William & Mary
Law Review

VOLUME 59 NO. 5, 2018

A REASONABLE BIAS APPROACH TO GERRYMANDERING: USING AUTOMATED PLAN GENERATION TO EVALUATE REDISTRICTING PROPOSALS

BRUCE E. CAIN,* WENDY K. TAM CHO,** YAN Y. LIU,*** AND EMILY R. ZHANG****

TABLE OF CONTENTS

INTRODUCTION 1523
I. OUTLIER ANALYSIS AND AUTOMATED PLAN GENERATION . . . 1528
A. Getting to a Reasonable Partisan Bias Standard 1538
B. Evaluating Existing Measures of Partisan Fairness . . . 1540

* Professor of Political Science, Stanford University; Spence and Cleone Eccles Family Director of the Bill Lane Center for the American West.

** Professor in the Departments of Political Science, Statistics, Mathematics, Asian American Studies, and the College of Law at the University of Illinois at Urbana-Champaign; Senior Research Scientist at the National Center for Supercomputing Applications; Faculty in the Illinois Informatics Institute; Affiliate of the Cline Center for Democracy, the CyberGIS Center for Advanced Digital and Spatial Studies, the Computational Science and Engineering Program, and the Program on Law, Behavior, and Social Science.

*** Senior Research Programmer, National Center for Supercomputing Applications, and the Department of Geography and Geographic Information Science at the University of Illinois at Urbana-Champaign.

**** PhD candidate at the Department of Political Science at Stanford University; Stanford Law School, J.D. 2016.

C. Optimizing Plans Along All Dimensions of Partisan Fairness 1548
D. Applications of PEAR in Redistricting 1555
CONCLUSION 1556

INTRODUCTION

Partisan gerrymandering has moved to the front burner of political reform once again. As the origin of the term "gerrymander" reminds us, partisan bias in redistricting has been a serious concern in the United States since the early nineteenth century.[1] When *Baker v. Carr* lowered the doctrinal barrier to more serious judicial scrutiny of redistricting efforts,[2] many initially hoped and expected that the Court would eventually find a way to prohibit, or at least severely limit, the practice of drawing district lines for partisan advantage.[3] That early optimism, however, faded when the Court failed to endorse any specific partisan bias test in *Davis v. Bandemer*[4] and *Vieth v. Jubelirer*,[5] even as it reaffirmed at the same time that partisan gerrymandering claims were justiciable.

The wildly fluctuating interests of politicians in limiting partisan gerrymandering illustrate the hoary political principle that "[w]here you stand depends upon where you sit"; that is, political actors tend to oppose redistricting reform when they control the line-drawing process and to favor it when they do not.[6] Republicans, historically more skeptical about political reform than the Democrats, nonetheless led a decade-long charge in California to fix the state's line-drawing process after a Democratic governor and state legislature imposed a particularly contentious redistricting on them in 1982.[7]

---

1. For history on Elbridge Gerry's original gerrymander, see ELMER C. GRIFFITH, THE RISE AND DEVELOPMENT OF THE GERRYMANDER 16-21 (Arno Press 1974) (1907).

2. *See* 369 U.S. 186, 209 (1962).

3. *See* Samuel Issacharoff, *Judging Politics: The Elusive Quest for Judicial Review of Political Fairness*, 71 TEX. L. REV. 1643, 1643-45 (1993) (discussing the "optimism of the 1960s" after the *Baker* decision).

4. *See* 478 U.S. 109, 119, 128-29 (1986).

5. *See* 541 U.S. 267, 310, 317 (2004) (Kennedy, J., concurring in the judgment) (noting that the plurality's holding of nonjusticiability was against controlling precedent).

6. *See* Rufus E. Miles, Jr., *The Origin and Meaning of Miles' Law*, 38 PUB. ADMIN. REV. 399, 399 (1978) (discussing the origins and general applications of the principle).

7. *See* Robert Lindsey, *Once Again, California Wrestles with the Problem of Redistricting*, N.Y. TIMES (Oct. 27, 1983), http://www.nytimes.com/1983/10/27/us/once-again-california-wrestles-with-the-problem-of-redistricting.html?mcubz=0 [https://perma.cc/2VKN-M8WN] (explaining the debate between California Republicans and Democrats over redistricting in the 1980s). *See generally* Bruce E. Cain & Janet C. Campagna, *Predicting Partisan Redistricting Disputes*, 12 LEGIS. STUD. Q. 265, 269 (1987) (noting state redistricting plans in the early 1980s).

By comparison, Democrats in the immediate post-*Baker* period were more conflicted about legislative redistricting, in part because they controlled more state legislatures than the Republicans.[8] Hence, in the 1980s, the Democrats were plaintiffs in *Davis v. Bandemer*[9] and defendants in *Badham v. Eu*.[10] By the 1990s, political conditions further dampened partisan redistricting concerns for both parties.[11] Divided government had become more prevalent in the states, blocking partisan designs and incentivizing incumbency protection plans.[12] The focus was redirected instead to how redistricting contributes to lower levels of competitiveness and bipartisan incumbent lockups.[13]

Political circumstances and party positions have changed again in recent years.[14] Partisan polarization has increased sharply, raising the stakes of political contestation for state and federal offices.[15] Thirty-seven states still allow their state legislatures to redraw congressional lines.[16] The Republicans, not the Democrats, currently dominate the state legislatures and governorships.[17] The concerns

8. *See* Philip Bump, *How Your State's Politics Have Shifted over the Years, in 49 Charts*, WASH. POST (Sept. 13, 2015), https://www.washingtonpost.com/news/the-fix/wp/2015/09/11/49-charts-that-tell-the-partisan-history-of-state-legislatures/?utm_term=.28bdc1c348df [https://perma.cc/P3DF-JFX6] (graphing the shifts in the members of state legislatures by party, both generally and specifically by state).

9. 478 U.S. at 113.

10. 694 F. Supp. 664, 665 (N.D. Cal. 1988), *aff'd*, 488 U.S. 1024 (1989).

11. *See generally* Mark D. Brewer, *The Rise of Partisanship and the Expansion of Partisan Conflict Within the American Electorate*, 58 POL. RES. Q. 219, 219 (2005) (noting the resurgence of partisanship in the 1990s).

12. *See id.* at 220; *see also* Mark A. Posner, *Post-1990 Redistrictings and the Preclearance Requirement of Section 5 of the Voting Rights Act*, *in* RACE AND REDISTRICTING IN THE 1990S 80, 106 (Bernard Grofman ed., 1998) (explaining the impact incumbency protection concerns had on almost all post-1990 plans).

13. *See generally* Michael Lyons & Peter F. Galderisi, *Incumbency Reapportionment, and U.S. House Redistricting*, 48 POL. RES. Q. 857 (1995) (arguing that bipartisan redistricting plans in the 1992 election benefitted incumbents more than partisan redistricting plans).

14. *See* Edward G. Carmines & Matthew Fowler, *The Temptation of Executive Authority: How Increased Polarization and the Decline in Legislative Capacity Have Contributed to the Expansion of Presidential Power*, 24 IND. J. GLOBAL LEGAL STUD. 369, 369-70 (2017).

15. *See id.*

16. *National Overview of Redistricting: Who Draws the Lines?*, BRENNAN CTR. FOR JUST. (June 1, 2010), https://www.brennancenter.org/analysis/national-overview-redistricting-who-draws-lines [https://perma.cc/ZH84-TRTF].

17. *State Partisan Composition*, NAT'L CONF. ST. LEGISLATURES (Aug. 1, 2017), http://www.ncsl.org/research/about-state-legislatures/partisan-composition.aspx [https://perma.cc/BKF8-D236].

of the Democratic Party's good government faction in reforming the redistricting process now more closely align with the interests of the Party's pragmatists in hedging against the likely Republican advantage in the 2021 round of legislative redistricting.[18] Since the probability of a partisan plan increases with single-party control of the executive and legislative branches and decreases with divided government,[19] any imbalance in single-party control at the state level could easily translate into decades-long electoral advantage for the Republicans in the future. The short-term political answer for the Democrats is to win back enough state legislative seats and governorships in order to block adverse partisan plans in 2021.[20] But given heightened partisanship and increasing exploitation of voting laws for electoral advantage in the current era, there is also a renewed interest in finding ways to curb partisan gerrymandering.[21] The long-term legal project of persuading the courts to be more interventionist in these matters requires diving back into and resolving the thorny controversy over whether there is a manageable standard to identify unconstitutional partisan plans.[22]

Despite several decades of efforts by academics, reformers, and a few political officials, there is still no consensus in the United States about how best to measure and judge the partisan fairness of any proposed districting plan.[23] While it is a straightforward calculation to identify seats-votes gaps at the end of a decade,[24] it is more problematic to project them with a high degree of certainty into the future when the district lines have just been drawn. Scholars have offered a number of suggestions for measuring political fairness,

---

18. *See generally* Amber Phillips, *The 2020 Redistricting War Is (Already) on*, WASH. POST (July 16, 2015), https://www.washingtonpost.com/news/the-fix/wp/2015/07/16/the-2020-redistricting-war-is-on/?utm_term=.565c9749e634 [https://perma.cc/8RH3-3Q2M] (discussing the contentious race to control the redrawing of district lines and the Democratic desire to block another Republican victory).

19. *See* Cain & Campagna, *supra* note 7, at 267, 271.

20. *See* Phillips, *supra* note 18.

21. *See* Brewer, *supra* note 11, at 219-20.

22. *See infra* Part I.

23. *See infra* Part I.B.

