

# **An Antidote for Gobbledygook: Organizing the Judge's Partisan Gerrymandering Toolkit into Tests of Opportunity and Outcome**

Samuel S. -H Wang, Brian Remlinger, Benjamin Williams

**SUMMARY:** Despite recent setbacks, litigation-based challenges to extreme partisan gerrymandering continue, and quantitative methods for detection are more important than ever. Many measurement tools have emerged that probe the question of whether a redistricting map is extreme or violates the principle of partisan symmetry. Such tools were used successfully in a lawsuit concerning Pennsylvania Congressional districts under that state's constitution. Here we provide a framework for categorizing these tests for future use in state and federal constitutional cases. Our framework explains how measures should be interpreted and identifies which tests will be most effective given the specific facts of a particular state. Broadly, the tests can be divided into two categories: those that identify inequality of opportunity, *i.e.* a systematic effort to deprive of a group's ability to elect representatives; and those that identify inequality of outcome, *i.e.* a durable distortion in the amount of representation. In each case, the measures examine the difference between the existing map and what would occur under a districting process in which partisan interests are not the overriding consideration. A general thread is that of "significance testing," in which a district or statewide districting scheme can be defined as more extreme than the great majority of possibilities that could arise through unbiased means. Such tests are most often done with well-established classical statistical tests but can also include recently-developed measures such as the efficiency gap. It is even now possible to evaluate, with mathematical rigor, whether a specific scheme is extreme relative to the virtually uncountable universe of possible maps. Taken together, these methods for detecting extremes comprise a statistical toolbox to address a wide variety of circumstances that may arise in any of the fifty states.

## **I. INTRODUCTION: THE TOOLS FOR GERRYMANDER DETECTION MUST BE ORGANIZED**

### **1. The search for a simple standard for partisan gerrymandering will live on in state courts**

To all appearances, recent events at the U.S. Supreme Court have dealt a setback to the movement for fair districting. In *Gill v. Whitford*,<sup>1</sup> the Supreme Court unanimously rejected the standing of individual voters to challenge a statewide plan on partisan grounds under the Fourteenth Amendment.<sup>2</sup> Furthermore, the Court's refusal to bless a First Amendment-inspired theory in *Benisek v. Lamone*<sup>3</sup> raised concerns about the vitality of that approach at a federal level. Finally, and most importantly, Justice Kennedy's retirement will, in all likelihood, make it more difficult to assemble a five-justice coalition willing to take any action at all.<sup>4</sup> If they do act, they

---

<sup>1</sup> 138 S.Ct. 1916 (2018).

<sup>2</sup> *See id.*

<sup>3</sup> 138 S.Ct. 1942 (2018).

<sup>4</sup> *See* Michael Wines, *Kennedy's Retirement Could Threaten Efforts to End Partisan Gerrymandering*, N.Y. TIMES (June 30, 2018), <https://www.nytimes.com/2018/06/30/us/kennedy-scotus-gerrymandering.html>.

may mandate a district-by-district evaluation, an approach which, if pursued in isolation, is prone to judicial bias and incomplete remedies. Although *Common Cause v. Rucho*<sup>5</sup> will likely force the Court to weigh in on partisan gerrymandering once more in the 2018-19 term, a ruling that meaningfully curbs partisan gerrymandering may be out of reach.

Despite these disappointments, we contend that from these sow's ears a silk purse can be made. Taken together, Chief Justice Roberts's opinion and Justice Kagan's concurrence in *Whitford* provide a conceptual framework for *state* court litigation, without the involvement of federal courts.<sup>6</sup> Roberts wrote that individual voters in specific districts may suffer Fourteenth Amendment harms by being either packed (stuffed into districts so that they win very few victories) or cracked (spread across districts so that they suffer many close losses). This approach may be used to identify single-district gerrymanders and specific groups of gerrymandered districts. Kagan, meanwhile, argued that voters at a statewide level may be injured on First Amendment grounds by that same pattern of packing and cracking. A majority of state constitutions include language similar to that in the First and Fourteenth Amendments, so state court litigation grounded in state constitutions can harness either or both strategies.<sup>7</sup>

Translating either of these theories into judicial intervention requires care. State Supreme Courts will be confronted by the limit-setting questions which have plagued federal courts. Courts must define acceptable and unacceptable levels of partisanship in a redistricting plan, and the resulting bounds must limit the most severe offenses while still allowing legislators latitude to satisfy competing interests in the redistricting process. Too much latitude will lead to repetition of the offense in the next round of redistricting. Moreover, an excessively vague standard would make even honest redistricting difficult, unleash a flood of lawsuits, and impair the orderly conduct of elections.

Quantitative measures of partisanship provide bright-line limits that will allow state courts to accomplish this task. In *Whitford* oral arguments, Chief Justice John Roberts referred to

---

<sup>5</sup> 279 F.Supp.3d 587 (M.D.N.C. Jan. 9, 2018), *vacated*, 2018 WL 1335403 (mem.) (following the District Court's finding that North Carolina's active congressional district plan was an unconstitutional partisan gerrymander under the First and Fourteenth Amendments, the Supreme Court vacated and remanded the case for reconsideration in light of *Gill v. Whitford*).

<sup>6</sup> Aside from the legal soundness of such claims, the rationale for regulating partisan gerrymandering suggested by Roberts has the additional appeal of presumably satisfying at least five U.S. Supreme Court Justices (the Chief Justice as well as the four signers of Justice Kagan's concurrence, who also signed onto the Chief Justice's opinion). This may somewhat insulate state court decisions regulating federal redistricting under such schemes from U.S. Supreme Court review. Avoiding federal court review is desirable for the obvious reason of avoiding reversal under the specific legal theory, but also because authority of state courts to regulate federal redistricting in general may be challenged under the Elections Clause. The outcome of such a challenge is far from clear, but a five-vote coalition including Justice Kennedy's replacement for stripping that authority from all state courts is not out of the realm of possibility, given the "muscular", originalist-powered dissent by Chief Justice Roberts in *Arizona v. Arizona Independent Redistricting Committee*, 135 S.Ct. 2652 (2015), and the Court's likely shift to more originalist-leaning jurisprudence accompanying Justice Kennedy's retirement.

