitous geographic boundaries, such as state-mandated requirements,[58] Section 2 of the Voting Rights Act, and other precedents that exist in federal law.

Before the judge (or other evaluator of a districting plan) chooses which test to apply, he or she should take the following advantages and disadvantages into account.

Test 1 has the advantage of simplicity: it can be worked out using a spreadsheet program such as Microsoft Excel that can perform a two-sample t-test. If such a program is not available, it can be done using a hand calculator and a table of statistical values. It directly tests for noncompetitive races, a mainstay of gerrymandering. It identifies partisan asymmetry, though not bipartisan gerrymanders in which individual candidates of both parties benefit. Test 1 has the disadvantage that it can only be used if both parties win at least two seats each, since this is required to calculate standard deviations, a necessary step of the test.

Test 2 measures the reliability of wins for the redistricting party. Like Test 1, it is simple to calculate. Test 2 can always be done, since it is calculated using most or all of a state's district-level results. In the case of the mean-median difference, it does not rely on any data from other states and is therefore self-contained. In the case of the chi-square test, national data must be used to provide a standard for comparison.

Test 3 quantifies effects. Its most powerful use is to obtain an exact range for the appropriate number of seats for a given vote share. It addresses whether

[56]Paul Cabilio and Joe Masaro, *A Simple Test of Symmetry About an Unknown Median*, 24 CANADIAN JOURNAL OF STATISTICS/LA REVUE CANADIENNE DE STATISTIQUE 349, 352 (1996); Tian Zheng and Joseph L. Gastwirth, *On Bootstrap Tests of Symmetry about an Unknown Median*, 8 JOURNAL OF DATA SCIENCE 397, 400–401 (2010).

[57]Karl Pearson, *On the Criterion that a Given System of Deviations from the Probable in the Case of a Correlated System of Variables Is Such That It Can Be Reasonably Supposed to Have Arisen from Random Sampling*, 50(302) PHILOSOPHICAL MAGAZINE SERIES 5, 157–175 (1900); GEORGE W. SNEDECOR AND WILLIAM G. COCHRAN, STATISTICAL METHODS (8th ed. 1989).

[58]The three tests proposed here address the overall apportionment plan but do not cover the case of individual self-dealing in single districts. Local laws may provide additional constraints. For example, the current congressional districts in Florida do not violate the three tests presented here. Nonetheless, the Florida Supreme Court has found the map to violate the Florida Constitution redistricting provisions (article III, section 20(a) that reads, "No apportionment plan or district shall be drawn with the intent to favor or disfavor a political party or an incumbent"). *League of Women Voters of Florida v. Detzner*, 2015 WL 4130852 (Fla. 2015). This stricter standard extends a mandate for competitive races to the level of single districts.

a redistricting scheme leads to an elected delegation that deviates from national districting norms. Test 3 can always be calculated for any set of election returns. Because it uses data from other states, it has the advantage of taking into account the overall nationwide demographic character of districts. Therefore it has the virtue of measuring effects that go beyond the natural effects of population clustering. However, because it requires computer simulation, it requires the use of a computer program, a version of which can be accessed at http://gerrymander.princeton.edu, or obtained separately by contacting the author.

Three examples: the original Gerry-mander, Maryland congressional districts, and Wisconsin State Assembly districts

To examine the general applicability of these tests, let us consider three examples: (1) the original Gerry-mander of 1812, (2) post-2010 Maryland congressional districts, which the Supreme Court recently remanded for consideration by a three-judge court,[59] and (3) post-2010 Wisconsin State Assembly districts, which are currently under review in the Western District of Wisconsin.

Example 1: The original "Gerry-mander," the Massachusetts State Senate election of 1812. For Test 1, the Federalists won five races (which accounted for 11 districts); in these races, their two-party vote share averaged 55.6%, with a standard deviation of 4.6%. The Democratic-Republicans won 13 races (which accounted for 29 districts), with an average vote share of 70.7% and a standard deviation of 5.3%. The resulting Delta (for a t-test, also called a "t-score") is 5.5, and therefore Test 1 is met to a standard of 5.5 sigma. This is an unusually high level of significance and is reached by chance 0.0025% of the time.