24. *See generally* Richard G. Niemi, *The Relationship Between Votes and Seats: The Ultimate Question in Political Gerrymandering*, 33 UCLA L. REV. 185, 191-92 (1985) (noting the *Bandemer* Court's "casual[]" reference to the seats-votes gap after the election).

including seats-votes bias, responsiveness, competitiveness, proportionality, and more recently, the efficiency gap.[25] Yet how these measures relate to one another is poorly understood: Are they duplicate measures that examine the same underlying political phenomenon, or do they tap different facets of partisan fairness? And if there is no silver bullet measure, what judicial framework is then possible for that messy reality?

We describe an important innovation in automated plan generation, aided by rapid computing advances, that holds enormous promise for the enterprise of redistricting. Automated map generation technologies will continue to improve and become more accessible in the future. We provide an example of how they could be used with the Parallel Evolutionary Algorithm for Redistricting (PEAR), the most advanced automated redistricting algorithm to date.[26] PEAR is able to utilize more than a hundred thousand processor cores on the Blue Waters supercomputer, the fastest research supercomputer in the world.[27] Though much progress has been made in the development and refinement of this important tool, its potential—and extensive—applications are still being developed.[28] But even now, it proves to be both a powerful and flexible tool. In this Article, we illustrate its enormous value with two applications: to better understand how the various partisan fairness measures interact with each other, and to draw maps that are optimized along many such measures.

---

25. *See infra* notes 100-04 and accompanying text.

26. Yan Y. Liu et al., *PEAR: A Massively Parallel Evolutionary Computation Approach for Political Redistricting Optimization and Analysis*, SWARM & EVOLUTIONARY COMPUTATION, Oct. 2016, at 78, 79. While we believe that PEAR is the most advanced approach, other well-respected scholars have also been working to develop automated redistricting procedures. *See* Micah Altman & Michael P. McDonald, *BARD: Better Automated Redistricting*, 42 J. STAT. SOFTWARE, June 2011, at 1, 2; Jowei Chen & Jonathan Rodden, *Cutting Through the Thicket: Redistricting Simulations and the Detection of Partisan Gerrymanders*, 14 ELECTION L.J. 331, 332 (2015) [hereinafter Chen & Rodden, *Cutting Through the Thicket*]; Jowei Chen & Jonathan Rodden, *Unintentional Gerrymandering: Political Geography and Electoral Bias in Legislatures*, 8 Q. J. POL. SCI. 239, 242, 248 (2013) [hereinafter Chen & Rodden, *Unintentional Gerrymandering*]; Benjamin Fifield et al., A New Automated Redistricting Simulator Using Markov Chain Monte Carlo 1-2 (Mar. 15, 2017) (unpublished manuscript), http://imai.princeton.edu/research/files/redist.pdf [https://perma.cc/QSH9-VDXD]. The friendly rivalry of these efforts gives us optimism about the future development of these tools.

27. *About Blue Waters*, NAT'L CTR. FOR SUPERCOMPUTING APPLICATIONS, http://www.ncsa.illinois.edu/enabling/bluewaters [https://perma.cc/KC3U-FK8H].

28. *See* Liu et al., *supra* note 26, at 90.

Our themes are as follows. First, we show how PEAR enables us to generate a very large set of constitutionally feasible redistricting plans in a given state.[29] We can then identify extreme outcomes in a reasonable bias framework by comparing the partisan bias scores of any proposed redistricting plan with the range of scores in the automated plans.[30]

Second, we use PEAR to study the intersection of the various proposed partisan fairness measures, including seats-votes bias, responsiveness, competitiveness, proportionality, and the more recently proposed efficiency gap.[31] It appears that the concept of political fairness, like compactness, is multidimensional and cannot be fully captured by a single number, which implies that any systematic evaluation of partisan bias should utilize several political fairness measures.[32] At a minimum, this would include at least one measure of partisan bias plus a measure of competitiveness (or responsiveness) as indicators of possible partisan or bipartisan gerrymandering.[33]

Third, we explore the value of a plural measures approach, more affectionately known as the "everything bagel" approach.[34] PEAR can obviate the need to choose one measure of political fairness to the exclusion of others.[35] That said, this does not mean that courts or any other redistricting bodies can escape the need to make value judgments about the relative weight and threshold values of political fairness measures, or trade-offs between political fairness and other redistricting goals such as preserving local district boundaries, respecting communities of interest, and the like.[36]

---

29. *See infra* notes 82-89 and accompanying text.

30. *See infra* notes 90-92 and accompanying text.

31. *See infra* Part I.B.

32. *See infra* Part I.B.

33. A partisan gerrymander can be biased toward one party with varying degrees of competitiveness. *See* Wendy K. Tam Cho & Yan Y. Liu, *Toward a Talismanic Redistricting Tool: A Computational Method for Identifying Extreme Redistricting Plans*, 15 ELECTION L.J. 351, 362-63 (2016). A bipartisan plan can be either noncompetitive (for example, an incumbent gerrymander) or competitive (the so-called fair, ideal plan). *See id.*

34. *See infra* Parts I.C-D.

35. *See infra* Part I.C.

36. *See infra* note 59 and accompanying text.

Lastly, in the Appendix to this Article, we demonstrate how a reasonable bias approach could be utilized to identify a top set of feasible alternatives with an example drawn from Minnesota.[37]

I. OUTLIER ANALYSIS AND AUTOMATED PLAN GENERATION

A manageable partisan bias test requires several elements. First and foremost, it needs an agreed-upon political fairness measure. Although scholars have proposed several, the Supreme Court has not chosen one, only ruling that proportional representation is not necessarily implied by the Equal Protection Clause of the Constitution.[38] The most commonly discussed measures derive from a symmetry concept: that is, in the same circumstances, partisan unfairness should not disadvantage one party more than the other.[39] King-Gelman bias scores[40] and the efficiency gap are the most prominent examples of partisan symmetry measures.[41]

Beyond having measures that are mathematically and operationally sound, the second element of a manageable partisan bias test is a specific threshold value that separates acceptable redistricting from unconstitutional partisan gerrymandering. When the Court decided that "one person, one vote" was the underlying principle for the weighting of each person's vote,[42] it had to determine how much

---

37. Bruce E. Cain et al., *Appendix: A Reasonable Bias Approach to Gerrymandering: Using Automated Plan Generation to Evaluate Redistricting Proposals*, 59 WM. & MARY L. REV. ONLINE 103 (2018).

38. *See* Whitcomb v. Chavis, 403 U.S. 124, 156-60 (1971); *see also* Davis v. Bandemer, 478 U.S. 109, 130-32 (1986) (plurality opinion) (reaffirming the *Whitcomb* holding).

39. *See infra* notes 52-58 and accompanying text (discussing the difficulties in achieving this symmetry concept).

40. *See generally* Andrew Gelman & Gary King, *Estimating Incumbency Advantage Without Bias*, 34 AM. J. POL. SCI. 1142 (1990) (explaining King-Gelman bias scores as a means of measuring incumbency advantage).

41. The plaintiffs in *Whitford v. Gill* proposed the efficiency gap (EG) as a viable test of partisan symmetry. 218 F. Supp. 3d 837, 854 (W.D. Wis. 2016) (three-judge court), *argued*, No. 16-1161 (U.S. Oct. 3, 2017). As Professor Cho has demonstrated elsewhere, however, this formulation has many odd properties. *See* Wendy K. Tam Cho, *Measuring Partisan Fairness: How Well Does the Efficiency Gap Guard Against Sophisticated as Well as Simple-Minded Modes of Partisan Discrimination?*, 166 U. PA. L. REV. ONLINE 17 (2017). For instance, because it includes both winning and losing wasted votes, it assigns the same efficiency gap value to very different distributions of partisan strength, and also counterintuitively penalizes plans with competitive districts. *See id.*

42. *See* Gray v. Sanders, 372 U.S. 368, 381 (1963) (Stewart, J., concurring) (setting out the

a district population could vary from perfect equality.[43] In the end, it settled on an "as nearly as practicable" rule for congressional districts[44] and a total population deviation of 10 percent for state and local districts.[45] Similarly, such a threshold value is required of a partisan fairness measure that can separate acceptable from unacceptable partisan bias. Even if one were to say that a particular political outcome measure should be optimized to its highest value—the equivalent of determining that population deviation should be as close to zero as possible—it would still be necessary to know what is possible in order to know whether any given plan is at or even near this optimal value.

In the case of population equality, the calibration of the exact threshold was worked out over time in a series of cases.[46] In other words, the congressional and state/local population deviation thresholds were not handed down on a stone tablet from on high, but rather evolved over a number of decisions into a gradual consensus

---

"one person, one vote" principle); *see also, e.g.*, Reynolds v. Sims, 377 U.S. 533, 558 (1964) (reiterating the same principle stated in Justice Potter Stewart's concurrence in *Gray*).

43. *See, e.g.*, Swann v. Adams, 385 U.S. 440, 443-44 (1967) (requiring a "satisfactory explanation" for deviations of 30 percent to 40 percent, but noting that "*[d]e minimis* deviations are unavoidable"); *Reynolds*, 377 U.S. at 579 ("[S]ome deviations from the equal-population principle are constitutionally permissible.").

44. *See* Kirkpatrick v. Preisler, 394 U.S. 526, 528 (1969).

45. *See* Brown v. Thomson, 462 U.S. 835, 842-43 (1983); Mahan v. Howell, 410 U.S. 315, 322, 324-25 (1973) (holding that the Equal Protection Clause merely requires a state to "make an honest and good faith effort to construct districts ... as nearly of equal population as is practicable" for its state and local districts, rather than the more stringent standard for congressional districts). *But see* Larios v. Cox, 300 F. Supp. 2d 1320, 1325-28, 1347 (N.D. Ga. 2004) (three-judge court) (per curiam) (discussing the 10 percent "safe harbor" but still striking down a "blatantly partisan and discriminatory" plan that fell within a 10 percent deviation), *aff'd*, 542 U.S. 947 (2004).

