IN THE UNITED STATES DISTRICT COURT
FOR THE MIDDLE DISTRICT OF NORTH CAROLINA

COMMON CAUSE, *et al.*, Plaintiffs, v. ROBERT A. RUCHO, in his official capacity as Chairman of the North Carolina Senate Redistricting Committee for the 2016 Extra Session and Co-Chairman of the Joint Select Committee on Congressional Redistricting, *et al.*, Defendants.	)	CIVIL ACTION NO. 1:16-CV-1026-WO-JEP THREE-JUDGE COURT
League of Women Voters of North Carolina, *et al.*, Plaintiffs, v. Robert A. Rucho, in his official capacity as Chairman of the North Carolina Senate Redistricting Committee for the 2016 Extra Session and Co-Chairman of the 2016 Joint Select Committee on Congressional Redistricting, *et al.*, Defendants.	)	CIVIL ACTION NO. 1:16-CV-1164-WO-JEP THREE JUDGE COURT

<u>Expert Report of Sean P. Trende</u>

I, Sean P. Trende, do hereby declare the following:

1. I am over 18 years of age and am competent to testify regarding the matters discussed in this declaration.

2. My areas of expertise include political history, United States voting laws, redistricting, and the study of campaigns and elections.

3. I have been retained in this matter to provide expert testimony explaining how the efficiency gap will likely operate in practice. I am compensated at a rate of $300 per hour, excluding travel time. All opinions contained in this declaration are offered to a reasonable degree of professional certainty.

4. My *curriculum vitae* is attached to this declaration as **Exhibit 1.**

EXPERT CREDENTIALS

5. I have studied and followed United States elections on both a part-time and full-time basis for almost two decades.

6. I received a B.A. from Yale University in 1995, with a double major in history and political science.

7. I received a J.D. from Duke University in 2001.

8. I also received an M.A. from Duke University in 2001, in political science. My coursework was entirely at the graduate level, meaning that I was evaluated under the same expectations as Ph.D. students. As part of this coursework, I took two semesters of graduate level statistics.

9. I am currently enrolled as a doctoral student in political science at The Ohio State University. My focus is on American Politics, with a minor in methodology. I have taken most of my coursework in the Department of Statistics, including courses in regression analysis, probability theory, non-parametric statistics, survey methodology, and experimental design.

10. I joined RealClearPolitics in January of 2009 as their Senior Elections Analyst. I assumed a fulltime position with RealClearPolitics in March of 2010.

11. RealClearPolitics is a company of around 60 employees, with offices in Washington D.C. It produces one of the most heavily trafficked political websites in the world, which serves as a one-stop shop for political analysis from all sides of the political spectrum and is recognized as a pioneer in the field of poll aggregation. It produces original content, including both data analysis and traditional reporting. It is routinely cited by the most influential voices in politics, including David Brooks of *The New York Times*, Brit Hume of *Fox News*, Michael Barone of *The Almanac of American Politics*, Paul Gigot of *The Wall Street Journal*, and Peter Beinart of *The New Republic*.

12. My main responsibilities with RealClearPolitics consist of tracking, analyzing, and writing about elections. I also am in charge of rating the competitiveness of House of Representatives races, and collaborate in rating the competitiveness of Presidential, Senate and gubernatorial races. As a part of carrying out these responsibilities, I have studied and written extensively about demographic trends in the country, exit poll data at the state and federal level, public opinion polling, and voter turnout and voting behavior.

13. As part of familiarizing myself with how parties have drawn lines over the decades, as well as learning the political geography of the United States, I drew, using Adobe Illustrator, complete maps of every congressional district ever drawn, dating back to 1789. These maps were plotted on county maps of each state, tracing over images of maps taken from various Almanacs of American politics, or from my copy of Kenneth Martis' *The Historical Atlas of United States Congressional Districts: 1789-1983* (1982).

14. I served as a Senior Columnist for Dr. Larry Sabato's "Crystal Ball" from January 2014 through the end of last year. I had to stop writing for the Crystal Ball because schoolwork was taking up too much of my time.

15. The overarching purpose of my writings, both at RealClearPolitics and the Crystal Ball, is to try to convey more rigorous statistical understandings of elections than are typically found in journalistic coverage of elections for a lay audience.

16. I am the author of *The Lost Majority: Why the Future of Government is up For Grabs and Who Will Take It.* The book offers a revisionist take on realignment theory. It argues that realignments are a poor concept that should be abandoned. As part of this analysis, it conducts a thorough analysis of demographic and political trends beginning around 1920 and continuing through the modern times. The book has been placed on syllabi at universities, and cited in peer-reviewed literature.

17. I also authored a chapter in Dr. Larry Sabato's *Barack Obama and the New America: The 2012 Election and the Changing Face of Politics*, which discussed the demographic shifts accompanying the 2012 elections. I further authored a chapter in Dr. Sabato's *The Surge: 2014's Big GOP Win and What It Means for the Next Presidential Election*, which discusses demographics and Electoral College shifts. I authored a chapter in Dr. Sabato's forthcoming *Trumped: The 2016 Election That Broke All The Rules.*

18. I co-authored the 2014 *Almanac of American Politics*. The Almanac is considered the foundational text for understanding congressional districts and the representatives of those districts, as well as the dynamics in play behind those elections. PBS's Judy Woodruff described the book as "the oxygen of the political world," while NBC's Chuck Todd noted that "[r]eal political junkies get two *Almanacs*: one for the home and one for the office." My focus was

researching the history of and writing descriptions for many of the newly-drawn districts, including those for North Carolina.

19. I have spoken on these subjects before audiences from across the political spectrum, including at the Heritage Foundation, the American Enterprise Institute, the CATO Institute, the Bipartisan Policy Center, and the Brookings Institution. In 2012, I was invited to Brussels to speak about American elections to the European External Action Service, which is the European Union's diplomatic corps. I was selected by the United States Embassy in Sweden to discuss the 2016 elections to a series of audiences there.

20. It is my policy to appear on any news outlet that invites me, barring scheduling conflicts, and I have appeared on both Fox News and MSNBC to discuss electoral and demographic trends. I have spoken on a diverse array of radio shows such as First Edition with Sean Yoes, the Diane Rehm Show, the Brian Lehrer Show, the John Batchelor Show, the Bill Bennett Show, and Fox News Radio. I have been cited in major news publications, including *The New York Times*, *The Washington Post*, *The Los Angeles Times*, *The Wall Street Journal*, and *USA Today*.

21. I sit on the advisory panel for the "States of Change: Demographics and Democracy" project. This three-year project is sponsored by the Hewlett Foundation and involves three premier think tanks: The Brookings Institution, the American Enterprise Institute, and the Center for American Progress. The group takes a detailed look at trends among eligible voters and the overall population, both nationally and in key states, in an attempt to explain the impact of these changes on American politics, and to create population projections, which the Census Bureau abandoned in 1995.

22. I previously authored an expert report in *Dickson v. Rucho*, No. 11-CVS-16896 (N.C. Super Ct., Wake County), which involved North Carolina's 2012 General Assembly and Senate maps. Although I was not called to testify, it is my understanding that my expert report was accepted without objection. I also authored an expert report in *Covington v. North Carolina*, Case No. 1:15-CV-00399 (M.D.N.C.), which involved almost identical challenges in a different forum.

23. I authored two expert reports in *NAACP v. McCrory*, No. 1:13CV658 (M.D.N.C.), which involved challenges to multiple changes to North Carolina's voter laws, including the elimination of a law allowing for the counting of ballots cast in the wrong precinct. I was allowed to testify at trial. My testimony was solely on the "effect" prong of the Voting Rights Act claim. I did not examine the issues relating to intent.

24. I authored reports in *NAACP v. Husted*, No. 2:14-cv-404 (S.D. Ohio), and *Ohio Democratic Party v. Husted*, Case 15-cv-01802 (S.D. Ohio), which dealt with challenges to a variety of Ohio voting laws. I was allowed to testify at trial. The judge in the latter case ultimately refused to consider one opinion, which is not relevant to this report.

25. I authored an expert report in *Whitford v. Nichol*, No. 15-cv-421-bbc, a partisan gerrymandering case. I was allowed to testify at trial.

26. Although I do not testify in defense of voter identification laws, I served as a trial consultant in *Lee v. Virginia Board of Elections*, No. 3:15-cv-357.

27. I authored an expert report in *Feldman v. Arizona*, No. CV-16-1065-PHX-DLR, which dealt with an attempt to ban the practice of "ballot harvesting" in Arizona.

I. **<u>There is no single "efficiency gap," making it difficult to choose a standard.</u>**

28. In the Wisconsin litigation, Dr. Jackman urged scrutiny when the absolute value of the efficiency gap (that is, the value of the efficiency gap when any negative signs are ignored) for state assembly districts in the first enacted year rose above .07; this contrasts with the .08 threshold recommended by the authors of the efficiency gap for such cases. *See* Nick Stephanolpoulos & Eric McGee, "Partisan Gerrymandering and the Efficiency Gap," 82 *U. Chicago L.R.*, 831, 837 (2015). If this case is successful, states with more than 15 congressional districts will trigger scrutiny if the absolute efficiency gap exceeds .075, while states with seven to 15 districts will trigger scrutiny if their absolute efficiency gap exceeded .12.

29. This contrasts with the two-seat threshold urged by Stephanopoulos and McGee. Stephanopoulos & McGee at 888 ("Since aggregate House seats are the parties' main objective, it follows that the efficiency gap should be measured in seats rather than in percentage points."). There is even an argument to be made that this sort of shift to seats is "necessary," given that "[t]he efficiency gap becomes lumpier when there are fewer seats, because each seat accounts for a larger proportion of the seat total, and the efficiency gap thus shifts more as each seat changes hands." Simon Jackman, Rebuttal Report, *Wisconsin v. Nichols* at 23, 24 (Dec. 21, 2015), Exhibit 2. There exists "no authority in the literature" for failing to convert efficiency gaps to seats at the congressional level or, for that matter, calculating efficiency gaps for states with fewer than eight seats. Jackman Rebuttal at 24.

30. In addition, states with fewer than *seven* congressional districts are, at least for now, beyond scrutiny, although a court could apparently employ an eight-state cutoff, which was the approach suggested by Stephanopoulos & McGee in their original article and by Dr. Jackman

17 months ago. *Id.*; *id.* at 25 ("Next take Trende's consideration of Alabama's congressional plan in 2002 (which had seven districts), Iowa's congressional plan in 2002 (five districts), and Colorado's congressional plans in 2002 and 2012 (seven districts each) (paragraphs 115-16, 119, 122). All four of these plans have fewer than eight districts, and so, based on the literature, should not be included in any efficiency gap analysis because of the measure's lumpiness when applied to so few seats."); Stephanopoulos & McGee at 836. For purposes of this case, of the 43 states that have more than one district, 19 states, consisting of 70 congressional districts, will have to have some additional test devised. Utilizing an eight-seat cutoff, that number grows to 22 states, containing 81 congressional districts.

31. Beyond that there is the question of the form of the efficiency gap. As discussed below, there are two different efficiency gap formulae. The formula utilized in the actual efficiency gap article is "Efficiency Gap = Seat Margin – (2 x Vote Margin)." This was also the formula utilized by Dr. Jackman in the Wisconsin litigation. Here, Dr. Jackman employs a more involved version of the efficiency gap, which is the total number of wasted Democratic votes, minus the total number of wasted Republican votes, divided by the total number of votes cast. In this version of the efficiency gap, wasted Republican votes are defined as the votes that a Republican wins when Republicans lose a particular seat. When a Republican wins a seat, wasted votes are the difference between the number of votes a Republican won in the seat and half the number of total votes cast in the district. The same is true for Democrats.

32. The choices made above are not immaterial. In their law review article, Mr. Stephanopoulos and Dr. McGee identify only four states that would trigger court scrutiny under the 2012 plan: Florida, Ohio, Pennsylvania and Virginia. Stephanopoulos & McGhee at 837. Given that the efficiency gap for the 2016 North Carolina plan was lower than the 2012 plan, this

map would not be a gerrymander under the standard used in the law review. Dr. Jackman's metric, by contrast, would subject thirteen plans – over half of the plans that were drawn that year – to court scrutiny. These are: Alabama, Florida, Indiana, Massachusetts, Michigan, New Jersey, North Carolina, Ohio, Pennsylvania, South Carolina, Texas, Virginia, and Wisconsin. Obviously the standard upon which the Supreme Court decides is highly material.

33. Some of these standards have stronger empirical bases than others, but most involve at a least a degree of arbitrary decision. It is perhaps true that a court could ultimately settle upon an efficiency gap, much as it eventually settled upon a 10 percent cutoff for population deviations of state plans and 0 percent for federal plans. This, however, is true of any statistic; the court could also have settled upon a Reock score that it found represented too much boundary distortion. Regardless, my point is just that courts will have to make these sorts of decisions, that those decisions are consequential, and that the bases for those decisions are not always cut-and-dried.

II. <u>The efficiency gap is not easy to calculate.</u>

34. Contrary to the claims of the main efficiency gap article, the efficiency gap is not easily calculable. Stephanopoulos & McGee at 837. More precisely, while calculating the version of the efficiency gap utilized by Mr. Stephanopoulos and Dr. McGee is easy, and calculating the more involved version of the efficiency gap can become easy with a bit of practice, setting up the underlying data is difficult.

33. For example, Dr. Jackman encounters a large amount of missing data. Missing data are hardly a novel problem in statistics – entire courses are offered on how to deal with it, and it can comprise an entire unit in a survey methodology course. There are also multiple different ways to deal with imputations (which courts will likely have to choose between in

future cases), and criteria for sometimes discarding the missing observations. As Stephanopoulos and McGee explain:

> The most defensible [imputation strategy] is to use variables that have been shown in the past to predict vote share, and then to impute values for uncontested races based on these variables. One might also examine how uncontested districts have turned out in previous years when those same seats were contested. Or one might simply assume that the opposing party would have received a certain vote share (for example, 25 percent) had it run a candidate in an uncontested district. Clearly, these imputation approaches can be more or less sophisticated, and can bring varying amounts of information to bear on the problem.

Stephanopoulos & McGee at 866.

34. For his part, Dr. Jackman utilizes Bayesian hierarchical modeling, using Markov Chain Monte Carlo runs, including 25,000 burn-in iterations, followed by an additional 150,000 iterations, saving every 30th iteration. While this isn't quite as exotic as it may sound, neither is it the sort of thing one encounters in a routine undergraduate statistics course. It is also not particularly simple to implement; it takes, by my count, twenty pages of computer code from Dr. Jackman's program to activate. Map drawers may or may not have the sophistication to run a model like this, and courts will have to eventually adjudicate between different imputation strategies, as they can yield different results.

35. Then there is the question of whether the "baseline" efficiency gap for a state should be zero. If random districts were drawn throughout the United States, according to Dr. Chen, we would ultimately end up with seat shares for parties that do not look dissimilar from what we have today on the national level. The problem for Democrats "is that they have overwhelming majorities not only in dense, poor urban centers, but also in isolated, far-flung college towns, historical mining areas and 19th-century manufacturing towns that are surrounded and ultimately overwhelmed by rural Republicans." *See* Jowei Chen and Jonathan Rodden, "Don't Blame the Maps," New York Times (Jan. 24, 2014).

36. So, while Indiana shows a strong first-year efficiency gap that is, in fact, more extreme than that of the present North Carolina map, Drs. Chen and Rodden find that randomly drawn maps would not produce drastic changes in seat allocations. In other words, if we accept the Chen and Rodden approach as applicable to gerrymandering cases, there may not be many votes wasted by any intent on the part of the legislature to create a gerrymander – though not all scholars accept that the Chen and Rodden approach as applicable. *See* Jackman Rebuttal at 20 ("While I respect Chen and Rodden's contribution, there are several issues with their work that make it inapplicable here.").

37. If there is spatial clustering, then where the baseline for an efficiency gap should be set is extremely difficult to calculate; indeed it may be impossible to translate a cluster-detecting algorithm to an efficiency gap. One could say in this circumstance that Indiana should be required to draw maps that minimize the efficiency gap, but this then effectively becomes a constitutional requirement for states to make "make up calls" for inefficient vote distributions, rather than a remedy for gerrymandering.

38. Likewise, although Illinois only shows a modest efficiency gap of .02, Drs. Chen and Rodden find that the number of seats won by Democrats far exceeds that which would be achieved under random maps. In other words, in Illinois, the baseline might need to be moved to the right to capture the gerrymandering of the Democrats there; otherwise the natural clustering of Democrats will mask what is, in fact, a heavy Democratic gerrymander. *Id.*

39. This is not an easy task. Dr. Chen's code does not run well on the latest version of Java, and seems to produce compiler errors until an earlier version is found and the proper packages are added to the class path. To even understand how the code actually works, a researcher who encounters problems is forced to wade through code like this:

```
private ArrayList MuDeltaTauRhoTau(int dnumber, double targetpop, ArrayList Upsallan,
int nextcty){

Random GammaEpsilonNu = new Random(); boolean complete=false;

ArrayList PhiCauTauSigma = new ArrayList(); PhiCauTauSigma.clear(); int
DeltaPhiOmegaPhi=0; ArrayList borders1 = new ArrayList(); borders1.clear(); ArrayList
lastctygroup = new ArrayList(); lastctygroup.clear();

int TauRhoIota=0; ArrayList ctys_reached = new ArrayList(); int splitcty = 0;

while(true){ TauRhoIota++; if (TauRhoIota>5000000){return null;}


lastctygroup.clear();

for (int i=Upsallan.size()-1; i>=0; i--){

int pct = (Integer)Upsallan.get(i); int pcty = (Integer)phichizeta.get(pct);

if (pcty==nextcty){

if (borders1.contains(pct)){borders1.remove((Object)pct);}

PhiCauTauSigma.add(pct); Upsallan.remove((Object)pct);
DeltaPhiOmegaPhi+=(Integer)phipiphi.get(pct); lastctygroup.add(pct);

boolean PhiCauTauSigma_cont0=IotaChiOmegaNu(PhiCauTauSigma,PhiChiBeta);
System.out.println("PhiCauTauSigma: "+PhiCauTauSigma.size()+" PhiCauTauSigma_cont0:
"+PhiCauTauSigma_cont0);

ArrayList PhiBetaOmegaRho = (ArrayList)PhiChiBeta.get(pct);

for (int j=0; j<PhiBetaOmegaRho.size(); j++){

int pb = (Integer)PhiBetaOmegaRho.get(j); System.out.println("pct: "+pct+"
PhiBetaOmegaRho: "+PhiBetaOmegaRho);

if(Upsallan.contains(pb) && !borders1.contains(pb)){borders1.add(pb); }
```

40. Beyond this, there exists a whole universe of techniques for discussing spatial clusters, all of which have their own problems, but which will have to be sorted through by courts to truly adjudicate these claims: nearest neighbor, Moran’s various tests, the index of dissimilarity, Oden’s I*pop, Tango’s MEET, and SaTScan, and the like.

III. <u>The efficiency gap is proportional representation for first-past-the-post systems.</u>

41. A system of proportional representation is easily defined and comprehended: Whatever share of the vote a party receives in a country corresponds more-or-less directly to the share of the seats they receive (I say more-or-less because countries frequently employ minimum vote share requirements to qualify for proportional representations). So, if we define VoteShare* as a party's share of the vote, centered at 50% (i.e., its vote share minus 0.5), and define SeatShare* as a party's share of seats, again centered at 50% (i.e., its seat share minus 0.5), we can see what I will call the bias of the plan. In equation form, it would look like this:

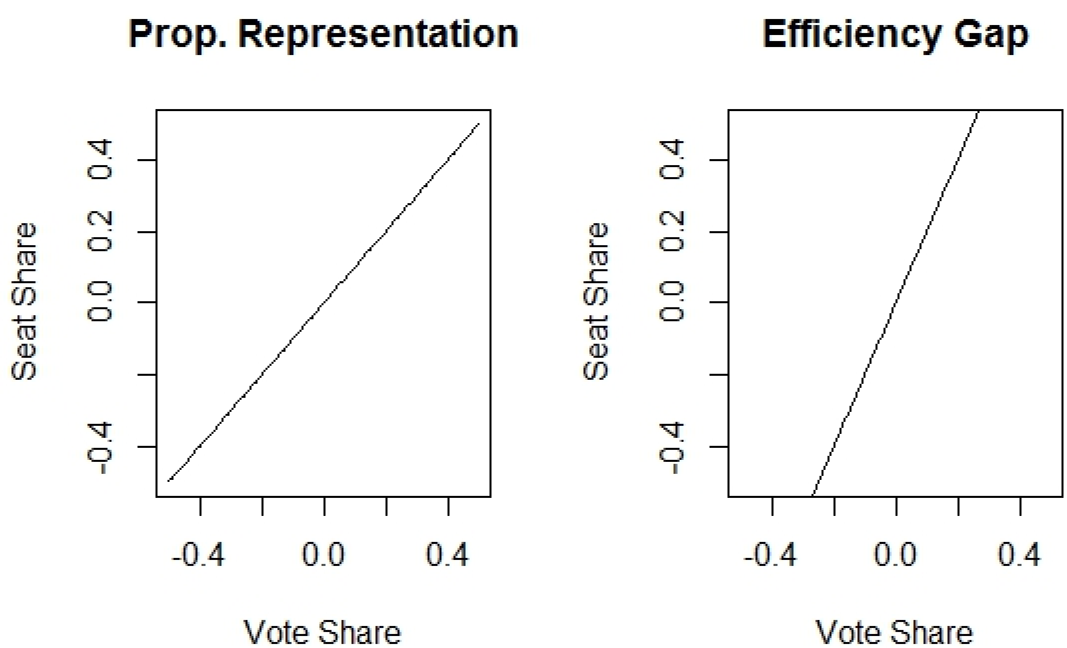
$$\text{Bias} = \text{VoteShare*} - \text{SeatShare*}$$

42. As Dr. Eric McGhee proved, when districts have equal populations, the complex version of the efficiency gap, utilized by Dr. Jackman here, simplifies to a simpler version utilized by Dr. Jackson in Wisconsin and Mr. Stephanopoulos and Dr. McGee in their law review article. Eric McGhee, Measuring Partisan Bias in Single-Member District Electoral Systems 39 *Leg. Stud. Q.* 55 (2014) [Ex. 3]; Stephanopoulos andMcGee at 853. ("Instead, if we assume that all districts are equal in population (which is constitutionally required), and that there are only two parties (which is typical in [single member district] systems), then the computation reduces through simple algebra to something quite straightforward."). Therefore, "[t]o produce partisan fairness—in the sense of equal wasted votes for each party—the bonus should be a precisely twofold increase in seat share for a given increase in vote share." Stephanopoulos and McGee at 854. In equation form, it looks like this:

$$\text{Efficiency Gap} = 2(\text{VoteShare*}) - \text{SeatShare*}$$

43. *See* Stephanopoulos & McGee at 853. The similarities are quite striking: A party's share of the seats above or below 50 percent will equal its share of the votes, similarly adjusted, in an ideal proportional representation scheme. Under the efficiency gap, it simply

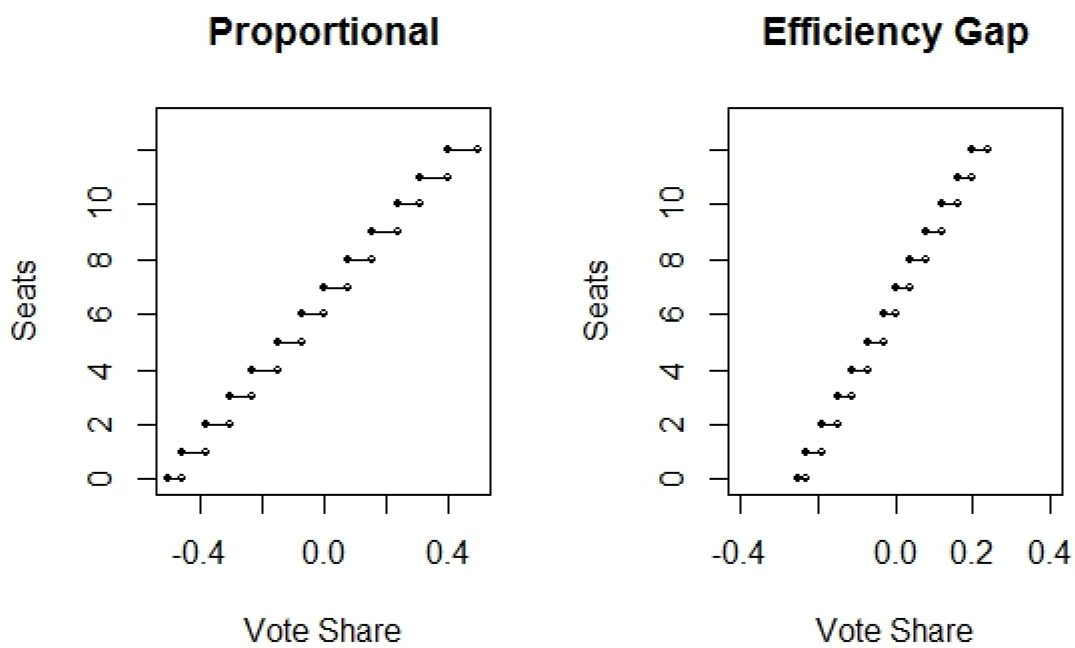
equals twice the adjusted vote share. Moreover, the efficiency gap is built upon direct proportionality in "wasted" votes. Stephanopoulos & McGee at 834 ("Algebraically, this means that Party A wins 20 percent (or 2) more seats than it would have had the parties wasted equal numbers of votes."). To put this into graphical form, below are the ideal share of seats for a party, given a particular vote share under both the efficiency gap and proportional representation. Indeed, proportional representation is related to partisan symmetry as much as the efficiency gap is, in that both have partisan biases of zero in a 50-50 election. For that matter, pure proportional representation can be thought of as a focus on wasted votes as well, at the national level rather than the district level.



44. In other words, the efficiency gap allocates seats with a proportion of 2% of the seats for every 1% of the vote, rather than 1% of the seats for every 1% of the vote.

45. Of course, seat allocations are discrete, not continuous (in plain English, that means you can't have a fraction of a seat share). To see how actual seat allocation in North

Carolina would work under proportional representation and the efficiency gap, see the charts below:



46. Notice, though, that the vote share for the efficiency gap truncates. That is, it does not return a value if a party receives, say, 10 percent of the vote (-0.4 on the above charts). Stephanopoulos and McGee make a practical argument here that very few plans reach this level of vote share. It nevertheless seems like an odd theoretical construct that for some conceivable values of United States elections, a federal standard of gerrymandering would demand negative seats.

47. For a final illustration of how this works in reality, I have calculated, using the simplified version of the efficiency gap (which, insofar as I can tell, is the version that has been used in the published literature), the following tables. The columns represent the number of congressional districts in a state. The rows represent a given vote share in the state. The left and right cell entries represent the minimum and maximum number of seats you can have under the

efficiency gap, while the middle value represents the ideal number of seats. I have calculated them for states with between seven and 15 districts, and between 16 and 24 districts, to illustrate both of Dr. Jackman's thresholds.

48. Note that the allowable seats progress in a linear manner; for example, in the case of North Carolina (13 districts) each category of seats (minimum, ideal, and maximum) increases by one seat roughly every four percentage points (occasionally, due to rounding, it is not exactly four seats). For a state with 22 districts, they increase by one seat roughly every three percentage points of popular vote. One could conceivably generate a similar table using the more complex version of the efficiency gap, but it would require a multi-dimensional space. Regardless, the differences between the simple and complex efficiency gap calculation are relatively small overall; if one exclude Dr. Jackman's imputations, the r-square is .9403.