Test 2 cannot be used because districts are not equal in size. In 1812 the number of votes per legislator ranged from Dukes/Nantucket (1,078 votes cast in total for one legislator) to Franklin (4,469 votes for one legislator).[60]

Test 3 is evaluated by starting from the fact that there were 18 races.[61] The average expectation of a nearly evenly divided popular vote is nine races for each party. The upper theoretical value to sigma is $0.5 * \sqrt{18} = 2.1$ races; computational simulation reveals a true value of sigma of 1.4 races. The

Federalists won only five races,[62] and therefore Test 3 is met to a standard of $(9-5)/1.4 = 2.9$ sigma, statistically significant.

Example 2: Maryland congressional districts. Maryland has eight congressional districts. Steven Shapiro and other plaintiffs filed suit in district court that the post-2010 districting plan violated their rights to political association and equal representation under the First and Fourteenth Amendments.[63] This complaint was dismissed, an outcome that was affirmed by the U.S. Court of Appeals for the Fourth Circuit.[64] However, in December 2015 the Supreme Court reversed the decision, remanding the case to a three-judge court for further consideration.[65]

In Maryland, Democrats typically win around 60% of the vote at a statewide level—the same as the margin needed for a safe victory. Artful arrangement is accomplished—and can be detected—in the form of many districts of near-identical partisan composition (Figure 3).

Test 1 cannot be applied because with only one Republican congressman, the standard deviation of the Republican winning vote share cannot be calculated.

Test 2 should be done for the case of partisan dominance, a situation that calls for the chi-square test to test whether Democratic votes are spread unusually uniformly across congressional districts. Figure 4 shows the classical measure of variability, the standard deviation.[66] The standard deviation of Maryland Democrats' winning vote share in seven districts was 6.6% in 2012 and 7.3% in 2014. I compared the variability of Maryland Democratic districts with the variability of Democratic districts nationwide. The values for Maryland fall outside the zone of chance.

Maryland's standard deviations would have arisen by chance in only 2.8% of cases in 2012

---

[59]*Shapiro v. McManus*, 136 S. Ct. 450 (2015).

[60]Lampi Collection of American Electoral Returns, 1787–1825, American Antiquarian Society (2016).

[61]In that election, multimember districts of unequal population were allowed. For the calculation of Test 3, each district election is used as one data value.

[62]Lampi Collection, *supra* note 60.

[63]*Shapiro*, brief of petitioners, at 12 and 35–39, <http://www.scotusblog.com/wp-content/uploads/2015/08/14-990-ts.pdf>.

[64]No. 14-1417 (4th Cir. 2014).

[65]*Shapiro*, 136 S. Ct. 450.

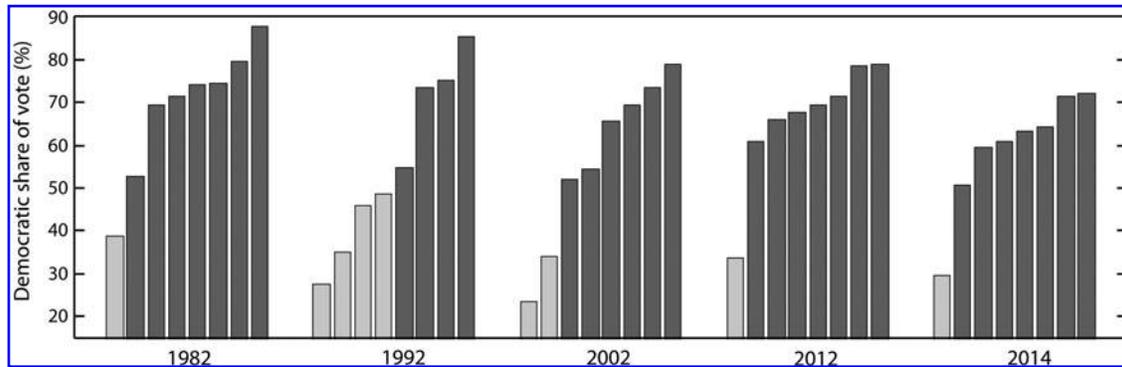[66]The standard deviation is the square root of the variance.