46. There was—and still is—disagreement with political implications over whether the right data in drawing equipopulous districts ought to be total population, voting age population, citizen voting age population, or even total registered voters. *See* Evenwel v. Abbott, 136 S. Ct. 1120, 1123 (2016) (holding that states can draw districts based on total population); Chen v. City of Houston, 206 F.3d 502, 505 (5th Cir. 2000) (holding that plaintiffs had not met their burden of proof in showing that counting based on potential eligible voters was the proper districting measure); Garza v. County of Los Angeles, 918 F.2d 763, 774-75 (9th Cir. 1990) (deciding that, due to discriminatory effects, drawing based on total population was more appropriate than using voting population, despite the argument that persons ineligible to vote due to age or citizenship should not be counted). There is more uncertainty in state and local population variance than in congressional election. *See Reynolds*, 377 U.S. at 577-78 (explaining the different population deviation standards for congressional and state districting, and noting that more flexibility is given to state districting).

that balanced the flexibility needed to accommodate other redistricting goals with a population constraint strict enough to prevent underpopulating or overpopulating districts to one party's advantage over another.[47]

Partisan fairness measures are more complex than population deviation measures. Ideally, the critical thresholds of various partisan bias measures would be uniform nationally, but such a task is made difficult by geographic and demographic realities.[48] Partisans are not randomly dispersed across geography.[49] Rather, they cluster in nonrandom ways, causing redistricting to produce natural partisan bias.[50] The clearest and most intuitive example of this is in the cities. The concentration of Democrats in urban areas makes districts comprising these usually compact areas "natural" Republican gerrymanders.[51] Democrats often win districts in cities by a supermajority of votes, thereby wasting votes that, if cast in more competitive districts, might help them win more seats.[52]

This problem is well documented.[53] It is also one that the Court has grappled with. Justice Anthony Kennedy's concurrence in *Vieth* raised the concern that even neutral redistricting principles like contiguity and compactness "would unavoidably have significant political effect, whether intended or not."[54] In the area of redistricting, even criteria "neutral enough on its face, would ... benefit one

47. Systematically overpopulating the opposing party's districts and underpopulating one's own party's districts is in effect the original malapportionment abuse that led to the Court intervening in *Baker v. Carr*. *See* 369 U.S. 186, 192 (1962). Systematically overpopulating and underpopulating districts along party lines within the allowable deviations is potentially an equal protection violation. *See Larios*, 300 F. Supp. 2d at 1334, 1338.

48. *See* Peter H. Schuck, *The Thickest Thicket: Partisan Gerrymandering and Judicial Regulation of Politics*, 87 COLUM. L. REV. 1325, 1353 (1987).

49. *See* Chen & Rodden, *Unintentional Gerrymandering*, *supra* note 26, at 240-41.

50. *See id.*

51. *See* Vieth v. Jubelirer, 541 U.S. 267, 289-90 (2004) (plurality opnion).

52. *See* Nate Cohn, *Why Democrats Can't Win the House*, N.Y. TIMES (Sept. 6, 2014), https://www.nytimes.com/2014/09/07/upshot/why-democrats-cant-win.html?mcubz=0 [https://perma.cc/G6F2-E72E] (discussing the Democratic tendency to win cities by a large margin and yet fail to take control of the House); *see also* Bruce E. Cain, *Assessing the Partisan Effects of Redistricting*, 79 AM. POL. SCI. REV. 320, 321 (1985).

53. *See* Micah Altman, *Modeling the Effect of Mandatory District Compactness on Partisan Gerrymanders*, 17 POL. GEOGRAPHY 989, 1000-06 (1998); Chen & Rodden, *Unintentional Gerrymandering*, *supra* note 26, at 240-41.

54. *Vieth*, 541 U.S. at 308-09 (Kennedy, J., concurring in the judgment).

political party over another."[55] Justice Antonin Scalia, in the plurality opinion declining to find discernable and manageable standards for partisan gerrymandering claims, also spoke to the inherent difficulty of detecting partisan influence in a redistricting plan.[56] He raised the example of the 2000 Pennsylvania congressional map that produced a partisan effect against the Democrats, even though the map was drawn "free from partisan gerrymandering."[57] "Whether by reason of partisan districting or not, party constituents may always wind up 'packed' in some districts and 'cracked' throughout others."[58]

While the Court could take the view that a constitutional partisan gerrymandering doctrine ought to correct for imbalances in the way partisans are distributed across space, it is more likely that the Court will find that natural gerrymanders are a permissible price of our redistricting regime. Our system of districting sometimes necessitates trading off values like seats-votes proportionality for other values, such as the ability to reflect localized and community interests.[59] If we redistrict, we will have to accept some extant demographic patterns as a given.[60] In that case, a measure of partisan effect must be able to distinguish between plans that are merely a product of geography and those that are the product of intentional partisan manipulation. In the vast majority of cases, both are likely to be in play. Therefore, a high-functioning measure of partisan

55. *Id.* at 309 ("District lines are rarely neutral phenomena. They can well determine what district will be predominantly Democratic or predominantly Republican, or make a close race likely." (quoting Gaffney v. Cummings, 412 U.S. 735, 753 (1973))); *see also* ROBERT H. BORK, THE TEMPTING OF AMERICA: THE POLITICAL SEDUCTION OF THE LAW 88-89 (1990) (documenting the author's service as a special master responsible for redistricting Connecticut, and noting that his final plan so benefited the Democratic Party—albeit unintentionally—that the party chairman personally congratulated him); Altman, *supra* note 53, at 1000-06 (explaining that compactness is not a neutral standard, especially when groups have distinct geographic distributions, as is the case with Democrats who are more likely to live in high-density regions).

56. *See Vieth*, 541 U.S. at 289 (plurality opinion).

57. *Id.*

58. *Id.* (citing ROBERT G. DIXON, JR., DEMOCRATIC REPRESENTATION: REAPPORTIONMENT IN LAW AND POLITICS 462 (1968)); Schuck, *supra* note 48, at 1359.

59. *See* Cho & Liu, *supra* note 33, at 355 (explaining the need for trade-offs to develop a plan that satisfies the conflicting voices within a geographic area).

60. *See, e.g.*, *Vieth*, 541 U.S. at 289 (plurality opinion).

effect must be able not only to parse out natural gerrymanders from unnatural ones, but also to *quantify* the effects of each.

Cross-sectional and historical applications of partisan bias fail in this crucial regard. The problem with both is that districts do not start out on the same playing field. Consider, for instance, the problem of comparing urban versus rural or suburban areas on a competitiveness measure: districts in cities are often naturally uncompetitive due to the predominance of city-dwelling liberals and minority groups that live there.[61] A plan that respects the urban community of interests may score poorly on the competitiveness metric through no fault of those who draw the lines. Indeed, given the circumstances, even the most evenhanded line-drawer would produce a highly uncompetitive district.[62]

No doubt line-drawers *could* be responsible for making an otherwise competitive district uncompetitive or vice versa. But, as it pertains to our discussion of what a functioning measure of partisan fairness must be able to accomplish, knowing that a district scores poorly on a competitiveness measure does not inform whether the outcome is a product of partisan manipulation or merely reflects the underlying political geography and communities of interest.[63] And while the two scenarios can be distinguished with the help of other evidence, courts are still unable to determine whether impermissible partisan gerrymandering is responsible for the observed partisan bias in the challenged plan, and if so, by how much.[64]

This problem is no different with any proposed measure of partisan fairness. Consider the more recently advocated efficiency gap.[65] An urban district likely uses Democratic votes inefficiently. Any Democratic votes above and beyond the 50 percent necessary to win are more efficiently spent in other more competitive districts, and all votes in a losing district are considered wasted as well.[66] Like the above example, a poor efficiency gap fulfills only a descriptive purpose, but not a diagnostic one. One is left, once again, to rely on

61. *See supra* notes 51-52 and accompanying text.

62. *See supra* note 55 and accompanying text.

63. *See supra* notes 49-58 and accompanying text.

64. *See supra* notes 42-45 and accompanying text.

65. *See* Whitford v. Gill, 218 F. Supp. 3d 837, 854 (W.D. Wis. 2016) (three-judge court), *argued*, No. 16-1161 (U.S. Oct. 3, 2017).

66. *See* Schuck, *supra* note 48, at 1359-60.

extrinsic evidence in determining whether the partisan outcome is a result of foul play or simply a natural consequence of the underlying demographics of the district.[67]

Indeed, the efficiency gap—or any other metric for that matter—is not only flawed in this absolute sense, but also uninformative even when used comparatively. Knowing that plan *A* has a higher or lower efficiency gap score than plan *B* does not inform whether the score in either plan is more likely to be produced by impermissible partisan gerrymandering. Without knowing about the natural level of bias in any given geographic area, the scores are unreliable as even a comparative measure of partisan bias.[68]

To be sure, there is a scenario in which the existing measures would be useful: comparing two maps with exactly the same base population. As we will later demonstrate, that is the intuition behind the logic of automated redistricting simulation.[69] However, recognizing the cross-state variance in political geography makes the task of finding a uniform national standard a bit more complex, but not necessarily impossible. It means looking for extreme departures from the mean of a given distribution rather than a magic number on any or all fairness scores that would apply to all jurisdictions across the country.

To elaborate, a viable partisan fairness measure sorts maps into two buckets—constitutionally permissible and impermissible—by producing a cutoff: plans that score extremely poorly on the metric are deemed constitutionally impermissible.[70] But applying the same cutoff to plans that contain vastly different political geography can be misleading. Some places in the country may naturally score poorly on any of the proposed partisan bias metrics. Applying a national cutoff to them would mean that line-drawers in those states would

---

67. *See supra* notes 57-58 and accompanying text.