	7	8	9	10	11	12	13	14	15
.30	0,1,2	0,1,2	0,1,2	0,1,2	0,1,2	0,1,3	0,1,3	0,1,3	0,1,3
.31	0,1,2	0,1,2	0,1,2	0,1,2	0,1,3	0,1,3	0,2,3	0,2,3	0,2,4
.32	0,1,2	0,1,2	0,1,2	0,1,3	0,2,3	0,2,3	0,2,3	0,2,4	0,2,4
.33	0,1,2	0,1,2	0,1,3	0,2,3	0,2,3	0,2,3	1,2,4	1,2,4	1,2,4
.34	0,1,2	0,1,2	1,2,3	1,2,3	1,2,3	1,2,4	1,2,4	1,3,4	1,3,5
.35	1,1,2	1,2,3	1,2,3	1,2,3	1,2,4	1,2,4	1,3,4	1,3,4	1,3,5
.36	1,2,2	1,2,3	1,2,3	1,2,3	1,2,4	1,3,4	1,3,4	1,3,5	1,3,5
.37	1,2,3	1,2,3	1,2,3	1,2,4	1,3,4	1,3,4	2,3,5	2,3,5	2,4,5
.38	1,2,3	1,2,3	1,2,3	1,3,4	2,3,4	2,3,5	2,3,5	2,4,5	2,4,6
.39	1,2,3	1,2,3	1,3,4	2,3,4	2,3,4	2,3,5	2,4,5	2,4,6	2,4,6
.40	1,2,3	1,2,3	2,3,4	2,3,4	2,3,5	2,4,5	2,4,5	3,4,6	3,5,6
.41	1,2,3	2,3,4	2,3,4	2,3,4	2,4,5	2,4,5	3,4,6	3,4,6	3,5,7
.42	2,2,3	2,3,4	2,3,4	2,3,5	2,4,5	3,4,6	3,4,6	3,5,6	3,5,7
.43	2,3,3	2,3,4	2,3,4	2,4,5	3,4,5	3,4,6	3,5,6	3,5,7	4,5,7
.44	2,3,4	2,3,4	2,3,4	3,4,5	3,4,6	3,5,6	3,5,6	4,5,7	4,6,8
.45	2,3,4	2,3,4	3,4,5	3,4,5	3,4,6	3,5,6	4,5,7	4,6,7	4,6,8
.46	2,3,4	2,3,4	3,4,5	3,4,5	3,5,6	4,5,6	4,5,7	4,6,8	5,6,8
.47	2,3,4	3,4,4	3,4,5	3,4,6	4,5,6	4,5,7	4,6,7	4,6,8	5,7,8
.48	2,3,4	3,4,5	3,4,5	3,5,6	4,5,6	4,6,7	4,6,8	5,6,8	5,7,9
.49	3,3,4	3,4,5	3,4,5	4,5,6	4,5,7	4,6,7	5,6,8	5,7,8	5,7,9
.50	3,4,4	3,4,5	3,4,6	4,5,6	4,6,7	5,6,7	5,6,8	5,7,9	6,8,9
.51	3,4,4	3,4,5	4,5,6	4,5,6	4,6,7	5,6,8	5,7,8	6,7,9	6,8,10
.52	3,4,5	3,4,5	4,5,6	4,5,7	5,6,7	5,6,8	5,7,9	6,8,9	6,8,10
.53	3,4,5	4,4,5	4,5,6	4,6,7	5,6,7	5,7,8	6,7,9	6,8,10	7,8,10
.54	3,4,5	4,5,6	4,5,6	5,6,7	5,6,8	6,7,8	6,8,9	6,8,10	7,9,11
.55	3,4,5	4,5,6	4,5,6	5,6,7	5,7,8	6,7,9	6,8,9	7,8,10	7,9,11
.56	4,4,5	4,5,6	5,6,7	5,6,7	6,7,8	6,7,9	7,8,10	7,9,10	8,9,11
.57	4,4,5	4,5,6	5,6,7	5,6,8	6,7,8	6,8,9	7,8,10	7,9,11	8,10,11
.58	4,5,5	4,5,6	5,6,7	5,7,8	6,7,9	6,8,9	7,9,10	8,9,11	8,10,12
.59	4,5,6	4,5,6	5,6,7	6,7,8	6,7,9	7,8,10	7,9,10	8,10,11	8,10,12
.60	4,5,6	5,6,7	5,6,7	6,7,8	6,8,9	7,8,10	8,9,11	8,10,11	9,10,12
.61	4,5,6	5,6,7	5,6,8	6,7,8	7,8,9	7,9,10	8,9,11	8,10,12	9,11,13
.62	4,5,6	5,6,7	6,7,8	6,7,9	7,8,9	7,9,10	8,10,11	9,10,12	9,11,13
.63	4,5,6	5,6,7	6,7,8	6,8,9	7,8,10	8,9,11	8,10,11	9,11,12	10,11,13
.64	5,5,6	5,6,7	6,7,8	7,8,9	7,9,10	8,9,11	9,10,12	9,11,13	10,12,14
.65	5,6,6	5,6,7	6,7,8	7,8,9	7,9,10	8,10,11	9,10,12	10,11,13	10,12,14
.66	5,6,7	6,7,8	6,7,8	7,8,9	8,9,10	8,10,11	9,11,12	10,11,13	11,12,14
.67	5,6,7	6,7,8	6,8,9	7,8,10	8,9,11	9,10,12	9,11,12	10,12,13	11,13,14
.68	5,6,7	6,7,8	7,8,9	7,9,10	8,9,11	9,10,12	10,11,13	10,12,14	11,13,15
.69	5,6,7	6,7,8	7,8,9	8,9,10	8,10,11	9,11,12	10,11,13	11,12,14	11,13,15
.70	5,6,7	6,7,8	7,8,9	8,9,10	9,10,11	9,11,12	10,12,13	11,13,14	12,13,15

	1	2	3	4	5	6	7	8	9
1	0,2,3	0,2,3	0,2,3	0,2,3	0,2,3	1,2,4	1,2,4	1,2,4	1,2,4
2	1,2,3	1,2,3	1,2,4	1,2,4	1,2,4	1,3,4	1,3,4	1,3,4	1,3,5
3	1,2,3	1,2,4	1,3,4	1,3,4	1,3,4	1,3,5	1,3,5	1,3,5	2,3,5
4	1,3,4	1,3,4	2,3,4	2,3,4	2,3,5	2,3,5	2,4,5	2,4,5	2,4,6
5	2,3,4	2,3,4	2,3,5	2,3,5	2,4,5	2,4,5	2,4,6	2,4,6	3,4,6
6	2,3,4	2,3,5	2,4,5	2,4,5	2,4,5	3,4,6	3,4,6	3,5,6	3,5,7
7	2,4,5	2,4,5	3,4,5	3,4,6	3,4,6	3,5,6	3,5,6	3,5,7	3,5,7
8	3,4,5	3,4,5	3,4,6	3,5,6	3,5,6	3,5,7	4,5,7	4,6,7	4,6,8
9	3,4,5	3,4,6	3,5,6	4,5,6	4,5,7	4,5,7	4,6,7	4,6,8	4,6,8
10	3,4,6	3,5,6	4,5,6	4,5,7	4,6,7	4,6,7	5,6,8	5,6,8	5,7,9
11	4,5,6	4,5,6	4,5,7	4,6,7	5,6,8	5,6,8	5,7,8	5,7,9	5,7,9
12	4,5,6	4,5,7	4,6,7	5,6,8	5,6,8	5,7,8	5,7,9	6,7,9	6,8,9
13	4,5,7	5,6,7	5,6,7	5,6,8	5,7,8	6,7,9	6,7,9	6,8,10	6,8,10
14	5,6,7	5,6,7	5,6,8	5,7,8	6,7,9	6,8,9	6,8,10	7,8,10	7,9,10
15	5,6,7	5,6,8	5,7,8	6,7,9	6,8,9	6,8,10	7,8,10	7,9,10	7,9,11
16	5,6,8	6,7,8	6,7,9	6,8,9	6,8,10	7,8,10	7,9,10	7,9,11	8,10,11
17	6,7,8	6,7,8	6,8,9	7,8,9	7,8,10	7,9,10	8,9,11	8,10,11	8,10,12
18	6,7,8	6,7,9	7,8,9	7,8,10	7,9,10	8,9,11	8,10,11	8,10,12	9,11,12
19	6,7,9	7,8,9	7,8,10	7,9,10	8,9,11	8,10,11	8,10,12	9,11,12	9,11,13
20	6,8,9	7,8,9	7,9,10	8,9,11	8,10,11	9,10,12	9,11,12	9,11,13	10,12,13
21	7,8,9	7,8,10	8,9,10	8,10,11	8,10,12	9,10,12	9,11,13	10,12,13	10,12,14
22	7,8,10	8,9,10	8,9,11	8,10,11	9,10,12	9,11,12	10,11,13	10,12,14	11,12,14
23	7,9,10	8,9,10	8,10,11	9,10,12	9,11,12	10,11,13	10,12,14	11,12,14	11,13,15
24	8,9,10	8,10,11	9,10,11	9,11,12	10,11,13	10,12,13	11,12,14	11,13,15	12,13,15
25	8,9,10	9,10,11	9,10,12	10,11,12	10,12,13	11,12,14	11,13,14	12,13,15	12,14,16
26	8,10,11	9,10,11	9,11,12	10,11,13	11,12,14	11,13,14	12,13,15	12,14,16	13,14,16
27	9,10,11	9,11,12	10,11,13	10,12,13	11,12,14	11,13,15	12,14,15	13,14,16	13,15,17
28	9,10,11	10,11,12	10,12,13	11,12,14	11,13,14	12,13,15	12,14,16	13,15,16	14,15,17
29	9,11,12	10,11,12	11,12,13	11,13,14	12,13,15	12,14,15	13,15,16	13,15,17	14,16,18
30	10,11,12	10,12,13	11,12,14	11,13,14	12,14,15	13,14,16	13,15,17	14,16,17	15,16,18
31	10,11,12	11,12,13	11,13,14	12,13,15	12,14,15	13,15,16	14,15,17	14,16,18	15,17,19
32	10,12,13	11,12,14	12,13,14	12,14,15	13,14,16	14,15,17	14,16,17	15,17,18	15,17,19
33	11,12,13	11,13,14	12,13,15	13,14,15	13,15,16	14,16,17	15,16,18	15,17,19	16,18,20
34	11,12,13	12,13,14	12,14,15	13,14,16	14,15,17	14,16,18	15,17,18	16,17,19	16,18,20
35	11,12,14	12,13,15	13,14,15	13,15,16	14,16,17	15,16,18	16,17,19	16,18,20	17,19,21
36	12,13,14	12,14,15	13,14,16	14,15,17	15,16,18	15,17,18	16,18,19	17,18,20	17,19,21
37	12,13,14	13,14,15	13,15,16	14,16,17	15,16,18	16,17,19	16,18,20	17,19,21	18,20,21
38	12,13,15	13,14,16	14,15,16	15,16,17	15,17,18	16,18,19	17,18,20	18,19,21	18,20,22
39	13,14,15	13,15,16	14,15,17	15,16,18	16,17,19	16,18,20	17,19,21	18,20,22	19,21,22
40	13,14,15	14,15,16	14,16,17	15,17,18	16,18,19	17,18,20	18,19,21	19,20,22	19,21,23
41	13,14,16	14,15,17	15,16,18	16,17,19	16,18,19	17,19,20	18,20,21	19,21,22	20,22,23

IV. Dr. Jackman's imputations appear to include unrealistic values

51. I agree with Dr. Jackman (as well as Dr. McGee and Mr. Stephanopoulos) that missing election results have to be somehow estimated if the efficiency gap is to work. The number of districts in a given plan is often small, and missing data are relatively frequent – about one-in-seven elections in the dataset are uncontested. Even if the misses were completely random (which we don't really have reason to suspect is the case) in the aggregate, ignoring an uncontested election altogether could have a significant effect upon the estimated efficiency gap in individual states, especially in those states with small numbers of districts. With that said, the fact that missing data must be dealt with for the efficiency gap to work can also lead to the conclusion that the efficiency gap cannot be made to work. With imputations accounting for roughly one-in-seven elections in the dataset, Jackman Report at 21, making mistakes can skew the results badly.

52. First, the imputations require the researcher to assess counterfactuals, something Dr. McGhee and Mr. Stephanopoulos identify, properly, as a problem with the partisan bias metric. Stephanopoulos & McGee at 835 (describing accessing the counterfactual as a "crucial problem" with partisan bias). In that instance we have to access the counterfactual of a 50-50 election; in this instance we have to attempt to access the counterfactual of a contested election. In some plans here, a majority of the districts are estimated off of hypothetical elections.

53. More importantly, a bad estimate may be worse than no estimate at all, and it isn't clear whether Dr. Jackman's estimates are better than no imputations at all. Some of his estimates for both turnout and election results seem unlikely to be accurate.

54. Counsel for plaintiffs produced a file entitled "districtLevel," which appears to include the election returns upon which Dr. Jackman relies to compute his efficiency gap scores. The data for turnout, including imputed values, appear to be contained in a column entitled "totalTwoParty_bar." In those cases where Dr. Jackman imputes values for turnout, he also includes estimates at the low end and high end for his selected "confidence level" (as a reasonable analogy, think of how political polling is often reported, with a given value, plus or minus a certain number of points, though the more precise analogy here would probably be internet polling). These uncontested elections, for which turnout and vote share are imputed, are identified as "FALSE" in a column labeled "contested."

55. Over the course of the elections Dr. Jackman surveys, the average district received 190,700 votes. Among the elections for which we have actual data, the lowest reported turnout came in New York's 12th Congressional District in 1978. In that election, Democrat Shirley Chisholm received 25,697 votes to Republican Charles Gibbs' 3,580 votes. But even this was unusual. About 95 percent of elections for which we have actual data show at least 100,000 votes cast, and 99 percent of elections show at least 67,556 votes cast.

56. But 59 of Dr. Jackman's 1,114 imputed values, or a little more than five percent of the elections for which he imputes values, show estimates that are lower than the lowest vote total that ever actually manifested in a contested election. Most of these are lower than the lowest vote total that ever actually manifested in a contested election by an order of magnitude. These are listed in the following table.

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
	stpost	year	district	dvote_bar	dvote_lo	totalTwoParty_bar	totalTwoParty_lo
1	CA	1972	15	1285.2	-1190	1938.3	-1845.6
2	CA	1982	3	1288.7	-1218	1974.7	-1977.6
3	CA	1982	25	1557.9	-1297	1993.9	-1770.5
4	CA	1982	43	729.5	-748	2016.9	-2010.8
5	FL	1992	14	712.4	-697	1998	-1868.6
6	FL	1992	17	1412.3	-1343	1972.8	-1865.6
7	FL	1992	21	775.6	-849	1976.3	-2268.4
8	FL	1994	14	672.7	-665	2002.9	-1920.6
9	FL	1994	17	1472.6	-1403	2004.6	-1945.9
10	FL	1994	21	602.7	-588	1967.1	-2030.5
11	FL	2012	4	723.7	-664	2066.2	-1833.5
12	FL	2012	21	1403.3	-1149	2070.6	-1929.6
13	FL	2012	25	856	-923	2008.1	-1917.3
14	FL	2014	4	615.1	-616	1946.9	-1801.3
15	FL	2014	21	1329.1	-1136	2070.1	-2057.7
16	FL	2014	25	823	-720	2079	-1863.4
17	GA	2002	5	1510.4	-1489	2042.2	-1596.2
18	GA	2002	10	467.4	-477	2011.2	-2025.2
19	GA	2004	5	1565.9	-1411	2007	-2125.7
20	GA	2004	10	484.1	-545	2022.2	-1696.3
21	GA	2008	4	1660.1	-1662	2013.4	-1715.5
22	NC	1998	7	1174.8	-1029	2070.1	-1804.8
23	NC	1998	10	652	-646	1996.9	-1900.4
24	NY	1972	9	1105.6	-1031	1989.7	-1799.7
25	OH	1982	3	1441.5	-1485	2005.8	-1851.3
26	OH	1984	3	1226.2	-1265	1953.3	-1763.8
27	TX	1972	1	1222.6	-1192	1957.6	-1922.4
28	TX	1972	10	1312.1	-1356	1957.9	-1900.1
29	TX	1972	11	1238.9	-1275	1971.5	-2006.8
30	TX	1972	12	1312.4	-1403	1979.2	-1864.2
31	TX	1972	14	1316.2	-1567	1959.9	-1848.2
32	TX	1972	15	1432.6	-1324	2046.9	-1907
33	TX	1972	16	1285.8	-1071	1971.4	-1695.1
34	TX	1972	17	1216.4	-1260	1984.4	-1950
35	TX	1972	19	1189	-1202	1957.6	-2215.6
36	TX	1972	20	1570.2	-1466	1985.9	-1925.4
37	TX	1972	23	1311.2	-1122	1976.7	-1834.3
38	TX	1974	20	17582.3	7075	20614.7	9009.1
39	TX	1982	1	1438.1	-1464	2006.3	-1874.9
40	TX	1982	2	1465.8	-1378	2036.2	-1839.8
41	TX	1982	6	1296.2	-1530	1916	-1794
42	TX	1982	10	1412.2	-1353	1960.7	-1870.2
43	TX	1982	11	1362.8	-1417	1992	-1766.7
44	TX	1982	20	1633.1	-1567	1996.7	-1806.2
45	TX	1982	22	626	-643	1986.4	-1978.5
46	TX	1992	3	519.5	-501	1998.5	-1781.2
47	TX	1992	7	484.7	-471	2022.8	-1989.4
48	TX	1994	3	459.5	-467	1991.7	-1658.4
49	TX	1994	7	415.9	-461	1912.4	-2121.4
50	TX	2002	7	620.7	-525	2051.4	-2001.3

	stpost	year	district	dvote_bar	dvote_lo	totalTwoparty_bar	totalTwoParty_lo
51	TX	2002	8	455.8	-470	1949.9	-2065.4
52	TX	2002	10	1329.9	-1292	2022.4	-1938.7
53	TX	2002	12	612.8	-631	1986.4	-1881.7
54	TX	2002	15	1357.6	-1358	2036.5	-2015.2
55	TX	2002	16	1387	-1398	1985.8	-2234.8
56	TX	2002	19	506.9	-506	2024	-1836
57	TX	2002	20	1387	-1194	2028.2	-1882.3
58	TX	2002	29	1407.9	-1476	1988	-2067.5
59	VA	1992	7	631.9	-712	2039.4	-1741.7

57. There may be something in the 3,000 lines of Dr. Jackman's code that corrects for this, but I was not able to identify it. As you can see, 59 of Dr. Jackman's imputed values show implausibly small results. The smallest reported result comes in Texas' Seventh District in 1994, where Dr. Jackman suggests 1,912 votes would have been cast had the Republican, Bill Archer, drawn an opponent. This cannot be correct, given that Archer actually garnered 116,873 votes in 1994. *See* Philip D. Duncan & Christine C. Lawrence, *Politics in America 1996: The 104th Congress* 1269 (1995) [Note: Politics in America will hereinafter be referred to as simply *PIA*, followed by the appropriate year. Note further that Politics in America is somewhat counterintuitively dated for the year of the cycle *after* the election it covers. In other words, *PIA 1996* covers the 1994 elections. The same is true for Almanacs of American Politics]. In fact, it is doubtful whether any of these results represent realistic imputations.

58. These apparent errors are problematic for four practical reasons. First, these results create implausible variance values for Dr. Jackman's estimates. As you can see from the table, Dr. Jackman's low-end estimates for turnout and democratic vote share (totalTwoParty_lo and dvote_lo, respectively) contain negative values, which cannot occur. While this does not affect his estimated efficiency gaps directly, which are keyed off of the variables ending in "_bar," they do affect his "error margins," which will be calculated in part on the basis of these impossible values.

59. Second, the errors clearly affect Dr. Jackman's efficiency gap calculations at the granular level. Consider North Carolina in 1998.The state had an efficiency gap of -.0168 that year. According to Dr. Jackman's estimates, Mike McIntyre would have defeated his Republican opponent by 1174 votes to 895 votes, had he drawn an opponent. In North Carolina 10, Republican Cass Ballenger would have prevailed 1,345 to 652 had he drawn a Democratic opponent. If we multiply these values by 85, to roughly approximate overall turnout in North Carolina in other districts in 1998, the efficiency gap flips signs, and has a value of .005.

60. The deviations can be more extreme. Consider Texas's 1972 map, which presents with a modest efficiency gap of .035. If we take Dr. Jackman's estimates and multiply them by 75 – which would bring turnout roughly into line with that found in elections in Texas with actual returns in that year, the efficiency gap balloons to .099, which would be actionable under Plaintiffs' standard.

61. Third, although this might not be a significant problem if errors distribute randomly – 59 erroneous observations out of 7,900 elections might well "come out in the wash" – a close inspection of these data show that the errors are not random. Instead, they almost all occur in the first years of plans. So not only will their impact be concentrated in particular plans, they will be concentrated in the most important plans for purposes of this litigation.

62. Fourth, this may turn out to be a relatively simple coding error to correct, but that misses a larger point. If a well-respected political scientist like Dr. Jackman makes mistakes like these, it puts a bit of a lie to the claim that these procedures are simple to implement.

63. There also seem to be some problematic imputations for vote shares. While it was not practicable to review all 1,114 imputations in the available time, in the course of my investigation I observed a few imputed vote shares that seem highly unlikely as well.

64. For example, Dr. Jackman shows Alabama Congressmen Bill Nichols winning with just 56% of the vote, Tom Bevill winning just 55% of the vote, and Robert Jones winning just 56.4% of the vote. Given the error on Democratic vote totals – ignoring any effect from changing *overall* vote totals – the projected vote shares for these members of Congress drop as low as 38.5%, 40.1% and 38.6%, respectively. (I note that combining a high estimate for Democratic votes and low estimate for vote totals would yield 128% of the vote for Alabama's Walter Flowers in 1972, though I presume Dr. Jackman truncates his output given how he wrote down his model).

65. During his 28-year career, the vote share that Robert Jones received ranged from between 71.3% of the vote and 91.62% of the vote against Republican opponents. Bill Nichols received between 58.39% of the vote and 83.71% of the vote during a 24-year career, with the 58.39% share coming the year he defeated a Republican incumbent to win his seat in Congress (1966). Setting aside that year, his lowest vote share was 75.57%. Tom Bevill's lowest vote share was 64.36%, also in 1966, the year he won his seat. *See* Michael J. Dubin, *United States Congressional Elections, 1788-1997: The Official Results* (1998). It seems extremely unlikely that the most likely outcome in 1974 -- probably the best Democratic election in the past fifty years – would be an extremely close election result, with a reasonable chance that these Democrats would lose.

66. Likewise, in 1972 and 1974, Dr. Jackman suggests that Bob Sikes of Florida would have won with 59% and 58% of the vote, respectively. Over the course of his 38-year career in Congress, Sikes was rarely opposed. His lowest vote total, 80%, came in 1970. Interestingly, even the high-ends of Dr. Jackman's estimates for Sikes' vote shares – 78% and 79%, respectively – would represent the worst showings of his career. *Id.*

67. Finally, in 1998, Dr. Jackman suggests that Mike McIntyre would have received 56.8% of the vote. The 1998 elections were good for North Carolina Democrats, who won back control of the North Carolina House and picked up five state Senate seats. Michael J. Dubin, *Party Affiliations in the State Legislatures: A Year-By-Year Summary, 1796-2006* (2007). Nationally, Democrats became the first party to pick up House seats in a midterm election while holding the presidency since 1934. But between 1998 and the Republican wave of 2010, McIntyre never fell below 69 percent of the vote against Republicans. It seems unlikely that he would have been in real danger of losing in an otherwise good Democratic year.

68. Of course, there is no real way of knowing for certain whether these estimates are valid. This is because they all suffer from the same problem from which the partisan symmetry standard suffers: we don't have access to the counterfactual, and are ultimately estimating the result.

V. **There appear to be errors in Dr. Jackman's sensitivity testing (and what the tests really mean).**

69. Dr. Jackman produces a series of charts on pages 44 and 45 that constitute his sensitivity analysis. That is, these charts show how, at various thresholds of efficiency gaps, how many false positives, false negatives, and so forth would be produced. I will explore these in depth later on, but for now would like to discuss what appear to be errors in Dr. Jackman's analysis. Because I was not able to reproduce his results here, I'm unable to tell whether the errors are in his charts or in his written analysis. But they do not appear to match.

70. Dr. Jackman explains that Figure 14 of his plan shows the sensitivity analysis for plans with three or more elections from 1972 to 2016. Figure 15 shows the results using only maps since 2000. Dr. Jackman writes: "[g]enerally, the precision of a prognostic test based on the 1st election EG is high, approaching 90% once the 1st election EG is greater than .03 in

magnitude." *Id.* at 46. But if one examines the relevant chart – the top-right pane on page 44 – at an efficiency gap of .03, the precision appears to below 75 percent. While Dr. Jackman's discussions of the percentage of plans that are flagged for having high efficiency gaps matches up with what is seen in the top left pane of page 44, he writes that "[a]t this threshold [.15] the precision of the test criterion remains high (or conversely, the false discovery rate stays low), but the false omission rate has climbed to 76%." This again seems wrong; the bottom right pane of his chart suggests a false omission rate that is below the 75 percent line.

71. Dr. Jackman continues: "[t]he overall accuracy of the test falls to around 32% if one were to adopt a very stringent threshold such as |1stEG| > .15." *Id.* But the accuracy at this rate, which is the bottom left corner, is above the 37.5 percent threshold. The rest of the numbers Dr. Jackman identify seem to match up with this page.

72. Dr. Jackman then switches to the districting plans enacted since the 2000s, displayed on page 45. Dr. Jackman writes that "[t]he false discovery rates in Figure 15 are all zero once the 1st election EG is .03 or greater in magnitude." *Id.* But this does not appear to be correct; the false discovery rate is the middle-right-lower pane, and the line there does not appear to reach zero until an efficiency gap of approximately .125. Nor does the false positive rate appear to be "zero once beyond a 1st election EG of .03." *Id.* At an efficiency gap of .03, the chart on page 45 seems to show a fairly substantial false positive rate of around 37 percent. Nor does the precision seem to reach 100 percent until the efficiency gap exceeds 12 percent.

73. Next, Dr. Jackman's way of describing the tests, while easily accessible for people with a statistical background, may not be so easily digested by courts or lawyers. Although average people perform conditional probabilities all of the time (whether they know it or not), they don't often encounter notation. Here is my interpretation of what Dr. Jackman's

charts show. Because I wasn't able to reproduce his findings here, my descriptions are approximate.

74. Utilizing all plans with three or more elections as our test:

- 45 percent of plans would be flagged with an EG standard of .07, while 25 percent would be flagged utilizing a standard of .12;
- If a plan has an average efficiency gap score that is consistent with the first election, there is a 50 percent chance it will be identified as actionable with an EG standard of .07, and a 25 percent chance it would be identified as actionable with an EG standard of .12;
- If a plan has an average efficiency gap score that is ***not*** consistent with the first election, there is a 67.5 percent chance it won't be identified as actionable with an EG standard of .07, and an 85 percent chance it won't be identified as actionable with an EG standard of .12;
- If a plan is flagged as actionable, there is a 75 percent chance it will have an average efficiency gap consistent with the first election, with an EG standard of .07, and a 75 percent chance it will have an average efficiency gap consistent with the first election , with an EG standard of .12;
- Around 60 percent of cases will either be a true positive or true negative with an EG standard of .07, while 45 percent of cases will either be a true positive or true negative with an EG standard of .12;
- If a plan has an average efficiency gap that is ***not*** consistent with first election, there is a 37.5 percent chance it will nevertheless be flagged as actionable with an

EG standard of .07, and a 10 percent chance it will be identified as actionable with an EG standard of .12;

- If a plan is flagged as actionable, there is a 25 percent chance it will not have an average efficiency gap consistent with the first election, with an EG standard of .07, and a 25 percent chance it will have an average efficiency gap consistent with the first election , with an EG standard of .12;
- If a plan is not flagged as actionable, there is a 65 percent chance it will nevertheless have an average efficiency gap with an average that matches the sign of the first-year plan, with an EG standard of .07, and a similar chance with an EG standard of .12;

75. Utilizing all plans with three or more elections as our test, confining ourselves only to 21st century redistricting:

- 50 percent of plans would be flagged with an EG standard of .07, while 25 percent would be flagged utilizing a standard of .12;
- If a plan has an average efficiency gap score that is consistent with the first election, there is a 67.5 percent chance it will be identified as actionable with an EG standard of .07, and a 25 percent chance it would be identified as actionable with an EG standard of .12;
- If a plan has an average efficiency gap score that is ***not*** consistent with the first election, there is a 67.5 percent chance it won't be identified as actionable with an EG standard of .07, and an 87.5 percent chance it won't be identified as actionable with an EG standard of .12;

- If a plan is flagged as actionable, there is an 87.5 percent chance it will have an average efficiency gap consistent with the first election, with an EG standard of .07, and an 87.5 percent chance it will have an average efficiency gap consistent with the first election , with an EG standard of .12;
- Around 67.5 percent of cases will either be a true positive or true negative with an EG standard of .07, while 40 percent of cases will either be a true positive or true negative with an EG standard of .12;
- If a plan has an average efficiency gap that is ***not*** consistent with first election, there is a 35 percent chance it will nevertheless be flagged as actionable with an EG standard of .07, and a 7.5 percent chance it will be identified as actionable with an EG standard of .12;
- If a plan is flagged as actionable, there is a 7.5 percent chance it will not have an average efficiency gap consistent with the first election, with an EG standard of .07, and a 7.5 percent chance it will have an average efficiency gap consistent with the first election , with an EG standard of .12;
- If a plan is not flagged as actionable, there is a 70 percent chance it will nevertheless have an average efficiency gap with an average that matches the sign of the first-year plan, with an EG standard of .07, and a similar chance with an EG standard of .12.

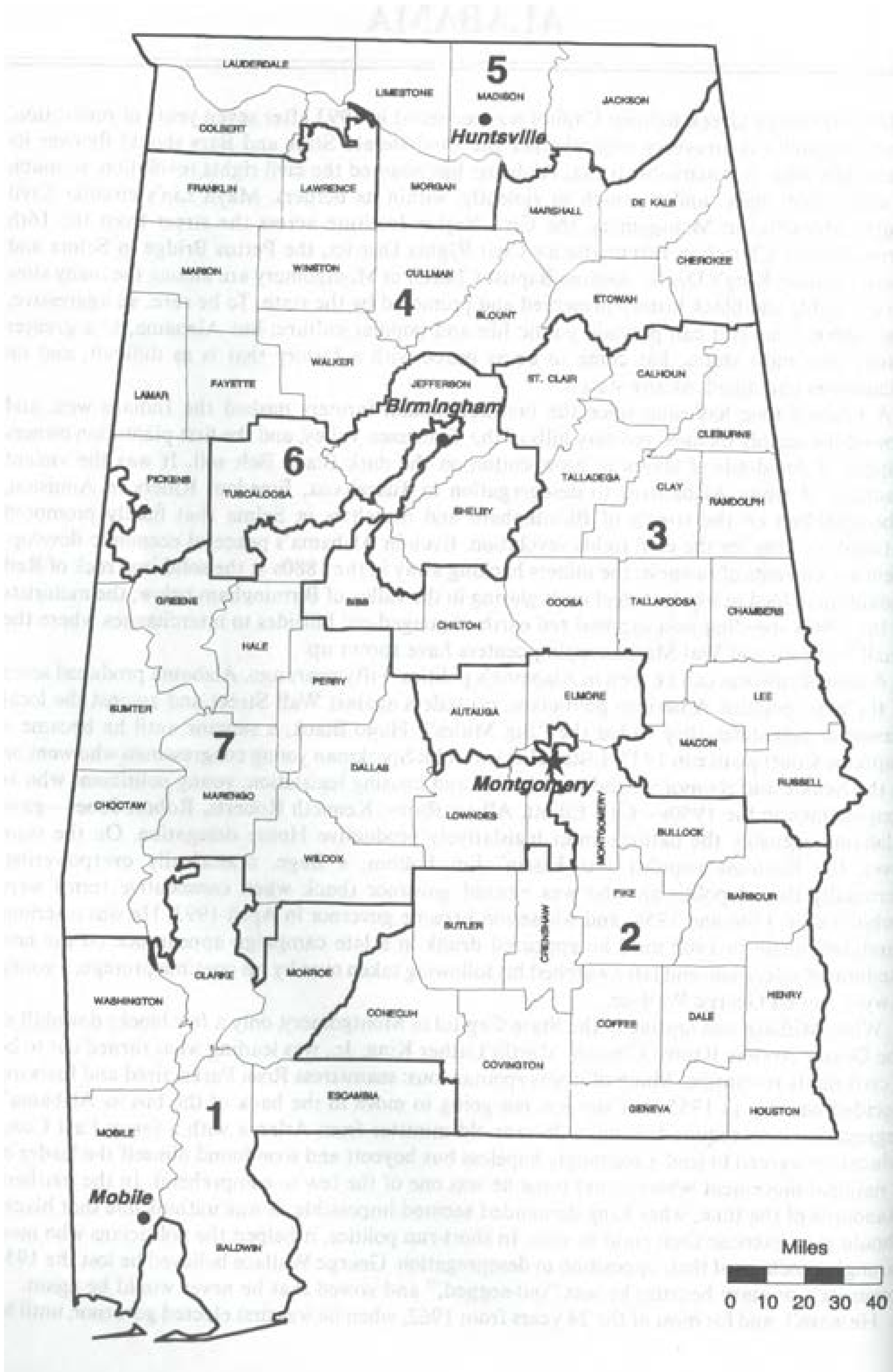
75. Some of these strike me as having high rates of "misses," but whether or not the Court is comfortable with these rates of misses is ultimately a question for the courts and lawyers, not the experts.

VI. <u>The Efficiency Gap is not clearly the hallmark of a gerrymander, as commonly understood.</u>

76. The efficiency gap should theoretically tell us something about gerrymandering on its own terms, if it is truly the "hallmark" of a gerrymander. If we have multiple instances of maps drawn by heavily Democratic legislatures, court-drawn maps, and commission-drawn maps that display large, or actionable, efficiency gaps, it calls into question whether the efficiency gap is as intimately tied to gerrymandering as plaintiffs suggest. At the very least, it provides further evidence that the baseline efficiency gap of zero that plaintiffs assume is not always appropriate.

77. Consider Alabama. In 1992, Democrats had a 27-8 lead in the state senate, an 81-23 lead in the state house, and control of the governorship. *The Almanac of American Politics 1994*, [hereinafter *AAP*] at 7 (1993). This would appear to be an unlikely candidate for a pro-Republican efficiency gap, to say nothing of an *actionable* Republican plan. Yet this is exactly what we see. The efficiency gap exceeded the threshold for an actionable efficiency gap in 1992, 1994, 1996, and 1998.