FIG. 3. Democratic two-party vote share in Maryland congressional districts, 1982–2014. For each year, the vote shares are sorted in ascending order of vote share. Republican districts are indicated in *gray*, Democratic districts in *black*. After the 2010 redistricting, vote share in Democratic-held districts became markedly less variable, as evidenced by the narrower range of Democratic win margins in 2012 and 2014.

and 1.7% of cases in 2014.[67] A third year, 2004, also showed an unusually low standard deviation.[68] These findings show that the Democrats' partisan advantage was achieved by spreading their partisan support in a highly even manner across their winning districts.

Test 3 quantifies the size of the effects of Maryland's gerrymander. In the pre-redistricting election of 2010, Democrats won 63.2% of the statewide vote and six seats,[69] compared with a simulated average of 6.1 seats—not statistically significant. After redistricting, in 2012 Democrats won 65.5% of the statewide vote and won seven seats,[70] compared with a simulated average of 6.1 seats. The value of Delta was 1.2 favoring Democrats, not quite statistically significant. In 2014, Democrats' vote share declined to 58.1%, but they retained all seven of their seats.[71] In this case, the simulated average was 5.1 seats, and the value of Delta was 2.4, statistically significant. These results indicate that redistricting gained Democrats a 1-seat advantage in a strong Democratic year, 2012, and that this advantage was retained in the national wave election of 2014 that swept dozens of Republicans into office in states outside Maryland.
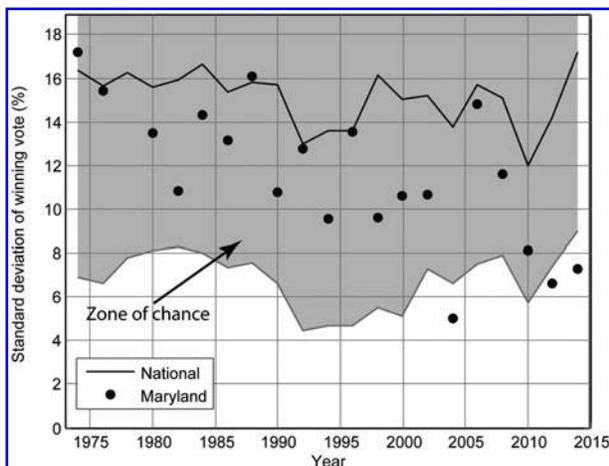

FIG. 4. Standard deviation of Democratic vote share over time. The *jagged line* at top indicates the standard deviation of the Democratic vote share nationally. *Black circles* indicate the standard deviation for Maryland districts. The *gray shaded area* indicates the zone of chance. Two years fall outside the zone of chance and pass an additional test for significance: 2012 and 2014.

[67]For a lower one-tailed test at significance level $p < 0.05$, the lower bound of the zone of chance is equal to sqrt(2.167/(N-1)) * (national s.d.). <http://www.itl.nist.gov/div898/handbook/eda/section3/eda358.htm>, <http://sites.stat.psu.edu/~mga/401/tables/Chi-square-table.pdf>. It should also be noted that the chi-square test assumes normally distributed vote shares. An additional test, the Ansari-Bradley test, does not make this assumption, and still identifies 2012 and 2014 (but not 2004) as being statistically significant departures from national Democratic districts. A.R. Ansari and R.A. Bradley, *Rank-Sum Tests for Dispersions*, 31 Annals of Mathematical Statistics 1174–1189 (1960).

[68]Without partisan intent, the Maryland standard deviation would still be expected to fall outside the zone of chance in five percent of cases—one in twenty. Maryland's 2004 congressional delegation was within the zone of chance by Test 1, indicating that the result of Test 3 is a chance result, i.e., a "false positive."

[69]Karen L. Haas, Statistics of the Congressional Election of November 2, 2010, at 22 (2011), <http://clerk.house.gov/member_info/electionInfo/2010election.pdf>.

[70]*Id.* at 27–28.

[71]New York Times (online), <http://elections.nytimes.com/2014/maryland-elections> (downloaded Feb. 18, 2016).