68. The same thing is true of the measure of partisan symmetry. Indeed, the efficiency gap is a measure of partisan symmetry. *See Whitford*, 218 F. Supp. 3d at 947. But once again, plans may be naturally asymmetrical or symmetrical. A comparison of partisan symmetry scores does not inform whether those scores are produced naturally, or as a result of impermissible gerrymandering.

69. *See* Cho & Liu, *supra* note 33, at 354.

70. *See, e.g.*, *id.* at 360 (explaining how a set of "legal maps" was culled to retain only "feasible map[s]").

have less room to maneuver between drawing a permissible or impermissible plan.

This could have perverse consequences in several directions. A national average that includes states with little or no minority population concentrations could serve as an excuse to split up some existing minority influence districts or to avoid creating new ones. Conversely, other places might naturally score well on a partisan bias metric, providing them with the knowledge and incentive to create mischief up to the allowable level. For these places, the cutoff effectively provides them with a partisan gerrymandering credit.

Setting a very high cutoff nationally for all states in order to provide more breathing space for states with concentrated urban populations is also no solution. Aside from creating yet another way to engage in partisan mischief by allowing packing levels that are not warranted by the underlying political geography, it offends the fundamental principle of fair treatment in a federal system by imposing unrealistic expectations that derive from the political conditions in other states.[71]

The stakes of developing a standard that accounts for natural gerrymanders are thus not merely academic, but rather, are at the core of our federal system of government. States, heterogeneous in their geography and demographics, must nevertheless be held to a uniform standard for redistricting purposes. Unconstitutional partisan gerrymandering, if recognized, must proscribe the same underlying bad conduct across the nation. Such is the challenge in developing a constitutionally cognizable standard for partisan fairness: How to both maintain a uniform national standard, while accounting for gross heterogeneity in circumstances faced by state legislatures?

The key to being able to do both is to develop a measure that accounts for the effects of natural gerrymanders. Put otherwise, if we can, in effect, control for the effects attributable to a particular district's underlying demographics, then districts can be compared with one another on the same level playing field.[72] In any case, being able to isolate the natural gerrymandering effect bias will

71. *See* Gray v. Sanders, 372 U.S. 368, 380 (1963).

72. *See, e.g.*, Cho & Liu, *supra* note 33, at 359.

help determine how much of the observed partisan bias is, in fact, attributable to impermissible gerrymandering.

Such an effect can be isolated if we can determine the counterfactual: Where would the district lines be drawn in the absence of partisan gerrymandering?[73] Knowing such a state of the world would allow us to calculate the distance between the counterfactual and the observed state of the world.[74] The difference between the observed and the counterfactual plan could then be attributable to artificial redistricting influences, such as the partisan biases of the line-drawers.[75] Deriving such counterfactual alternatives is therefore crucial to attributing and quantifying the observed bias in any proposed plan.

Before we consider automated plan generation and the role it could play in determining the counterfactual, what alternative tools currently exist in redistricting litigation that might do the job? A neutral expert, preferably blind to the actually produced map, could draw a counterfactual plan.[76] The goal of such an exercise would be to determine what the plan *would have looked like* had there not been any partisan influence. Indeed, courts commonly employ special masters in redistricting cases to draw remedial maps.[77] Their expertise ensures that they are competent in drawing maps, and their theoretic neutrality ensures that partisan motivations do not influence those maps.[78]

---

73. *See id.* (defining the counterfactual set of maps as "the set of plans that are at least as good or better on non-partisan factors because there are known considerations, but do not consider partisanship" (emphasis omitted)).

74. *See id.*

75. *See id.*

76. A variant of this that we will not discuss in any detail is to have the neutral redistricting team draw maps that reveal the extremes with respect to different redistricting values. So with respect to partisan bias, this would mean purposefully trying to create the most plans that most favor one party and then the other. *See, e.g.*, Bruce E. Cain et al., *Sorting or Self-Sorting: Competition and Redistricting in California?*, *in* THE NEW POLITICAL GEOGRAPHY OF CALIFORNIA 245, 247 (Frédérick Douzet et al. eds., 2008). This approach has the advantage of revealing the outside limits of partisan bias, but the disadvantage of not revealing the overall distribution of bias across all possible plan options. Moreover, automated plan generation could do this exercise much more efficiently and quickly than individuals.

77. *See, e.g.*, DeWitt v. Wilson, 856 F. Supp. 1409, 1410-11 (E.D. Cal. 1994) (three-judge court), *aff'd in part, dismissing appeal in part*, 515 U.S. 1170 (1995); Puerto Rican Legal Def. & Educ. Fund, Inc. v. Gantt, 796 F. Supp. 681, 685 (E.D.N.Y. 1992) (three-judge court) (per curiam).

78. *See Puerto Rican Legal Def. & Educ. Fund, Inc.*, 796 F. Supp. at 685 (discussing the

Yet, relying on an expert to determine such a counterfactual has crippling pitfalls. As the experience with independent redistricting commissions and even court panels demonstrates, neutrality is especially elusive when it comes to redistricting.[79] More critically, a data point of one can be highly unreliable and idiosyncratic, especially in an enterprise as complex as redistricting. Many—easily millions of viable and neutral maps—could be drawn.[80] How would a court know whether the particular map drawn by a neutral expert is typical of the kind of map that neutral experts, as a whole, would draw?[81]

Ideally, then, a court would have at its disposal a survey of all the maps that all neutral experts would draw. Instead of finding the single-but-elusive counterfactual, the spectrum of viable maps better constitutes the baseline or natural set of maps against which to compare the observed map.[82] But hiring one special master, let alone several, comes at a steep cost.[83] Hiring a group of them to draw maps and then to take an average, therefore, only makes sense in principle and certainly not in practice.

This is where automated plan generation can offer some unique advantages. A computer program essentially substitutes for a very large body of neutral experts and the viable, neutral maps they draw.[84] By programming neutral redistricting criteria, such as the preservation of extant communities, compactness, contiguity, and adherence to one-person, one-vote guidelines,[85] a computer algorithm can generate a very large set of neutral redistricting plans that by design are not influenced by partisanship, and take as given the natural gerrymandering effect—if any—of the underlying

---

appointed special master's recognized need for particular expertise from a university professor).

79. *See generally* Bruce E. Cain, *Redistricting Commissions: A Better Political Buffer?*, 121 YALE L.J. 1808 (2012) (arguing that, while successful to some degree, independent citizen commissions have not eliminated partisan suspicions).

80. *See* Liu et al., *supra* note 26, at 78 (noting the significant number of plans that could be drawn within legal parameters).

81. Formulating a baseline with the help of many, as opposed to a single or a small set of maps, is analytically crucial.

82. *See* Cho & Liu, *supra* note 33, at 353-54.

83. *Cf. Puerto Rican Legal Def. & Educ. Fund, Inc.*, 796 F. Supp. at 685 (noting the need to hire experts beyond the appointed special master due to significant time restraints).

84. *See* Liu et al., *supra* note 26, at 79, 89.

85. *See, e.g.*, *id.* at 79-81, 83, 91.

demographics.[86] Indeed, the algorithm is better than a large group of experts. Humans, despite all attempts to be neutral, may inject subconscious and latent biases. The algorithm operates purely based on the parameters that are imposed. And, as long as the algorithm is transparent, courts, scholars, and litigants can examine and critique the soundness of the parameters used in the map generation algorithm.

When we refer to a very large set of redistricting plans, we truly mean a *very* large set of such plans. Redistricting is an activity containing many degrees of freedom: there is an astronomically large number of redistricting plans that satisfy basic redistricting principles.[87] Most of this set is uninteresting to us because while these redistricting plans meet minimal legal requirements, they do not represent the set of plans a human might draw.[88] PEAR, however, is able to produce not simply large numbers of random plans, but random *high-quality* redistricting plans.[89] Using this corpus of reasonably imperfect plans (in other words, plans that meet or exceed the threshold values on a set of redistricting fairness criteria),[90] we can derive a good understanding of what the viable counterfactual options are in a given state at a given time, and begin to define how far out on the tail of the distribution a plan can go before it is deemed extreme.[91] This same exercise can be used for any of the criteria, including partisan bias and competitiveness.[92]

The core contribution of an automated plan generation approach is producing a large set of legally viable maps with respect to multiple criteria. Once the corpus has been generated, there are many potential ways to use it in determining the partisan effects of the challenged plan. Without claiming to exhaust all possibilities, the next Section of this Article attempts to map out some of the possible ways forward.

---

86. *See id.* at 91.
87. *See* Cho & Liu, *supra* note 33, at 354-55.
88. *See id.* at 355.
89. *See* Liu et al., *supra* note 26, at 79.
90. *See* Cho & Liu, *supra* note 33, at 355.
91. *See id.* at 363-64.
92. *See* Liu et al., *supra* note 26, at 89.

A. Getting to a Reasonable Partisan Bias Standard

Developing a rule for courts to detect gerrymandering in legislative plans will likely have to follow the path of the population deviation rules: that is, a case-by-case evolution of a benchmark definition of the distribution tail given the natural disparities in any particular geopolitical setting.[93] Let us say, for example, that a court was looking at a rough seats-votes proportionality measure as one of its indications of bias. One possibility is simply to compare the seat share in the challenged plan to those in the counterfactuals. The counterfactual plans may contain plans with a wide variety of projected seat shares. But knowing where the challenged plan stands compared to the counterfactuals will be informative in discovering whether the seat share observed in the challenged plan is unusual or pedestrian. If the vast majority of plans generated by neutral redistricting principles would produce the same seat share as that in the challenged plan, then one cannot claim that the challenged plan has a partisan effect above and beyond what the geography and demographics would produce. If, on the other hand, the challenged plan has a seat share that is rarely seen in the computer-generated plans, that is convincing evidence that the observed bias in the challenged map is unlikely to be a product of the underlying geography.[94] One could do a similar comparison with the other measures of partisan bias[95] by asking the question of how far the proposed plan score is from the mean score of the automated plan distribution, expressed perhaps in terms of standard deviations.