(Alabama 1992 redistricting. Maps from the *Almanac of American Politics* unless noted)



78. Interestingly, the 1994 efficiency gap favored *Democrats*. This underscores a consistent feature of the efficiency gap, developed in subsequent sections: It is highly dependent on factors that the metric itself does not take into account: incumbency, challenger quality, and the national environment. Because of this, on a very practical level, mapmakers are challenged with guessing what the environment and candidate choices will look like in the first year of implementation.

79. The Alabama example (as well as later ones) puts this in stark relief. Had the 1994 wave election come in 1992, we would see the map as a heavy Democratic gerrymander, because the wave would have inflated Republican vote scores without overcoming legendary Southern Democrats like Tom Bevill and Glen Browder. But 1992 was, generally speaking, a good Democratic year. Longtime Republican incumbent William Dickinson retired in the Second District, and George C. Wallace, Jr., ran for the seat. Wallace lost, but only by 3,000 votes, generating a huge number of wasted Democratic votes. *AAP 1994* at 14. In the 6th district, Democratic incumbent Ben Erdreich found himself placed in a district that was giving George H.W. Bush a 40-point win, due to the need to turn Erdreich's district into a minority-majority district to comply with the Voting Rights Act. But, given Erdreich's incumbency and the generally good Democratic environment, he lost by only seven points, creating a large number of wasted Democratic votes. *AAP 1994* at 22.

80. Then consider 1996. Browder and Bevill had held the third and fourth districts for years (decades in Bevill's case). In a good Republican year like 1994, Republicans might receive a fair number of votes, but they would almost always be wasted. A mapmaker would then have to decide how to deal with these popular incumbents. Then, in 1996, Browder decided to run for Senate (he did not clear the primary), while Bevill retired at 75 years young. The districts were

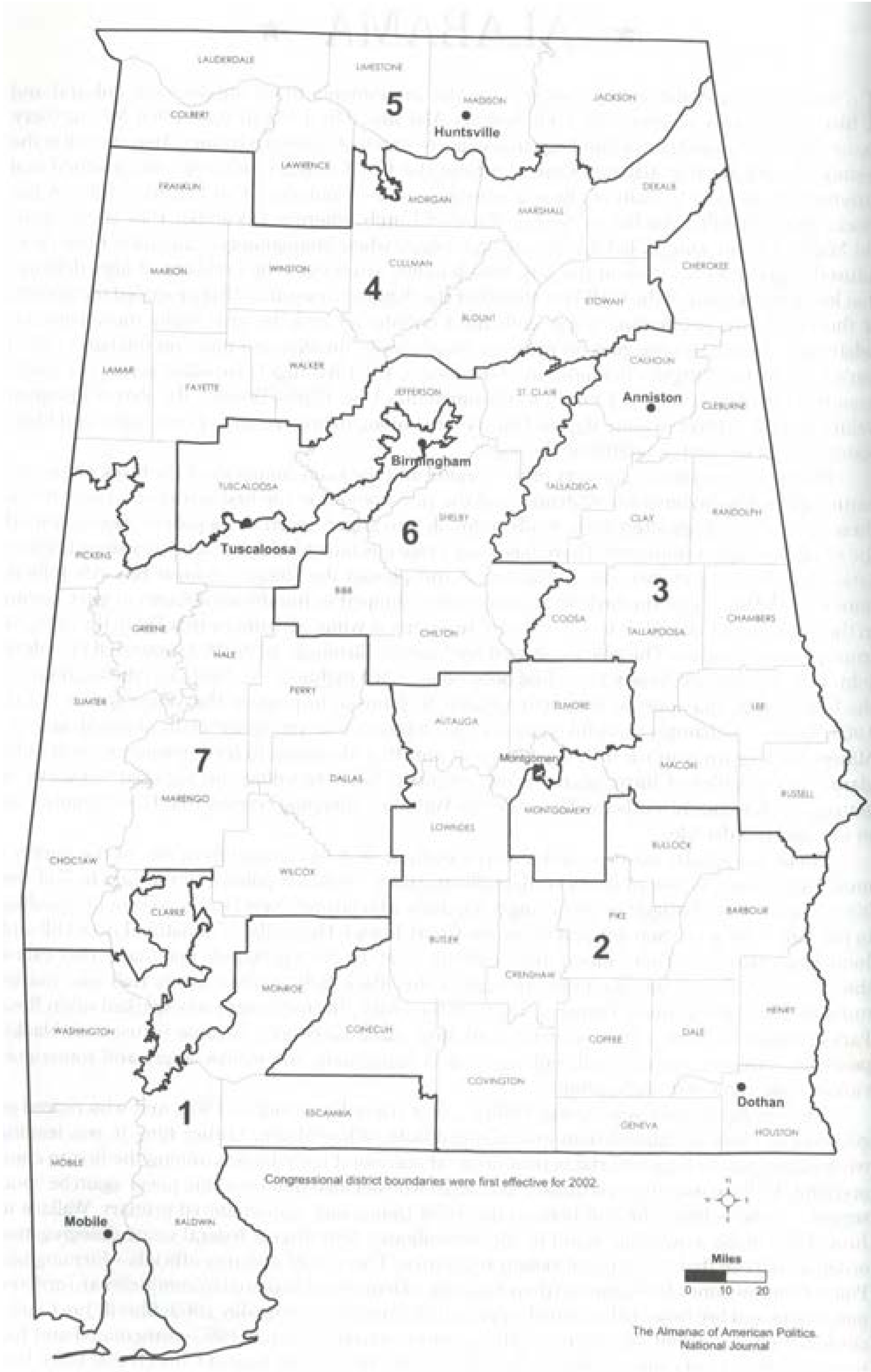
competitive; Bob Riley won by four points, while Robert Aderholt won by two, *PIA 1998* at 14, 16, which generated a huge number of wasted Democratic votes and a map that presents as an actionable Republican gerrymander. The seats were competitive again in 1998, creating a significant efficiency gap that year as well. But by 2000, Democrats had pretty well given up, and Republicans began to waste votes in these seats, and the efficiency gap disappeared.

81. To summarize, in Alabama in 1992, a map drawn by Democrats presents as an actionable Republican gerrymander three times, an actionable Democratic gerrymander once, and a weak Republican gerrymander once. This is a poor performance for something touted as the hallmark of a gerrymander. Moreover, given that this map can clearly present in different ways, a mapmaker would be charged with guessing what the national environment would be like in the first year of implementation, when popular incumbents will retire, when incumbents will find themselves saddled by scandal, and when parties will field strong challengers.

82. In 2002, Democrats attempted to maximize their precarious position in the state. As the Almanac of American Politics describes: "The Democrats in control of redistricting in Alabama in 2002 did a pretty good job of helping their party in drawing the boundaries of the state's seven congressional districts, but not quite good enough of a job to add to the two seats they have held in 1994." *AAP 2004* at 54. Democrats strengthened Bud Cramer in the 5th District, while making the 3rd District substantially more Democratic by increasing the African-American percentage from 25 percent to 32 percent. *Id.* In other words, the 2000 map in Alabama was a Democratic gerrymander; it was just an unsuccessful one. Mike Rogers won with just 50 percent of the vote, creating 87,351 wasted Democratic votes and just 1,909 wasted Republican votes. The result? A map that was created to maximize Democratic opportunities by

a Democratic legislature with a Democratic governor presents with a Republican-leaning efficiency gap that is almost actionable: -.114.

Alabama 2002 redistricting



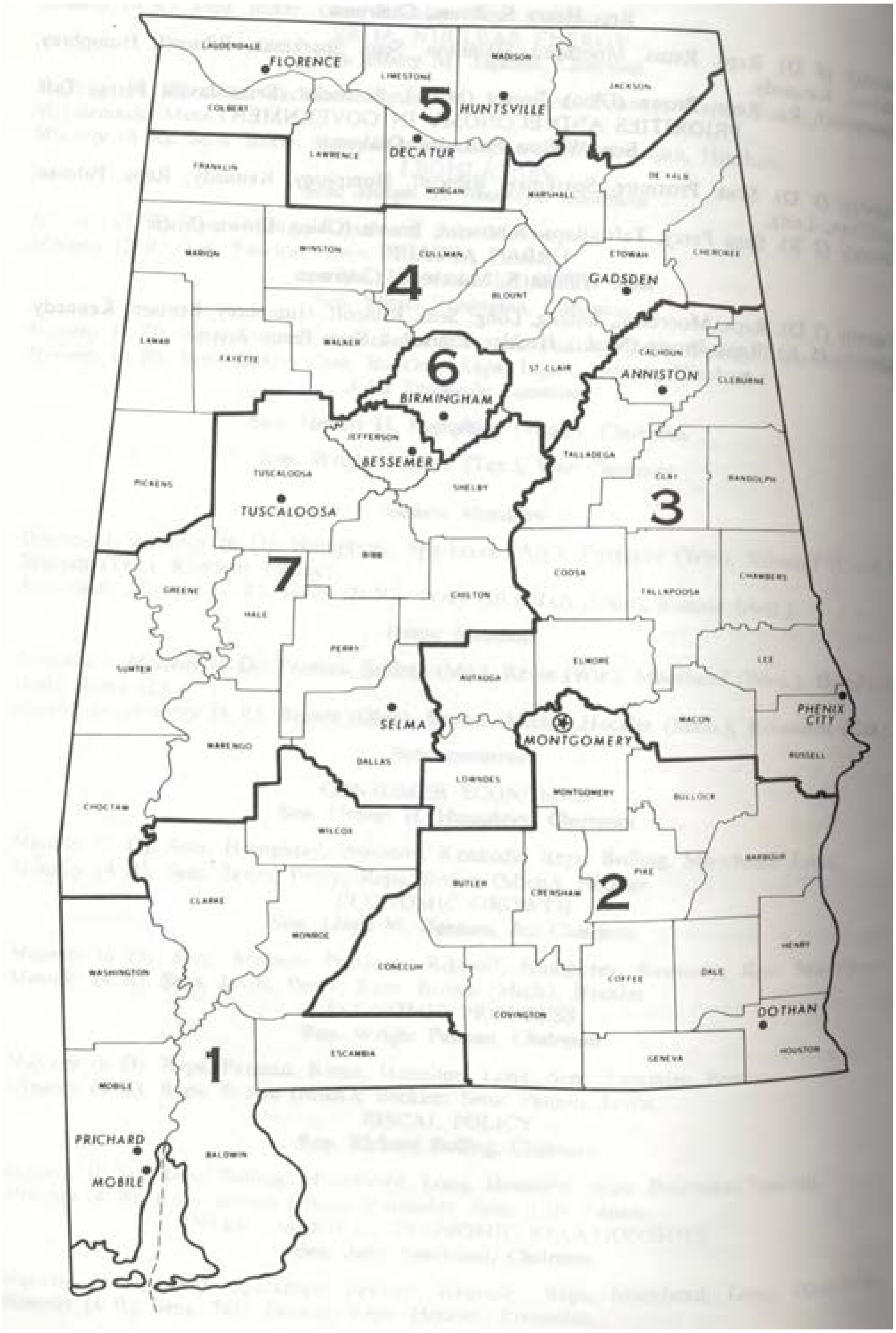
83. In fact, in 2010, when Republicans managed to capture the Fifth, the map appears to be a Republican gerrymander. Had the environment from 2010 presented in 2002, and had Cramer retired in 2002, it probably would have been an actionable gerrymander that year. The Democratic nature of the plan wasn't on full display until the Democratic wave to 2008, when Democrats managed to capture the 2nd district, which likely was possible only because of the national Democratic wave and the retirement of Republican incumbent Terry Everett. *AAP 2012* at 49. Again, we might ask ourselves, what if a Republican legislature had drawn the exact same map, and a Democratic wave had occurred? What might otherwise be considered a strong Republican gerrymander would be interpreted by courts as a Democratic map.

84. In fact, the history of Alabama politics during this time period shows that rather than being a clear-cut "hallmark of a gerrymander," interpreting the efficiency gap requires a fact-intensive inquiry. Across the four maps drawn by Democrats, the average efficiency gap leans Republican, with an average gap of -.039. It only leans Democratic 25 percent of the time.

85. Perhaps Alabama Democrats were particularly beneficent, but it is more likely that this is a function of the long-standing effects of a freak election: 1964. This was a combination of the fact that Southern cities and coastal plains had been trending Republican for a while, the fact that Democrats had failed to pass a redistricting plan in 1962 after losing a seat, so incumbents were running in new districts after having been elected at large in 1962, and, of course, the backlash to the passage of the Civil Rights Act of 1964. Republicans captured five of Alabama's eight congressional districts, wiping out 86 years of congressional seniority. Democrats defeated one of these representatives in 1966, and picked up an open seat that same year, but the remaining three members managed to survive.

86. It was not for lack of trying on the part of Democrats. During the 1970 redistricting, there were no Republican state senators, and just two Republican state representatives (out of 106). Dubin at 17. Alabama lost a seat in 1972, and Democrats attempted to eliminate William Dickinson by combining his district with that of longtime Democratic incumbent George Andrews, a race the Almanac of American politics expected Andrews "can be expected to win fairly easily." *AAP 1972*, at 8; *AAP 1974*, at 8. This was probably a reasonable assumption, but then Andrews died unexpectedly in late 1971 at the relatively young age of 64. So, instead of running against a popular incumbent with 28 years of congressional seniority, Dickinson ran against a young state legislator. A race he probably would have lost, generating a large number of wasted Republican votes, turned into a narrow Republican win, generating a large number of Democratic votes.

(Alabama 1972 redistricting)



87. At the same time, the remaining two Republicans – Jack Edwards and John Buchanan – continued to win elections, although their vote shares were rarely above 60 percent.

The result? Relatively few wasted Republican votes, lots of wasted Democratic votes, and a negative efficiency gap in every year of the map's existence, with the exception of 1974 (when some of the imputation issues above play a role).

88. Notice too that efficiency gaps are close to zero in the 1980s. This is not a result of redistricting – the lines were barely changed in 1982. *AAP 1984* at 4. Instead it was a result of the sort of random effect that mapmakers couldn't predict with regularity: Buchanan lost the Republican primary in 1980 due, in part, to a series of votes cast that were perceived as being friendly to African-Americans. The man who defeated him, Albert Lee Smith, won narrowly in the Reagan landslide, creating a large number of wasted Democratic votes. Smith then lost in the good Democratic year of 1982, and a district that generated a large number of Democratic wasted votes began to generate wasted Republican votes.

89. I dwell at length on the Alabama example because it is a state whose politics I happen to know well, and also because it illustrates nicely just how reductionist it is to suggest that gerrymandering can be summarized by a single statistic. Gerrymandering has no particular hallmark; it is inherently a fact-intensive inquiry. This is especially true if we are going to measure it in terms of elections, which are frequently beset by unpredictable effects that have long-lasting consequences.

90. Or consider the 1982 California map. This is widely recognized as one of the more egregious partisan gerrymanders in history. It was designed by Rep. Phil Burton and the brother of future Rep. Howard Berman; the former famously referred to it as his "contribution to modern art." It was so egregious that it was repealed by referendum, and Stephanopoulos and McGee take particular pride in the fact that their metric identifies the Burton-mander as having one of the largest efficiency gaps in history. But it barely crosses Dr. Jackman's threshold for

actionability, at .0796. Moreover, Burton's response to the court ruling was to tweak some of the lines; contemporary sources refer to the 1984 map as "similar," AAP 1984, and "slightly revised." AAP 2004, at 155. Yet the map is no longer actionable in any of the remaining years.

91. Remaining California maps illustrate how the efficiency gap can go astray in the correct circumstances. The bipartisan "incumbent protection" map of 2002 actually shows as an actionable Republican gerrymander in 2006 and 2008, notwithstanding the fact that Democrats had controlled the process; had 2002 been a Democratic wave, the map might be actionable. The plan drawn by the independent redistricting commission displays as a Democratic gerrymander in 2014. This makes sense, as maps that seek to be fair by failing to threaten the other side's incumbents will tend to appear as gerrymanders in wave elections.

92. Colorado's 2002 map isn't actionable (it barely misses), but it illustrates nicely how a quite substantial efficiency gap can have nothing to do with gerrymandering, and everything to do with random effects. The map is actually a Democratic gerrymander that was selected by a state court judge; it drew the new 7th district in the older, more Democratic inner suburbs of Denver instead of the newer suburbs to the south, as Republicans preferred. *AAP 2004* at 304. So why does it produce a Republican-leaning efficiency gap of -.1095 in its first year? Because the Democrat for whom this map was designed, Ed Perlmutter, declined to run in 2002, and a Republican, Bob Beauprez, won by 121 votes. That meant that all 81,668 votes cast for his opponent were wasted, while Republicans wasted just 61 votes. *Id.*

93. Those 61 votes were the difference between an almost-actionable efficiency gap (7,000 votes spread across the remaining districts are all that prevent it from being actionable)remaining districts are all that prevent it from being actionable), and a Democratic-leaning efficiency gap of .012. Put differently, a switch of 61 votes here would result in a swing

of 12% in the efficiency gap. But that is the sort of things that map drawers will have to plan for somehow. It also seems like an exaggeration to refer to those 61 votes switching as the hallmark of a gerrymander.

94. One of the more egregious examples of the efficiency gap at work is the fate of the Georgia plans. Here is how the Almanac of American Politics 2004 describes the processes:

> "After the 1990 and 2000 Censuses, Georgia Democrats, led by Speaker Thomas Murphy, pushed through convoluted redistricting plans – arguably the most convoluted in the nation each time – to guarantee majorities for their party in the state's House delegation. Both times they failed. In the 1990s Murphy tried to end the career of Newt Gingrich and strengthen incumbent Democrats. Instead, what was a 9-1 Democratic delegation in October 1992 was 8-3 Republican in April 1995, and Gingrich was Speaker of the House. A court-ordered redistricting in 1995 left virtually all incumbents with safe seats, and the balance remained 8-3. In 2001 the Democrats tried again, drawing several plans and negotiating among themselves. This time the boundaries were even more convoluted, and Democrats had a bit more success. But only a bit – plus some unintended consequences."

AAP 2004, at 454. The Almanac continues:

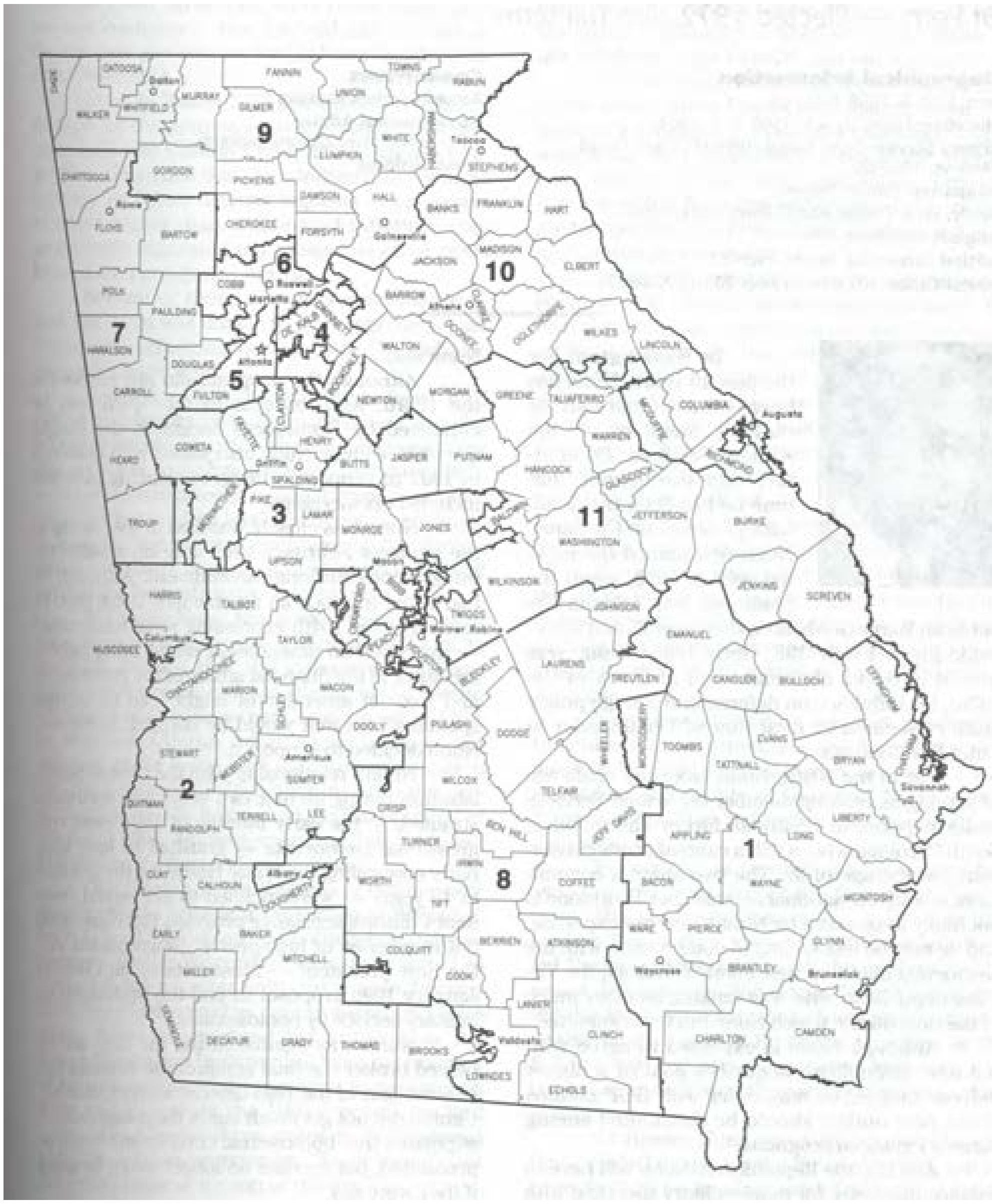
> "Still, the Democrats' plan must be admired for its creativity. The new 13th District sense narrow tentacles into 11 metro Atlanta counties to unite black neighborhoods along strip highways or in town centers with majority-black Clayton County just south of Atlanta. The new 11th District made a stab, though unsuccessful in 2002, at creating a Democratic district in Republican northernwest Georgia by excluding fast-growing mostly white areas and sending

in tentacles to south Cobb County with its increasing population . . . . The new 12th District for the most part has a regular shape and yet connects black neighborhoods in cities as distant as Savannah, Augusta and Athens. Heavily Republican areas are packed into five districts."

Id.

95. This would seem to represent a classic candidate for an illegal Democratic gerrymander: Heavy partisan intent to gerrymander matched with bizarre lines. But this is not what happens. In its first year of implementation – the map designed to create a 9-1 Democratic advantage produces, narrowly, a Republican leaning efficiency gap, as Republicans won surprise victories in the 1st and 3rd districts. in two (Congressman Ben Jones, who played "Cooter" on the Dukes of Hazzard, lost the Democratic primary in the 10th district). Had Democrats won the races that they were supposed to win, the efficiency gap would likely have been actionable as a Democratic gerrymander. But because they were bad at gerrymandering, they receive a pass.

Georgia 1992 Redistricting (Politics in America 1996)

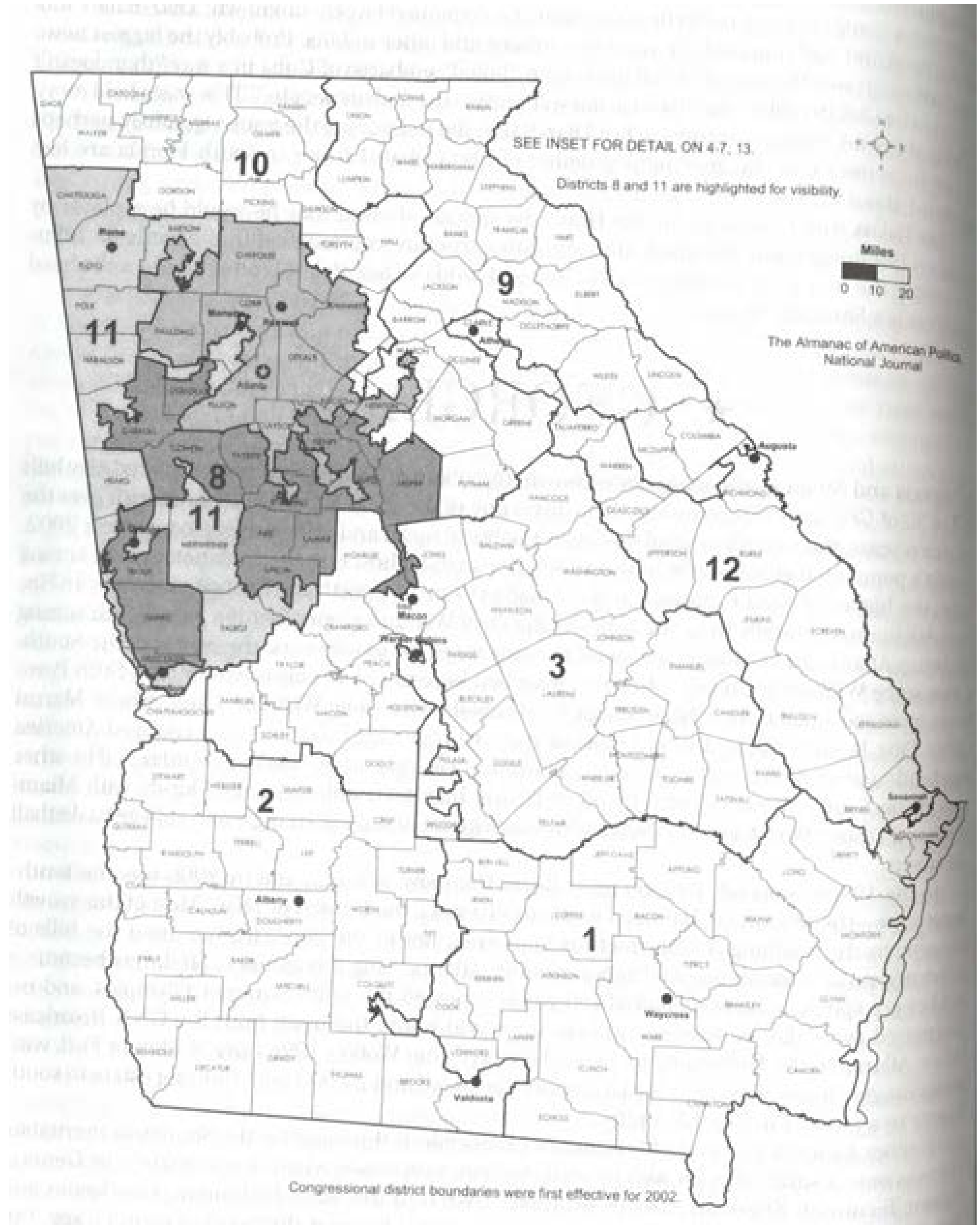


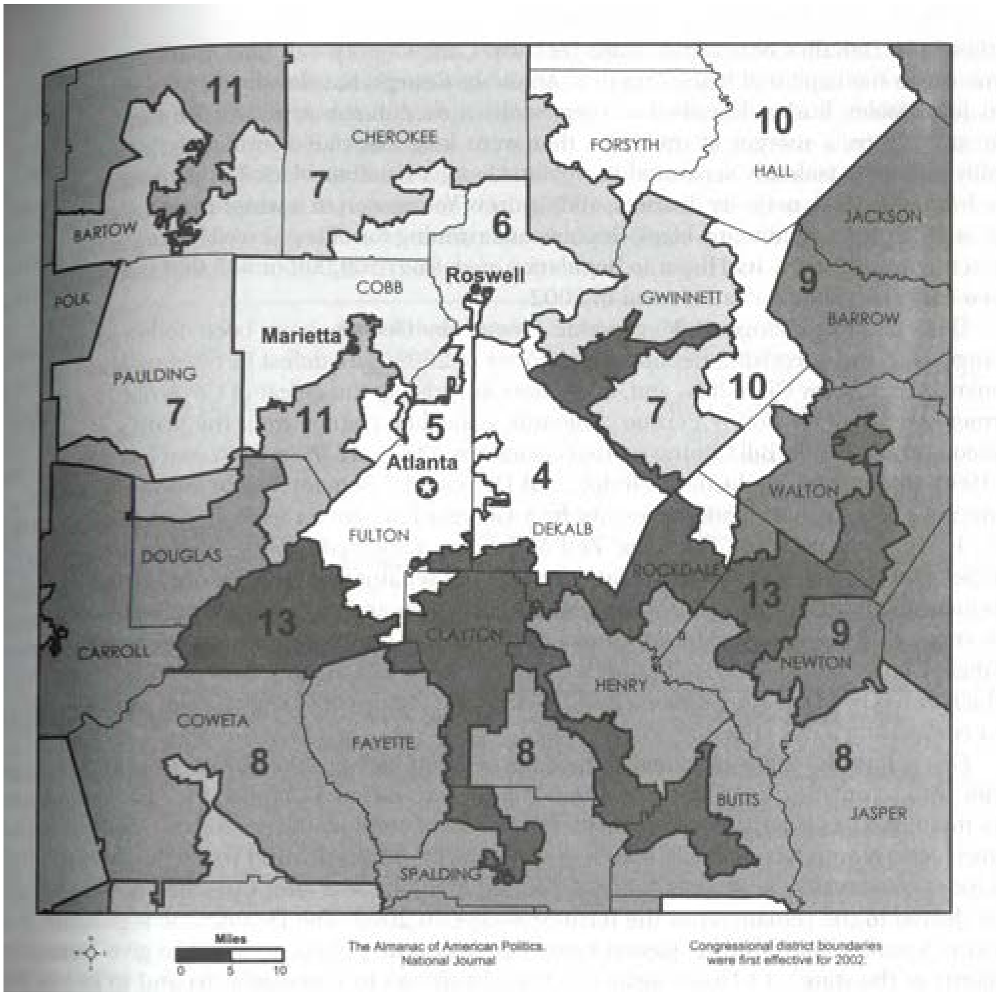
96. The map was eventually struck down as a *racial* gerrymander, and the subsequent map was drawn by Democrats. This map, however, does produce an actionable efficiency gap, that favors Republicans. According to the efficiency gap standard, the Democrats drew a map that was heavily in their favor; yet the efficiency gap was -.161 that year. Moreover, the map was negative over the duration of its implementation, so this shows up in Dr. Jackman's sensitivity analysis as an example of the efficiency gap performing exactly as it should perform. This shows

that it is possible to achieve a first-year efficiency gap that almost rivals that found in North Carolina, even with a map drawn by the opposing party with intent to favor that party. Far from being the hallmark of a gerrymander, the efficiency gap here is just a fluke outcome.

97. In 2002, Democrats drew an even more aggressive gerrymander. It at least shows up as Democratic-leaning, but not actionable. But again, this is a result of a few tight races not breaking their way. Had Democrats won in the 11th and 12th districts as intended, the efficiency gap would probably have been somewhere on the order of .16, and been actionable. Those two districts were decided by fewer than 10,000 votes, combined. Had 2002 not been a good Republican year overall (it was the second time the party holding the presidency picked up seats in a midterm since 1934), and had Democrats not nominated the son of the Senate majority leader for one district, who turned to have been arrested four times and been involved with multiple failed business ventures, it probably would have been actionable.