<sup>7</sup> State courts may also apply a third approach, which was taken by the Pennsylvania Supreme Court in its anti-gerrymandering decision in January 2018, where the Congressional districting scheme was invalidated because it violated the "elections shall be free and equal" language in the state constitution. Twenty-four states have similar free-and-equal clauses. Claims under this theory will still need quantitative evidence of unacceptable levels of partisanship, and so the remainder of this article is equally applicable to those cases. See Simon Jackman, *The Predictive Power of Uniform Swing*, PS: Political Science & Politics, 47, no. 2 (2014); Laura Royden, Michael Li, and Yuri Rudensky, *Brennan Ctr. for Justice, Extreme Maps and the 2018 Midterm* (2018), <https://www.brennancenter.org/publication/extreme-gerrymandering-2018-midterm>.

recently-developed standards for partisan gerrymandering as “sociological gobbledygook.”<sup>8</sup> He subsequently retreated from this statement, acknowledging that several mathematical approaches are sound.<sup>9</sup> Far from being gobbledygook, statistical methods offer straightforward mathematical representations of fairness that clarify the differences between nonbiased and extreme redistricting schemes.

The retreat from federal to state court removes one of the major difficulties to bounding partisanship in redistricting. In federal litigation, plaintiffs have sought a meta-standard that would work in every state, whether dominated by a single party or evenly split, whether urbanized or rural, and everything in between. When bringing claims in a state court, however, plaintiffs need only argue for a standard that works in that state. State courts can choose the metrics that work for their state without worrying how those metrics work elsewhere.

## 2. Identifying neutral districting plans

A recurring theme in the identification of partisan gerrymandering is a quest for an ideal districting plan. Gerrymandering claims allege that a districting scheme is unacceptable along some axis because it harms the individual or group rights of voters, but such claims imply the existence of a better alternate plan, which the state has a legal obligation to adopt. In some cases, like in Pennsylvania’s 2011 Congressional map, a court may use straightforward traditional redistricting criteria like compactness or number of split political units as indicia of an overly partisan map. However, such straightforward proxies when used by themselves can easily fail.<sup>10</sup> While such traditional “floor criteria” can often help identify egregiously engineered maps, redistricters can still achieve a high degree of partisanship despite achieving geographic compliance. In addition, current technology has extreme difficulty showing that a map is maximally compliant along any single criterion, let alone balancing multiple criteria, and no state has seriously attempted to use geographic metrics to remove all subjectivity from the map-drawing process.<sup>11</sup>

The redistricting community has presented several major approaches to answer the ideal districting question. One longstanding approach relies on the principle of partisan symmetry, a normative idea of how a party’s voting power should translate to seat share. More recently, another approach, outlier analysis, sidesteps normative votes-to-seats questions, instead identifying plans

---

<sup>8</sup> Philip Rocco, *Chief Justice Roberts Said Political Science is ‘Sociological Gobbledygook.’ Here’s Why He Said It, and Why He’s Mistaken*, Wash. Post (Oct. 4, 2017), [https://www.washingtonpost.com/news/monkey-cage/wp/2017/10/04/justice-roberts-said-political-science-is-sociological-gobbledygook-heres-why-he-said-it-and-why-hes-mistaken/?utm\\_term=.35279b583247](https://www.washingtonpost.com/news/monkey-cage/wp/2017/10/04/justice-roberts-said-political-science-is-sociological-gobbledygook-heres-why-he-said-it-and-why-hes-mistaken/?utm_term=.35279b583247).

<sup>9</sup> *Gill v. Whitford*, 138 S. Ct. 1916, 1933 (2018) (“We need not doubt the plaintiffs’ math”).

<sup>10</sup> Indeed, the remedial maps presented by the Republican-dominated state assembly, as well as an example Democratic-favoring maps on file with the authors of this paper, display the limits of using traditional redistricting criteria to bound gerrymandering. The 2011 plan split 28 counties, while the Republican remedial plans split only 15 counties and Democratic maps split 12 counties. In 2016, Republicans earned 48% of the statewide vote, but 13 of the state’s 18 Congressional seats under 2011 map which was still in effect. The Republican remedial plan, however, would have still elected 13 Republican Congressmen and 5 Democratic Congressmen in 2016, while the Democratic example plan would have elected 9 Republican Congressmen and 9 Democratic Congressmen that year, despite Democrats losing the statewide vote. The plan the state Supreme Court ultimately adopted, in contrast, would have elected 10 Republican and 8 Democratic Congressmen in 2016, while splitting 13 counties. The Republican remedial plan, Democratic example plan, and Court-ordered remedial plans showed comparable improvements on compactness, number of township splits, and other metrics, demonstrating the range of partisan outcomes that can occur even with maps that perform nominally equally on traditional redistricting criteria.

<sup>11</sup> Wendy Tam Cho, *Towards a Talismanic Redistricting Tool: A Computational Method for Identifying Extreme Redistricting Plans*, 15:4 Election L. J. 351 (2016).

that fall outside a reasonable range of unbiased outcomes. In either case, election results taken across the districts of a state play a central role in determining whether an offense has occurred.