Example 3: Wisconsin State Assembly districts. After the 2010 election, the Republican Party controlled the Wisconsin State Senate, Assembly, and governorship, bringing post-Census redistricting into its control. The resulting State Assembly map was challenged by a group of Wisconsin Democratic voters who have alleged partisan gerrymandering under the First and Fourteenth Amendments.[72]

The Wisconsin Assembly has 99 seats. To evaluate its partisan asymmetry in historical context, I applied Test 1 (the lopsided-outcomes test) and Test 2 (the reliable-wins test). I analyzed state elections from 1984 to 2014. During this period, the average two-party vote across districts was between 45% and 55% for both parties. This condition of near-parity provides the greatest potential advantage to the party that can impose a partisan gerrymander. Over the entire 30-year period, the difference in winning vote share between the two parties (Test 1) was at its greatest in the 2012 election (Figure 5). Democrats won 39 seats with an average vote share of 68.8% (standard deviation 8.3%), while Republicans won 60 seats with an average vote share of 59.7% (standard deviation 6.5%). The difference, 9.1%, is statistically significant: this outcome would have arisen from a partisan-symmetric process by chance with a probability of less than 1 in 10 million (i.e., a two-sample t-test shows that $p < 10^{-7}$, or a 1 in 10 million chance that the outcome arises by nonpartisan mechanisms).[73] Of particular note is the fact that this partisan advantage appeared immediately after redistricting. Such a sudden jump would not be expected from population-clustering effects, which should change more gradually over time.

From 1984 to 2010, the overall results of Test 1 did not show a consistent pattern of partisan disadvantage. In 1990, Democrats and Republicans jointly controlled redistricting, leading to an impasse and a court-ordered redistricting. In the following five elections from 1992 to 2000, the difference in average winning vote share was not statistically significant and never exceeded 2% in either direction. Then, in 2000, redistricting was again court-ordered, and in the following five election cycles from 2002 to 2010, the median value of the lopsided-outcomes test was a 5.0% advantage in favor of Republicans, reaching statistical significance three times.

In 2014, a majority of Assembly seats were uncontested: 29 out of 63 Republican seats and 23 out of 36 Democratic seats. In this situation, the average winning vote share is dominated by imputed values. For example, if all races were uncontested, the difference in average winning vote share would be defined as zero. Therefore an abundance of uncontested races tends to underestimates of the degree of partisan asymmetry. In this case, the difference in average winning vote share was 2.0% favoring Republicans, or 6.4% not counting imputed support (Figure 5A, open symbol). This case demonstrates that when many races are uncontested, an additional measure of partisan asymmetry is needed.

As a second test for gerrymandering, I used Test 2, the mean-median difference. The mean-median difference is applicable since the parties are closely matched in statewide strength. After redistricting, the average Democratic vote share in 2012 was 51.5% and the median vote share was 45.7%. The difference, 5.8% favoring Republicans, was statistically highly significant at $p < 10^{-5}$, meaning that under symmetric conditions, the mean-median difference would reach 5.8% by chance less than once in one hundred thousand cases. In 2014, Democrats' average vote share declined to 46.0%, and their median vote share was 41.1%. The difference, 4.9% favoring Republicans, was again statistically significant ($p < 0.01$). Both 2012 and 2014 had a higher mean-median difference than the pre-redistricting election of 2010, in which the mean-median difference was 3.2% favoring Republicans. These findings are consistent with the idea that partisan asymmetry increases suddenly when a new gerrymandering scheme is put into place.

Test 3 (quantifying the number of excess seats) was not done because it optimally requires a population of districts from the same year for purposes of

---

[72]*Whitford v. Nichol*, No. 3:15-cv-00421 (W.D. Wis. 2015).