Georgia Redistricting 2002





98. Of course, in 2006, Republicans drew a gerrymander of their own. Their mid-decade redistricting was aimed at dislodging Jim Marshall in the third district and John Barrow in the 12th District. But this map actually has a Democratic lean, according the efficiency gap, although it is not actionable. Again, this shows how the interaction of the national environment and a few key candidates can radically alter efficiency gap scores. Both Marshall and Barrow survived, by a combined 1,308 votes. So there were almost no wasted Democratic votes in this district, but a great many wasted Republican votes. Had 655 votes flipped, the efficiency gap would have been -0.116, which would make more sense. Whether it would have been actionable then probably turns on the quality of Dr. Jackman's imputation in the 5th district. If actual

turnout would be at the lower end of Dr. Jackman's estimation, then the map would have been actionable.

99. The efficiency gap results for Illinois are particularly bad. In 1982, a federal court selected a partisan Democratic plan, and Democrats picked up two seats. *AAP 1984*, at 326. Yet the map shows as an actionable Republican gerrymander, then has a Democratic efficiency gap for the remaining four years of the plan's life.

100. Much more egregious is the sequence from 2002 to 2012. In 2002, Republicans controlled the state senate and governorship, while Democrats controlled the state House. While the resulting plan was a "nightmare for those who believe redistricting plans should have compact and competitive districts," it nevertheless was a bipartisan plan that was aimed at protecting incumbents of both parties. AAP 1994, 528-29. It, however, shows up as an actionable Republican gerrymander in every year of the plan's existence, including in 2008, by which point Democrats held 12 of 19 seats.

101. Then came 2012. Here is how the Almanac describes the process:

> "In 2011, the tables turned dramatically. Democrats had hung onto the Illinois legislature and governor's office in 2010, awarding them their only free hand in the country to give a large state's existing map a total makeover. Under heavy pressure from party leaders desperate to offset Republican gains in other states, Democrats in May 2011 released a map designed to eliminate up to six Republican seats. . . . In the Chicago suburbs, Democrats recrafted tea party crusader Joe Walsh's marginal 8th District into a Democratic stronghold anchored by Schaumburg. They also dismantled moderate Republican Judy Biggert's seat to forge a new

strongly Democratic 11th District anchored by Auroraa and Joliet . . . Along Chicago's Nroth Shore, Democrats removed freshman Republican Robert Dold's Kenilworth home from the 10th District and pushed him north into more Democratic areas of Lake County.

AAP 2014 at 541.

The Almanac continues:

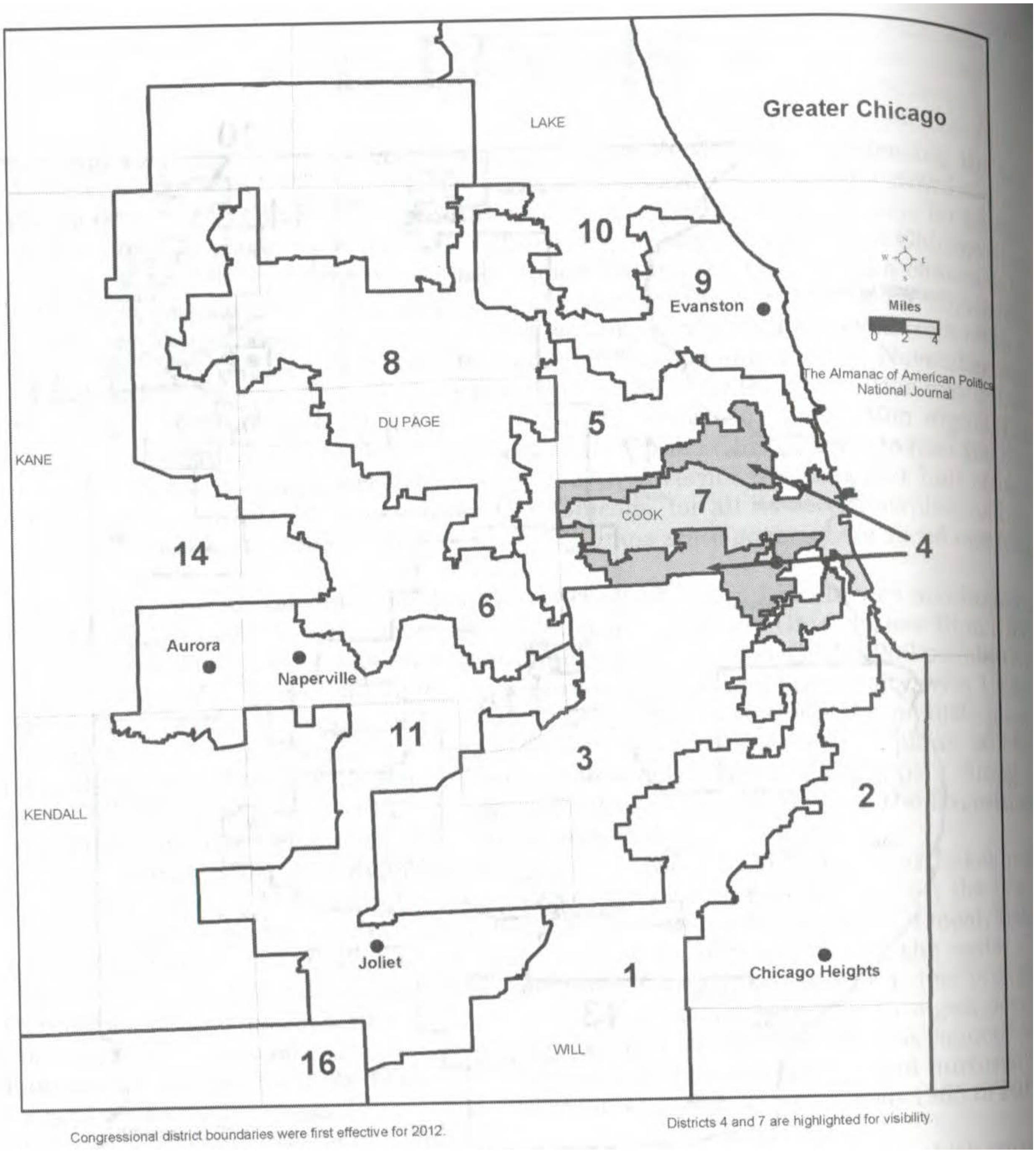
The bloodbath wasn't limited to Chicago. Along Illinois's northwestern border, Democrats stuffed Republican freshman Bobby Schilling's Quad Cities home base into a reconfigured 17th District with tentacles stretching into minority neighborhoods in Rockford and Peoria. And downstate, Democrats endangered Republican Tim Johnson by stretching his Champaign-based seat southwest to link up with other college enclaves such as Bloomington -Normal and Edwardsville. . . . In hindsight, Democrats' strategy largely paid off and generated a rare triumph in an otherwise wrenching redistricting year."

Id.

102. Not only that, but as discussed above, Jowei Chen, plaintiffs' own expert, identifies Illinois as a Democratic gerrymander in 2012. What does the efficiency gap show? A slight Democratic gerrymander of .023. In fact, over the course of the plan's implementation, its efficiency gap has actually been negative, indicating a slight Republican lean.

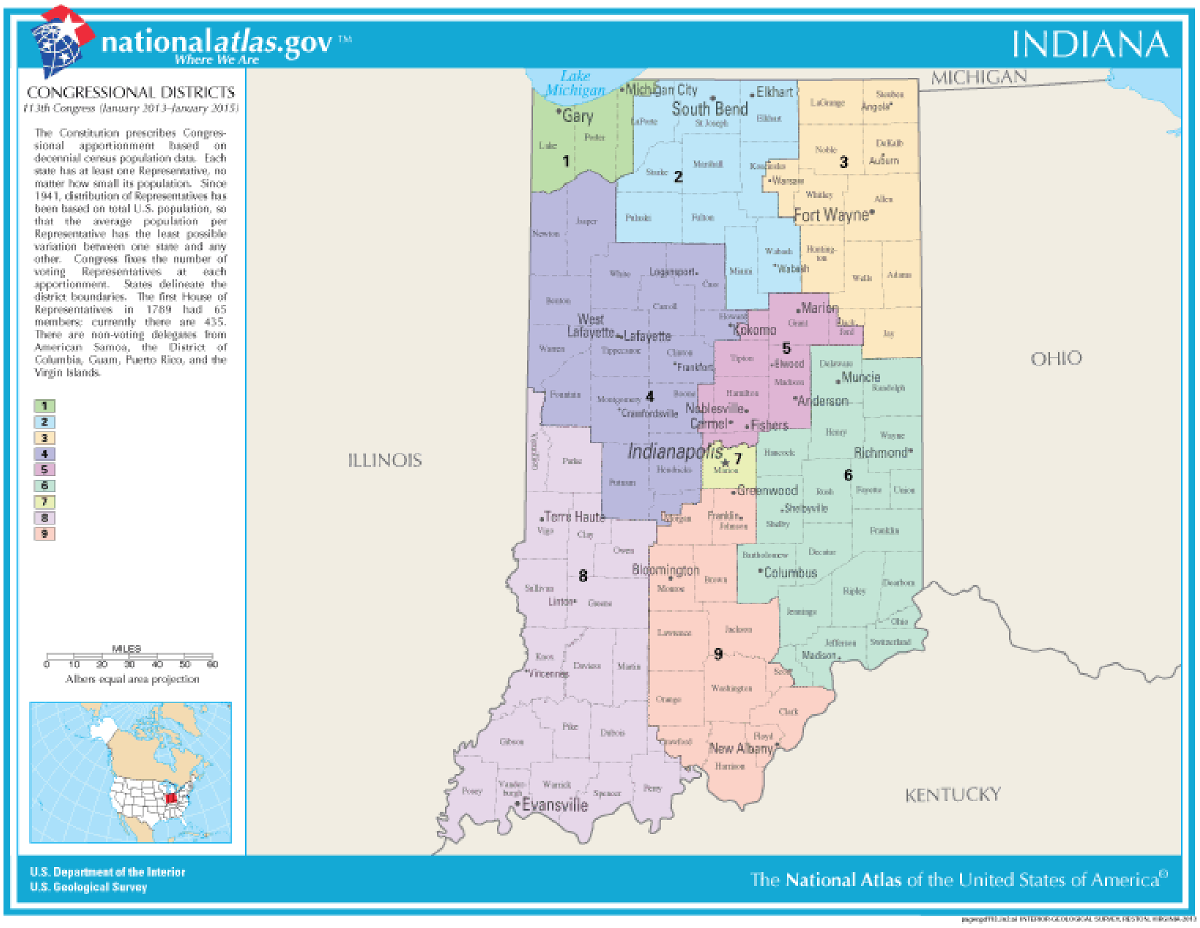
103. In short, the 2002 Illinois map, which was not a gerrymander, shows the supposed hallmark of a gerrymander, while the 2012 Illinois map, which was a gerrymander, lacks its supposed hallmark.

Illinois Redistricting, Chicago, 2012

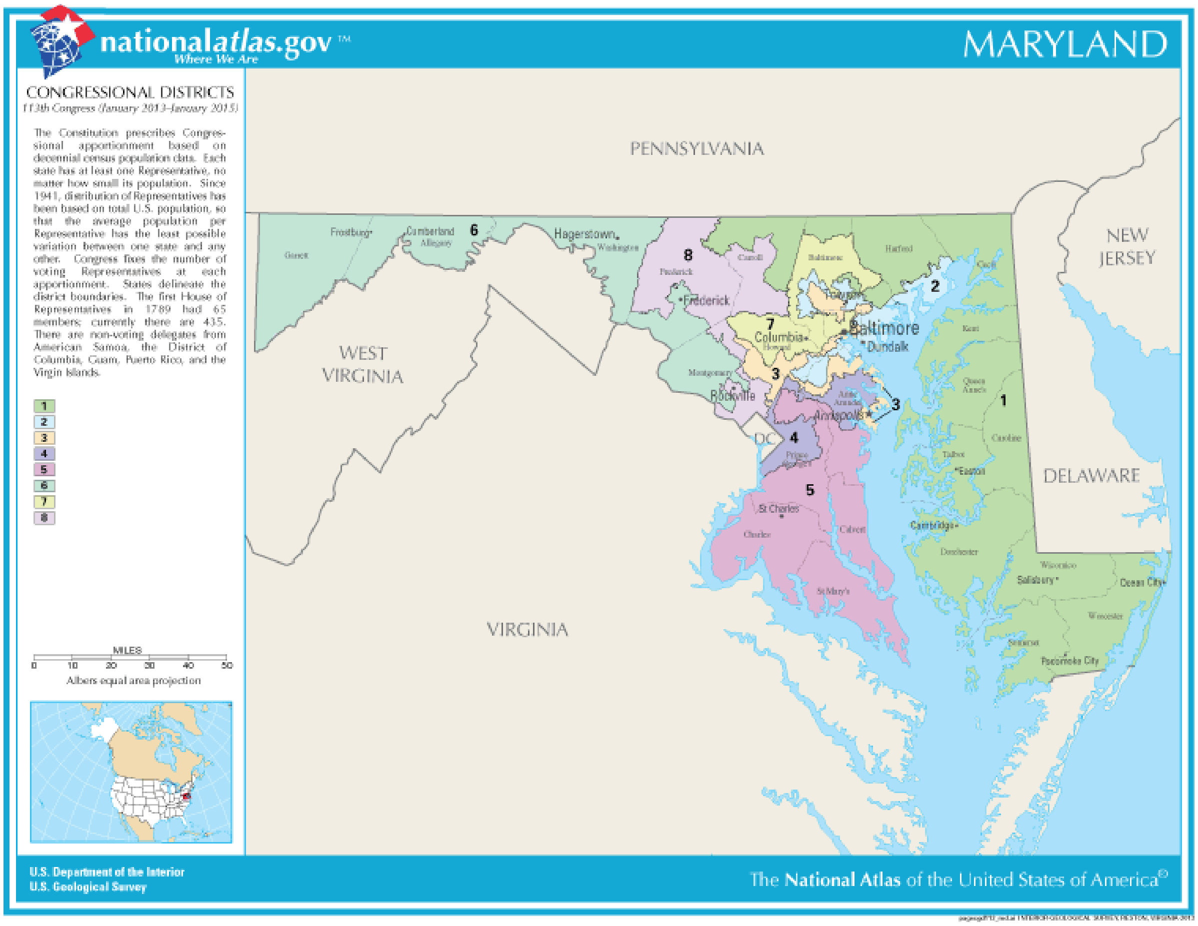


104. The Kentucky map in 1982 was drawn with overwhelming Democratic majorities in both chambers, as well as Democratic control of the governorship. Yet it bears the hallmark of a *Republican* gerrymander, with an efficiency gap of -.129.

105. The 2012 Indiana map also presents as Republican gerrymander, with an efficiency gap of -.201. But it is almost a paragon of regular lines, and, as mentioned above, Dr. Chen suggests that geography might be the culprit:



106. Indiana shows up as a gerrymander despite its regular lines. Yet *Maryland* shows up as a map with only a modest Democratic lean, despite the fact that it was drawn to defeat one of the two remaining Republican congressmen (after Democrats drew an actionable map in 2002 that resulted in the loss of two Republicans), and has lines that look like this, which exhibit almost no respect for traditional redistricting principles:



107. Stranger still is the 1992 efficiency gap. Democrats had control of the governorship, a 38-9 edge in the state senate, and a 116-25 edge in the state house. Yet the map they produced has a Republican lean in every year the plan was in effect, and was just 16 hundredeths of a percentage point away from being flagged as a Republican gerrymander in 1992. Indeed, had the 1996, 1998 or 2000 elections occurred in 1992, it would have been flagged as a Republican gerrymander.

108. Again, this is not an instance of strange Democratic beneficence. Connie Morella was an unusually liberal Republican member of the House, who was able to win her district while Bill Clinton was carrying it by 18 points. *AAP 1994* at 591. In addition, longtime Democratic representative Beverly Byron lost a surprising primary against a liberal Democratic

representative; this enabled Republican Roscoe Bartlett to narrowly claim a district that had been Democratic since 1970. Finally, Republican Wayne Gilchrest won a narrow, four-point victory in a member-versus-member race against Democratic representative Thomas McMillen. Had Gilchrist and Bartlett lost – their combined margin was about 10,000 votes – the map would present as an actionable Democratic gerrymander. Again, about 10,000 votes is the difference between a map that is almost flagged as a Republican gerrymander.

109. New Jersey's 2012 map would be flagged as a Republican gerrymander, with an efficiency gap of -.174, almost as large as that found in North Carolina. Yet New Jersey's map is drawn by a bipartisan redistricting panel; its tiebreaking vote was cast in favor of a 6-6 map in 2012. Yet it presents as a heavy Republican gerrymander, with an efficiency gap of -.17373.

110. The 1982 map in Ohio is a Republican gerrymander according to the efficiency gap, with a score of -.104. This contrasts with contemporary descriptions of the process. The *Almanac of American Politics* writes: "Congressional redistricting was truly a bipartisan exercise in Ohio in 1982, not because its politicians are altruistic, but because the Democrats controlled the state House of Representatives and the Republicans the state Senate and governoship." *AAP 1984* at 908. The map borders on becoming an actionable Democratic gerrymander in 1984, with a Democratic efficiency gap of .0692. Dr. Jackman considers this a separate map, but the changes made to the 1982 lines were marginal.

111. Then there is Texas, home to some of the nastiest redistricting wars in recent times. The efficiency gap metric, to its credit, properly identifies the 1992 map in Texas as being a political gerrymander, with an efficiency gap of .158.

112. The 2002 redistricting is less successful. The first map escapes scrutiny as a Democratic gerrymander, albeit by a slim margin. The 2004 redistricting, upheld in *LULAC*,

would receive scrutiny, despite Justice Kennedy's observation that the map merely made "the party balance more congruent to statewide party power."

113. Indeed, this is an interesting case. The efficiency gap met plaintiffs' proposed threshold and there was ample evidence of Texas' intent to gerrymander aggressively. Yet Justice Kennedy rejected the challenge.

114. Washington State holds the distinction of holding the largest Democratic efficiency gap in the history of the series -- .282 in 1978. Indeed, that map also produced the 8th, 17th and 19th most heavily Democratic maps in the series.

115. Yet little in the initial plan would have suggested that such a large waste of Republican votes was forthcoming, as the initial map was almost perfectly balanced: the gap was .013. What happened? The initial map hid the tendency of the underlying maps to produce massive Democratic efficiency gaps. It is a classic illustration of how small, unpredictable effects in the electorate can cause huge changes in the efficiency gap.

116. In 1972, Democrats won six of the state's seven congressional seats. They came up just short in a seventh, as Republican Joel Pritchard narrowly held the seat held by retiring Republican congressman Thomas Minor Pelly. The 2,602-vote margin resulted in almost all of the Democrats' wasted votes: 104,959 of the Democrats' 315,739 wasted votes came from this district. Republicans also fielded nominal candidates in the 5th, 6th and 7th districts, resulting in a large number of wasted votes. Interestingly, had the Democrat been successful in that race, the map would have shown an efficiency gap of .17, in line with other maps.

117. But Democrats never came close to capturing the first again. Offsetting this, however, Republicans held the Democrats to around 60 percent of the vote in the second through

fifth districts, and to seventy percent in the 6th and 7th. The result was a substantial reduction in wasted Democratic votes, and inflation in Republican votes. This repeated in 1976.

118. Then, in 1978, the perfect storm hit. Republicans fielded competitive candidates in all of the districts, falling just short in most of them. Democrats won 51 percent in the 2nd district, 59 percent in the 3rd district, 61 percent in the 4th district, 53 percent in the 5th district, 62 percent in the 6th district, and just 53 percent in the heavily Democratic 7th District, where a Republican won a fluke victory in a special election and acquitted himself well in the general election. In the process, Republicans racked up huge numbers of wasted votes, while their strong showing in the 1st district hurt them further.

119. But even this pales in comparison to what happened in the 1990s. That map opened with the fourth-largest Democratic efficiency gap in history; .2355, eclipsing all of North Carolina's maps. It then produced the largest Republican efficiency gap in the time series, of -.25887. The next year it repeated the feat, with an efficiency gap of -.2297. Then, suddenly, in 1998 it produced an almost negligible efficiency gap of .008, before then producing a healthy Democratic gap of .077.

120. How could this be? In the first year of implementation, Democrats won eight of the states nine seats, but did so very narrowly. Only two of those wins were with 57 percent of the vote or more. So Republicans racked up a lot of wasted votes, while Democrats' wasted votes were fairly contained, coming almost entirely from the eighth district. This, under plaintiffs' theory, is exactly what a gerrymander looks like.

121. Then the wave hit. In 1994 Republicans won seven of these districts, but did so very narrowly; they wasted only 30,068 votes total across the six seats they picked up. The Democratic incumbents who lost those districts, however, accounted for 528,755 wasted votes.

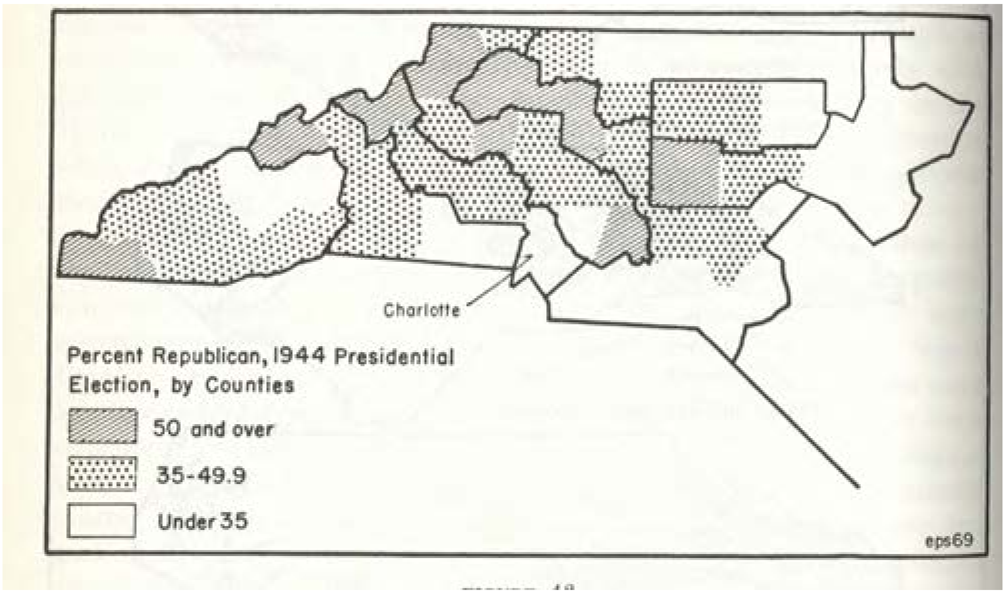
In 1996, we had a similar scenario. Democrats managed to defeat just one Republican, Randy Tate, narrowly, but came up short in the remaining districts. Republicans wasted just 30,586 votes across five districts this time. Democrats wasted 568,813.

122. By 1998, however, things reverted to normal. Rick White was defeated in the 1st Congressional District. Linda Smith ran for Senate; her district flipped. At the same time, Adam Smith became entrenched in the 9th District. He won only 51.5 percent of the vote when defeating Tate in 1996, but his margin ballooned to 65 percent in 1998. Overall, districts 1-5 produced just 30,500 wasted Republican votes and 569,000 wasted Democratic votes in 1996. They produced 234,000 wasted Republican votes in 1998, and 269,000 wasted Democratic votes in 1998. In 2000, Democrats picked up the 2nd district, which had been vacated by Jack Metcalf. This resulted in more wasted Republican votes, and a Democratic leaning efficiency gap.

123. These maps bear heavy indicia of gerrymandering – indeed, some of the heaviest on record. Yet they are probably not gerrymanders. The 1972 maps were drawn by a Democratic legislature and Republican governor. The 1992 maps were produced by an independent commission.

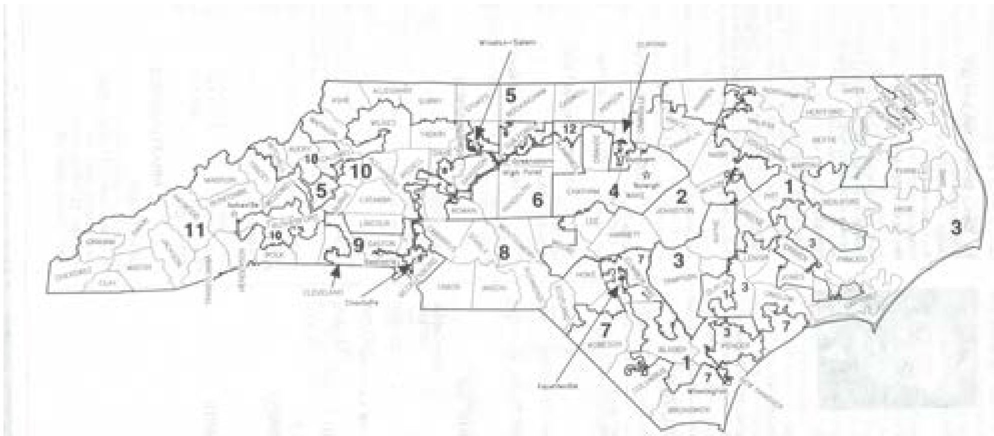
124. Finally, we come to North Carolina. North Carolina has a long history of partisan gerrymandering; in his classic work on the political dynamics of southern states, V.O. Key singled out North Carolina Democrats for drawing what we would today label wryly as "baconmanders": Maps with stretched districts that pair Republican areas with Democratic areas. So, for example, one district in the 1940s stretched from Mecklenburg County to Yancey and Mitchell counties, while another wound from Stanley County up to Watauga, before hooking around to Alleghany County.

North Carolina, 1940s (Key, *Southern Politics* 226 (1949))



125. But of course, none of this compared to what came in 1992. The Democrats drew a map whose only competitor was Texas for bizarre district lines. It actually had a lengthy history: The first map Democrats drew created a majority black district in the east, but was struck down by the Bush Justice Department, which demanded a second minority-majority district. Republicans believed that this would help them weaken Charlie Rose and Bill Hefner in the 7th and 8th districts. But instead, Democrats drew the infamous 12th Congressional district, which snaked up I-85, connecting black populations in Charlotte, Winston-Salem, Greensboro, and Durham. *AAP 1994* at 942. These districts forced irregular lines, but Democrats went further, attaching a tendril into South Durham in an attempt to shore up Tim Valentine, utilized touch-point contiguity to prevent the Third District from extending down into Wilmington, and again to avoid wasting Democratic precincts in Howard Coble's district, and extended the Fifth down to Burke County. They utilized touchpoint contiguity a third time in the western portion of the state, extending Cass Ballenger's district to pick up Mitchell and Avery Counties, which were the two most Republican counties in the state in 1992.

North Carolina District, 1992 (*PIA 1996*)



126. But in 1994, this map actually produces a Republican lean, one that is a few votes shy of the threshold for being flagged. In fact, Dr. Jackman imputes values for the 6th District; if actual turnout proved to be slightly lower than Dr. Jackman's projections, or if the challenger's vote share were slightly higher than Dr. Jackman's projections, this map would be flagged.

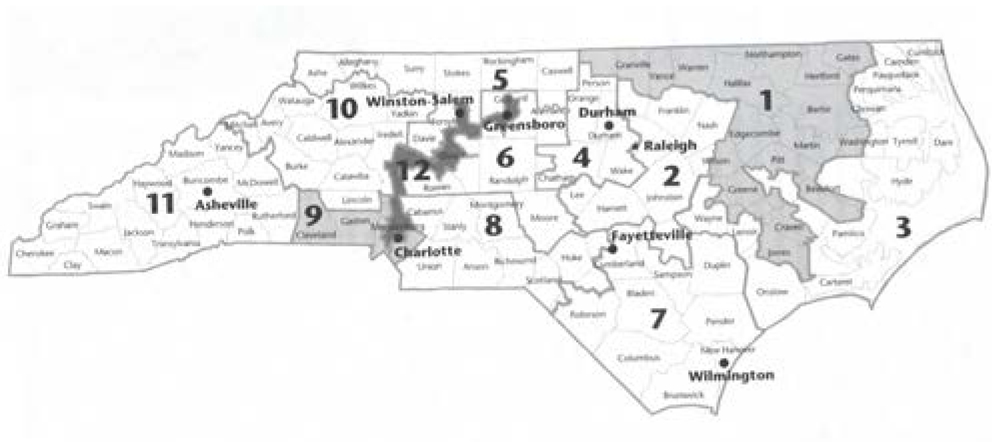
127. Whatever the 1992 map is, it is not a Republican gerrymander. Yet if there had been a Republican wave in 1992 rather than 1994, that is exactly what the efficiency gap would suggest. Once again, a map that is plainly not a gerrymander bears the supposed hallmark of a gerrymander.

128. What happened in 1994? Republicans performed better. In 1992, the 2nd, 3rd, 4th, and 5th districts gave Democrats 55, 56, 66 and 64 percent of the vote. This resulted in around 69,000 wasted Democratic votes, and 366,000 wasted Republican votes. But in 1994, Republicans won those seats, though not by overwhelming margins. Suddenly *Democrats* wasted 267,000 votes in those seats, while Republicans wasted just 24,000 votes. Again, a shift in the environment produced a large efficiency gap.

129. The 1998 and 2000 maps contain more regularized lines, but both produce slight Republican efficiency gaps. This is odd, because the state sought to defend these maps in court

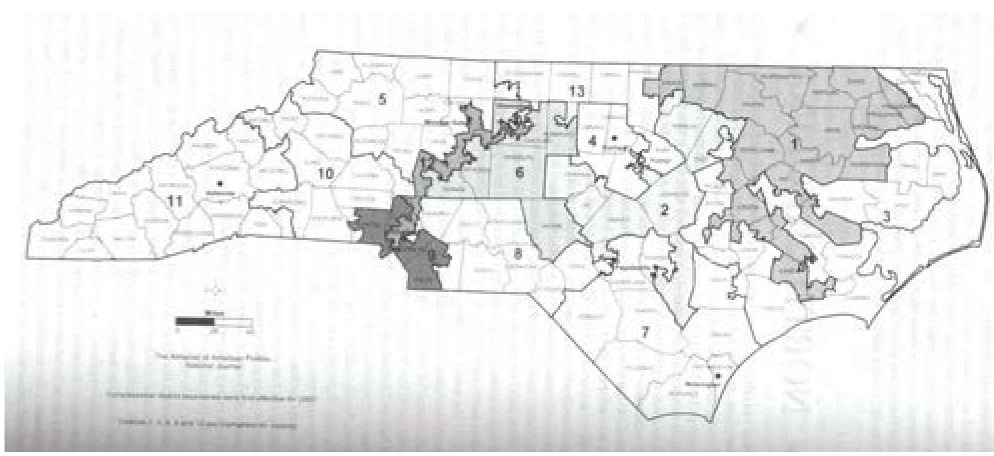
in part on the grounds that they were a Democratic gerrymander; Justice O’Connor ultimately agreed.

North Carolina 2000 (*PIA 2002*)



130. In 2002, the state returned to more irregular lines, in an attempt to create yet another Democratic district while strengthening the remaining Democratic incumbents. According to the Almanac, the map created a new Democratic district for Brad Miller (the Senate redistricting committee chairman), weakened Robin Hayes in the 8th districts, and shored up the 2nd and 7th districts. AAP 2004 at 1187.