The concept of partisan symmetry is nominally simple to define: a set of districts is symmetrical if reversing the outcome of the election—flipping each party’s average district vote totals—would also reverse the number of seats won. The idea that both major parties should be treated similarly is a principle that dates to antiquity, including the Golden Rule and the Bible.<sup>12</sup> According to one amicus brief in *Whitford*, “[p]artisan symmetry, the principle put forward by the [*Whitford*] plaintiffs...is highly intuitive, deeply rooted in history, and accepted by virtually all social scientists. Tests for partisan symmetry are reliable, transparent, and easy to calculate without undue reliance on experts or unnecessary judicial intrusion on state redistricting judgments.”<sup>13</sup>

Partisan symmetry, however, has not gained universal acceptance. Some opposition to the idea comes from the nonuniform geographic distribution of voters in the United States, which does not lend itself naturally to plans that award an equal number of seats to both major parties when they each win half the vote.<sup>14</sup> In order for a plan to be partisan-symmetric in most states, a plan drawn blind to partisanship would have to be adjusted to account for this geographic advantage. In other words, redistricting schemes that respect partisan symmetry might not arise from a partisan-blind redistricting process, making the principle potentially unattractive to some legislators and jurists.

While partisan symmetry has floundered, outlier detection has taken center stage. Outlier detection combines two ideas: first, that for any measure of partisan fairness, a state may take on many different values because it can be redistricted in many different ways, and second, in most cases, these scores of most plans evaluated using these measures fall within reasonable limits. Limits on reasonableness can be identified in a number of ways, including by examining historic districting plans or by drawing many alternative districting plans. A baseline limit on partisanship can then be determined according to standard statistical criteria or other means.

Outlier analysis has several benefits. First, courts need not settle on an ideal plan to set bounds on acceptable levels of partisanship – as Justice Kagan observed in *Benisek*, by setting boundaries a court simply says, “however much you think is too much, this case is too much.”<sup>15</sup> Second, outlier analysis lets courts allow some partisanship in redistricting, so long as it is not excessive. Finally, outlier analysis may be tailored to each state’s political geography and the natural distribution of districting plans, as opposed to being tied to an appealing but unnatural partisan-symmetry standard.

How partisan symmetry violations or excessive outliers are detected will vary according to each state’s particular circumstances. As explained in Appendix A, different metrics for partisan gerrymandering are effective in different political geographies. A state which is highly urbanized cannot be evaluated with the same tools as a state that is more rural in nature, and an evenly-divided state requires different tools from a party-dominated state. Luckily, as discussed above, the “metrics problem” is much reduced with a switch to state courts. The search for a federally-

---

<sup>12</sup> *Matthew* 7:12; *Genesis* 13:8-9.

<sup>13</sup> Brief for Heather K. Gerken et al. as Amici Curiae Supporting Appellees at 18, *Gill v. Whitford*, 138 S.Ct. 1916 (2018) (No. 16-1161).

<sup>14</sup> Jonathan Rodden, *The Geographic Distribution of Political Preferences*, ANNU. REV. POLIT. SCI. 13:321–40 (2010).

<sup>15</sup> Oral Argument at 24:08, *Benisek v. Lamone*, 138 S.Ct. 1942 (2018) (No. 17-333), <https://www.oyez.org/cases/2017/17-333>.

imposed meta-standard can be exchanged for state-bound standards grounded in the political and geographic realities of the specific state under examination.

Crafting a framework for outlier analysis requires organizing the mathematical detection tools available and harmonizing them with legal principles. Once courts are able to decode math into law, they can better understand how the many measures of partisan gerrymandering correspond to specific legal rights. Such organization overcomes two problems. First, without such a framework, it is unclear which test is applicable. Second, without a clear understanding of each test's usefulness, a judge may be deceived by "gobbledygook" from competing experts.

We do not present any single mathematical standard as a magic bullet. Instead, we propose that the standards should be taken together as a toolkit, sorted into clear conceptual categories. With that in mind, what follows is a categorization of some of the best tools offered to courts.

## **II. A TWO-PART DOCTRINE FOR PARTISAN GERRYMANDERING: UNEQUAL OPPORTUNITY AND DURABLE OUTCOMES**

### **1. Partisan gerrymandering doctrine should test for unequal opportunity and durable outcomes**

We propose that mathematical tests for partisan gerrymandering fall into two broad categories: (1) tests of unequal opportunity and (2) tests of durable outcomes. This duality mirrors the framework of intents and effects outlined in *Davis v. Bandemer*.<sup>16</sup> To develop this categorization, we borrow concepts from the Chief Justice's and Justice Kagan's theories of partisan gerrymandering litigation, as presented in *Whitford*.

### **2. Inequality of electoral opportunity**

Competitive elections are a core principle of democracy. By engineering reliable wins, gerrymandering deprives clearly identifiable, geographically connected groups of voters of the opportunity to translate their votes into electoral victories. Partisan gerrymanders distort representation by treating voters of one party differently than voters of the other based on their prior voting behavior. Supporters of the party disadvantaged by the gerrymander are cracked or packed to reduce their party's overall strength. Whereas many of these voters may have been represented by a member of their party or lived in districts where a member of their party may plausibly have won an election before redistricting, they no longer have the opportunity to meaningfully translate their votes to electoral victories after redistricting. Measures of *electoral opportunity* detect these differences by examining the indicia of uncompetitive races.

Importantly, measures of electoral opportunity are flexible and can be adapted to a state's particular political geography. Uncompetitive wins can arise incidentally, through compromise or through natural patterns, but tests for inequality of opportunity distinguish partisan misdeeds from neutral accidents. The tests do not require external information, in the form of national norms, historic results, or alternative election maps. Instead, they rely only on the election results in question to measure the fairness of a map.