[73]In such a calculation, provision must be made for how to score uncontested races. The calculation in the main text was done counting uncontested races as 75%–25% victories. This assumption is established in previous literature (Andrew Gelman and Gary King, *A Unified Method of Evaluating Electoral Systems and Redistricting Plans*, 38 AMERICAN JOURNAL OF POLITICAL SCIENCE 541, 550 [1994]) as a means of evaluating likely imputed amounts of support in a situation where one party is dominant. In the case of the 2012 election, 23 Democratic seats and 4 Republican seats were uncontested. If these 27 races were counted as 100%–0% splits, the average vote share would be 83.5% for Democrats and 61.4% for Republicans, with even greater statistical significance ($p < 10^{-9}$). Generally, imputed support is a conservative assumption that tends to reduce differences between the two parties.
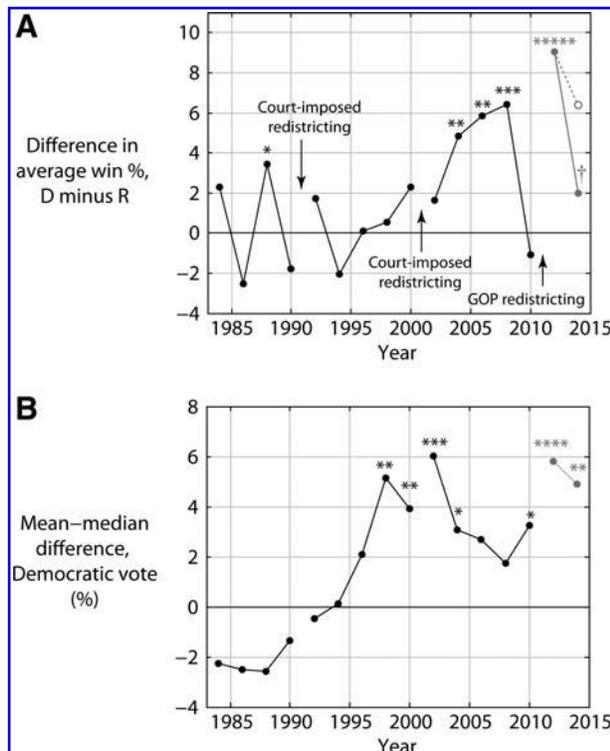
FIG. 5. Application of gerrymandering tests to the Wisconsin State Assembly, 1984–2014. **(A)** Application of Test 1, the difference between average Democratic win margin and average Republican win margins. Statistical significance was tested with a two-tailed unpaired *t*-test. The open symbol indicates the value calculated without imputing support. **(B)** Application of Test 3, the mean-median difference. Significance levels: $*p<0.05$; $**p<0.01$; $***p<0.001$; $****p<0.0001$; $*****p<10^{-7}$. [†]a majority of Assembly seats were uncontested, diminishing the numerical and statistical value of Test 1.

simulation. For analysis of a state legislature, this information is not available. If necessary, the test could potentially be done using Wisconsin election data from a different year or by drawing districts from a symmetric distribution with a realistic standard deviation.

DISCUSSION

In this article I have presented three tests for rapid identification of partisan gerrymanders. These tests can be used to evaluate intents and effects, the two prongs articulated in *Davis v. Bandemer*. The two intents tests can be done with computing resources already available on a judge's or clerk's desk, and the effects test requires some additional software.[74] All three tests rely on well-

established statistical principles. The tests measure different aspects of partisan asymmetry and therefore fall within the scope of principles that have been expressed by the Supreme Court. I suggest that these tests may constitute a manageable standard for courts to evaluate the impact of a state's districting scheme on its residents' Equal Protection and First Amendment rights.

The broader implications of this article are twofold. First, I have used statistical science to express the idea that a pattern of election results might have arisen by chance and therefore not warrant judicial intervention. By establishing "zones of chance" in which the partisan impacts of a districting plan are ambiguous, the three tests presented here can help a judge evaluate whether an identifiable injury has occurred in the first place. Second, an intents-and-effects standard based on the tests is unambiguous and may mitigate the need to demonstrate predominant partisan intent. For these reasons, these statistical tests comprise a valuable and timely addition to the judge's toolkit for rapid and rigorous identification of partisan gerrymanders.

Zones of chance and the First Amendment

My statistical analysis of the effects of gerrymandering may be of particular relevance to First Amendment analysis, which "allows a pragmatic or functional assessment that accords some latitude to the States."[75] By allowing for a normal amount of statistical variation, the three tests proposed in this article build in zones of chance where any of a range of outcomes would lead to an acceptable amount of asymmetry.