North Carolina 2002



131. But the nature of the gerrymander wasn't apparent until 2010. According to the efficiency gap, the map actually had a slight Republican bias in 2002 and 2004. Even though Robin Hayes' Eighth District flipped in 2006, it does not change the efficiency gap much. Due to the Democratic wave and heightened African-American turnout (the largest increases in turnout statewide vis-à-vis 2004 were in the 1st and 12th districts), Democrats ended up wasting a lot of votes in their safe districts. Heath Shuler's win in 2006 does not change the efficiency gap for the same reason.

132. As the environment moved leftward, these maps performed as they "should," adding Democrats to the map as their statewide vote share grew. The reason these maps were Democratic maps, though, is that this function was a one-way ratchet. When the Republican wave election hit in 2010, they found themselves able to defeat just one Democratic incumbent. The Democrats' vote share in the 1st, 4th, 7th, 8th, 11th and 13th districts were all below 60 percent. The result was that Republicans wasted 560,000 votes in these districts, while Democrats wasted around 73,000 votes. Also, because the Democratic vote shares fell in Republican incumbent districts, there were few offsetting Democratic votes there.

133. The result would have been even more extreme had a longtime Democratic incumbent not made a key mistake. The Almanac describes the events: "[T]wo young Republican operatives approached the incumbent outside the House office buildings and asked him whether he supported the 'Obama agenda.' Etheridge asked them repeatedly in angry tones, who they were, and grabbed one by the wrist and the other, briefly, by the neck." AAP 2012 at 1246. Had Etheridge not done this, he probably would have won.

134. Had Etheridge taken a different route back to his office that day and won by the same margin by which Renee Ellmers actually won, the efficiency gap would have been .199,

larger than the efficiency gap produced by the present North Carolina map. So, while the 2002 Democratic gerrymander did not appear to be an extreme gerrymander in its first year, the potential was certainly there. All it took was the defeat of two Republican congressmen (one of whom was intentionally endangered by the map), and then a Republican wave election. This is the scenario that creates the cluster of Republican votes just shy of the 50 percent mark that plaintiffs describe as their hallmark of the gerrymander. It is just that in this map, this scenario doesn't unfold until the last year of enactment. Again, small events can have significant effects on the efficiency gap.

135. The 2012 maps also illustrate the impact small changes can have. Mike McIntyre was able to survive the 2012 election because he was a good congressman who tended to his district. Had Heath Shuler opted to run for re-election, or had Etheridge not imploded in 2010 and opted to run in 2012, there is really no reason to suspect they could not have performed as well as McIntyre (especially Shuler). Had this happened, the efficiency gap that year would have been -.09: not actionable. If those three had won narrow victories under the 2016 map, the efficiency gap would have been zero. Of course, as they retired, their districts would have flipped, and the efficiency gap likely would have swung rightward (although that was true of the 2002 map as well). But this again illustrates how the efficiency gap is not some intrinsic feature of plans, but is often dependent upon slight changes in elections.

136. Plaintiffs would likely respond that partisan intent acts as a further screen on gerrymandering. First, misses the point. The point is that the efficiency gap is not what it is being sold as: clear evidence of gerrymandering. Second, I am skeptical that this operates as a meaningful bar; plaintiffs have been able to take cases to trial alleging partisan motives underlying maps drawn by independent commissions in the past, see *Harris v. Arizona*

Independent Redistricting Comm'n, 578 U.S. __ (2016), so I am not sure why a state could not be haled into court in other circumstances. Of course, the issue is not just that there are things that are plainly not gerrymanders that present as gerrymanders under the efficiency gap; it is that things that are plainly gerrymanders often appear not to be gerrymanders.

137. Finally, Plaintiffs may emphasize North Carolina's performance. But we are trying to ascertain a national standard, one that is manageable and easy to apply. The more that plaintiffs emphasize the particulars of North Carolina's current map, the more we see what a fact-intensive, arbitrary standard the plaintiffs' test is in practice.

138. The plaintiffs claims that the efficiency gap is the hallmark of a gerrymander. In reality, the meaning is much more ambiguous. I understand the intuition behind it. In fact, if elections occurred in the lab, they would probably be correct; under controlled conditions of similar incumbency, candidate quality, electoral environment, scandal, or Voting Rights Act requirements, and assuming there is no spatial clustering or other factors that might produce a baseline other than zero, they may even be correct.

139. But reality is messier. Incumbents survive races they were supposed to lose, or lose races that a map drawer might believe they were going to win. Wave elections hit in the first year of plan implementation, and skew the produced efficiency gap. Candidates grab reporters by the neck, and lose elections they were supposed to win. Right-wing Republican candidates win low turnout special elections in historically Democratic areas, and help to skew the efficiency gap in the subsequent elections. As the above analysis shows, these sorts of things are not one-off events; they occur with some regularity.

140. What the above really shows is that map drawers don't have nearly as much control over where efficiency gaps emerge as plaintiffs imagine. While I might agree with the

suggestion that what map-drawers are *trying* to do when gerrymandering is to waste the other sides' votes, to accomplish the feat they have to control facts on the ground to an unrealistic degree. In reality, actors with the intent to waste votes will end up failing, or even wasting their own votes with some regularity, while actors with a command to avoid wasting votes will often fail in their mission.

VII. It is unclear exactly why "Average Efficiency Gap" is of interest. It does not mean a voter's influence is consistently degraded.

141. In the Wisconsin litigation, Dr. Jackman inquired whether "over the life of a redistricting plan," the efficiency gap "remain[ed] on one side of zero or the other." Jackman, WI Report, at 53. Indeed, this was a "key" inquiry, *id.* because it helped determine whether large values of the efficiency gap were likely to be repeated over the life of the plan. This is because "plan's gaps vary substantially over the plans' lifetimes," Stephanopoulos & McGee at 836, indeed, this variation caused McGee to conclude initially that the efficiency gap demonstrated why Courts did not need to get involved with settling gerrymandering claims at all. McGee at 56.

142. From a theory standpoint, this made at least some sense. If a plan resulted in a sign *never* switching, it would mean that one side or the other would always have some votes wasted. While I'm not clear what the significance in practice would be if those numbers ended up small – having one wasted vote in four elections may not add up to a constitutional inequity – I understood the argument.

143. Here, Dr. Jackman switches his inquiry to the sign of the "average" efficiency gap. I am unsure what the significance of this is. The fact that a map might, on average, disadvantage one party or the other does not seem like clear evidence that a wrong has been committed. First, practically speaking, a plan *has* to favor one side or the other, at least on average. Second, as we've seen, maps can show strong efficiency gaps due to freakish results in

single elections. Averages are sensitive to outliers, especially with small numbers of observations. It's entirely possible for a map to favor the opposing party a majority of the time, yet show an average efficiency gap that favors the map-drawing party.

144. There are a large number of maps where the efficiency gap does not switch sign. This would seem to augur well for the efficiency gap. But when one digs down, the pattern does not match well with plaintiffs' system.

145. The following table displays the maps that, by my count, display consistent efficiency gaps across plans (Note: I disagree with Dr. Jackman's coding of Ohio in 1984, New York in 1974, and Washington in 1984, because the changes to the maps were trivial). It also details whether the plan is actionable – that is, whether you would be able to bring a cause of action – as well as whether the cause of action would be sensible – that is, whether the sign of the efficiency gap matches the party that drew the map.

State	Year	n	Actionable?	Sensible?
CA	1992	5	No	No
CA	2002	5	No	No
CO	2012	3	No	No
FL	1996	3	No	No
GA	1996	3	Yes	No
GA	2012	3	No	Yes
IL	2002	5	Yes	No
MD	1992	5	No	No
MD	2012	3	No	Yes
MA	1992	5	No	Yes
MA	2002	5	No	Yes
MN	1972	5	No	Yes
MN	2002	5	No	No
MO	2012	3	No	Yes
NJ	1972	5	No	No
NJ	2002	5	No	No
NJ	2012	3	Yes	No
NY	1972	5	No	Yes
NY	1982	5	No	No
NY	1992	3	No	No
NC	1972	5	No	Yes
OH	2002	5	No	Yes
TX	1974	4	No	Yes
TX	1996	3	Yes	No
WA	1972	5	Yes	No
WA	2012	3	No	No
AL	2012	3	Yes	Yes
FL	2002	5	Yes	Yes
FL	2012	3	Yes	Yes
IL	1972	5	Yes	Yes
IN	2012	3	Yes	Yes
MA	1982	5	Yes	Yes
MA	2012	3	Yes	Yes
MI	2002	5	Yes	Yes
MI	2012	3	Yes	Yes
OH	2012	3	Yes	Yes
PA	2012	3	Yes	Yes
SC	2012	3	Yes	Yes
TX	1984	4	Yes	Yes
TX	2006	3	Yes	Yes
TX	2012	3	Yes	Yes
WI	2012	3	Yes	Yes

146. In fact, it performs quite poorly. Of the 42 maps with consistent efficiency gaps, only about 38 percent occur in situations where both the party that benefits from the efficiency

gap had control of the process, and the map would be actionable under Dr. Jackman's suggested standard. About a quarter of the maps that produce consistent efficiency gaps occur in situations where the benefitting party did not have control of the process *and* the initial map isn't actionable. Further, there are only four maps in the dataset where the party that controlled redistricting benefitted from a consistent efficiency gap, where the gap was actionable, and where a map went through a complete five cycles.

VIII. **<u>The Efficiency Gap requires mapmakers to guess what the political environment will be in implementation year 1.</u>**

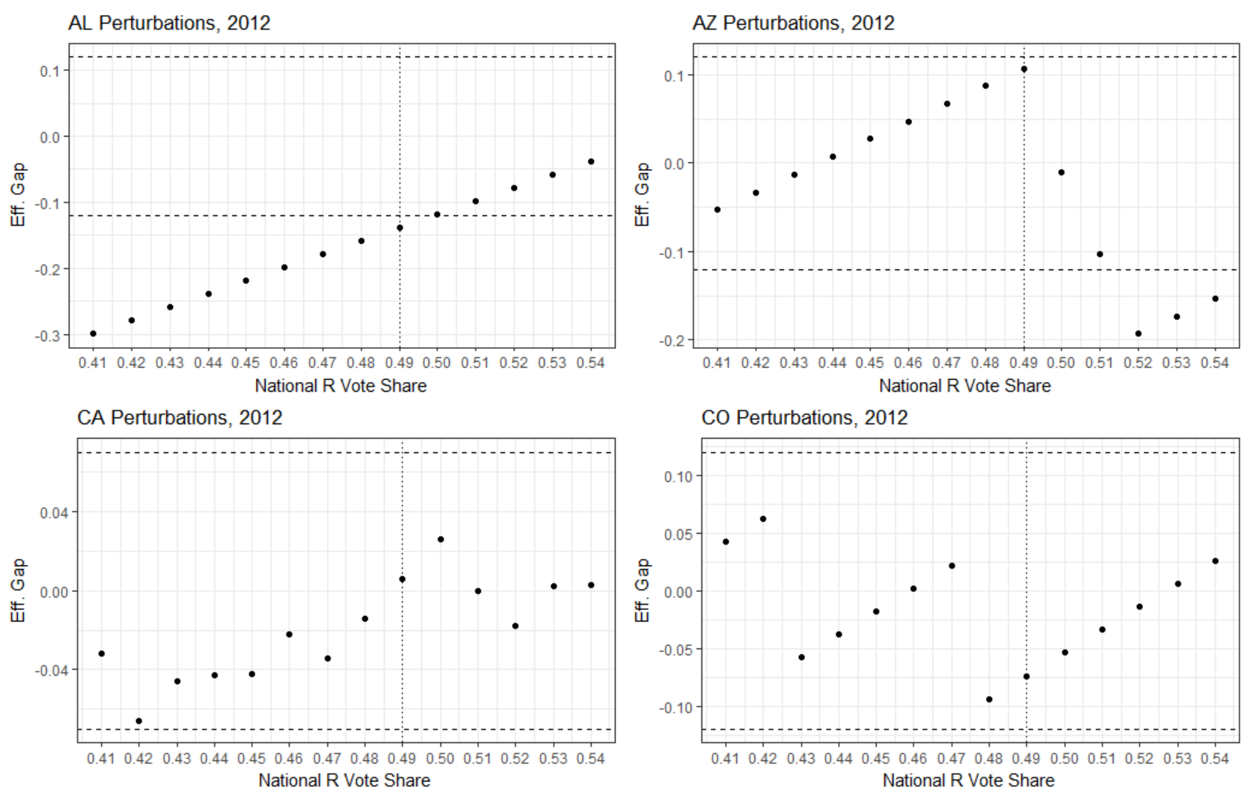
147. As we've seen, from a purely practical perspective, implementing the efficiency gap can be a difficult thing. A mapmaker must try to get a sense of what the national environment will be like in the first year of implementation, which lawmakers are likely to suffer a scandal, which lawmakers are likely to retire, and which lawmakers are likely to suffer a challenge. As we've seen, a large number of maps produce large efficiency gaps even though they are unlikely candidates for extreme gerrymanders, while others produce large effects that are fleeting. As we saw in Washington, the national environment can account for swings of over 50 percent with respect to efficiency gaps in sequential elections. Courts evaluating what mapmakers are doing will "get it right" a fair number of times with the efficiency gap, but they will get it flatly wrong a fair number of times as well.

148. To show how dependent mapmakers are on guessing what the first year environment would look like, we can explore more fully the "perturbations" Dr. Jackman utilizes. To be clear, I'm not at all certain this is permissible, as this is the exact same exercise Dr. Jackman, Stephanopoulos and McGee, and courts have rejected with respect to measuring partisan symmetry. *See also* Jackman, 24 British J. Poli. Sci, at 335, cited in Stephanopoulos &

McGee at 860 n. 139. But, if it is somehow permissible to use to test a statistic, but impermissible to create the statistic itself, we should look at some more perturbations.

149. My approach is as follows: During the time period Dr. Jackman explores, the Republican share of the two-party vote has varied between 41 percent in 1974 and 54 percent in 1994. In 2012, Republicans won 49 percent of the two party vote. So for each state in Dr. Jackman's database, I added one percentage point to the Republican vote total, and recorded what the efficiency gap would be. I then added two percentage points, three percentage points, and so forth, until we simulated a 1994 vote share. I then did the reverse, to simulate a 1974 environment.

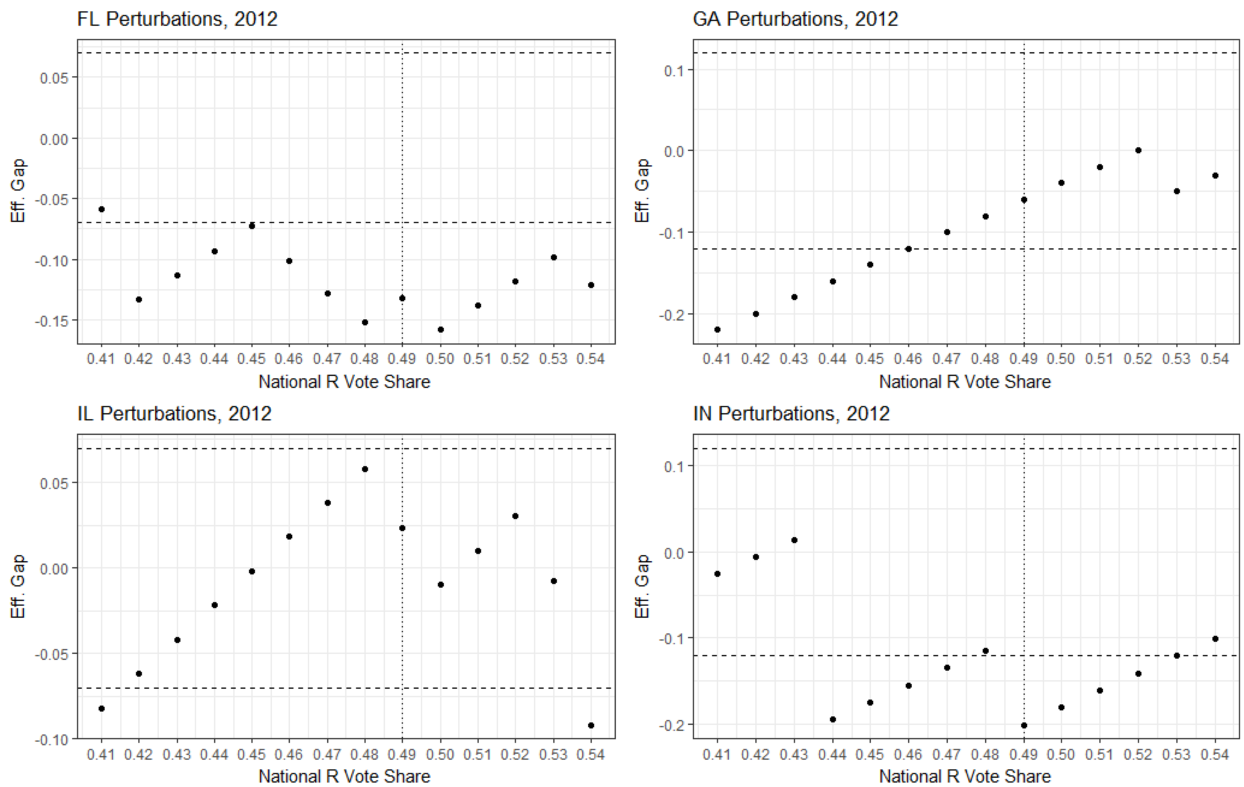
150. I then plotted these results. The two horizontal lines indicate the range where a map is "safe." The vertical line shows 49 percent, which is the baseline for 2012.



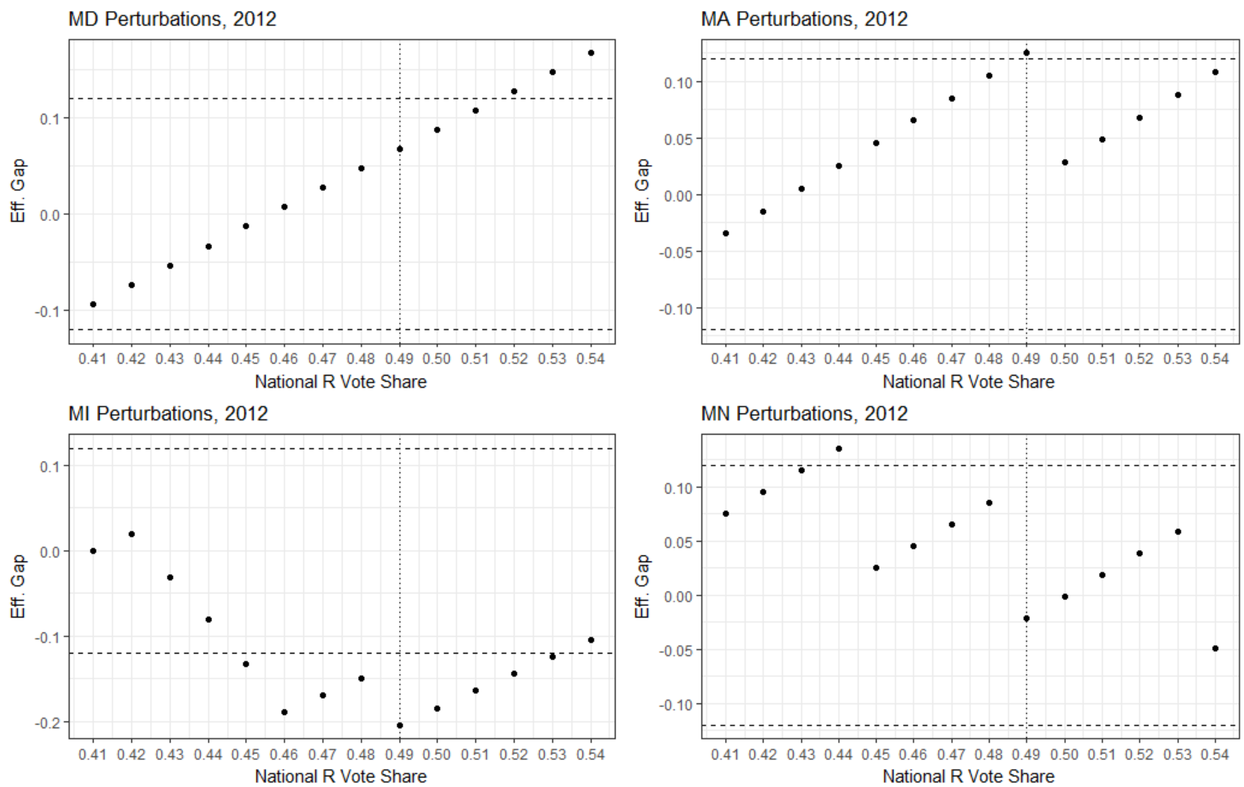
151. First, notice the linear relationship here. This further underscores the fact that the relationship between the efficiency gap and vote share is a proportional one. Notice also the large effect that flipping a seat has. These are the sorts of things mapmakers will have to predict.

152. Alabama's map is actionable, but only because the Republican national vote share fell below 49 percent of the two-party vote. Had Republicans done a bit better, the map would not be actionable.

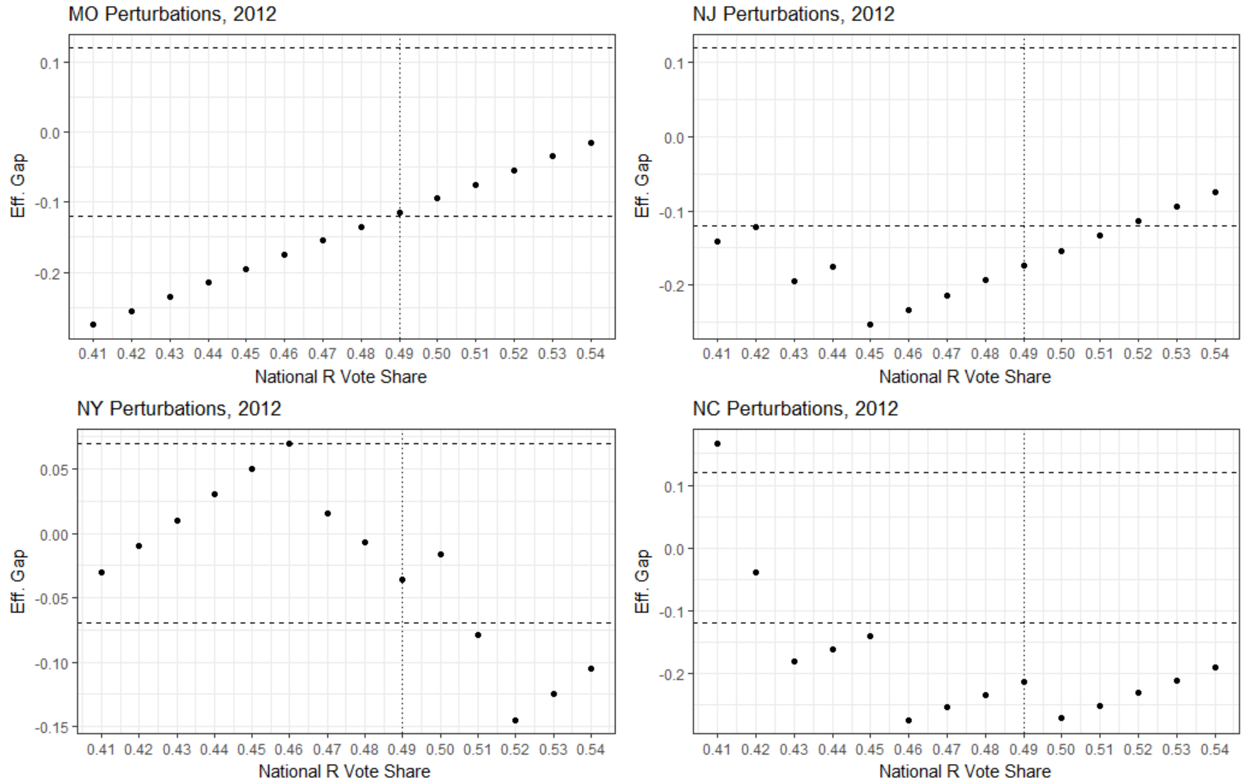
153. In Arizona, we see the efficiency gap's quirks on display. Between 41 percent and 49 percent of the two-party vote, the map is safe, although it approaches an actionable Democratic gerrymander (even though it was drawn by an independent commission). But in a marginally better environment than 2012, Ron Barber in AZ-02 loses, and the map goes from an almost-actionable Democratic map to a Republican efficiency gap. With another percentage point, Ann Kirkpatrick loses to Jonathan Paton, and the map has become an almost-actionable Republican gerrymander. In wave elections, utilizing Dr. Jackman's approach, we'd suddenly have an actionable Republican gerrymander, as Kyrsten Sinema would lose.



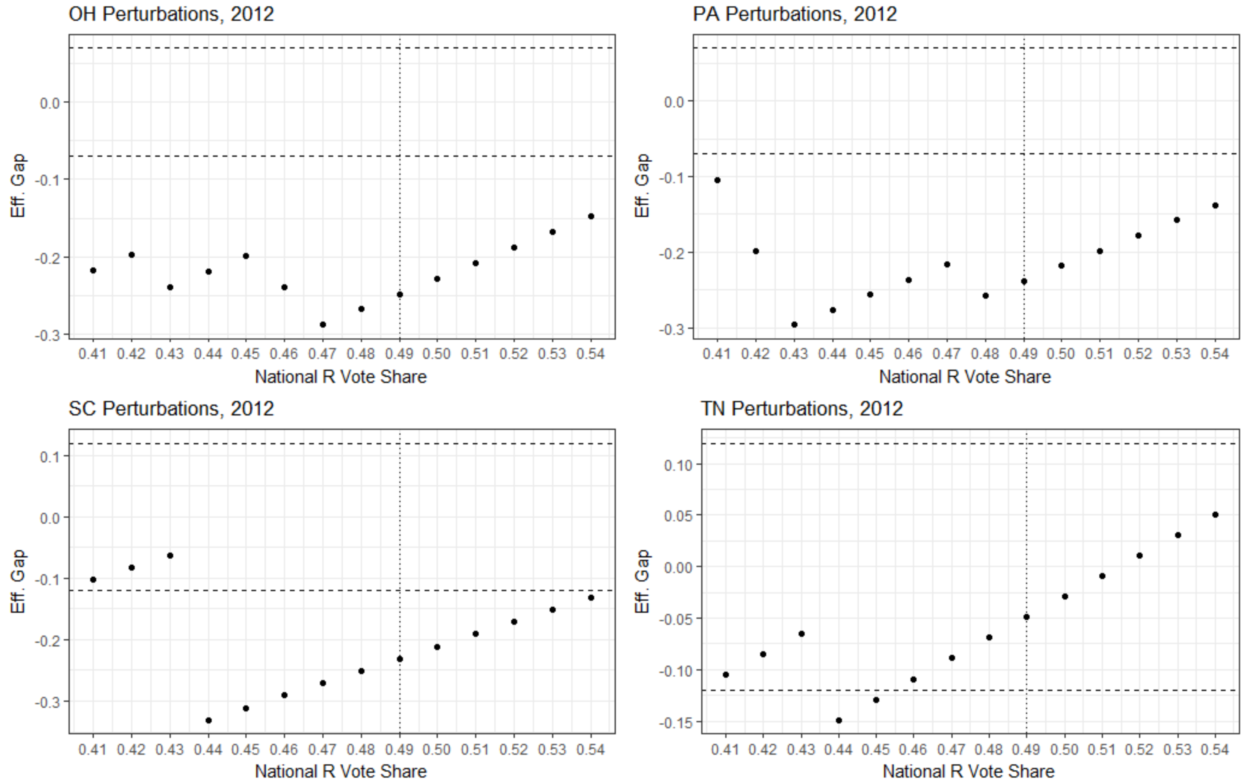
154. Every map in the above four panels produces actionable efficiency gaps under certain circumstances. Illinois is actually identified as a *Republican* gerrymander in sufficiently good and bad GOP years, while Florida gets a pass in a truly bad Democratic year. Georgia becomes a Republican gerrymander in an environment like 2006 or 2008, while Indiana likewise is very fact-dependent. If Republicans had done a fraction worse, Jackie Walorski would have lost, and the map would be fine. Or, if the national environment had matched 2014 or 2010, the seat share would have matched the vote share sufficiently.



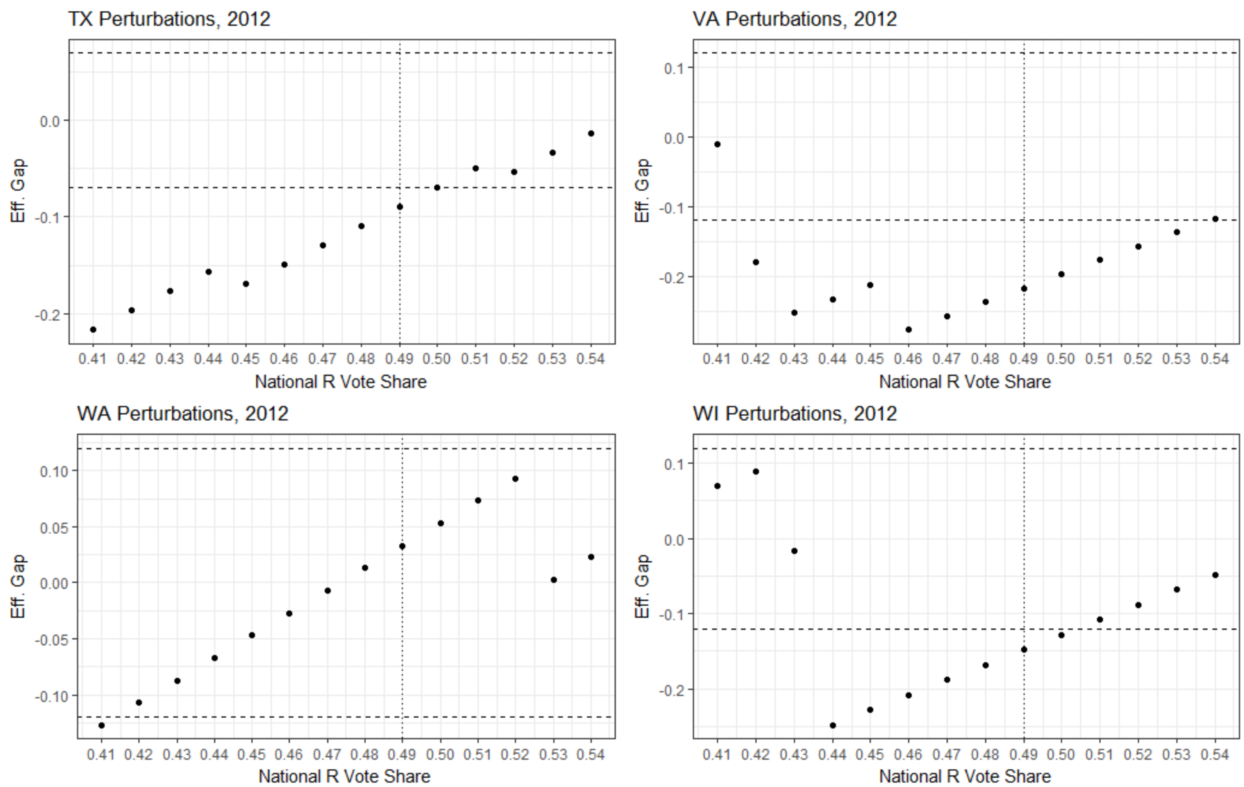
155. Once again, whether maps trigger their efficiency gap thresholds depends on the environment. Lest one believe that the maps with one value outside the safe harbor are truly safe, consider Massachusetts, whose map would be brought into Court because the single value that would trigger scrutiny is what actually happened in 2012.