The effects of unequal opportunity can be identified both in single districts and at a statewide level. In states where one party is dominant, evaluation is best done at the single-district level. This is the case in Maryland, where Republicans are outnumbered considerably at the statewide level, and are specifically deprived of representation in Maryland's Sixth District, the

---

<sup>16</sup> 478 U.S. 109 (1986).

focus of *Benisek v. Lamone*.<sup>17</sup> These cases can proceed via examination of the specific boundaries and demographics in question, reminiscent of racial-gerrymandering cases. But in states like Wisconsin, where the two parties are at near-parity and the entire map, not just a single district, targets one party, different tests are required, so that statewide election results may be evaluated to identify a pattern of systematic discrimination.

A cracked district represents a loss of electoral opportunity, because that district is drawn to bring one side's numbers below viability. A packed district also represents a loss of opportunity, because voters for one side have lost the opportunity to have their voices heard in neighboring districts. Cracked and packed districts can be identified individually, or the overall pattern can be identified as constituting a combination of cracking and packing. For example, the lopsided-wins test, also known as the t-test, measures the difference between cracked and packed districts across a state.

### **3. Durable, inequitable outcome**

Supporters of a political party victimized by a partisan gerrymander will also experience harm in the form of durably reduced representation at the statewide level. The Constitution does not guarantee a specific electoral outcome, but a repeated failure of a party's supporters to achieve electoral representation reveals the effects of partisan gerrymandering. Partisan gerrymandering distorts electoral outcomes by manipulating the relationship between the votes earned by a party and the number of seats won by that party. As an example of durability, from 2012 to 2018 only 1 out of 183 Congressional elections has led to a switch of parties in the most egregiously gerrymandered states, Pennsylvania, Ohio, Michigan, and North Carolina. By contrast, in 2010, before redistricting, 13 of the 65 seats in these four states changed hands in a single election.<sup>18</sup>

To achieve durable asymmetries in representation, the districting plan must "lock in" electoral outcomes; in other words, the plan must produce results which are demonstrably durable and insulate elected officials from any realistic changes in voter preference. Tests of durable outcome determine whether a particular districting scheme would lead to the same electoral outcome under a variety of situations. One method for probing durability is to examine results across multiple elections, which can happen when partisan gerrymandering cases move slowly, as they have in the current redistricting cycle.

Judges need not (and indeed, should not) wait for multiple election cycles to investigate the durability of maps. Another approach to durability analysis is to evaluate distortions in representation that are evident after a single election, using tools outlined below in conjunction with simple concepts like uniform swing analysis.<sup>19</sup> In both cases, single districts and statewide maps can be evaluated to test whether discriminatory partisan effects are durable.

## **III. CONCLUSIONS**

### **1. Using metrics for inequitable opportunity and durable outcome and the legal theories set forth in *Whitford*, state courts can limit partisan gerrymandering**

---

<sup>17</sup> See *supra* note 3.

<sup>18</sup> Calculation on file with the authors.

<sup>19</sup> Simon Jackman, The Predictive Power of Uniform Swing, *PS: Political Science & Politics*, 47, no. 2 (2014); Laura Royden, Michael Li, and Yuriy Rudensky, Brennan Ctr. for Justice, *Extreme Maps and the 2018 Midterm* (2018), <https://www.brennancenter.org/publication/extreme-gerrymandering-2018-midterm>.

By declining to decide the substantive issues in *Whitford* and *Benisek*, the U.S. Supreme Court has signaled a hostility to partisan gerrymandering claims that will likely deepen with the confirmation of Justice Kennedy's successor. Even so, *Whitford* still produced two coherent theories encouraging judicial intervention that can be pursued in at least 31 states.<sup>20</sup> The mechanics of the two approaches are different, but the end result is the same: voters cannot be penalized for their past voting behavior. They cannot be denied the ability to translate their votes to electoral victories for their candidate or their party, at the district or the statewide level.

The 2010 redistricting cycle saw extreme partisan gerrymanders in almost one in five states.<sup>21</sup> Left unchecked, the next redistricting cycle will see, in the words of *Whitford* litigator Paul Smith, a "festival of copycat gerrymandering the likes of which this country has never seen."<sup>22</sup> State courts are the best remaining avenue of attack for litigation-minded reformers, and the time to act is now. The tools exist, and the quantitative questions are easier in the state court room. Properly used, these tools can make partisan gerrymanders obvious and give judges confidence that their evaluations are objective. Furthermore, the existence of a manageable standard provides a way for states to avoid litigation: they can apply the tests to proposed districting plans before implementation. In this way, judges and lawmakers can use the statistical toolkit to strengthen democracy.

Gerrymandering is as creative as it is nefarious. Each state has its own distinctive pattern of communities and of partisans, and in response, redistricters have devised a wide range of tools to hem their opponents into relative powerlessness. In the face of such variety, choosing a single standard to identify partisan asymmetry might seem hopeless, but a simple categorization of the metrics for partisan gerrymandering detection makes intervention manageable. Judges must simply sort the tests into categories, according to (a) whether one wants to test inequality of opportunity or durable inequity of outcome; and (b) whether the case concerns a statewide or single-district misdeed. With these categories in mind, the plethora of approaches can be transformed from "gobbledygook" into a well-organized toolkit for the judge's use.

## **APPENDIX A: ORGANIZING THE TOOLKIT: INEQUALITY OF OPPORTUNITY AND OUTCOME**

---

<sup>20</sup> Of particular interest is North Carolina, a heavily gerrymandered state whose constitution guarantees freedom of elections, freedoms of speech and expression, equal protection of the laws, and substantive due process. N.C. CONST. art. 1, §§§ 10, 14, 19.

<sup>21</sup> Sam Wang and Brian Remlinger, *Slaying the Partisan Gerrymander*, AMERICAN PROSPECT (Sept. 22, 2017), <http://prospect.org/article/slaying-partisan-gerrymander>.

<sup>22</sup> See *supra* note 8.