Any statistical approach contains some possibility of accidentally identifying gerrymandering where it does not exist (in statistical terminology, "false positives") or missing cases where it did occur (false negatives). Tests may also sometimes not be usable, for instance Test 1 when one party only wins one seat. For these reasons, I have provided two separate tests of intents. These tests are oriented toward the outcomes of elections rather than the specifics of map boundaries or district

---

[74]A version of this software is available on GITHUB at <http://github.com/SamWangPhD/gerrymandering>; it is also available for use at <http://gerrymander.princeton.edu>.

[75]*See, e.g., Eu v. San Francisco County Democratic Central Comm.*, 489 U.S. 214. (1989).

procedures. The tests hew closely to the electoral goals of redistricters and do not rely on geographically oriented approaches which require normative assumptions of what constitutes good districting procedure.

The transparency of well-known statistical standards

If statistical tests for gerrymandering are sufficiently complex, the use of expert witnesses becomes necessary. However, complex arguments are subject to challenge on technical grounds,[76] creating the secondary question of the credibility not just of the statistical method but of the experts themselves. Although the use of expert testimony and statistical reasoning is commonplace in courts,[77] for constitutional questions where statutory guidance is lacking a judge may wish to conduct his or her own evaluation in a more direct manner.

Whitford v. Nichol provides an example of the complications that may arise. In *Whitford*, the districting plan was evaluated using a recently developed measure of asymmetry, the efficiency gap.[78] Expert witness Prof. Simon Jackman established the statistical properties of the efficiency gap in a presentation that included 36 figures.[79] This report was challenged by the state's expert witness, who focused on the question of how much asymmetry came from population clustering; that expert was, in turn, counter-challenged.[80]

While such challenges are an inevitable part of complex litigation, the use of longstanding and simple statistical tests may reduce the need for expert witnesses and detailed presentations. In particular, Tests 1 and 2 proposed here use well-known statistical tests with established procedures for significance testing, can be explained succinctly,[81] and can be worked out by hand. These qualities confer transparency to my proposed analysis of intents.

In addition, this article's tests can be used to separate the contributions of gerrymandering and population clustering. Since gerrymandering relies on the ability to sequester voting populations, the geographic patterns that give Republicans a naturally occurring advantage can also be used to construct further artificial advantages. Conceptually, this addresses the concern about natural clustering expressed in the *Whitford* testimony.

What is the role of intent?

The intent prong in *Bandemer* initially required that the intent be predominantly partisan.[82] This presented a higher bar to proving injury than simply showing that partisanship was one of multiple factors. It is a far higher bar than the evaluation of disparate impact alone. Such a stringent standard may have been appropriate in the absence of legislative guidance or a large body of court precedent. In the *Bandemer/Vieth* framework, the lack of simple and reliable tests made it necessary to assess the link between redistricters' actions and the injury. Indeed, current approaches to proving gerrymanders focus on intent, are diverse in approach, and sometimes do not agree with one another.[83]

An example of ambiguous intent is found in *LULAC v. Perry*.[84] The Republican majority was able to involve individual Democratic legislators in the districting process.[85] However, in matters of redistricting, a party as a whole has motivations

[76]*Daubert v. Merrell Dow Pharmaceuticals*, 509 U.S. 579 (1993).

[77]*Tex. Dep't of Housing and Comm. Aff. v. Inclusive Communities Project*, 136 S. Ct. 2507 (2015).

[78]E. McGhee, *Measuring Partisan Bias in Single-Member District Electoral Systems*, 39 LEGISLATIVE STUDIES QUARTERLY 55–85 (2014); Nicholas Stephanopoulos and Eric McGhee, *Partisan Gerrymandering and the Efficiency Gap*, 82 UNIV. OF CHICAGO L. REV. 831 (2015).

[79]S. Jackman, *Assessing the Current Wisconsin State Legislative Districting Plan.* Exhibit 3 in *Whitford v. Nichol*, <http://www.fairelectionsproject.org/s/Exhibit-3.pdf> (last downloaded Mar. 27, 2016).