156. Again, all of the maps above are actionable under some environments, and not actionable under others. The New York map could be a Republican gerrymander in a very good Republican year, even though it was drawn by a political science professor and redistricting expert, Nathan Persily. North Carolina's 2012 map appears to be a Republican gerrymander in most years, but appears to be a Democratic gerrymander in a blowout Democratic year.

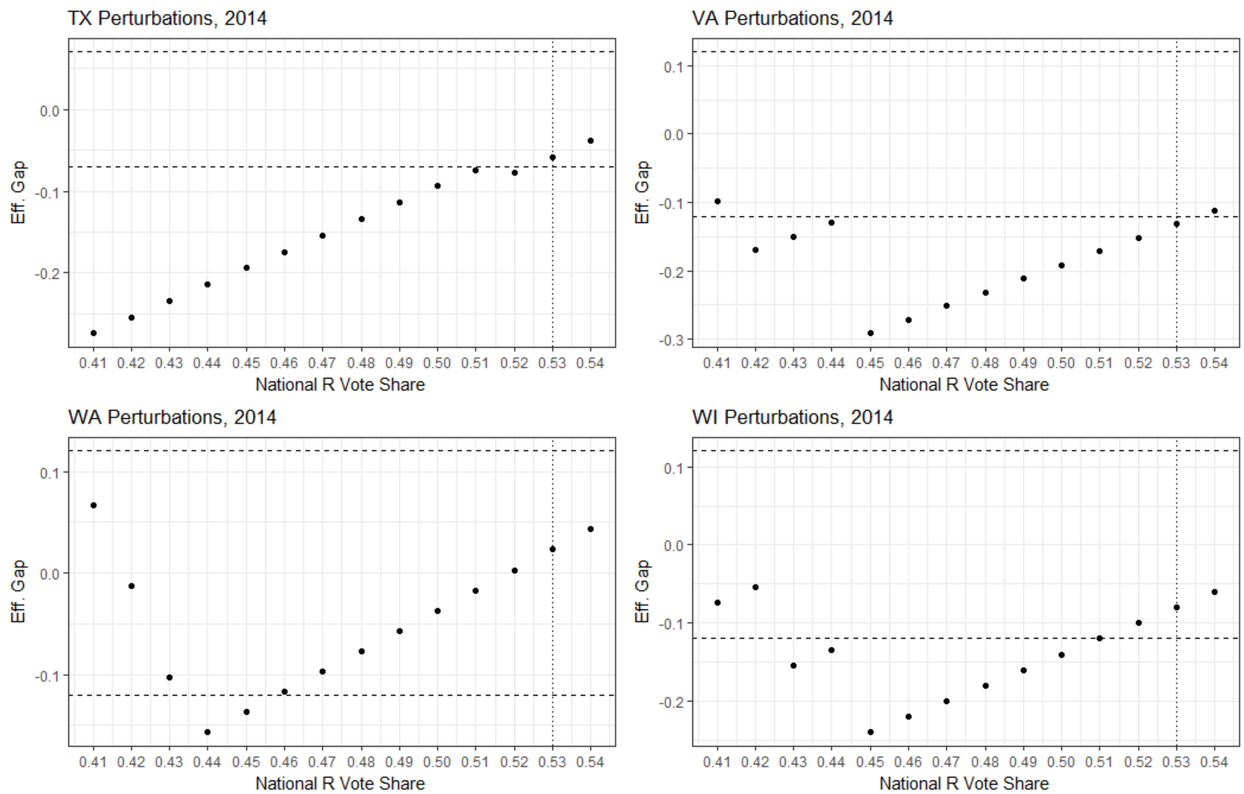


157. The efficiency gaps in Ohio and Pennsylvania are fairly stable, but Tennessee and South Carolina are environment-specific.

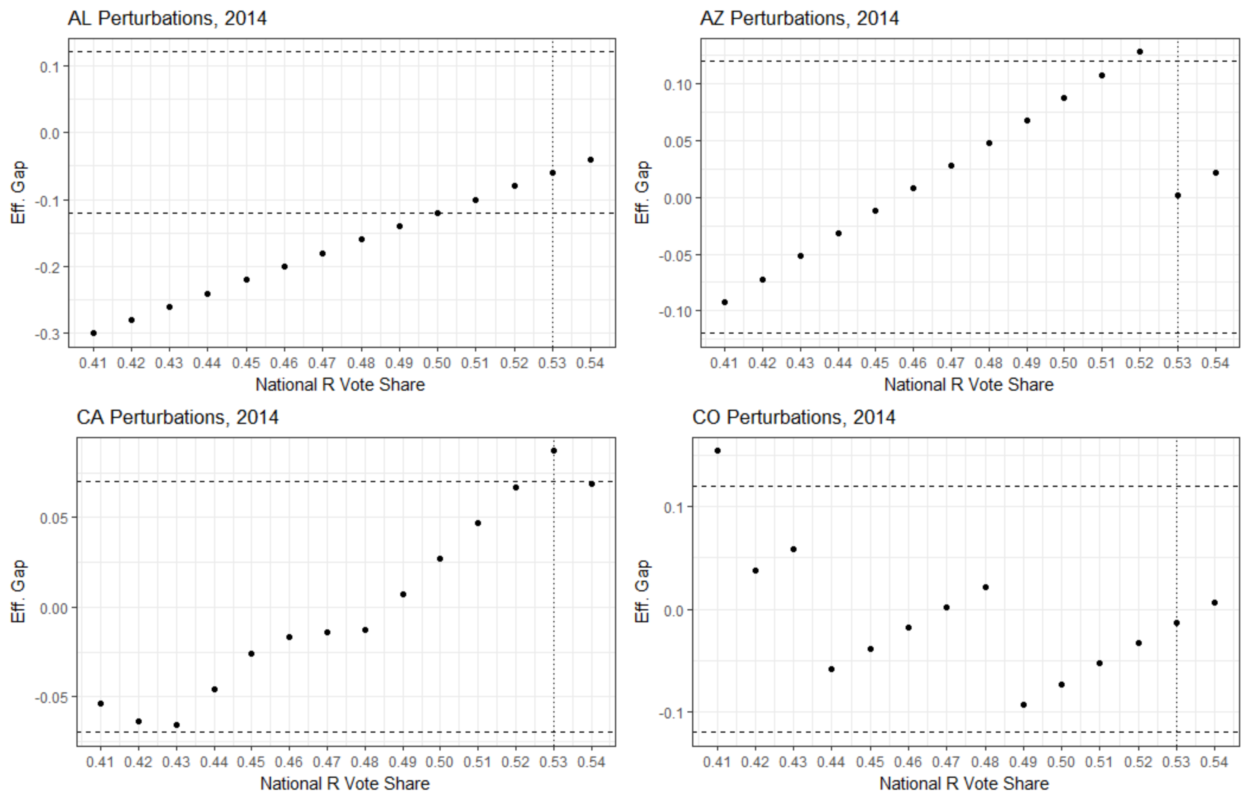


158. All of these maps are actionable under the right set of circumstances, and are safe under the right set of circumstances.

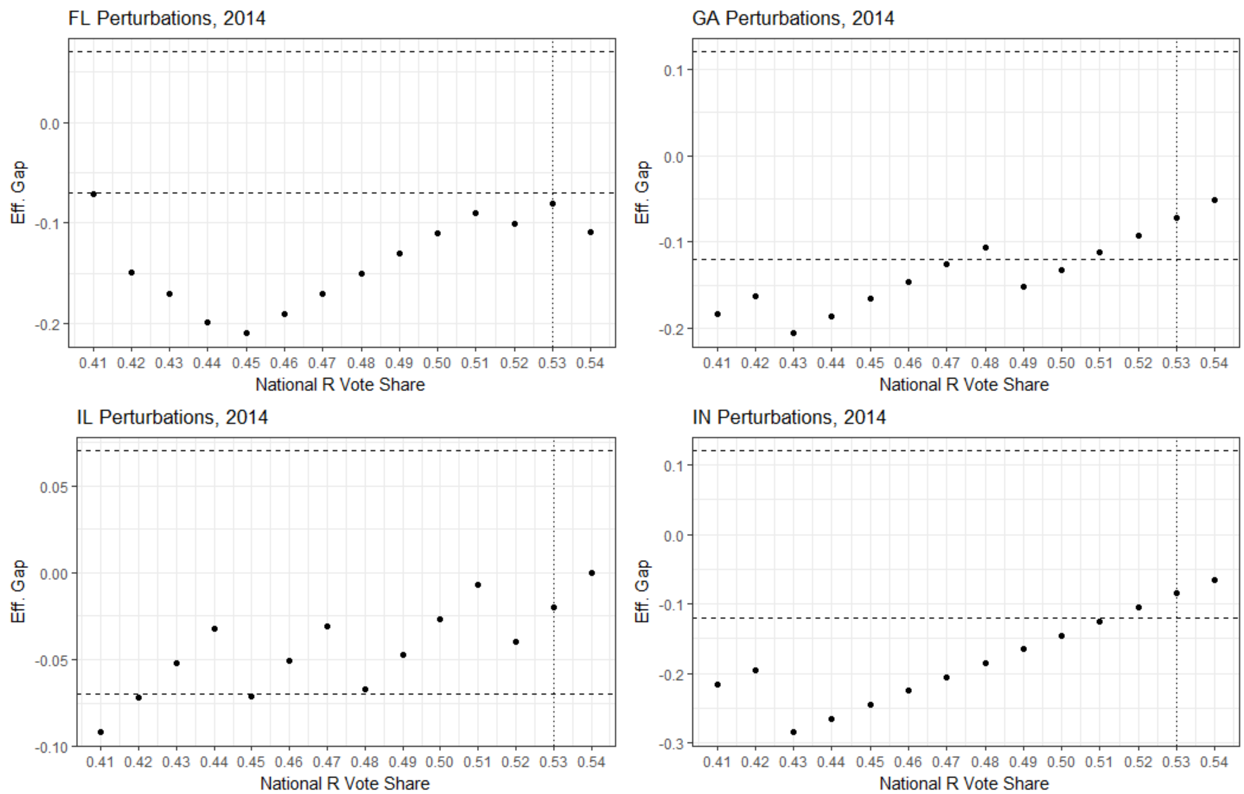
159. But, of course, this is about more than the national environment. Efficiency gaps can be affected by retirements, candidate quality, scandals, and a host of other issues. If we run the perturbations for the 2014 maps, a different set of outcomes emerges. Compare the charts above with the charts below. Note that none of the maps receive scrutiny at the same values of the popular vote share. If Republicans had done a few points better under the 2012 set of candidates, imputations, retirements, and so forth, the Texas map would have escaped scrutiny. But under the 2014 values, it is not:



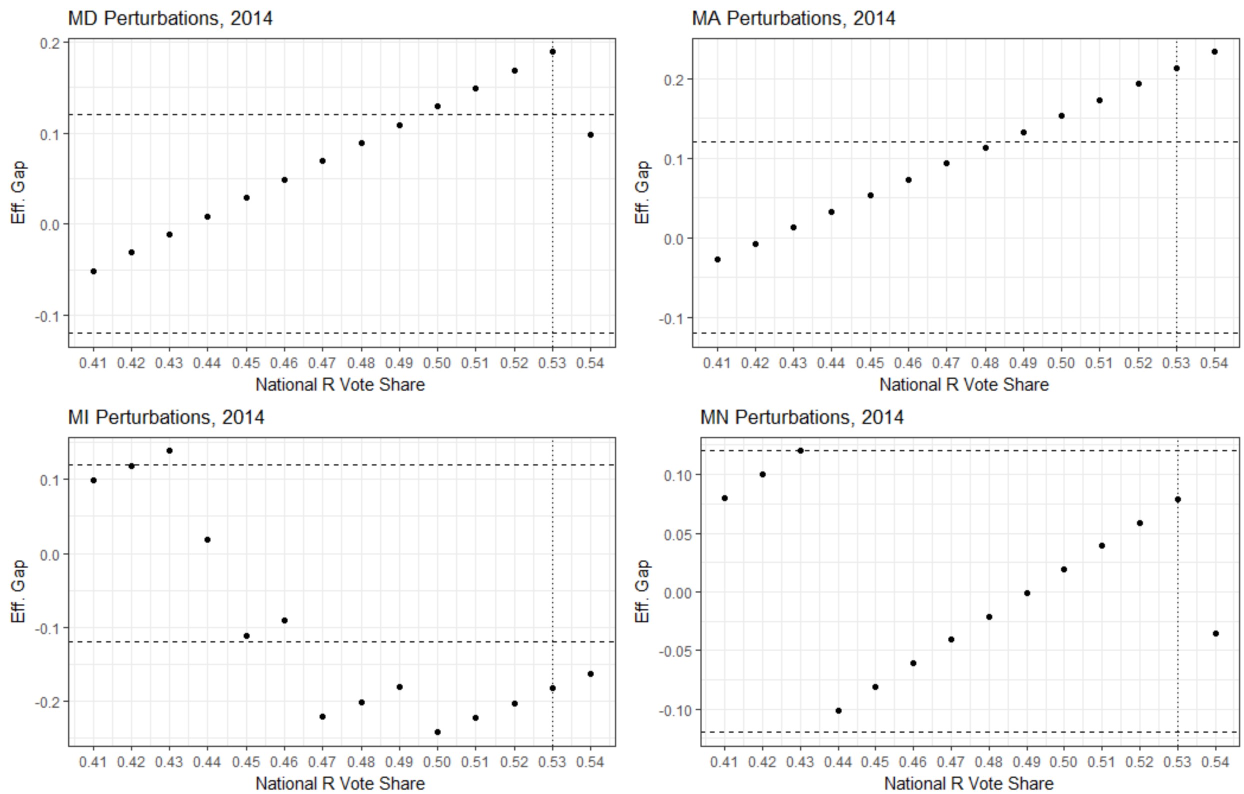
160. Notice that while Arizona triggered scrutiny under a fair number of popular vote values in 2012, under the 2014 set of candidates, it only does so under one. Also, while Arizona became a Republican gerrymander at high values, here it becomes a Democratic gerrymander at one high value.



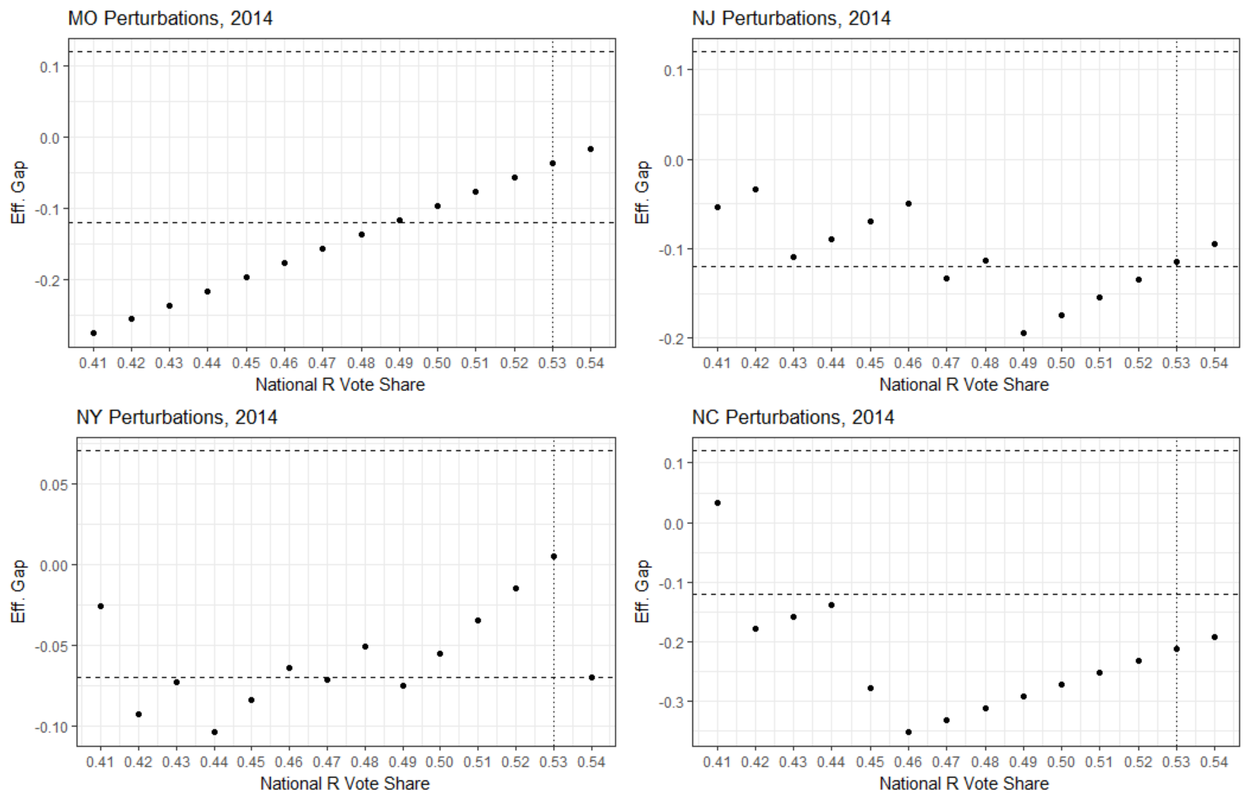
161. Here are Florida, Georgia, Illinois and Indiana. Note, for example, that while Indiana received scrutiny in most scenarios in 2012, here it receives scrutiny in all but three. Georgia escaped scrutiny most of the time in 2012, here it receives scrutiny most of the time.



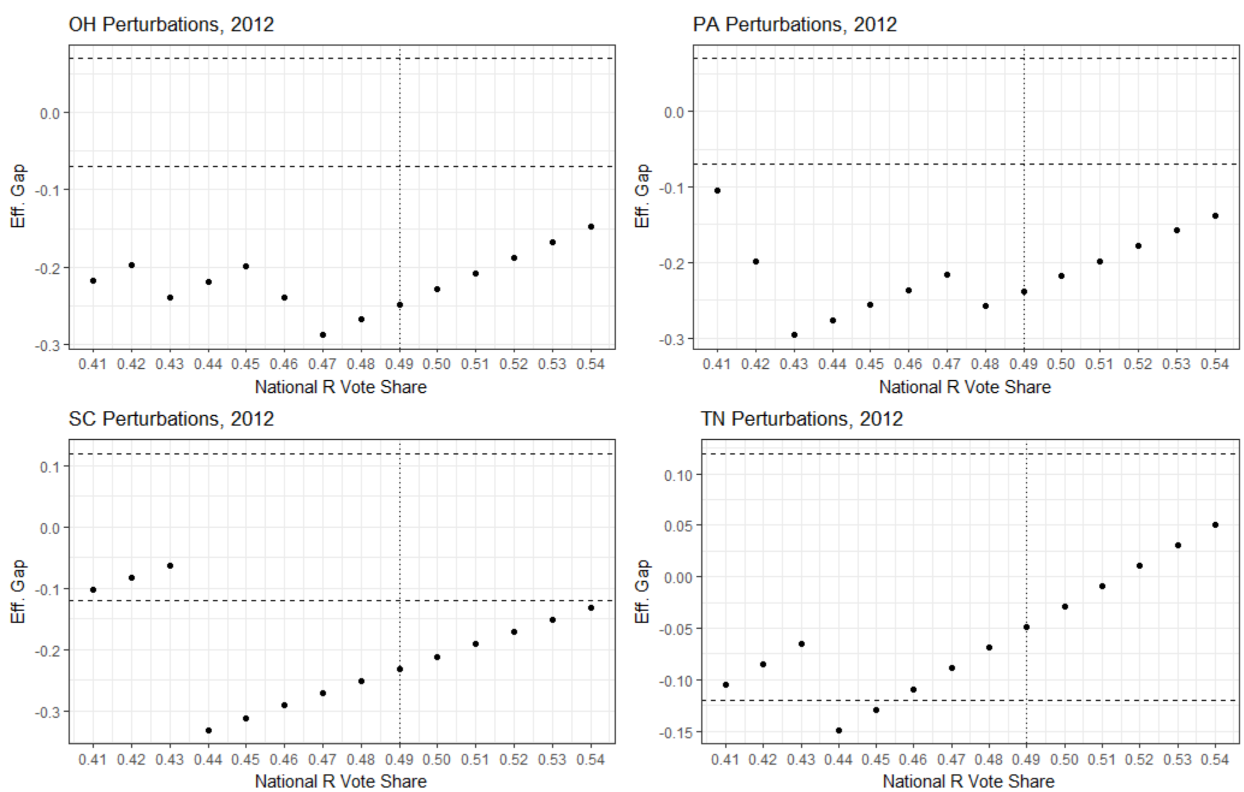
162. Massachusetts almost always avoids scrutiny under 2012, because Republicans fielded a good candidate in the 6th District that year against a scandal-plagued incumbent. He almost won. So in slightly better environments than actually played out in 2012, Richard Tisei wins, and Democrats waste a lot of votes. Tisei was back in 2014, but Tierney lost in the Democratic primary, and Tisei lost by 14 points. Therefore, improving the environment for Republicans only wastes more Republican votes.



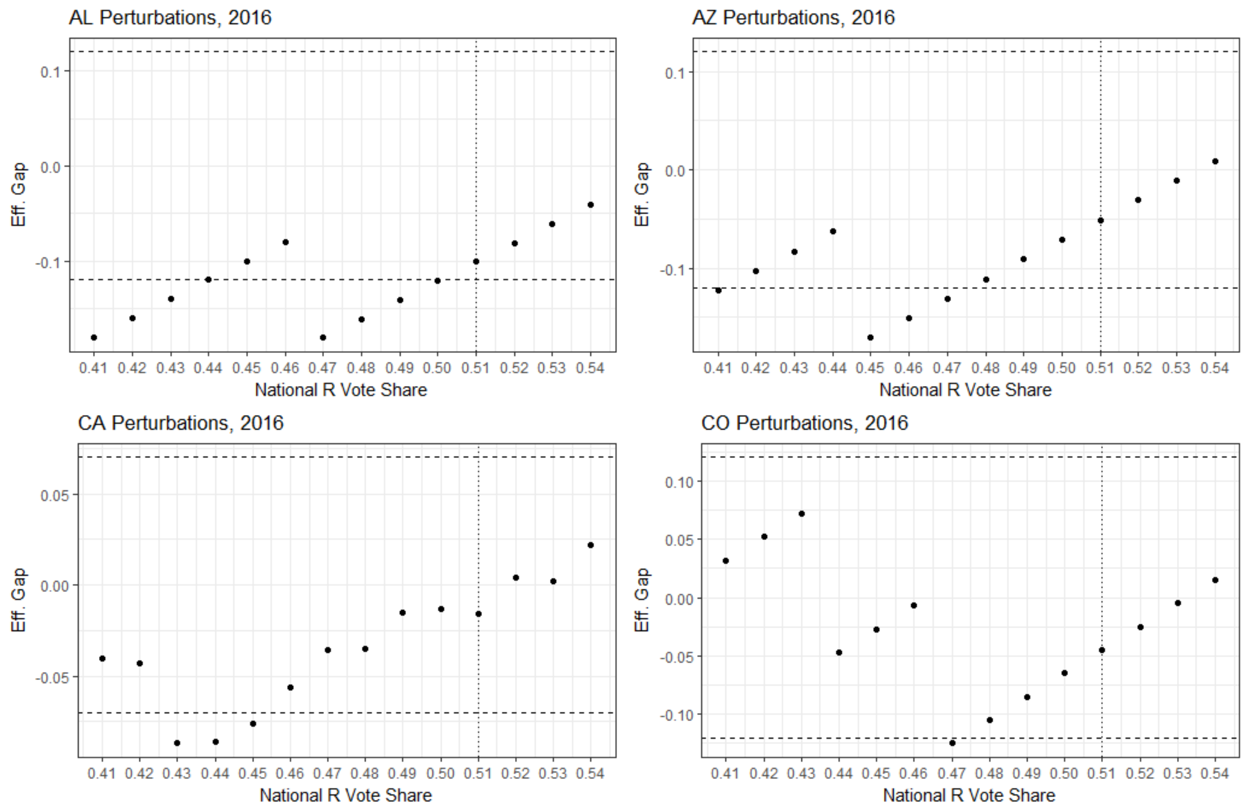
163. Here is 2014 for Missouri, New Jersey, New York and North Carolina. Recall that on the basis of 2012 outcomes, we expected New York to be a Republican gerrymander if Republicans received 53 percent of the vote nationally. As it turns out, an *actual* election where Republicans received 53 percent of the vote nationally resulted in an efficiency gap near zero (and that leaned a bit to the left, in fact).

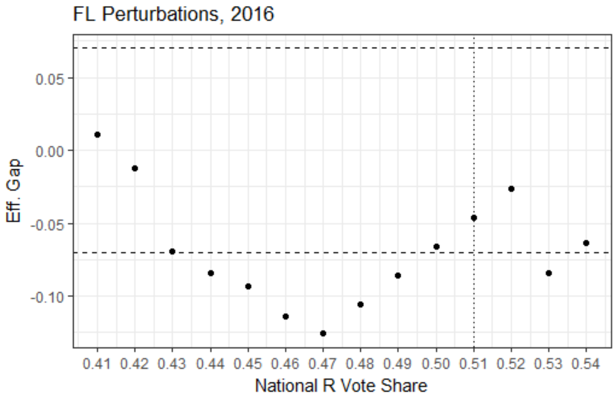


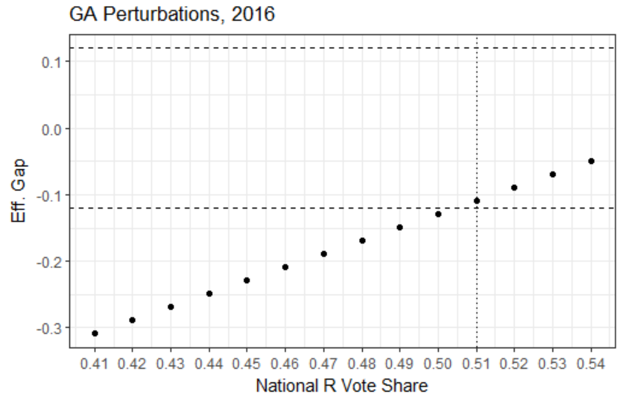
164. Here are Ohio, Pennsylvania, South Carolina and Tennessee.

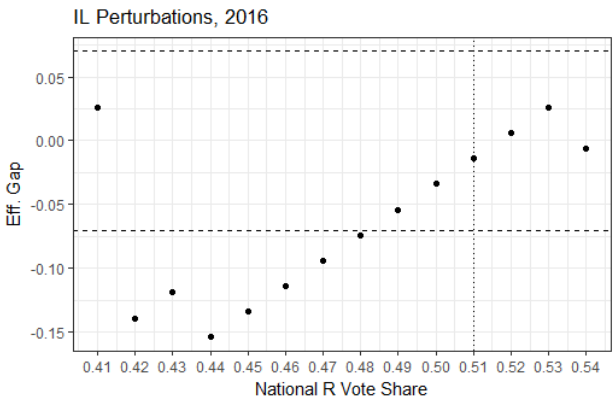


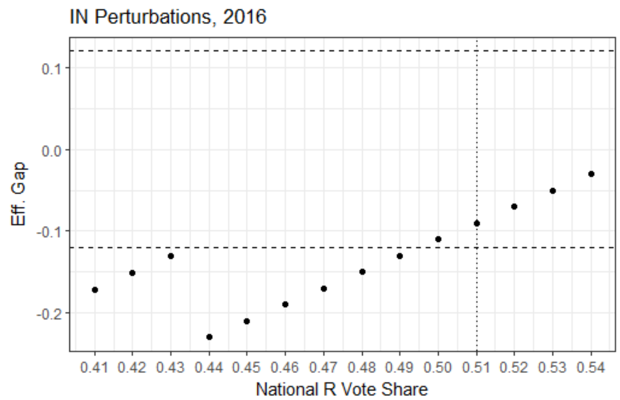
165. Here are the 2016 values. Note that the differences in within-state distributional shape are even more extreme here. Note too that a state like Ohio, which had actionable efficiency gaps in every scenario in 2012 and 2014, now has efficiency gaps that are not actionable in environments that are only modestly better than what occurred in 2016. Had the candidate matrix looked in 2016 like it did in 2014, Ohio's map would survive scrutiny.