**Tests for inequality of opportunity (intents)**

Test	Level of Analysis	Operational Circumstances*	Applicable federal cases (states)
Examination of district lines	individual districts	all states	Benisek (MD)
Lopsided wins	statewide	large states	Whitford, Rucho (WI, NC)
Consistent advantage	statewide	large, closely-contested states	Whitford, Rucho (WI, NC)
Uniform wins	statewide	large, party-dominated states	Benisek (MD)
Map fragility	statewide	all states	Whitford, Benisek, Rucho (WI, NC, MD)

**Tests for inequality of outcome (effects)**

Test	Level of Analysis	Operational Circumstances*	Applicable federal cases (states)
Repeated losses	individual districts	all states	Benisek (MD)
Distributional comparisons to simulated maps	statewide	large states	Whitford, Benisek, Rucho (WI, NC, MD)
Efficiency gap	statewide	large states	Whitford, Rucho (WI, NC)
Simple Monte Carlo	statewide	large states	Whitford, Rucho (WI, NC)

Table 2: A summary of tests of electoral opportunity. While all tests work in large, closely divided states, smaller states and states dominated by one party require present larger challenges.

**Tests of Inequality of Opportunity**

Measures of electoral opportunity quantify the size of the unfairness without commenting on underlying reasons. A state with equal (50/50) partisan balance and 10 Congressional seats should not be flagged by measures of electoral opportunity if it elects 8 members of one party and 2 of the other, or vice-versa, so long as either an 8-2 or 2-8 outcome can plausibly occur. Because of their flexibility under a wide variety of electoral scenarios, measures of electoral opportunity allow direct examination of the ways voters were sorted based on partisanship. If a map is an extreme outlier on any of these tests compared to alternative maps judged using the same criteria, it may be evidence that the map was constructed with partisan intent. In this way, statistical testing can determine the likelihood that a pattern would have arisen incidentally. In conjunction with an examination of the events surrounding the legislative redistricting process, such tests can help reveal partisan intent.

Statistical testing also addresses a concern of Chief Justice John Roberts, who asked in *Whitford* arguments whether a decision in a partisan gerrymandering case would be interpreted as a ruling for the prevailing political party. If a districting scheme is extreme by these metrics over the range of possibilities, this can distinguish a gerrymander from a normal product of political give-and-take.

*Examination of district lines*

The most obvious way to scrutinize potential discriminatory purpose in a map is to examine the map itself, line-by-line. A mainstay of racial gerrymandering litigation, the line-by-line approach forces the map drawer to explain each choice made. Why was a town put in one district, rather than another? Why was this county split, rather than that one? The mapmaker must account for every detail.

This approach has obvious appeal, especially in scenarios like the Maryland Congressional gerrymander where only the construction of a single district is at issue. However, it breaks down when applied to statewide challenges. Unlike racial gerrymandering, which is typically contained in areas with substantial populations of minority voters, partisan gerrymandering can span an entire



state. Additionally, interrogating political operatives will not always lead to an admission that lines were drawn for partisan gain. A skilled mapmaker, restricted from using partisan data when drawing a map, can create maps that advantage one party using either partisan voting information or proxies like income distribution, consumer habits, or even web search patterns. Even population density alone can help predict partisan tendency.

#### *Lopsided wins (Student's t-test)*

At a statewide level, inequality of opportunity can be measured in various ways. The simplest measure, requiring no theoretical concepts from social science, is the comparison of average wins. If Democrats within a state win their races with an average of 60% of the vote, and Republicans also win with 60% on average, then they could be said to have similar opportunities to win seats. In contrast, if Democrats win races with an average of 80% while Republicans still win with 60%, then this might suggest that Democrats had been packed into relatively few districts. Such "lopsided wins" would be evidence that the side with larger wins was packed and therefore had fewer opportunities to win elections.

It is possible to calculate the probability that the difference between the two averages arose incidentally from natural variation that occurs when nonpartisan factors are the predominant criteria for drawing districts. This is done using Student's t-test, an old method in statistical science that is broadly accepted across all the sciences. Originally proposed in 1908 by experimental brewer "Student" (a.k.a. William Sealy Gosset) of the Guinness Beer Company to ensure the quality of hops, the t-test is taught to college freshmen and even high school students. As the probability that the difference arose by incidental means decreases, the likelihood that the map is a partisan gerrymander correspondingly increases.<sup>23</sup>

#### *Consistent advantage (the mean-median difference)*

An even older way to measure unequal opportunity is a test for consistent advantage, originally developed by Gosset's mathematical mentor Karl Pearson in 1895.<sup>24</sup> To carry out this test, compare the average (or mean) statewide vote captured by each party with that of the median district — the district that falls in the middle when they are ranked by one party's vote share. When both parties are treated similarly, this difference is close to zero. If the mean-median difference is large — with the median district tilted strongly toward one party — it means one party gained a consistent advantage at the district level. Call it the Lake Wobegon test: the redistricting party has ensured that a majority of its districts perform above average.<sup>25</sup>

In a closely-divided state, engineering a representational advantage for one party usually leads to a large mean-median difference. The lopsided-wins test and the consistent-advantage test are useful at detecting the kind of gerrymandering that is done in states that are closely politically divided, as is the case in Wisconsin, Pennsylvania, or North Carolina. However, this approach does not work in strongly partisan states such as Maryland.<sup>26</sup>

---

<sup>23</sup> Samuel S.-H. Wang, *Three Practical Tests for Gerrymandering: Application to Maryland and Wisconsin*, 15 ELECTION L.J. (2016).

<sup>24</sup> George W. Snedecor and William G. Cochran, STATISTICAL METHODS (Iowa State University Press, 8th ed., (1989); Paul Cabilio and Joe Masaro, *A Simple Test of Symmetry About an Unknown Median*, 24 CAN. J. STAT.. 349, 352 (1996).