Adopting this approach, critical decisions would have to be made about how to draw the line between permissible and impermissible deviations from the counterfactual plans. Any partisan gerrymandering doctrine that the Court adopts will presumably allow states to draw maps that deviate *some* from the counterfactual plans.[96] Strict adherence is not likely to be required. The critical question in

93. *See* Cho & Liu, *supra* note 33, at 364.
94. *See* Liu et al., *supra* note 26, at 89.
95. *See id.* at 80-82, 89-90.
96. *See* Cho & Liu, *supra* note 33, at 363-64.

applying this method then becomes: How much deviation is too much?

Like in many areas of law, courts face an age-old dilemma in whether to choose rules or standards.[97] While any rule would be inevitably arbitrary, it has the virtues of clarity, transparency, and predictability, enabling line-drawers to conform ex ante. One example of a rule for identifying partisan extremity could be: a redistricting plan cannot yield an expected seats-votes share or partisan bias score that occurs in only 5 percent or less of the machine-generated maps.[98] An expected seat share or bias score that far from the mean would be good evidence that the map has a partisan effect that is both extreme and unnecessary given the large number of viable alternatives. Moreover, the partisan bias in the challenged map would demonstrably not be a result of the underlying political geography, as the comparison plans would be based on the same area.[99]

What if courts decide to go with standards instead? For instance, the height of the cutoff could depend on the existence or persuasiveness of additional evidence. Where the challenged map stands relative to the simulated maps could be considered simply as one piece of evidence in a more holistic analysis of partisan intent, actions line-drawers may have taken, and features of the challenged maps. While standards provide less predictability, risk aversion on the part of line-drawers may in fact provide an incentive for more responsible redistricting outcomes.

---

97. For a proposed definition of rules versus standards, see Louis Kaplow, *Rules Versus Standards: An Economic Analysis*, 42 DUKE L.J. 557, 559-60 (1992).

98. One could also imagine having a safe-harbor percentile within which a map could not be deemed to have partisan effects.

99. Another possibility is to focus not on the percentile, but rather, on the absolute distance between the vote share of the median map generated by the simulation approach and that of the challenged map. Suppose the median map generated by simulation generates five Republican seats and five Democratic seats, and the challenged map generates one Republican seat and nine Democratic seats. Courts would determine that a partisan effect exists if a certain distance between the simulated and observed outcome is breached. But any such rule devised must be scaled to treat districts of different sizes with different numbers of seats similarly.

B. Evaluating Existing Measures of Partisan Fairness

In addition to using PEAR as a simulation tool, it can also serve the scholarly community in providing better analyses of redistricting. We present one such application by using automated plan generation to ascertain how different measures of political fairness that scholars have proposed over the years compare to one another.

There is a rich and growing academic interest in the development of a measure of partisan fairness. Scholars have offered a number of suggestions, including seats-votes bias,[100] responsiveness,[101] competitiveness,[102] proportionality,[103] and more recently, the efficiency gap.[104] This is in part fueled by the Court's sustained ambivalence over whether to accept any existing measure as a standard of partisan fairness for partisan gerrymandering claims.[105] Over the years, this has produced a panoply of partisan fairness measures, some purporting to have solved the golden riddle.

However, the growth in the number of measures proposed is not matched by the size of an evaluative literature on how the measures relate to one another. Are they duplicative or overlapping measures, or do they tap different facets of partisan fairness? We find that, like compactness measures, the choice of different measures results in widely varying outcomes.[106] There is not as much overlap between efficiency, bias, and proportionality, or responsiveness and competitiveness as one might expect.[107] Each individual measure, instead of singularly capturing the concept of partisan fairness, may simply be picking up on only one facet of the ideal.

---

100. *See* Cain, *supra* note 52, at 321-23; Bernard Grofman, *Measures of Bias and Proportionality in Seats-Votes Relationships*, 9 POL. METHODOLOGY 295, 296 (1983); Gary King & Robert X. Browning, *Democratic Representation and Partisan Bias in Congressional Elections*, 81 AM. POL. SCI. REV. 1251, 1251-53 (1987); Niemi, *supra* note 24, at 186.

101. *See, e.g.*, Cho & Liu, *supra* note 33, at 360, 362, 364.

102. Cain & Campagna, *supra* note 7, at 266-67, 271, 273; Cho & Liu, *supra* note 33, at 362-63.

103. *See, e.g.*, Cho & Liu, *supra* note 33, at 360-62.

104. *See* Nicholas O. Stephanopoulos & Eric M. McGhee, *Partisan Gerrymandering and the Efficiency Gap*, 82 U. CHI. L. REV. 831, 834, 837-38 (2015).

105. *See supra* note 38 and accompanying text.

106. *See infra* Figure 1.

107. *See infra* Figure 1.

While there are doubtless other ways to study the interaction between the various measures of partisan fairness, we employ PEAR for this exercise. Again, PEAR is able to draw a set of maps that comport with neutral standards of redistricting, such as compactness, contiguity, and equal population.[108] We run the algorithm on Minnesota's congressional redistricting. What is new here compared to the use of PEAR in producing simulated counterfactuals is that in addition to drawing maps that satisfy neutral criteria, we explicitly include political fairness measures.

What the algorithm produces, as a result, is a very large set of viable maps that perform well on the one measure of partisan fairness that we have included in the algorithm.[109] Say, for instance, we choose competitiveness as the measure of interest: the algorithm then generates a very large corpus of maps that score well on competitiveness. We can repeat this process for other fairness measures: seats-vote bias, responsiveness, proportionality, and the efficiency gap. We then have five sets of maps, each of which scores high on one of the aforementioned measures.

In order to determine whether two different measures measure similar phenomena, we look at each corpus in turn. What we are interested in is whether maps that score highly on one measure also score highly on another measure. To take an example, if every map that does well on competitiveness also scores highly on the efficiency gap, that would suggest that the two measures are duplicates. If, on the other hand, every map that scores highly on proportionality also has a poor efficiency gap, it would mean that the two measures tap different features.

Since the efficiency gap has generated much interest through its litigation in *Whitford*,[110] we present our findings comparing each of the other partisan fairness measures with the efficiency gap. In Figure 1, we have four histograms showing the frequency of plans that score at various points along the measure on the vertical axis.

---

108. *See* Liu et al., *supra* note 26, at 79.

109. *See id.* at 89.

110. *See* Whitford v. Gill, 218 F. Supp. 3d 837, 854 (W.D. Wis. 2016) (three-judge court), *argued*, No. 16-1161 (U.S. Oct. 3, 2017).