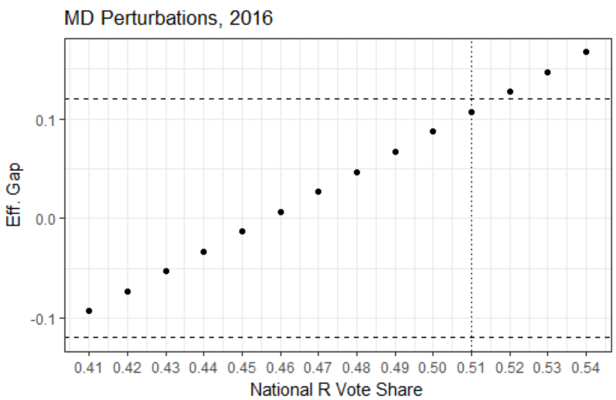


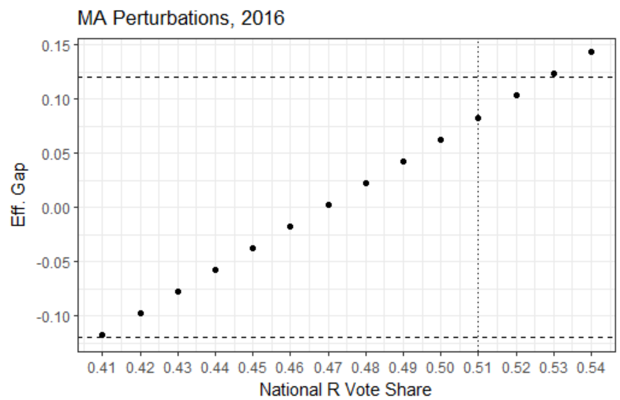

FL Perturbations, 2016
Eff. Gap
0.05
0.00
-0.05
-0.10
0.41 0.42 0.43 0.44 0.45 0.46 0.47 0.48 0.49 0.50 0.51 0.52 0.53 0.54
National R Vote Share

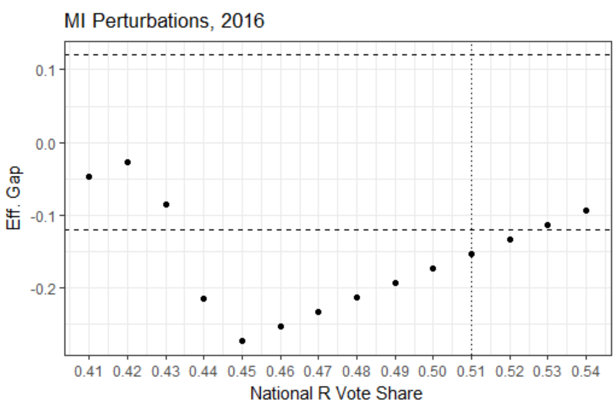

GA Perturbations, 2016
Eff. Gap
0.1
0.0
-0.1
-0.2
-0.3
0.41 0.42 0.43 0.44 0.45 0.46 0.47 0.48 0.49 0.50 0.51 0.52 0.53 0.54
National R Vote Share

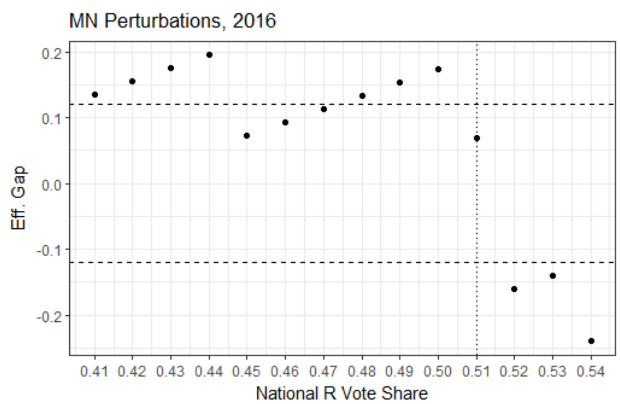

IL Perturbations, 2016
Eff. Gap
0.05
0.00
-0.05
-0.10
-0.15
0.41 0.42 0.43 0.44 0.45 0.46 0.47 0.48 0.49 0.50 0.51 0.52 0.53 0.54
National R Vote Share

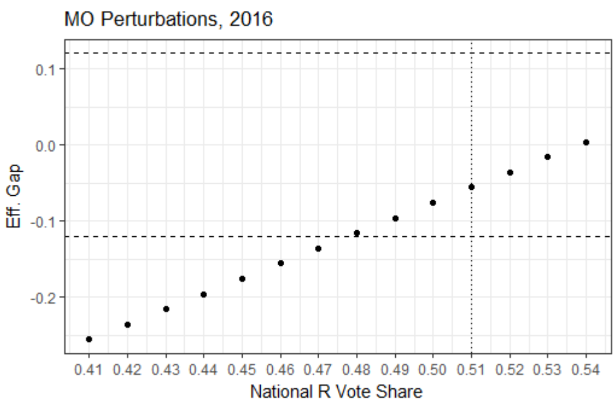

IN Perturbations, 2016
Eff. Gap
0.1
0.0
-0.1
-0.2
0.41 0.42 0.43 0.44 0.45 0.46 0.47 0.48 0.49 0.50 0.51 0.52 0.53 0.54
National R Vote Share

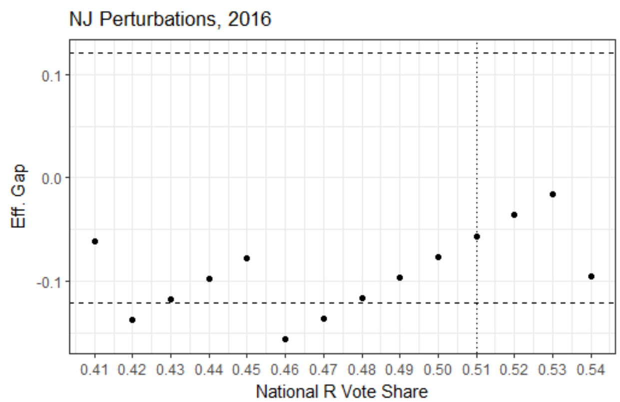

MD Perturbations, 2016
Eff. Gap
0.1
0.0
-0.1
0.41 0.42 0.43 0.44 0.45 0.46 0.47 0.48 0.49 0.50 0.51 0.52 0.53 0.54
National R Vote Share

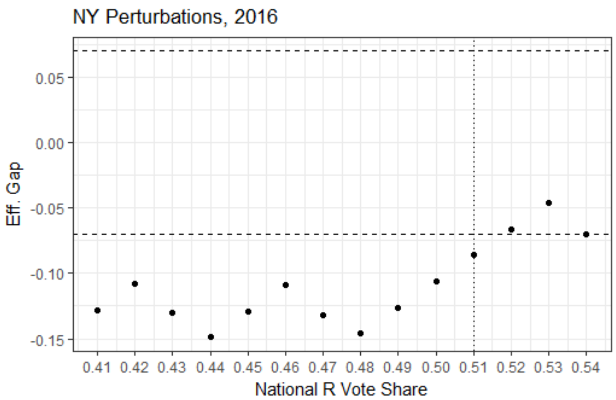

MA Perturbations, 2016
Eff. Gap
0.15
0.10
0.05
0.00
-0.05
-0.10
0.41 0.42 0.43 0.44 0.45 0.46 0.47 0.48 0.49 0.50 0.51 0.52 0.53 0.54
National R Vote Share

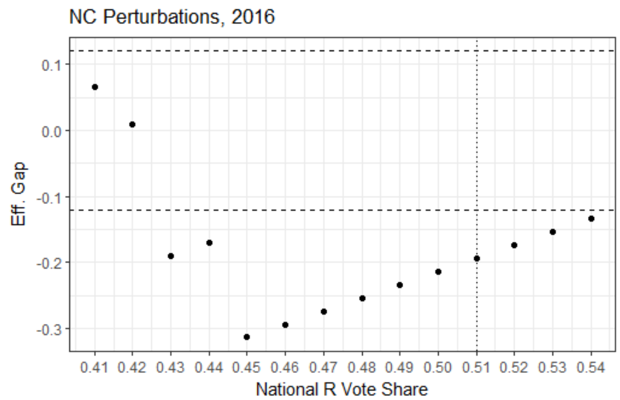

MI Perturbations, 2016
Eff. Gap
0.1
0.0
-0.1
-0.2
0.41 0.42 0.43 0.44 0.45 0.46 0.47 0.48 0.49 0.50 0.51 0.52 0.53 0.54
National R Vote Share

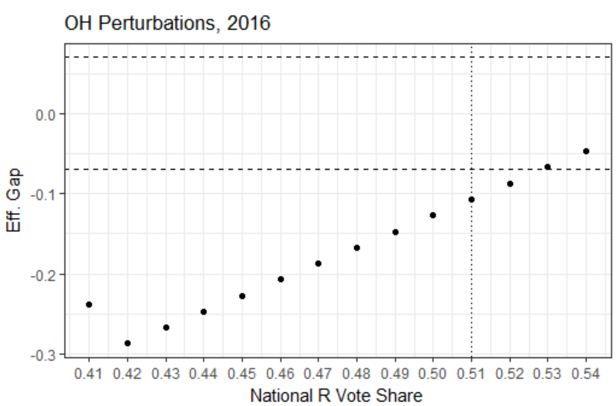

MN Perturbations, 2016
Eff. Gap
0.2
0.1
0.0
-0.1
-0.2
0.41 0.42 0.43 0.44 0.45 0.46 0.47 0.48 0.49 0.50 0.51 0.52 0.53 0.54
National R Vote Share

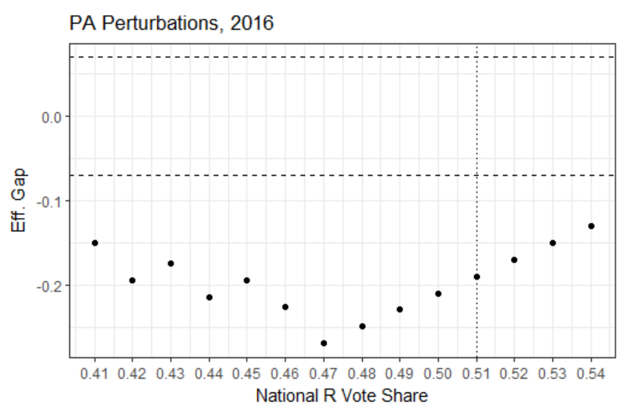

MO Perturbations, 2016
Eff. Gap
0.1
0.0
-0.1
-0.2
0.41 0.42 0.43 0.44 0.45 0.46 0.47 0.48 0.49 0.50 0.51 0.52 0.53 0.54
National R Vote Share

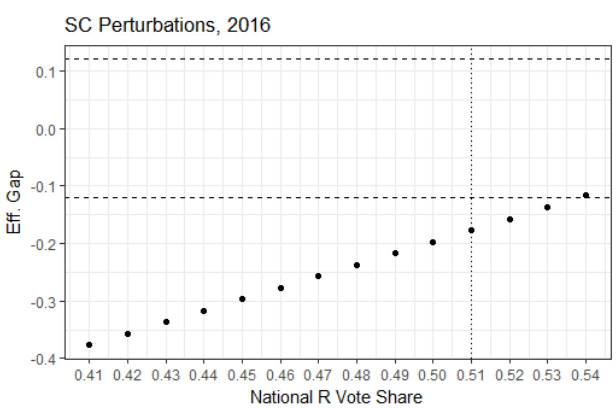

NJ Perturbations, 2016
Eff. Gap
0.1
0.0
-0.1
0.41 0.42 0.43 0.44 0.45 0.46 0.47 0.48 0.49 0.50 0.51 0.52 0.53 0.54
National R Vote Share

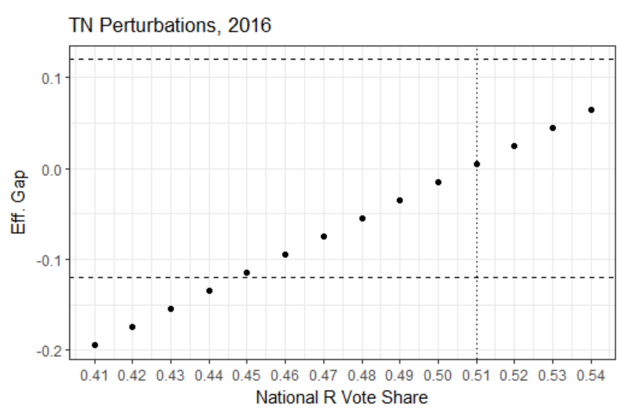

NY Perturbations, 2016
Eff. Gap
0.05
0.00
-0.05
-0.10
-0.15
0.41 0.42 0.43 0.44 0.45 0.46 0.47 0.48 0.49 0.50 0.51 0.52 0.53 0.54
National R Vote Share

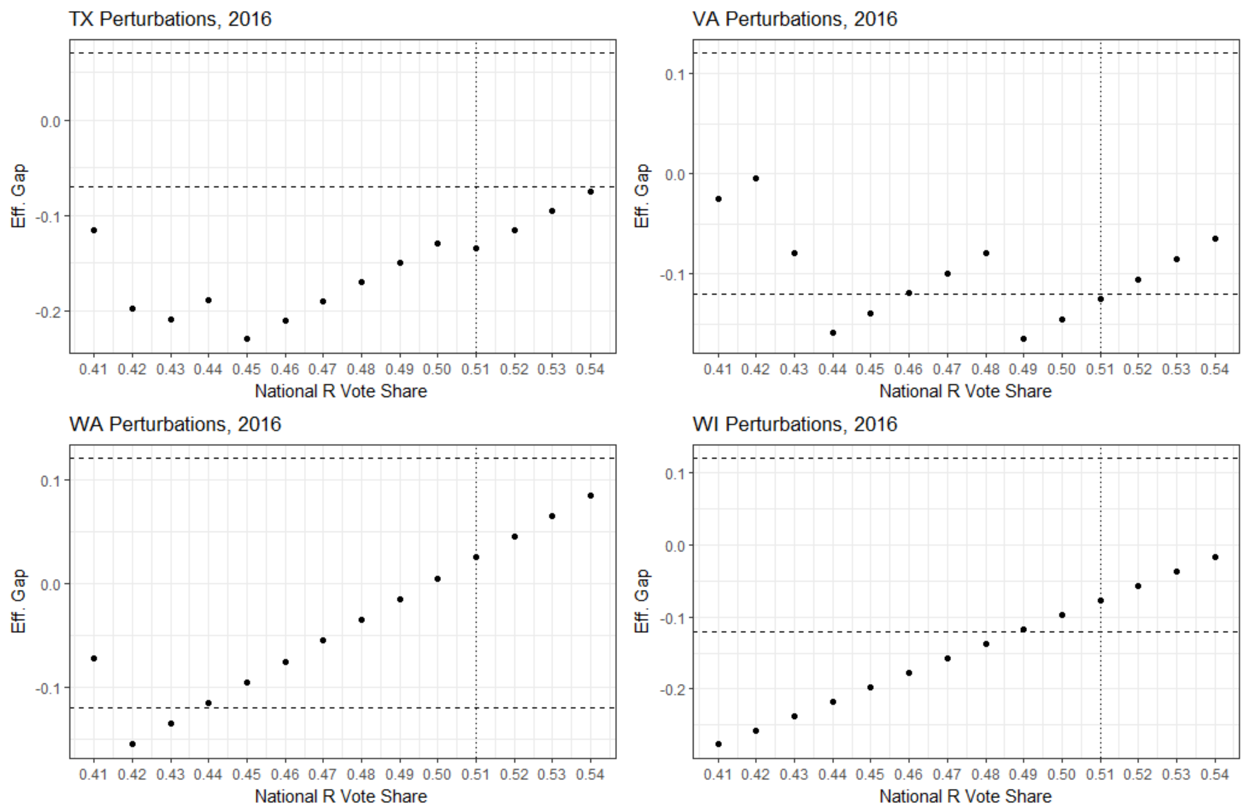

NC Perturbations, 2016
Eff. Gap
0.1
0.0
-0.1
-0.2
-0.3
0.41 0.42 0.43 0.44 0.45 0.46 0.47 0.48 0.49 0.50 0.51 0.52 0.53 0.54
National R Vote Share


OH Perturbations, 2016
Eff. Gap
0.0
-0.1
-0.2
-0.3
0.41 0.42 0.43 0.44 0.45 0.46 0.47 0.48 0.49 0.50 0.51 0.52 0.53 0.54
National R Vote Share


PA Perturbations, 2016
Eff. Gap
0.0
-0.1
-0.2
0.41 0.42 0.43 0.44 0.45 0.46 0.47 0.48 0.49 0.50 0.51 0.52 0.53 0.54
National R Vote Share


SC Perturbations, 2016
Eff. Gap
0.1
0.0
-0.1
-0.2
-0.3
-0.4
0.41 0.42 0.43 0.44 0.45 0.46 0.47 0.48 0.49 0.50 0.51 0.52 0.53 0.54
National R Vote Share


TN Perturbations, 2016
Eff. Gap
0.1
0.0
-0.1
-0.2
0.41 0.42 0.43 0.44 0.45 0.46 0.47 0.48 0.49 0.50 0.51 0.52 0.53 0.54
National R Vote Share



166. The mechanics of this are quite difficult for mapmakers to game out. Most maps have scenarios in which they produce actionable efficiency gaps, and most maps have scenarios in which they do not. Moreover, those scenarios change from year-to-year. Yet map drawers will have to figure out a way to anticipate these changes.

IX. <u>**The Efficiency Gap does nothing to address distorted boundaries, violations of communities of interest, or "Congressmen choosing their constituents," and could encourage uncompetitive districts and entrenched majorities.**</u>

167. Notably, the efficiency gap does not address some of the more common objections to gerrymandering: distorted boundaries and "congressmen choosing their constituents." Plaintiffs' approach is largely indifferent to process and to traditional redistricting principles, so long as a large efficiency gap is not produced in the first year of operation.

168. Perhaps more importantly, the efficiency gap could actually worsen two of the biggest complaints about gerrymanders: That they entrench majorities, and that they reduce competitive districts.

169. Consider an extreme example from North Carolina, which illustrates nicely how the efficiency gap can incentivize things. Assume that in a particular year, Democrats win 2,000,005 votes, while Republicans win 2,000,004 votes. These are split up into evenly-matched seats, such that in a completely neutral year, with evenly matched and evenly funded incumbent candidates, Democrats will win seven districts with one more vote than Republicans, and vice versa. This results in an efficiency gap of .05.

170. If this scenario repeated itself, Democrats would control the congressional delegation by one seat in every election, on the basis of just one vote. This is what I mean when I say the efficiency gap is indifferent to party control.

171. But now assume that a Democratic incumbent stumbles down the stretch, and a Republican manages to win one of these districts narrowly. Suddenly, the efficiency gap shifts by a tenth point to -.054.

172. As an even more extreme example, imagine that there is a uniform national swing of 1 percent. Suddenly, the efficiency gap is an astonishing -.675. This is because the swing districts held by Democrats all flip, and suddenly, Republicans control the delegation with just 51 percent of the vote.

173. This may seem far-fetched, but it should sound familiar. This is basically what happened in Washington in the 1990s. The independent redistricting panel drew a large number of competitive seats, that went back-and-forth between the parties during the 1990s. The efficiency gap ping-ponged back-and-forth accordingly.

174. The bottom line is that competitive districts generate a large amount of uncertainty with respect to the efficiency gap, because small perturbations in the national environment, or freak candidate effects produce outsized swings in the gap. You see this in the perturbation charts above, where changes in the national vote produce small shifts in the efficiency gap, until a seat flips, in which case a substantial movement ensues.

175. There are many responses to this, but one possible response is to stop drawing competitive districts, and instead draw a large number of entrenched seats that guarantee single party control, but are unlikely to violate the efficiency gap.

176. Instead, a mapmaker could simply look at the chart at the beginning of the report. For a state with 13 districts, 8 safely Republican districts are acceptable when Republicans receive between 48 and 63 percent of the vote. If a party was fairly comfortable its statewide vote share was unlikely to fall below 48 percent statewide, this would be a safe bet. True, the first-year efficiency gap may fall below 48 percent in a bad enough Republican year. But in the unlikely event that this occurred, mapmakers could just redraw the map. and hope for a better environment for Republicans in the following year. Likewise, if Wisconsin Republicans were to redraw the general assembly map with 50 safe Republican seats and 49 safe Democratic seats, they would be more-or-less guaranteed control of the chamber, provided their vote share statewide didn’t fall below 47 percent. Given that it would be more difficult to convince quality challengers to run in districts where they would have no chance of winning, the odds of this occurring could be slight.

177. Of course, parties would probably not take a course of action this extreme. But it would be wholly permissible for them to do so and it does illustrate the perverse incentives that flow from the efficiency gap.

X. **<u>It is not clear why we would assume that Party Control Drives Efficiency Gaps</u>**

178. Two final points can be addressed briefly. First, Dr. Jackman claims that party control is a primary driver of efficiency gaps. While I don't doubt that Republicans draw maps that lean Republican compared to those drawn by independent redistricting commissions and Democratic legislatures, and while I appreciate the power of regression analysis, I am unconvinced that the correlation Dr. Jackman finds has causal power.

179. Seat share over time does not look like the chart Dr. Jackman creates in Figure 9. It is a stepwise pattern, with horizontal lines extending, generally speaking, over the course of the decade. Party control is constant.

180. Yet the efficiency gap, as described by Stephanopoulos and McGee, does not follow this stepwise pattern. It rises gradually, over the course of multiple decades, with hardly any notable movement in most redistricting years.

181. More importantly, it rises fairly substantially during the 1990s, when Republicans were mostly shut out of redistricting. I do not see how Republicans can be responsible for a phenomenon related to a process in which they did not participate is beyond me.

XI. **<u>Dr. Chen's approach did not perform well in Wisconsin</u>**.

182. Second, while I have generally avoided commenting on Dr. Chen's approach, I note that he filed an amicus brief in the Wisconsin litigation, employing a similar technique. See Exhibit 4. Although I am favorably disposed toward Dr. Chen's work, it has remained mostly theoretical. Its application in Wisconsin, however, reveals problems.

183. In his report, Dr. Chen's simulations produced efficiency gaps between .03 and -.06. The problem is that Wisconsin maps produced absolute efficiency gaps in excess of .06 in

1998, 2000 and 2002, and in the neighborhood of .1 in 2004 and 2006. *See* Jackman report in 72, Exhibit 5.

184. This isn't a problem in and of itself, but the 1992 and 2002 maps were drawn by a court, with the 1992 map drawn by a panel that included one of the judges from the eventual *Nichol* majority. It seems highly unlikely that those panels acted with partisan intent. The upshot of this is that for whatever reason, Dr. Chen's algorithm does not appear to sample from the full universe of plans drawn without partisan intent, and falling outside of the intervals of his results does not seem to prove that partisan intent is responsible for that disparity.

Pursuant to 28 U.S.C. § 1746, I declare under penalty of perjury under the laws of the United States that the foregoing statements are true and correct.

This the 3rd day of April, 2017.

____

Sean P. Trende

SEAN P. TRENDE
1146 Elderberry Loop.
Delaware, OH 43015
strende@realclearpolitics.com

EDUCATION

B.A., Yale University, with distinction, History and Political Science, 1995.

M.A., Duke University, *cum laude*, Political Science, 2001. Thesis titled *The Making of an Ideological Court: Application of Non-parametric Scaling Techniques to Explain Supreme Court Voting Patterns from 1900-1941*, June 2001.

J.D., Duke University School of Law, *cum laude*, 2001; Duke Law Journal, Research Editor; Moot Court Board.

PROFESSIONAL EXPERIENCE

Law Clerk, Hon. Deanell R. Tacha, U.S. Court of Appeals for the Tenth Circuit, 2001-02.

Associate, Kirkland & Ellis, LLP, Washington, DC, 2002-05.

Associate, Hunton & Williams, LLP, Richmond, Virginia, 2005-09.

Associate, David, Kamp & Frank, P.C., Newport News, Virginia, 2009-10.

Senior Elections Analyst, RealClearPolitics, 2009-present.

BOOKS

Larry J. Sabato, ed., *Trumped: The 2016 Election that Broke all the Rules* (2017), forthcoming.

Larry J. Sabato, ed., *The Surge:2014's Big GOP Win and What It Means for the Next Presidential Election*, ch. 12 (2015).

Larry J. Sabato, ed., *Barack Obama and the New America*, ch. 12 (2013).

Barone, Kraushaar, McCutcheon & Trende, *The Almanac of American Politics 2014* (2013).

The Lost Majority: Why the Future of Government is up for Grabs – And Who Will Take It (2012).

REAL CLEAR POLITICS COLUMNS

Full archives available at http://www.realclearpolitics.com/authors/sean_trende/

PUBLICATIONS FROM LAST 10 YEARS

"The GOP and the Latino Vote," *National Review*, June 15, 2012.

"Political Economy," *National Review*, Special 2010 Election Issue.

"It's 1974 All Over Again," *The Weekly Standard*, Apr. 26, 2010.

(with Christian C. Burden), "The Economic Loss Rule and Franchise Attorneys," 27 Franchise L.J., 192 (2008)

SELECTED PRESENTATIONS AND APPEARANCES

Panelist, "Independent Experts on Republican Candidates" (with Michael Barone and Josh Kraushaar), American University, Washington, DC, November 2011.

Panelist, "2011 Mortimer Caplin Conference on the World Economy" (with Bill Schneider, John Sides, and Sarah Binder), The National Press Club, Washington, DC, December 2011.

"The State of the Presidential Nominating Process: A Debate" (with Jay Cost), Berry College, December 2011.

"The Lost Majorities: 2008, 2010 and America's Political Future," Bradley Lecture, American Enterprise Institute, January 2012.

Panelist, "Collective Bargaining, Public Pensions and Voters: The Policy and Politics of Public-Sector Employees in the 2012 Elections," (with Karlyn Bowman, Ruy Teixeira, and Henry Olson), American Enterprise Institute, January 2012.

"The People's Money: How Voters Will Balance the Budget and Eliminate the Federal Debt.," (with Michael Barone and Scott Rasmussen), CATO Institute, March 2012.

Panelist, "Republican Primaries: Explaining the Results and Assessing What they Mean for the Future of the GOP," (with Dante Scala and Kate Zernike), Chaire Raoul-Dandurand en Etudes Strategiques et Diplomatiques, Montreal, March 2012

"Obama's Vanished Coalition," (with Lance Tarrance and Emily Ekins), CATO Institute, April 2012.

Panelist, "The Future of Red and Blue," (with Ruy Teixeira), Bipartisan Policy Center, Washington, DC, April 2012.

"The 2012 Elections: Trends, Prognostications and What's at Stake," 3rd Annual Family Office Wealth Management Forum, Greensboro, Georgia, May 2012.

“2012 U.S. Elections Series,” with Bruce Stokes and Alexandra de Hoop Scheffer, German Marshall Fund, Brussels, Belgium, Oct. 4, 2012.

Panelist, “The Power of Pundits,” (with John Sides, Linda Vavreck, and Melissa Harris-Perry), American Political Science Ass’n, Aug. 29, 2013.

Panelist, “Post-Election Coverage” (with Raul Avillar, Dan Balz, Robert Collins, Jen O’Malley Dillon, Alex Isenstadt, Nathan Klein, Joe Lenski, John McLaughlin, and Patrick O’Connor), University of Kansas, Dec. 11-12, 2014.

Panelist, “Evenwel v. Abbott: What Does One Person, One Vote Really Mean?” (with Andrew W. Grossman and Hans A. von Spakovsky), Heritage Foundation, Sept. 15, 2015.

Appeared in countless radio and television appearances including appearances on Fox News, MSNBC, ABC News Australia, Fox News Radio, Beijing Radio, CNN Radio, NPR, and other outlets.

Rebuttal Report

Simon Jackman

December 21, 2015

Introduction

In this rebuttal report, I respond to criticisms made by Sean P. Trende and Professor Nicholas Goedert in their respective expert reports. I also conduct new empirical analyses further confirming the validity of the efficiency gap as a measure of partisan gerrymandering and the reasonableness of the proposed 0.07 threshold. More specifically, my principal contributions are the following:

- *First*, I respond to Goedert's various critiques of the efficiency gap and of the proposed efficiency gap threshold. Among other things, he misunderstands the relevance of efficiency gap data, cherry-picks information from my initial report while ignoring its broader context, and wrongly claims that plaintiffs' test would mandate "hyper-responsiveness" or prevent states from pursuing goals such as competitiveness or proportional representation.

- *Second*, I calculate several widely accepted prognostic measures—all based on the rates of true positives, false positives, true negatives, and false negatives—with respect to the odds of a district plan's efficiency gap changing signs over the plan's lifetime given a certain efficiency gap value in the plan's first election. Based on these measures, I conclude that the proposed 0.07 threshold is highly conservative. In fact, this threshold *sacrifices* some accuracy (which would be maximized at a lower threshold) in order to reduce the proportion of false positives.

- *Third*, I calculate the same prognostic measures with respect to the odds of a district plan's *average* efficiency gap, over its lifetime, having a different sign than that observed in the first election under a plan, given a certain efficiency gap value in this first election. Under this method, the proposed 0.07 threshold appears even more conservative, driving down the share of false positives to below 5%.

- *Fourth*, I compare the values of the efficiency gap in the *first* election under a plan and *on average* over the plan's lifetime. This relationship is impressively tight (r^2=0.73), indicating that a plan's initial bias is a very good predictor of its overall lifetime bias. For Act 43, this analysis allows us to predict that it will *average* a pro-Republican efficiency gap of almost 10% over the 2010 cycle as a whole.

- *Fifth*, I examine to what extent changes in party control over redistricting are responsible for the pro-Republican trend in the efficiency gap since the 1990s. In the current cycle, about *four times* more state house plans were designed by Republicans in full control of state government than in the 1990s. Had the distribution of party control over redistricting remained unchanged, essentially *all* of the pro-Republican movement in the efficiency gap over the last two decades would not have occurred. It is thus changes in party control, and *not* changes in the country's political geography, that primarily account for Republicans' growing redistricting advantage over the last generation.

- *Sixth*, I address recent work by Chen and Rodden (2013), cited by both Trende and Goedert for the proposition that Republicans enjoy a natural geographic advantage over Democrats. Chen and Rodden's simulated maps are not *lawful* because they ignore the Voting Rights Act and state redistricting criteria; they are based on presidential election results rather than more relevant state legislative election results; they do not constitute a representative sample of the entire plan solution space; and they are contradicted by other recent work (Fryer & Holden 2011) finding that randomly drawn plans *reduce* bias and *increase* electoral responsiveness.
- *Lastly*, I comment on Trende's analysis of particular state legislative and congressional plans. This analysis is marked by conceptual and methodological errors severe enough to render it useless. For example, Trende ignores two of the three prongs of plaintiffs' proposed test; he calculates congressional efficiency gaps without converting them from percentage points to House seats and for House delegations too small to generate reliable estimates; and he simply *substitutes* presidential election results for congressional election results whenever the latter are missing due to uncontested races. None of this work meets accepted standards of social science rigor.

1 Responses to Goedert's criticisms

In his report, Goedert offers several critiques of the efficiency gap and of the 0.07 threshold I recommended in my initial report, based primarily on the alleged instability of the efficiency gap. None of these critiques have merit. In this section, I respond to Goedert's points relying only on the analysis of my initial report and on the existing literature. My new empirical analyses appear in subsequent sections.

First, Goedert appears to believe that a plan's efficiency gap is only relevant to the extent that it sheds light on the partisan intent (or lack thereof) underlying the plan. He writes that "such intent cannot be inferred" from a large efficiency gap, that "a durable bias . . . is not even a sign of deliberate partisan intent," and that the "efficiency gap [is] a standard to measure partisan intent" (pp. 11, 13, 19). But this is not at all the legal function of the efficiency gap in plaintiffs' proposed test. Rather, partisan intent is its own independent inquiry, and the efficiency gap then comes into play at the *second* stage of the test, to determine if a plan's electoral *consequences* are sufficiently severe that it should be deemed presumptively unconstitutional. To put it simply, the efficiency gap is plaintiffs' measure of partisan *effect*, not of partisan *intent*. Goedert's misunderstanding of this basic point infects all of his discussion.

Second, Goedert observes that of *all* plans, anytime in the decade, with a *pro-Democratic* efficiency gap of greater than 0.07, a substantial proportion of them switch signs over their lifetimes (p. 11). In making this observation, Goedert cherry-picks a single bit of data from my initial report, and an irrelevant piece of data at that. This fact is irrelevant because it applies to plans no matter when their elections were held, while the appropriate universe for plaintiffs, defendants, and courts is limited to the *first* elections held under plans. It is the first elections that typically will be used in litigation, given Justice Kennedy's admonition in *Vieth* that plans should not be struck down based on a "hypothetical state of affairs," but rather "if and when the feared inequity arose" (*Vieth v. Jubelirer* (2004), p. 420). And the fact is misleading because it applies only to pro-Democratic efficiency gaps above 0.07, and not to the larger set of pro-Republican efficiency gaps above this threshold.

If we consider only plans that exhibit a pro-Democratic efficiency gap above 0.07 in their *first* elections, the probability that they will switch signs over their lifetimes drops by about five percentage points (Jackman Report, p. 61). And if we then turn to plans that exhibit a *pro-Republican* efficiency gap above 0.07 in their first elections—a more sizeable set, for which more accurate estimates are possible—this probability drops all the way to about 15% (Jackman Report, p. 61). In other words, of plans that open with large pro-Republican efficiency gaps, close to 85% of them continue to favor Republicans in every election for the remainder of the cycle. *This* is the most pertinent data point in my report, not the one cherry-picked by Goedert, and it reveals the persistence of many gerrymanders.

Third, Goedert discusses *congressional* district plans throughout his report, even though this case is exclusively about state legislative redistricting (pp. 7-8, 10, 12, 20). In doing so, he makes some of the same errors as does Trende: namely, not converting the efficiency gap from percentage points to House seats, and improperly handling uncontested races (in his case, by not adjusting for the uncontestedness *at all*, and simply treating the races as if all of the vote went to one party and none to the other). I discuss these errors in more detail later in this report.