<sup>25</sup> See Samuel S.-H. Wang, *Three Tests for Practical Evaluation of Partisan Gerrymandering*, 68 STAN. L. REV. 1263 (2016).

<sup>26</sup> See *supra* note 23.

### *Uniform wins (the chi-square test)*

In partisan gerrymandering, the specific mechanisms of discriminatory maps change when a state's voters favor one political party more heavily. In these situations, the dominant party seeks to make its electoral wins more uniform. A common statistical tool, the chi-squared test, was designed to detect exactly this kind of uniformity.

Under a map drawn with nonpartisan intent, each election would be expected to be won and lost by differing margins. Some elections would be close, while others would be blowouts. In a partisan gerrymander in a state that favors one party, however, electoral results are more uniform; each favored politician wins by very similar margins, because their districts are filled with just enough reliable partisans that their wins are guaranteed. Such an arrangement is to be found in Maryland's Congressional map.<sup>27</sup>

### *Map fragility (Pegden's universal sampling theorem)*

To create the most advantageous and durable gerrymander possible, redistricters assign precincts to a district in a precise manner. The resulting maps are fragile: because the mapmaker's overriding goal was partisanship, they choose the districting scheme that creates the most partisan outcome, and all similar maps must necessarily be less partisan. This fragility can be tested rigorously with the help of a recently-proven mathematical theorem that allows the concept of significance testing to be applied to maps. This remarkable concept has been applied to evaluate proposed district maps in a recent redistricting case in Pennsylvania.<sup>28</sup>

The theorem, developed by Wesley Pegden and others, demonstrates that by exploring many possible modifications of a proposed map, it is possible to show that a map is extreme was designed to reach nakedly partisan ends. Pegden's theorem can be used in conjunction with any measure of partisan gerrymandering, including the efficiency gap, the mean-median difference, lopsided wins, or the recently-developed declination.<sup>29</sup> It can also be applied to a wide range of districting rules, and can identify whether a districting plan is an extreme outlier, despite complying with traditional redistricting criteria. Such an approach gives flexibility that lower courts will want in the months and years ahead.

Testing fragility requires the ability to perform computer simulations, so it will be necessary to call upon a mathematical expert. The procedure is to make many minute changes to a suspect districting scheme by switching a small number of precincts between districts. Such exploration can be repeated on a desktop computer billions of times in a matter of a few hours. If the resulting slightly different districting schemes would have resulted in less-extreme electoral outcomes than the plan as originally drawn, the districts may have been engineered to create partisan advantage.<sup>30</sup> Using Pegden's theorem, it is possible to calculate how extreme the map is in the entire universe of all possible maps without needing to explore all of them.

## **Tests of Durable Outcome**

---

<sup>27</sup> See *supra* note 23.

<sup>28</sup> Moon Duchin, Outlier Analysis for Pennsylvania Congressional Redistricting (Feb. 15 2018), <https://www.governor.pa.gov/wp-content/uploads/2018/02/md-report.pdf>.

<sup>29</sup> Gregory Warrington, Introduction to the Declination Function for Gerrymanders (March 13, 2018), <https://arxiv.org/abs/1803.04799>.

<sup>30</sup> Benjamin Fifield et al., A New Automated Redistricting Simulator Using Markov Chain Monte Carlo (January 20, 2018), <http://imai.princeton.edu/research/files/redist.pdf>.

Measures of unexpected outcomes have two essential components. First, given a statewide vote share, they predict a corresponding seat share. Second, they can quantify the difference between the expected outcome and the actual outcome. By examining the differences between actual outcomes and broader trends, measures of unexpected outcomes can clearly identify maps drawn to create maximal partisan advantage. Using these tools, a court can establish the nature and durability of a map's electoral consequences.

Importantly, measures of electoral outcome must account for the underlying political geography and nonpartisan criteria driving redistricting in a state. Identical voting patterns may yield different results in different states, not for nefarious reasons, but simply because of differences in how voters are arranged. In addition, jurisdictions governed by the Voting Rights Act may be forced to draw districts in certain ways to comply with requirements set forth there. Such districts may cause unexpected partisan outcomes but are not necessarily evidence of partisan malfeasance.

### *Monte Carlo simulation*

Monte Carlo simulation is a basic statistical tool with a long history. First used by in the eighteenth century by the French scientist Georges-Louis Leclerc as a means of estimating the value of  $\pi$ , the modern Monte Carlo simulation was created and refined by Stanisław Ulam, Nicholas Metropolis, and John von Neumann as part of the Manhattan Project at Los Alamos in the 1940s.<sup>31</sup> Since its introduction in modern statistics, Monte Carlo simulation has become ubiquitous in the quantitative sciences and is commonly applied in problems as disparate as experimental physics, financial modeling, and naval search and rescue operations.

The Monte Carlo method as applied to partisan gerrymandering is a comparison of a state's electoral results to national norms.<sup>32</sup> For a state with 12 congressional districts and a given statewide vote share, the expected electoral outcome is determined by creating many "fantasy" delegations of 12 districts drawn from all 435 congressional elections across the country. This pool of fantasy delegations is then pared down to include only delegations which have the same vote split as the state in question. The resulting delegation constitutes a distribution of electoral outcomes, which indicate the likelihood of a state's election outcomes under neutral maps. Given this distribution, the actual outcome can be converted to a percentile rank, giving the probability that an outcome equal to or more extreme than the actual outcome would have happened by chance under neutral maps.<sup>33</sup> The lower the percentile rank, the less likely the plan was to arise incidentally via a non-maximal mapping process.

### *The efficiency gap*

Of the many ways of measuring asymmetry, one has taken center stage: the efficiency gap. The efficiency gap measures asymmetry using a formula that examines how many votes were cast for either party and the seats that were won as a consequence. The efficiency gap measures the portion of votes each party has "wasted."<sup>34</sup> When a party perpetrates a partisan gerrymander, it seeks to use its votes more "efficiently" than the opposing party, by packing and cracking opposing

---

<sup>31</sup> Richard Rhodes, *THE MAKING OF THE ATOMIC BOMB* (1986).