Figure 1. Comparing the EG Against Other Measures

Figure 1a. EG v. Bias

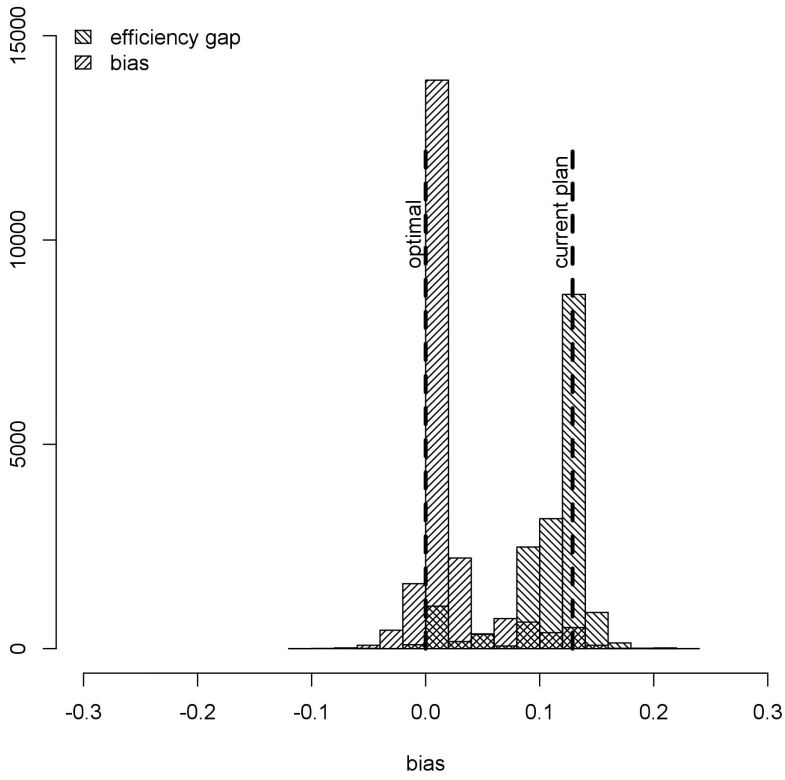


Figure 1b. EG v. Responsiveness

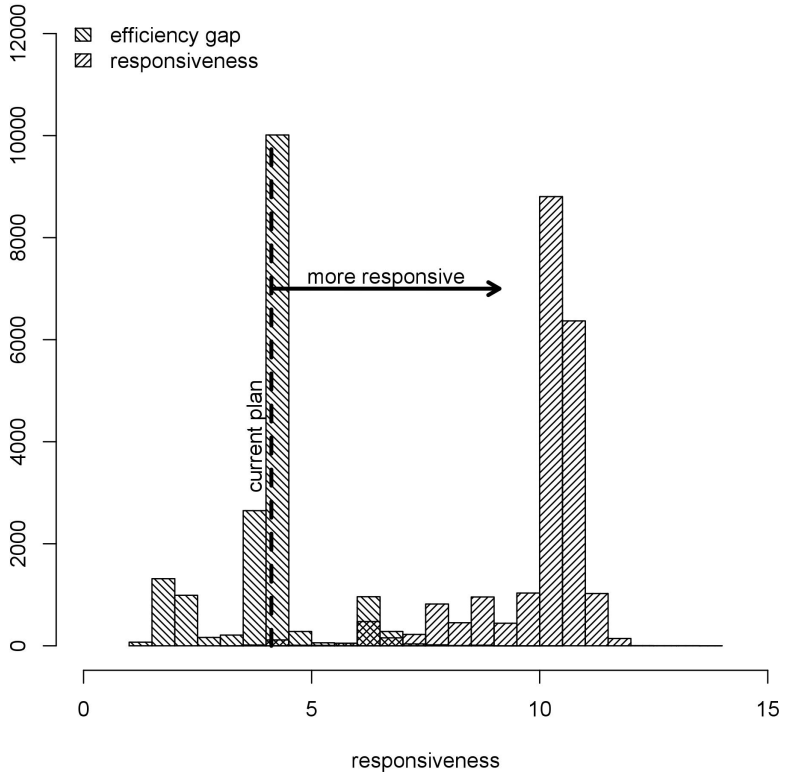


Figure 1c. EG v. Competitiveness

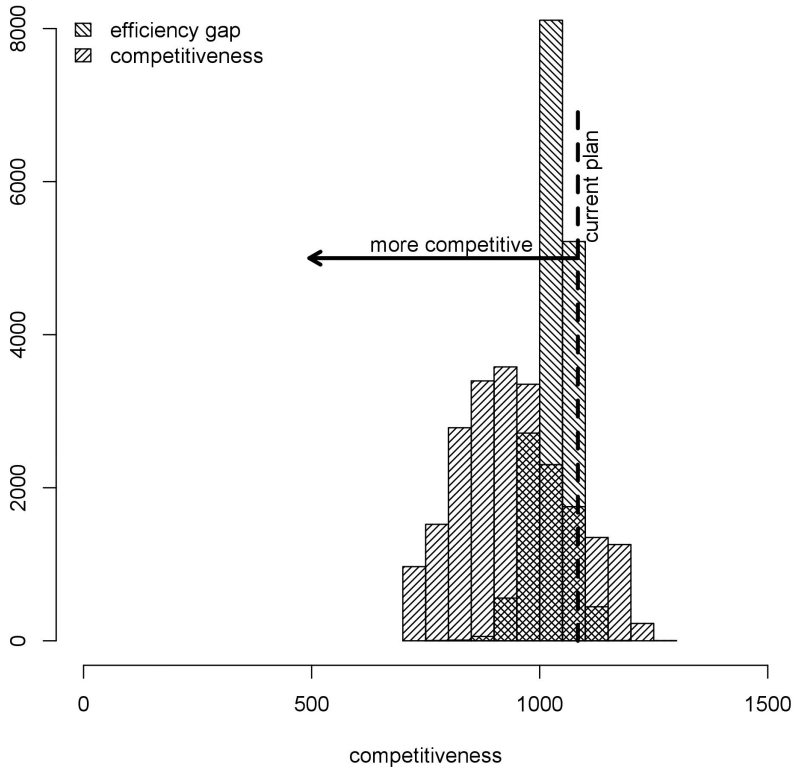
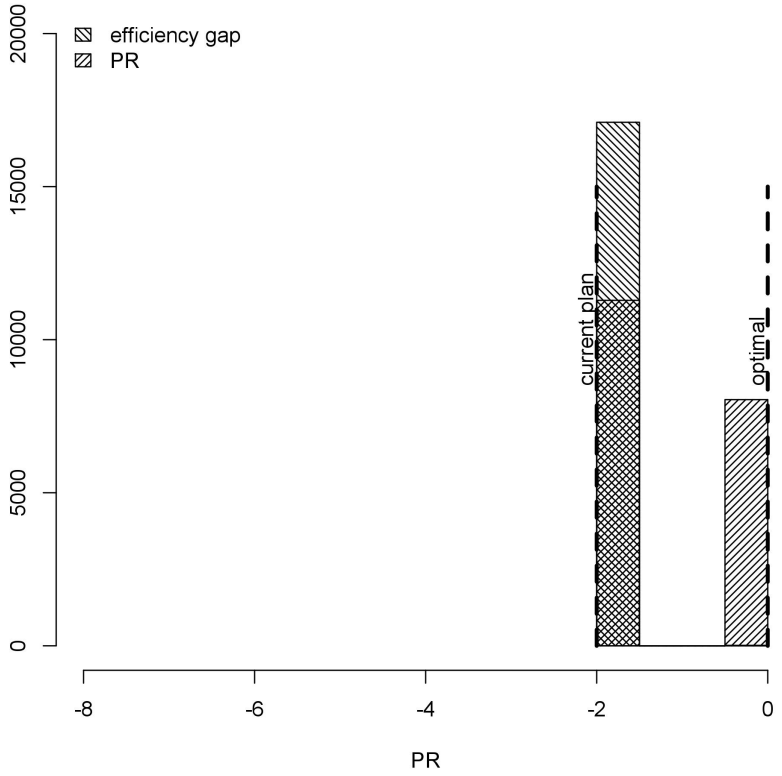


Figure 1d. EG v. Proportional Representation



The first histogram compares the bias scores of the maps that were optimized using the bias measure with the bias scores of the maps that were optimized using the efficiency gap measure.[111] Each histogram also shows the optimal and the current plan scores for purposes of comparison. We then repeat the exercise for responsiveness, competitiveness, and seats-vote proportionality.[112] We have also done a full comparison of all measures with each other, but in

111. *See supra* Figure 1a.
112. *See supra* Figures 1b-d.

the interest of space, we will only allude to what we found in those histograms. The full set of histograms is available online.[113]

Since political fairness has at least two dimensions (corresponding to partisan plans that favor one party over the other and bipartisan plans that favor incumbents over challengers),[114] we might expect more overlap between measures from the same category; that is, bias, proportionality, and the efficiency gap for identifying partisan advantage versus responsiveness and competitiveness for bipartisan lockup potential. The point of this exercise is not to weigh in on the comparative merits of any particular measure versus the other, but only to test the simple question of whether optimizing on these measures yields different or very similar results. If they yield different outcomes, then it suggests that courts and other entities will have to choose between these measures or develop an approach that utilizes some combination of them.[115] We advocate for the latter.

What do we discover? We find that choosing to optimize on any given measure as opposed to another does indeed lead to different conclusions about the best districts. This is true even within the two categories of fairness (partisan advantage versus bipartisan lockup). The bars with lines slanting up to the northeast in the above graphs show the distribution of scores on the indicated measures.[116] In the first graph, those bars display the distribution of the bias scores for all the maps optimized by the bias measure plus the compactness, contiguity, and equal population measures.[117] The bars with lines slanting down toward the southeast are the bias scores for the maps that were optimized according to the efficiency gap.[118] The northeast slanting bars in the other histograms correspond to the maps optimized in the same way with the measure listed on the horizontal axis, and the bars with lines slanting southeast are as before, the efficiency gap optimized maps.[119]

113. Cain et al., *supra* note 37.
114. *See, e.g.*, Cho & Liu, *supra* note 33, at 352.
115. *See infra* Part I.C.
116. *See supra* Figure 1.
117. *See supra* Figure 1a.
118. *See supra* Figure 1a.
119. *See supra* Figures 1b-d.

As Figure 1 shows, there is only some overlap between all of the measures and the efficiency gap, and in all cases, optimizing with any given measure produces better scores in terms of that measure than the efficiency gap. This is true even of bias and the efficiency gap, even though they both purport to measure partisan symmetry.[120] The same holds true when the exercise is flipped and we optimize on the efficiency gap and then score the bias optimized maps by the efficiency gap scores: there is only a partial overlap in the top sets of the two measures.[121] The same conclusion holds for the overlap between all of the measures, not just with the efficiency gap.[122]

In short, this exercise demonstrates what should surprise no one. Partisan fairness is a multifaceted concept. It encompasses a variety of values that can be measured in several different ways. The person who programs the computer has to choose or prioritize both among these core values and the measures of them. Some may care more about competitive elections where both parties have a chance at winning a seat.[123] They may want changes in votes to reflect changes in seats. Others may care more about whether the party affiliations of the elected representatives reflect the underlying distribution of votes cast for each party.[124] Still others may put more of a premium on whether a particular map is fair enough to both sides: whether both sides would have a similar chance at winning the seats if they have similar levels of statewide support.[125] While all of these ideas are embedded into the concept of political fairness, they are also, in many cases, distinct. Perhaps no single measure will capture all the richness of what partisan fairness means.

---

120. *See supra* Figure 1a.

121. *See* Cain et al., *supra* note 37, at 107.

122. *See id.* at 104-19.

123. *See, e.g.*, Navajo Nation v. Ariz. Indep. Redistricting Comm'n, 230 F. Supp. 2d 998, 1002 (D. Ariz. 2002) (three-judge court) (noting that the Arizona Constitution requires the Independent Redistricting Committee to "create competitive districts").

124. *See, e.g.*, Adam Cox, Commentary, *Partisan Fairness and Redistricting Politics*, 79 N.Y.U. L. REV. 751, 765-66 (2004) (attempting to adapt the single-member district model to the benefits of proportionality by reflecting the vote percentages in the seats gained).