[80]S. Trende, State Expert's Declaration, *Whitford v. Nichol*, <http://www.fairelectionsproject.org/s/Declaration-of-Sean-Trende-Dkt-55.pdf> (last downloaded Mar. 27, 2016); S. Jackman, Rebuttal Report, *Whitford v. Nichol*, <http://www.fairelectionsproject.org/s/Jackman-Rebuttal-Report-Dec-21-2015.pdf> (last downloaded Mar 27, 2016).

[81]S. Wang, Amicus Brief, *Harris v. Arizona Independent Redistricting Commission*, 993 F. Supp. 2d 1042 (D. Ariz. 2014), <https://www.brennancenter.org/sites/default/files/legacy/CGR%20Reprint%20Single%20Page.pdf> (last downloaded March 27, 2016); *see also* Wang, *supra* note 5.

[82]*Bandemer*, 478 U.S. at 128.

[83]Micah Altman, Brian Amos, Michael P. McDonald, and Daniel A. Smith, *Revealing Preferences: Why Gerrymanders Are Hard to Prove, and What to Do about It*, SOCIAL SCIENCE RESEARCH NETWORK, Mar. 22, 2015, at 11–36 (enumerating the strengths and weaknesses of different approaches to evaluating partisan gerrymanders), <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2583528> (last accessed Aug. 27, 2015).

[84]*LULAC*, 548 U.S. 399, 417–418 (describing cooperation of individual Democratic legislators).

[85]*Id.*

that can be at odds with those of some of their own party's individual legislators.[86] Therefore intent is most fairly evaluated at the state level or at the individual level, but not both at the same time. In addition, the majority in *Crawford v. Marion County Election Board* held that partisan intent is insufficient as a reason to strike down voting restrictions.[87]

The identification of intent begins with a fact-specific inquiry into the state of mind of the legislature and/or the entity that drew the district lines. Statistical testing such as my proposed Tests 2 and 3 allows the identification of patterns of districting that are highly unlikely to have arisen by chance, thereby providing concrete evidence that a legislature or other district-drawing body acted specifically to produce partisan outcomes. This rigorous standard should aid tremendously in the identification of intent.

Furthermore, I suggest that districting can impose a burden on a group's representational rights whether or not the effects (as measured by Test 3) are intentional. Even where intentions are non-partisan, bipartisan, or unknown, the effect of a districting plan with partisan asymmetry is to produce legislative blocs whose size is unrepresentative of the popular will. The construction of a reliable measure of effect provides clear guidance when an injury has taken place and a template for how the injury can be repaired. Just as a road worker may act to right an upended orange traffic cone even if she or he does not know how the cone came to be tipped over, a court may act when effects are sufficiently strong, as in disparate impact cases in racial discrimination cases.[88] Although partisan gerrymandering cases are governed by different doctrine (constitutional) from racial discrimination cases (statutory interpretation), both types of case concern the issue of intent.

Evaluating the partisan impact of district maps before implementation

Although in this article I used election results to calculate the three tests, the tests could alternatively use other inputs. For example, to rule out the possibility that the tests may be influenced by variations in the quality of specific candidates, it would be possible to use district-level presidential vote shares as inputs.[89]

In current federal precedent, the need for redrawing a set of districts often relies on forensic evidence; that is, on elections that have already occurred.[90]

However, by that time an injury to voters has already occurred. To preempt such an injury from occurring, the three tests could be calculated using information that is available before an election. Under the First Amendment rationale of not penalizing groups for their partisan preference, party registration might be used as an input to calculate the three tests. Political scientists, redistricters, and commercial redistricting software also use other variables to predict overall partisan preference; these predictions could also serve as inputs to the tests. Doing so would allow a hypothetical districting scheme to be assessed before it has passed into law.

The standards presented here can quantify the benefits of reform efforts directed at reducing the likelihood of partisan gerrymandering. One such route is the establishment of nonpartisan districting commissions that remove districting from the direct control of legislators. In California, a voter referendum in 2008 established the formation of the California Citizens Redistricting Commission.[91] The commission is composed of 14 members who are drawn from members of the general public, including five Democrats, five Republicans, and four members who decline to state a

---

[86]*See* discussion of mixed partisan motivations, *LULAC, supra* note 84.