<sup>32</sup> Sam Wang, Opinion, *The Great Gerrymander of 2012*, N.Y. TIMES (Feb. 2, 2013), <http://nyti.ms/WMCC7Q>.

<sup>33</sup> See *supra* note 25

<sup>34</sup> Nicholas O. Stephanopoulos and Eric M. McGhee, Partisan Gerrymandering and the Efficiency Gap, 82 U. CHI. L. REV. 831 (2015); Eric McGhee, *Measuring Partisan Bias in Single-Member District Electoral Systems*, 39 LEGIS. STUD. Q., (Feb. 2014).

voters. The larger the efficiency gap, the more extreme the difference in each party's ability to win elections. The gap between the number of seats won and the number of seats expected given the popular vote is the efficiency gap multiplied by the number of seats in the state. For example, in a district where party A defeats party B by a 60-40 margin, party A wasted 10 percent of the votes cast, since they were in excess of the bare 50 percent plus one vote needed to win. All of party B's 40 percent were wasted.

While a partisan gerrymander will usually have a large efficiency gap, the efficiency gap can take on large values when there are very few districts. Thus, it is best suited for the analysis of large states. While it can be used to compare maps, it cannot be used to calculate statistical significance.

#### *Distributional comparisons to simulated maps*

Using computer simulations, researchers can draw many possible redistricting schemes for a state. By examining precinct-level election returns, the electoral outcome for each plan can be calculated, and then the total number of plans that result in a given delegation split can be counted. When this is done for many possible plans, the difference of a state's electoral results from the average result can be quantified using simple statistics. In an early version of this work, Chen and Rodden have compared hundreds of maps.<sup>35</sup> It is now possible to draw millions or even billions of maps.<sup>36</sup> This approach has been accepted as evidence in a variety of cases. An objection to the automated-maps approach is that any automated procedure must make choices as to what shapes of districts to draw, which jurisdictions and communities of interest to keep together, and so on. These choices may bias the range of outcomes. This is potentially a serious problem, because the number of possible districting schemes for even a small state is astronomical, but this concern may be addressed using Pegden's universal districting theorem to test for map fragility.<sup>37</sup>

#### **4. Applying the tests: How much partisanship is too much?**

Statistical measures of partisanship do not become tests until a bright line threshold is drawn. When the measurement fails to cross that threshold, the test is negative. When the measurement returned is beyond the threshold, the test is positive. Ideally, thresholds answer the question "how much is too much?" by correctly identifying many true positives without incorrectly flagging many false negatives. Different thresholds are suited for different tasks. In physics, a very well-established discipline, detection thresholds are set extremely high to avoid spurious claims.<sup>38</sup> In medicine, thresholds for individual diagnosis are set lower because false negatives can adversely affect patient health.<sup>39</sup>

Although an ideal test would perfectly call balls and strikes, statisticians, like judges, necessarily must settle on tests that remove reasonable doubt. In the field of statistical science, the process of choosing the correct threshold for a test is known as sensitivity testing. In voting rights, the Supreme Court has sometimes engaged in threshold setting, including the one person, one vote

---

<sup>35</sup> Jowei Chen and Jonathan Rodden, *Unintentional Gerrymandering: Political Geography and Electoral Bias in Legislatures*, 8 Q. J. Pol. Sci. (2013).

<sup>36</sup> Gregory Herschlag, et al., Evaluating Partisan Gerrymandering in Wisconsin (Sept. 5, 2017), <http://arxiv.org/abs/1709.01596>; Wendy K. Tam Cho and Yan Y. Liu, *Toward a Talismanic Redistricting Tool: A Computational Method for Identifying Extreme Redistricting Plans*, 15 ELECTION L.J. 351 (2016).

<sup>37</sup> See *Map Fragility* in Appendix A.1.

<sup>38</sup> Louis Lyons, *Discovering the Significance of 5 Sigma* (Oct. 4, 2013), <https://arxiv.org/abs/1310.1284>.

<sup>39</sup> Paul D. Ellis, *THE ESSENTIAL GUIDE TO EFFECT SIZES: AN INTRODUCTION TO STATISTICAL POWER, META-ANALYSIS AND THE INTERPRETATION OF RESEARCH RESULTS* (U.K.: Cambridge University Press, 2010).

principle to a requirement that Congressional districts be of perfectly equal population to within one person. With partisan-gerrymandering law in its infancy, sensitivity testing may help lower courts set useful thresholds for the many measures. One possible threshold might be that measures of outcome should classify positive outcome as a distortion of at least one seat, and tests of opportunity should flag states where the probability is less than 0.05 that a plan arose by incidental means. This is called the “ $p < 0.05$  standard” in the sciences. For example, in the post-2010 redistricting cycle, these thresholds, in conjunction with partisan control of the process, would have identified approximately 7 state Congressional maps as extreme partisan gerrymanders.<sup>40</sup>

## **APPENDIX B: HOW TO MATCH A TEST TO PARTICULAR CIRCUMSTANCES**

### **1. A partisan gerrymander will fail both tests of unequal opportunity and durable outcome**

Partisan gerrymanders are drawn to maximize the electoral strength of one party at the expense of the other by treating supporters of one party differently than supporters of the other. Gerrymandered maps discriminate by reducing the opportunities for voters of one party to translate their votes to electoral victories and consequently by creating statewide patterns of unequal effects and electoral imbalance.

We propose that a partisan gerrymander should only be found unconstitutional if it meets two criteria. First, the gerrymander must be constructed to deny a party’s voters the ability to translate their votes to electoral victories, an abridgement of their opportunity to elect. Second, the gerrymander must have arisen by non-incidental means – it should not be a predictable, durable effect of the underlying political geography.