125. *See, e.g.*, Whitford v. Gill, 218 F. Supp. 3d 837, 852 (W.D. Wis. 2016) (three-judge court), *argued*, No. 16-1161 (U.S. Oct. 3, 2017) (disputing a map that would have allowed the Republicans to maintain a majority even with only 48 percent of the statewide vote, compared to the 54 percent of the vote that Democrats would need to win the majority).

C. Optimizing Plans Along All Dimensions of Partisan Fairness

Since partisan fairness is not aptly captured by one of these proposed measures, how should redistricting bodies and courts choose among them? This Article does not purport to answer that question definitively. Our hope is merely to begin to evaluate ways to proceed when we recognize that political fairness has multiple dimensions and different ways of measuring them. What we have shown to date suggests that the different measures make unique contributions in capturing the complex concept of partisan fairness. Selecting one measure over all of the others necessitates sacrifices.

However, it is possible to combine these measures to produce a fuller picture. PEAR allows us to select plans that are optimized over multiple criteria: that is, an "everything bagel" approach to fair redistricting. In addition to so-called formal redistricting criteria such as equal population, contiguity, and compactness, this approach seamlessly handles multiple measures of political fairness, optimizing over all of them simultaneously.[126] This process involves several basic steps. First, we have to choose the redistricting criteria we seek to optimize and the relative weights we choose to assign to each. In this case, we assign equal weight to each, recognizing that they could instead be assigned different weights if that was agreed upon.

Second, we designate the starting value for the optimization, which in this example is the redistricting index score associated with Minnesota's current congressional district lines. An advantage of starting with the status quo districts is that they were presumably minimally valid in terms of state and federal criteria at the time they were adopted. The algorithm then optimizes by creating a set of new plans that meets or exceeds the redistricting criteria score of the current plan. In this setup, we are optimizing over the total score of all of the criteria.[127]

---

126. *See* Liu et al., *supra* note 26, at 82.

127. It is thus possible for the value of any specific criteria to be lower on one of the generated plans than for the status quo plan, as we are optimizing over the score of all of the criteria. However, it would also be possible to set up the process so that no plan scores lower than the current plan on any specific redistricting criterion.