[87]*Crawford v. Marion County. Elec. Bd.*, 553 U.S. 181, 203–204 (2008).

[88]In one recent example, in a racial discrimination case the Supreme Court ruled that demonstration of disparate impact was sufficient to prove discrimination, and that a demonstration of intent was not necessary. *Tex. Dep't of Housing and Comm. Aff. v. Inclusive Communities Project*, 136 S. Ct. 2507 (2015). This case held that in light of results-oriented statutory language in the Fair Housing Act, determination of disparate impacts was sufficient to warrant a remedy, even without discriminatory intent. I argue that if gerrymandering has a sufficiently large effect on a party's supporters, such an injury should still be remedied even when redistricters are not motivated purely by partisan intent.

[89]*LULAC*, 548 U.S. at 420 (Justice Kennedy, joined by Justices Souter and Ginsburg), states in regard to a partisan gerrymandering claim that "such a challenge could be litigated if and when the feared inequity arose." Redistricting software is capable of using quantities such as the presidential vote share to estimate the partisan tendency of a hypothetical district. Redistricters use such measures to judge the likely outcome of a district, and could use them as inputs to the three tests in this Article to evaluate a districting plan before it is implemented.

[90]*LULAC*, 548 U.S. at 2638 (n.9) (opn. of Stevens, J.P., joined by Breyer, S).

[91]CALIFORNIA CITIZENS REDISTRICTING COMMISSION, <http://wedrawthelines.ca.gov/regulation_archive.html> (last visited Aug. 24, 2015).

partisan loyalty.[92] The commission's mandate is to draw districts that respect principles of contiguity, compactness, and representation of a community's interests.[93] The resulting congressional districts have become more competitive: margins of victory have become smaller, and incumbents have lost their re-election races at higher rates than before the formation of the commission.[94] Like the Arizona commission, the work of the California commission has led to closer races and more euproportional overall outcomes.

These tests could also be used in approaches that leave districting under the control of state legislators, but place constraints on how and what they produce. Such an approach has been taken in Florida; ballot initiatives known as Amendments 5 and 6 were passed in 2010, becoming Article III, §§ 20 and 21 of the Florida Constitution.[95] Together with Article III, § 16,[96] the Florida Constitution stipulates that district lines "must be contiguous, compact, and use existing political geographical boundaries where available."[97] Districts also may not be drawn to "favor or disfavor a political party or incumbent."[98] The resulting plans are subject to review by the Florida Supreme Court for review, leading either to approval or return to the legislature for a further attempt to meet districting criteria.[99] The tests described in this article could be useful in identifying statewide partisan favor. Individual districts would still need to be evaluated separately, for example to comply with Voting Rights Act restrictions and other principles set down in federal or state law. These tests, which address the properties of combinations of districts, can complement these other constraints without conflict.

CONCLUSION

Partisan gerrymandering distorts relationships between voting and representation that would otherwise arise naturally, generates seats that are unresponsive to shifts in public opinion, and chills the freedom of voters to associate with a political party of their choosing. The health of democratic processes would be considerably improved by reducing the ability of legislative processes to impose partisan distortions on redistricting maps. The three tests for asymmetry presented here may contribute to a manageable standard for identifying partisan gerrymanders, with the eventual goal of reducing or eliminating them.

Address correspondence to:
Samuel S.-H. Wang
Princeton Neuroscience Institute
Princeton University
Washington Road
Princeton, NJ 08544

E-mail: sswang@princeton.edu

---

[92]Calif. Const. art. XXII, § 2(c)(2).
[93]Calif. Const. art. XXII, § 2(d).
[94]*Id.*
[95]Justin Levitt, *Florida,* ALL ABOUT REDISTRICTING, <http://redistricting.lls.edu/states-FL.php> (last visited Aug. 24, 2015).
[96]Fla. Const. art. III, § 16.
[97]Fla. Const. art. III, §§ 20–21.
[98]*Id.*
[99]Fla. Const. art. III, § 3(b).