Both criteria were met from 2012 onwards in Wisconsin, North Carolina, Maryland, and many other states. In the case of inequality-of-outcome criteria, effects were durable, giving identical or close-to-identical seat outcomes in 2012, 2014, and 2016.<sup>41</sup> In the case of inequality-of-opportunity criteria, again different indicia were met.<sup>42</sup> However, understanding this point requires an understanding of the proper uses of the various detection tools.

As described in Appendix A, a number of tools exist to assess the partisanship of a gerrymander. This multiplicity is useful. In jurisprudence, as with as home repair, it can be handy to have a kit containing more than one tool. Partisan gerrymanders are to be found in small states and big states, party-dominated states and closely-contested states, and with varying geographies and communities of interest. Any manageable approach to curbing partisan gerrymandering must use tools that are suited to the particular situation to be found, even if the underlying principle is the same.

When examining a districting plan for evidence of partisan bias, a court must be conscious of both the plan’s effects on electoral opportunity and electoral results. When combined with documentary evidence that partisanship predominated in the redistricting process, these tests of electoral opportunity and electoral outcome provide a court with the information needed to identify extreme partisan gerrymanders. This framework can also reduce the likelihood that a lawsuit will be brought in the first place. With the help of clear guidelines, states can ensure that their districting plans do not create or perpetuate unintended partisan outcomes.

---

<sup>40</sup> See *supra* notes 21 and 25; see also Laura Royden and Michael Li, Brennan Ctr. for Justice, *Extreme Maps* (2017), <https://perma.cc/6V52-RPHZ>.

<sup>41</sup> See *supra* note 18.

<sup>42</sup> See *supra* note 21.

We now examine the measures defined in Appendix A and categorize them according to the situations in which they are most useful.

## **2. Statistical testing of statewide patterns can be carried out for all state legislatures and Congressional delegation with seven or more seats**

Many recently proposed metrics for measuring partisan gerrymandering share a common feature: they examine patterns of statewide results. This represents a departure from other forms of gerrymandering litigation, including racial gerrymandering, which focus on district-by-district analysis. Detailed examination of individual districts is well-suited to the harms in racial cases, where specific groups of voters are harmed by the placement of district lines.<sup>43</sup> Partisan gerrymanders, however, do not focus on individual communities in one region alone. Instead, the goal of a partisan gerrymander is to produce an accumulated statewide advantage emerging from many small decisions to treat voters of one party differently than voters of another party. Simply, partisan gerrymanders are designed to maximize a party's legislative power by winning as many elections as possible. The injury of partisan gerrymandering – the disenfranchisement of one party – is felt statewide. Thus, the most powerful tests for detection of statewide harms operate at a statewide level. At the Congressional level, statewide testing can be conducted on states with seven or more congressional seats.<sup>44</sup> Due to their relatively large size, all state legislatures may be analyzed using statewide tests.

The proper application of tests vary by the characteristics of the particular state, as follows:

### **3. Closely-divided states and partisan-dominated states must be treated differently**

In a closely divided state, partisans gain advantage by building districts that are packed with their opponents. Such skew can be detected in two ways: the lopsided-wins test, which reveals a difference between average wins, and the mean-median difference, which reveals an overall skew.

In states dominated by one party, tests of skew are ill-suited to detecting partisanship. The reason is that in these cases, a durable advantage does not require packing; it can be constructed by creating safe districts of a highly uniform nature.<sup>45</sup> For this reason, strongly partisan states require a subtler detection approach such as the chi-squared test, which identifies an unexpectedly uniform set of outcomes.

### **4. In smaller Congressional delegations, inspection of specific districts is required**

Measurements for partisan gerrymandering developed over the last several years have focused on the extreme gerrymanders of 2012 in state legislatures and medium and large states relatively many Congressional seats, but an exclusive focus on such offenses overlooks less extreme gerrymanders or gerrymanders in states with smaller Congressional delegations. States where one party earned a single extra seat, or simply moved the lines to make elections more favorable to their candidates, still present constitutional problems under most proposed theories of partisan gerrymandering. However, the small number of districts makes it difficult to perform statistical testing, a process that requires large numbers of districts. In these cases, a more detailed examination is needed.

---

<sup>43</sup> Justin Levitt, *Quick and Dirty: The New Misreading of the Voting Rights Act*, 43 FLA. ST. U. L. REV. 573 (2016).

<sup>44</sup> See *supra* note 40.

<sup>45</sup> See *supra* note 23.

Because states with small Congressional delegations have fewer districts, partisan redistricters must by necessity focus on fewer boundaries. Such a situation is manageable for inspection on a district-by-district basis. Generally, when the number of districts is small, the district-by-district analysis used in other cases to identify “pinpoint gerrymanders” may be a suitable approach. Concepts from racial gerrymandering may be modified to suit the specific needs associated with the impairment of a political party's representational rights. To examine a single district for impairment of a party's rights, it is possible to examine whether voters of that party had the opportunity to elect a representative of their choice before redistricting, and whether that opportunity changed as a consequence of a new district map.

Maryland's Sixth District represents a pinpoint partisan gerrymander. The replacement of a district that was likely Republican to with a district that was likely Democratic required the artful drafting of boundaries to encompass Democratic-leaning suburbs from far-flung regions. In conjunction with the statewide representational harm, identification of a partisan drafting process presents an unambiguous demonstration of a partisan gerrymander. Such an examination may take some guidance from the procedures that are used to examine racial gerrymandering.<sup>46</sup>

---

<sup>46</sup> Cooper v. Harris, 137 S. Ct. 1455 (2017); Ala. Leg. Black Caucus v. Alabama, 135 S. Ct. 1257 (2015).