Figure 2. Optimizing Plans on All Measures

Figure 2a. Combined Fairness Measures v. Bias

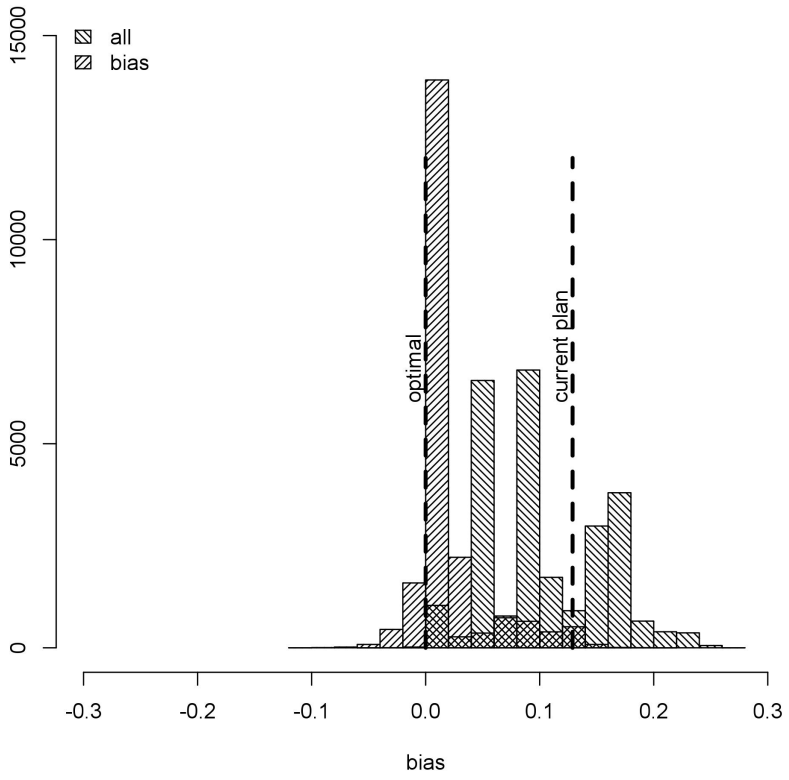


Figure 2b. Combined Fairness Measures v. Responsiveness

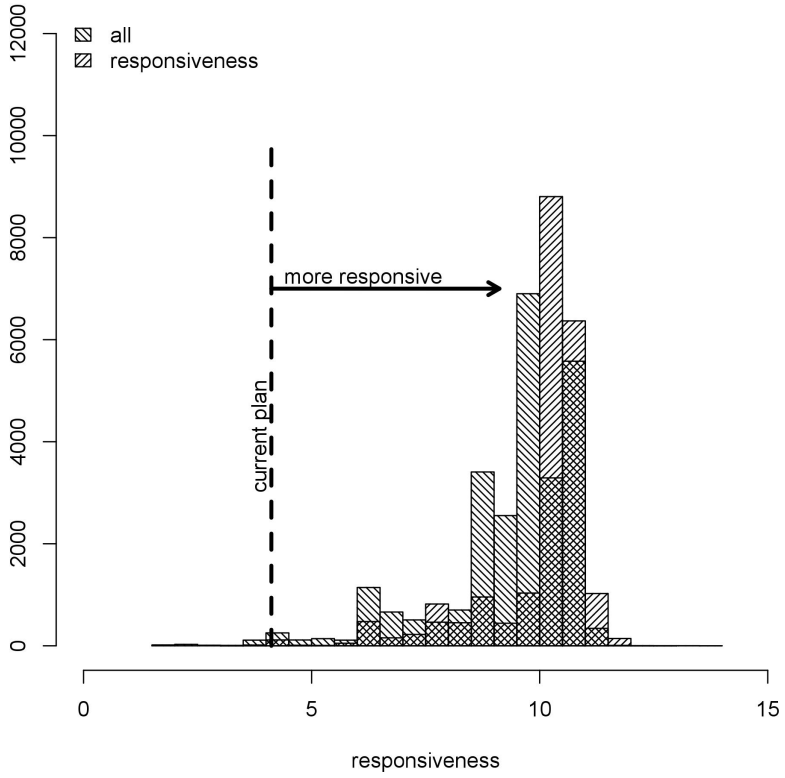


Figure 2c. Combined Fairness Measures v. Competitiveness

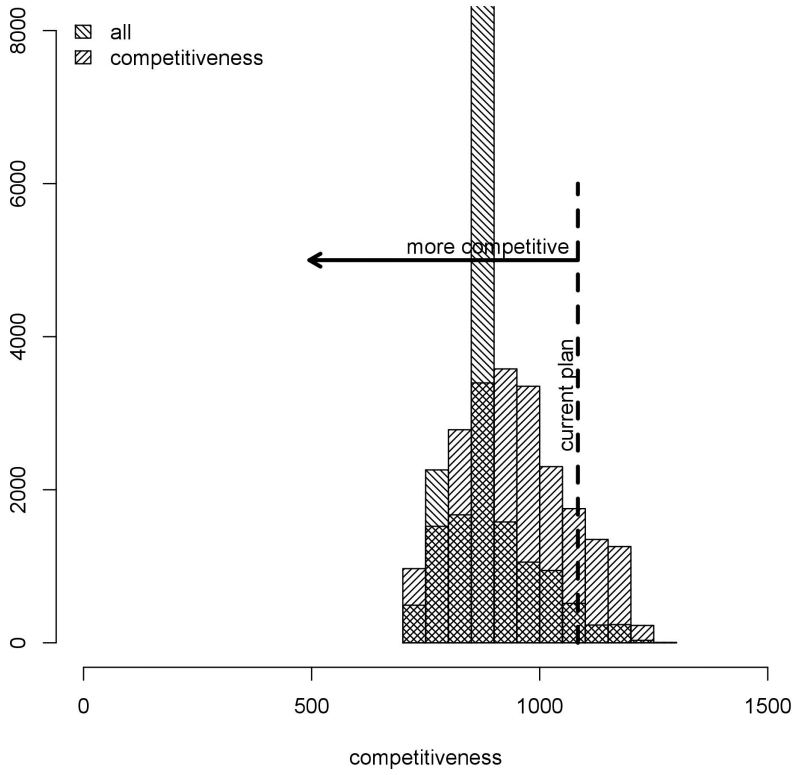


Figure 2d. Combined Fairness Measures v. Proportional Representation

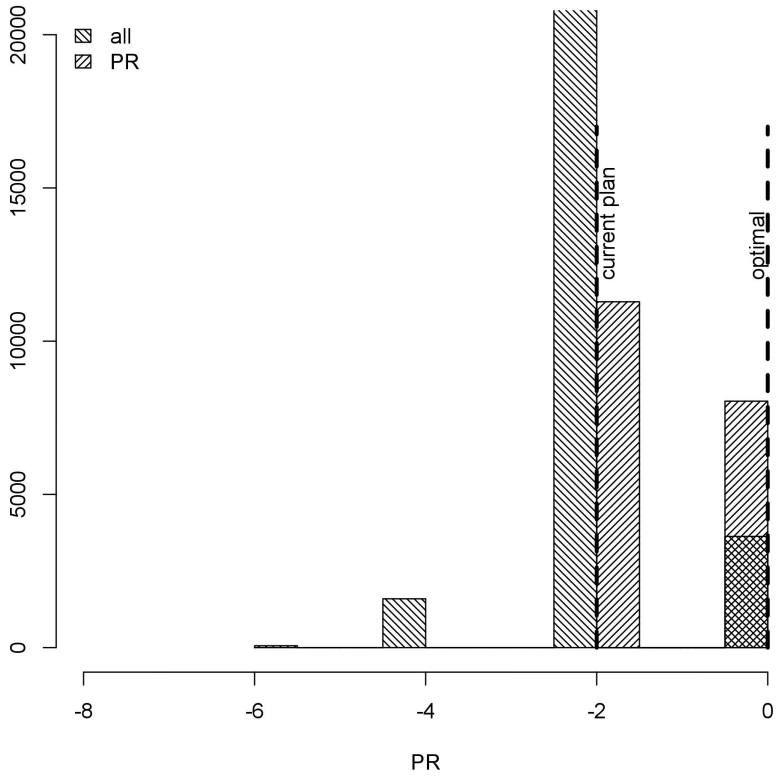
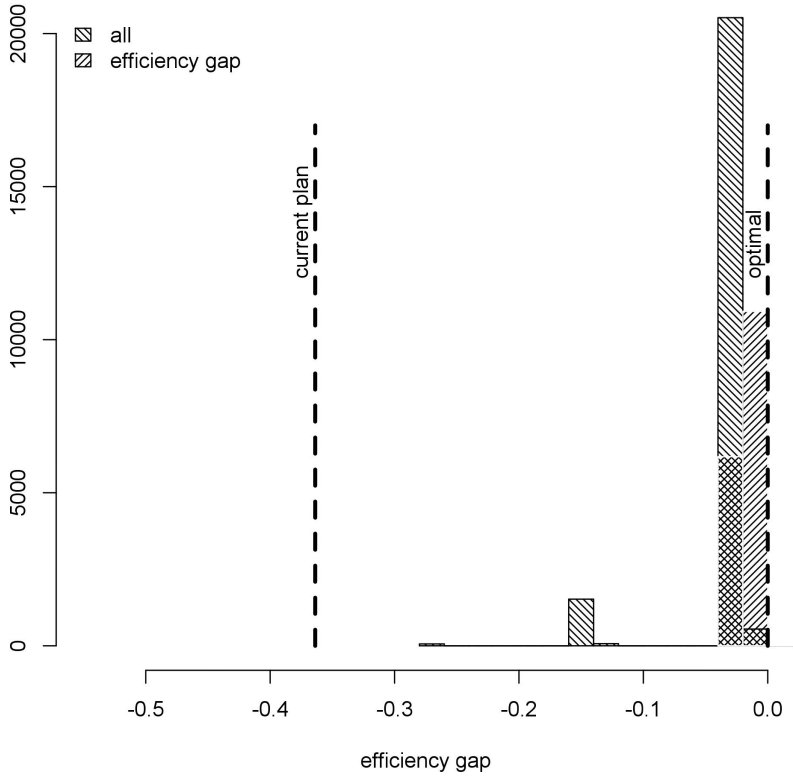


Figure 2e. Combined Fairness Measures v. EG



In Figure 2, the bars with lines slanting down to the southeast in the graphs represent the maps generated by the algorithm that optimizes along all of the measures plus the two formal criteria (the everything bagel plans), while the bars with lines slanting up to the northeast represent the scores of those maps optimized along only a single political fairness measure plus the formal criteria. The vertical dimension represents the number of plans that attain the various values of the measure as displayed along the horizontal axis.

Figure 2a illustrates an important point. It compares the set of maps using the combined fairness measure with the maps that were optimized using bias only. As we might expect, the bias scores of the maps that only optimize on the bias measure cluster much closer to the optimal value than the ones that optimize over all of the political fairness criteria. This is because, as noted earlier, the different political fairness criteria measure several dimensions of fairness.[128] In other words, the combined measure trades off bias with the other fairness and formal criteria measures to achieve an overall higher score.

These trade-offs seem particularly strong for the bias measure as a number of the everything bagel plans in the first histogram fall below the current plan's bias scores.[129] This appears to be less of a problem with the other fairness criteria.[130] Note also that because all of the single fairness criteria plans are traded off with the formal nonpolitical criteria, some of those automated plans also fall below the current plan's score.[131] A simple way to handle this is to say that a plan must fall in the reasonable bias interval between the current and optimal lines in order to ensure that any chosen plan is at least as good or better in every dimension of fairness as the current districts.

The histograms for competitiveness and responsiveness show the closest correspondence between the single criteria optimization and the everything bagel plans.[132] This suggests less trade-off with the other criteria than we observe with partisan-bias-only plans.[133] Most of the everything bagel plans in all the histograms are located in the reasonable bias, safe-harbor zone between the current plan and the optimal plans.

In short, automated plan generation offers the option of combining both political fairness and formal criteria to generate maps that meet or exceed the values of the maps that were previously approved by the courts. Because the maps derive from the same natural geography, they provide an accurate representation of what

128. *See supra* notes 106-10 and accompanying text.
129. *See supra* Figure 2a.
130. *See supra* Figures 2b-e.
131. *See supra* Figure 2.
132. *See supra* Figures 2b-c.
133. *See supra* Figure 2a.

is feasible given the particular demographic circumstances in a given state.

D. Applications of PEAR in Redistricting

Aside from providing a tool for the courts and others to judge the merits of various proposed plans on a basis that takes into account natural demographic concentrations, PEAR can also be used by redistricting commissions, court masters, and other bodies charged with the task of generating redistricting plans. One straightforward application of PEAR would be assisting courts when they are tasked with drawing remedial maps. In this case, a court could simply follow the process outlined above, taking the existing lines as a starting point and generating alternatives that are at least as good or better. This would greatly reduce time and expense for the courts.

Alternatively, courts could also decide simply to choose the plan with the best combined score from the set of plans that satisfy all other criteria thresholds. Such an approach would have the advantage of uniformity in standards for remedial plans across jurisdictions. Or, courts could choose blindly and randomly from the top set (that is, the set of plans that meets the minimum criteria). Such uncertainty might incentivize parties in the state legislature to come to an agreement rather than face the vagaries of a plan chosen from a large distribution of reasonably fair outcomes.

Similarly, generating a feasible set of plans could save time and trouble for citizen redistricting commissions. The Arizona Redistricting Commission is required by the state constitution to initiate its redistricting process by using compact, contiguous, and equally populated districts (but disregarding the other constitutional and statutory requirements) as the baseline of all subsequent negotiations and modifications.[134] This has spawned two decades of litigation about how far the final plan can deviate from the initial bare-bones foundations.[135] They might consider a process that starts, as our method did, from the criteria values of the status quo

134. ARIZ. CONST. art. IV, pt. II, § 1, cl. 14.

135. *See, e.g.*, Harris v. Ariz. Indep. Redistricting Comm'n, 136 S. Ct. 1301, 1306-07 (2016); Ariz. Minority Coal. for Fair Redistricting v. Ariz. Indep. Redistricting Comm'n, 121 P.3d 843, 869-70 (Ariz. Ct. App. 2005) (per curiam).

districts. Or, to take another example, the California Redistricting Commission starts from a staff-generated plan after receiving considerable public input.[136] Having staff-generated districts raises partisan suspicions about the true neutrality of the staff.[137] Starting with one or more automated plans would avoid this problem.

Still, neither a court nor a redistricting commission can avoid making the determination of a threshold level of the partisan fairness score.[138] But it should not be difficult for either the court or redistricting commissions to do so, provided they do it early in the process. They could either start from the values of the plan they designed in the previous decade, or they could take an outlier approach, setting the threshold values to exclude the tails of the automated plan distribution—for example, the plans that fall into the 5 percent most extreme partisan scores in either party's direction.

CONCLUSION

This Article has two main takeaway points. First, automated generation of a large number of viable plans is a better way to judge the fairness of proposed alternative districting plans than any standard based on time series or cross-sectional data. Political demography is dynamic, changing over time, and varying from state to state. Shifts in immigration policy, migration patterns, and differential birth rates across various racial and ethnic groups both within and across states make out-of-time and out-of-place comparisons highly problematic. Without taking into account natural inefficiencies in the geographic concentration of different segments of the population, it is impossible to gauge whether a plan that scores well or poorly by some nationally derived average is natural or intentionally imposed by those drawing the lines. It is not enough to say that a plan has a high or low score. It is necessary to determine whether that score is representative or reasonable within the context of possible plans, or whether it is an extreme outlier that needs to be examined closely.

136. *See* Cain, *supra* note 79, at 1824-27.
137. *See id.* at 1829.
138. *See supra* notes 42-46 and accompanying text.

Second, the Court needs to avoid locking in any one particular measure of political fairness because these measures not only have to be traded off with other formal criteria but also with other fairness measures that tap other dimensions of what is considered fair overall. Measures of partisan symmetry alone do not capture bipartisan lockup plans for incumbents. Different measures of either partisan bias or lockup can lead to different outcomes if they are used to guide plan generation. We have shown that it is possible to use all of the measures simultaneously in order to gain a more complete perspective of competing alternative plans. We also show that one sensible way to proceed is to take the values of the current plan as the starting point, and then accept an alternative that sits in the reasonable bias interval between the current plan and one with an optimal score. Of course, it is also possible to insist on a plan that has the optimal score. That is a choice for the courts, redistricting commissions, and citizens of a given state to decide.

It is natural and human for any measure's proponents to sell their product enthusiastically. After all, it is the ticket to fame and glory in academic circles and the reform community. But setting into stone a flawed and partial measure only invites gaming and manipulation by savvy partisan actors. Rather, we suggest that the standards should develop organically through adjudicating actual cases and controversies, as the "one person, one vote" standards did, so that we might discover more about which measures are most applicable to political geographies and citizen expectations. Over time, we can also expect automated plan generation to improve in efficiency and accessibility. The computing power of the Blue Waters supercomputer will be on laptops in the future. It is better to build the legal architecture around that future than to settle for and ensconce partial and flawed approaches that draw on soon-to-be passé technologies.