Fourth, Goedert claims that it is "arbitrary" to focus on the first election after redistricting, and that doing so "biases toward a finding of *EG* durability" by ignoring wave elections (p. 14). As noted above, the first election after redistricting is the critical one for purposes of litigation, since under *Vieth*, it is after this election that a lawsuit will typically commence and have to be decided by the courts. Later elections are largely irrelevant for litigation purposes, since it is unreasonable to expect suits to be brought six or eight or even ten years into a cycle. Moreover, my analysis in no way ignored wave elections; to the contrary, I determined the odds that a plan's efficiency gap would switch signs by examining *all* elections held under the plan, waves and non-waves alike. If anything, the fact that most wave elections over the last forty years have not taken place in the first election after redistricting biases *against* a finding of durability, since these elections may well cause the efficiency gap to flip signs.

Fifth, Goedert is wrong that an efficiency gap of zero represents "'hyper-responsive' representation" (p. 2). In fact, as he has recognized in his own prior work, an efficiency gap of zero corresponds almost exactly to the responsiveness actually displayed by American elections over the course of the twentieth century, under which "a 1% increase in vote share will produce about a 2% increase in seat share" (Goedert 2014, p. 3). Indeed, this correspondence is one of the efficiency gap's most attractive properties, and it explains why Goedert himself calculated a quantity nearly identical to the efficiency gap in his work (Goedert 2014; Goedert 2015).

And sixth, Goedert is wrong as well that plaintiffs' proposed test might discourage states from pursuing worthwhile goals such as competitiveness or proportional representation (pp. 6-10). If a state's aim in redrawing districts was to make them more competitive or to produce more proportional representation, then the partisan intent required by the first prong of plaintiffs' test would not be present. Even if partisan intent were somehow found, the state would likely be able to show that its plan's large efficiency gap was necessitated by its pursuit of competitiveness or proportional representation. And in any event, competitiveness and proportional representation are extremely rare objectives in American redistricting. Only *one* state, Arizona, has a competitiveness requirement, and not a *single* state has a proportional representation criterion. (And needless to say, line-drawers do not tend to seek out either of these goals on their own.)

2 Reliability of a district plan's first efficiency gap

Having rebutted Goedert's criticisms using preexisting data, I now provide further analysis of the reliability of the first efficiency gap (*EG*) observed in the life of a district plan. This played a key role in the determination of the threshold *EG* value in my initial report. In that report, I focused on the probability of a "sign-flip": that is, given the magnitude of the efficiency gap observed in the first election under a district plan, what can we infer about the likelihood that all subsequent efficiency gaps observed under that plan will have the same sign as that from the first election.

Under this approach, just one election that produces an efficiency gap with a different sign from the efficiency gap in the first election will generate a "failure," in the sense we would say that the plan has generated an efficiency gap that conflicts with that from the first election. In short, the "constant sign" analysis in my original report considers the most extreme set of efficiency gap estimates produced under a plan and insists that they have the same sign. In this sense, the "constant sign" analysis I performed is a quite stringent and conservative test of what we can or ought to infer from the efficiency gap observed in the first election under the district plan. Another approach would be to inquire as to the *average* efficiency gap over the life of the district plan. A summary statistic such as the average is—by definition—less sensitive to extreme values. At the same time—and again, by definition—the average measures central tendency or typicality, and is the most widely used summary statistic in existence. I thus consider how well the first *EG* observed under a district plan predicts the average *EG* observed over the life of the plan.

But I first provide some additional analysis of the prognostic properties of the first efficiency gap observed under a district plan. In each instance the test is whether the first *EG* observed under a plan exceeds a given threshold value. The outcome of interest is whether the plan's remaining efficiency gaps have the same sign as the *EG* from the first election. For purposes of this exercise, plans are classified as "positive" (all *EG* scores under the plan have the same sign) or "negative" (*EG* scores differ in sign). With these definitions in place, we can then classify plans according to the accuracy of the prediction implicit in the first *EG* observed under the plan:

	Actual	
Test	Positive	Negative
Positive	True Positive	False Positive
Negative	False Negative	True Negative

The prognostic measures I rely on are conventional measures of predictive or classification accuracy used throughout the quantitative sciences:

1. sensitivity, or the *true positive rate*: proportion of positives that test positive, TP/(TP + FN)
2. specificity, or the *true negative rate*: proportion of negatives that test negative, TN/(TN + FP)

3. *balanced accuracy*, the average of the sensitivity and the specificity

4. *accuracy*, the proportion of cases that are true positives or true negatives, (TP + TN)/(TP + FP + FN + TN).

5. the *false positive rate*; proportion of negative cases that test positive, 1 minus the specificity or FP/(TN + FP).

6. the *false discovery rate*; proportion of cases testing positive that are actually negative, FP/(TP + FP).

7. the *false omission rate*; proportion of cases that test negative that are actually positive, FN/(FN + TN).

Figure 1 shows how these prognostic performance indicators vary as a function of the absolute *EG* threshold (on the horizontal axis in the figure). That is, as we move to the right in each panel of the graph, the test is becoming increasingly stringent: larger absolute values of the efficiency gap in the first election under a district plan are required to trip the increasingly higher threshold. When the threshold is set to zero, all plans trip the threshold (all first-election *EG*s are greater than zero in magnitude, by definition) and so all cases test positive; in this case the sensitivity is 1, while conversely the specificity is 0 and the false positive rate is 1 (all negatives test positive).

The test has better properties as the threshold grows, with the accuracy measures maximized around absolute values of .03 to .04. Yet accuracy is not all in this context. The rate of false positives is quite high at thresholds where the accuracy is high, as is the false discovery rate. At a threshold of .03, for example, over half of plans that would go on to exhibit sign flips in their *EG*s would test positive and be flagged for inspection; of the plans selected for scrutiny, more than a third would turn out to have *EG* sign flips over the life of the plan. The .07 threshold is thus a conservative standard, the point at which the rate of false positives is becoming reasonably low (25%), without letting the false omission rate go above 50%.

It is worth noting the weight being put on false discoveries or false alarms versus the weight on false omissions in this context, which in turn reflects the conservatism and caution of the thinking underlying the .07 threshold. We propose accepting *twice* the rate of false omissions (plans that should have been scrutinized but were not) than the rate of false discoveries (plans that would be flagged for scrutiny given the *EG* observed in the first election, but would then go on to display sign flips). To reiterate: the proposed standard for judicial scrutiny is cautious and conservative, erring on the side of letting even durably skewed plans stand.

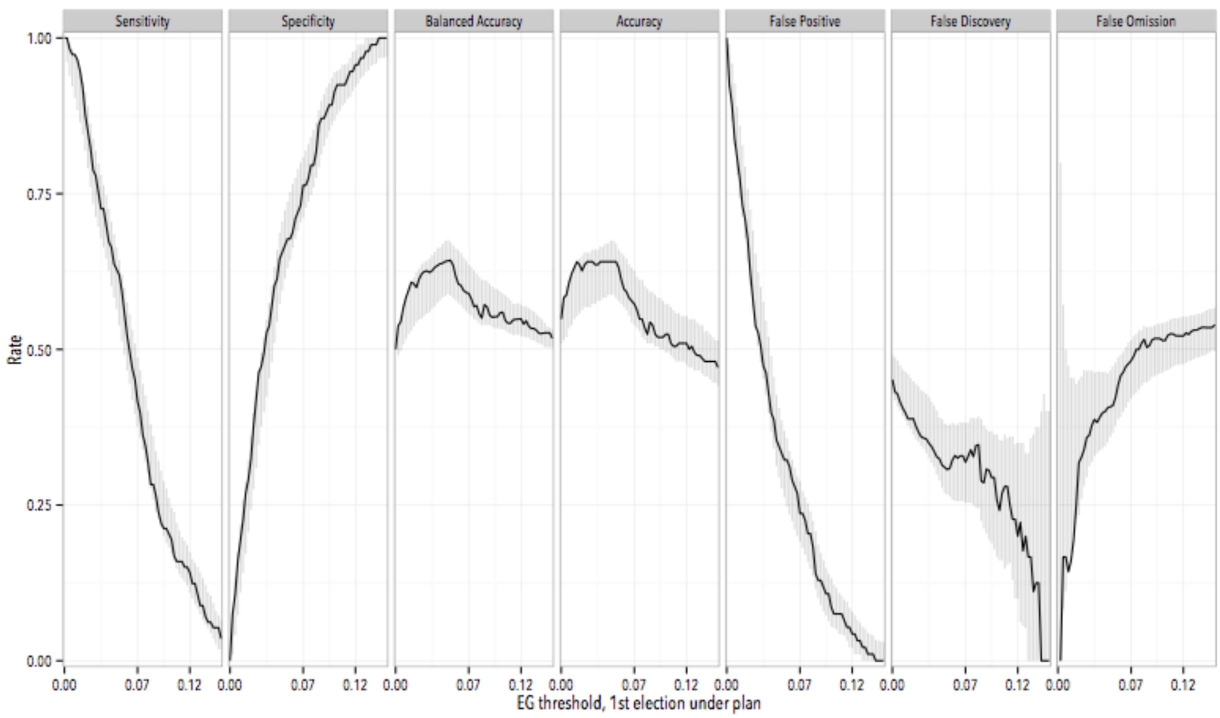


Figure 1: Prognostic performance measures, first efficiency gap under a district plan more extreme than threshold (horizontal axis) as a predictor of whether the subsequent efficiency gaps recorded under the district plan all have the same sign as the first efficiency gap. Vertical lines indicate 95% confidence intervals. Analysis spans all state legislative elections and district plans as per my initial report, 1972-2014.

Figure 2 repeats this analysis, but only considering the performance of *negative* values of the first-election efficiency gap threshold, consistent with Republican advantage (and more relevant to the Wisconsin plan at issue). Here the threshold becomes less stringent as we move across the horizontal axis from left to right, from larger negative thresholds to closer to zero at the right hand edge of each panel. With a large negative threshold (left hand edge of each panel), almost all plans test negative and so the sensitivity is close to zero, the specificity is 1, and the false positive rate is zero. The accuracy measures increase as the threshold becomes less stringent, attaining maxima in the range -.05 to -.02. Again—and consistent with the cautious approach we take—we emphasize that accuracy is not the sole criterion we use to evaluate a decision rule. At low values of the threshold, where accuracy is maximized, the false positive and false discovery rates are relatively high. On the other hand, at the proposed threshold value of -.07, the false positive rate is under 10% (fewer than 10% of plans with efficiency gaps changing signs would be scrutinized), and the false omission rate is about 35% (close to

35% of plans would not be flagged despite having *EG*s of the same sign over their lifetimes). The proposed threshold again errs on the side of restraint, tolerating a higher rate of false omissions than false discoveries.

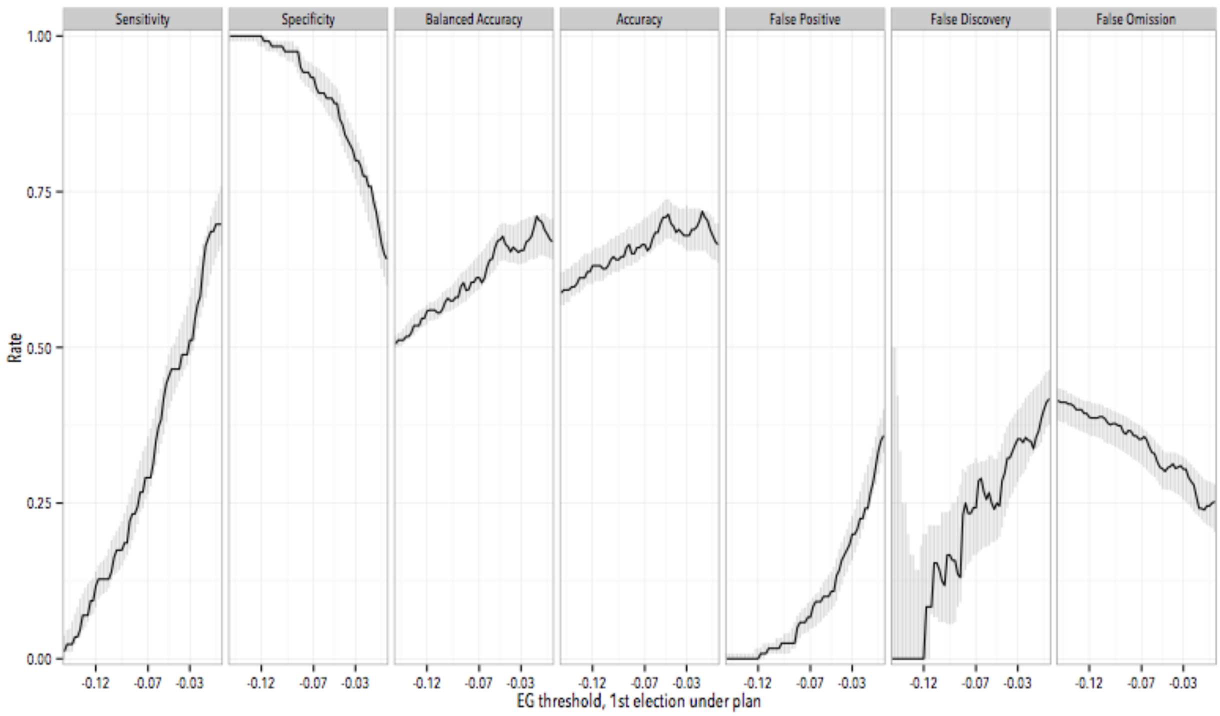


Figure 2: Prognostic performance measures, first efficiency gap under a district plan more extreme than threshold (horizontal axis) as a predictor of whether the subsequent efficiency gaps recorded under the district plan all have the same sign as the first efficiency gap. Vertical lines indicate 95% confidence intervals. Analysis examines negative, first-election threshold values of the efficiency gap, consistent with Republican advantage.

Figure 3 presents the corresponding analysis of *positive* values of the first-election *EG* threshold, consistent with Democratic advantage. Here the proposed threshold becomes more stringent as we move to the right of each panel, in the sense that fewer plans trip the threshold. At high values of the threshold (the right hand edge of each panel), no plans trip the threshold and all are classified as "negatives," leading to a specificity of 1, and false positive and false discovery rates of zero. Once again, accuracy is maximized at a less stringent threshold than the proposed .07 standard, around .03. The false positive rate is much lower at the proposed threshold of .07 than at the accuracy-maximizing threshold of .03. Note that the false discovery rates are moderately large but unstable and estimated with considerable imprecision; this is because there are so few plans exhibiting high (pro-Democratic) levels of *EG* in their first election. Moreover, of the few plans that do trip a given pro-Democratic threshold in their first election, it is reasonably likely that they will record efficiency gaps that will change sign over the life of the plan; this sign-flip or "false discovery" probability is about 35% at the proposed threshold of .07.

Comparing the analyses in Figures 2 and 3, we see an asymmetry in the results. The .07 threshold is more permissive with respect to plans that begin life exhibiting Democratic advantage than it is for plans that initially exhibit Republican advantage. At a +/- .07 threshold, the false discovery rate for plans initially exhibiting Republican advantage is under 10%, but around 35% for plans initially exhibiting Democratic advantage. As Figure 3 shows, it is difficult to find a threshold for apparently pro-Democratic plans that drives the false discovery rate to reliably low levels, if only because the historical record has relatively few instances of these types. We also note that the .07 threshold generates false omission rates of about 30% for both sets of plans.

Because the preceding discussion is somewhat technical, it is worth restating its principal conclusion: It is that an efficiency gap threshold of 0.07 is quite conservative, in that it sacrifices some accuracy (which would be maximized at a threshold of around 0.03) in order to drive down the false positive and false discovery rates. At a threshold of 0.07, in fact, the false positive and false discovery rates are about *half* of the false omission rate, indicating that there are about twice as many plans that are *not* being flagged even though their *EG* signs would remain one-sided throughout the cycle, than there are plans that *are* being flagged even though their *EG* signs would flip. This is further powerful confirmation of the reasonableness of the 0.07 efficiency gap threshold.

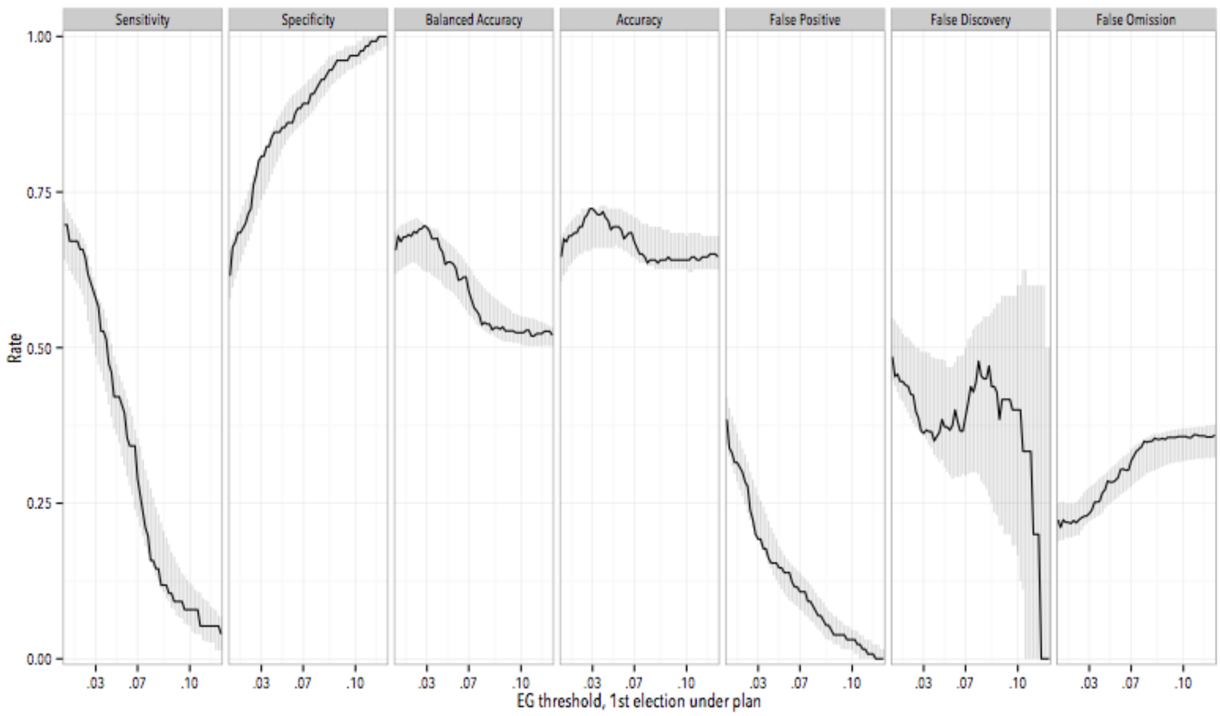


Figure 3: Prognostic performance measures, first efficiency gap under a district plan more extreme than threshold (horizontal axis) as a predictor of whether the subsequent efficiency gaps recorded under the district plan all have the same sign as the first efficiency gap. Vertical lines indicate 95% confidence intervals. Analysis examines positive, first-election threshold values of the efficiency gap, consistent with Democratic advantage.

3 First-election efficiency gap reliability with respect to the plan-average efficiency gap sign

Next we consider a slightly different kind of test; given that the first election under a district plan produces a value of the efficiency gap above or below a given threshold, how likely is it that the *average* value of the efficiency gap produced over the life of the plan lies on the same side of zero as that of the first election? Recall that the sign of the efficiency gap speaks to the corresponding direction of partisan advantage ($EG < 0$ is consistent with Republican advantage; conversely for $EG > 0$). We expect that this will be a less strenuous test than asking if *any* EG has an opposite sign to the first EG observed under a district plan.

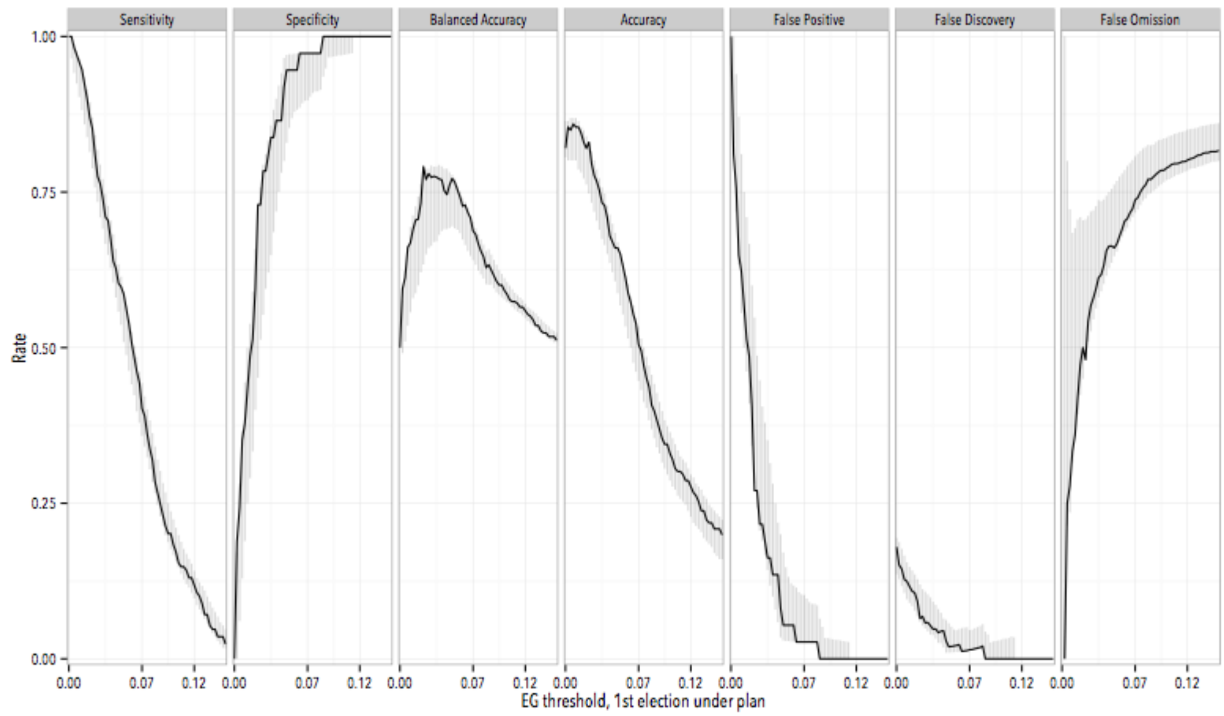


Figure 4: Prognostic performance measures, first efficiency gap under a district plan more extreme than threshold (horizontal axis) as a predictor of whether the average efficiency gap recorded under the district plan has the same sign as the first efficiency gap. Vertical lines indicate 95% confidence intervals. Analysis spans all state legislative elections and district plans as per my initial report, 1972-2014.

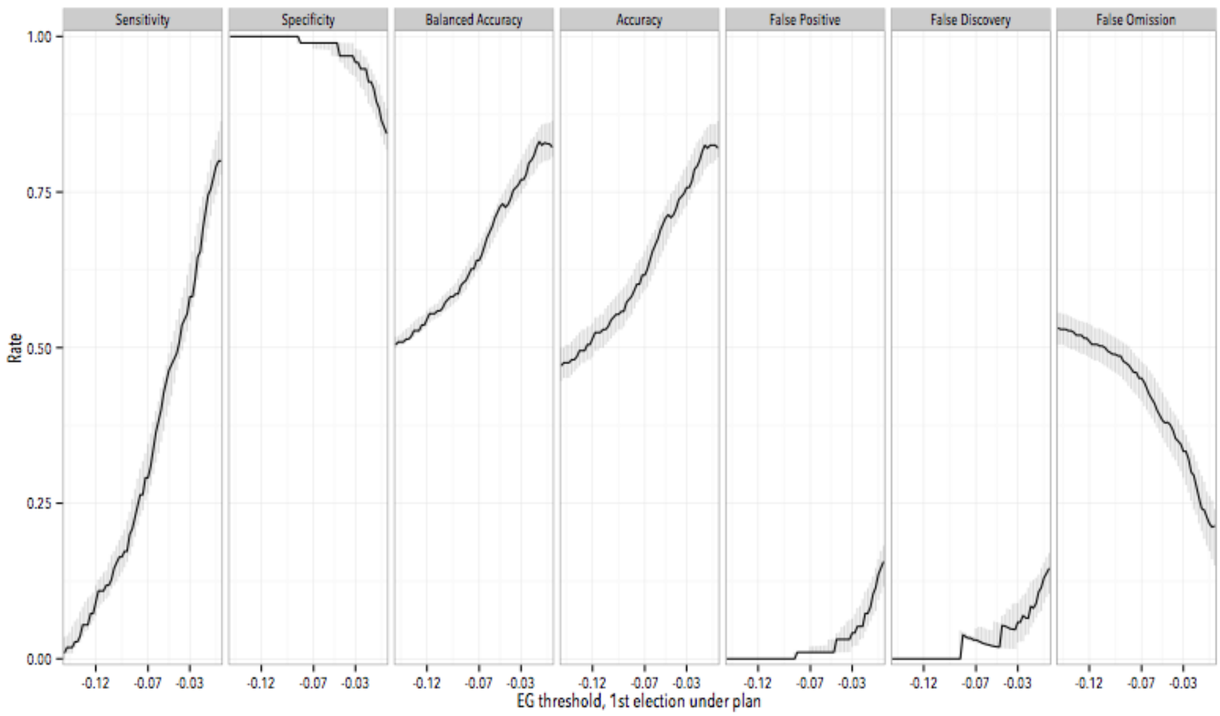


Figure 5: Prognostic performance measures, first efficiency gap under a district plan more extreme than threshold (horizontal axis) as a predictor of whether the average efficiency gap recorded under the district plan has the same sign as the first efficiency gap. Vertical lines indicate 95% confidence intervals. Analysis examines negative, first-election threshold values of the efficiency gap, consistent with Republican advantage.

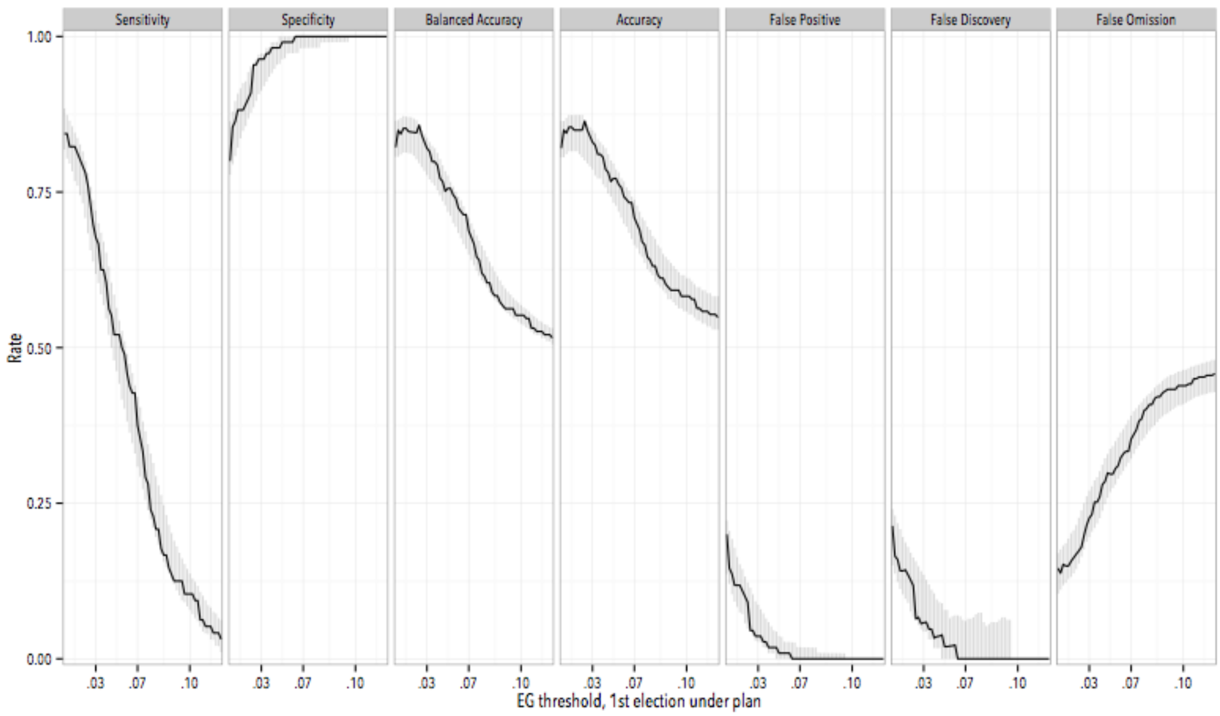


Figure 6: Prognostic performance measures, first efficiency gap under a district plan more extreme than threshold (horizontal axis) as a predictor of whether the average efficiency gap recorded under the district plan has the same sign as the first efficiency gap. Vertical lines indicate 95% confidence intervals. Analysis examines positive, first-election threshold values of the efficiency gap, consistent with Democratic advantage.

Figures 4, 5 and 6 show the prognostic performance of the first-election *EG* with respect to the sign of the corresponding plan's average *EG*, looking at the absolute value of the first-election *EG* (Figure 4), negative first-election efficiency gaps (Figure 5) and positive first-election efficiency gaps (Figure 6). The first thing to observe is the generally superior prognostic performance when it comes to forecasting the sign of the *plan-average* efficiency gap, relative to the prognostic performance with respect to *all* of the plan's efficiency gaps having the same sign. As anticipated, the former is better predicted by the plan's first-election efficiency gap than the latter. Second, the accuracy-versus-caution tradeoff noted earlier is also apparent. The proposed threshold of +/- 0.07 trades away accuracy for very low false positive and false discovery rates, below 5%, at the cost of higher false omission rates, a pattern we observed earlier. Finally, note that at the proposed threshold of +/- 0.07, almost one-half of all plans with a negative (pro-Republican) average *EG* would *not* be candidates for scrutiny (right-hand panel of Figure 5); about one-third of plans with a positive (pro-Democratic) average *EG* also would not trigger the threshold for scrutiny.

4 Relationship between the first-election efficiency gap and the plan-average efficiency gap

I next present analysis on a related issue, the relationship between the magnitudes of the *first* efficiency gap observed under a plan and the *average* efficiency gap we observe over the life of the plan. Does a larger or smaller first-election efficiency gap portend anything for the average value of the efficiency gap generated over the life of a district plan?

Clearly the first value of the efficiency gap and the plan-average efficiency gap are related; the former contributes to the calculation of the latter, and after the first election under a district plan we observe at most four more elections under the plan (given elections every two years in most states and redistricting once a decade). Accordingly we expect a positive correlation between the two quantities. The interesting empirical question—and one with considerable substantive implications for the issue at hand—is *how strong* the relationship is between the first-election efficiency gap and the corresponding plan-average efficiency gap. This speaks to the reliability of the first-election *EG* measure as a predictor of *EG* over the life of the plan.

Figure 7 shows the relationship between the first-election *EG* and the average *EG* observed over the entire plan. Note that we restrict this analysis to plans with at least three elections, so that the first election does not unduly contribute to the calculation of the average; this restriction has the consequence of omitting elections from the most recent round of redistricting after the 2010 Census, which have contributed at most two elections. The black diagonal line on the graph is a 45-degree line: if the relationship between first-election *EG* and plan-average *EG* were perfect, the data would all lie on this line. Instead we see a classic "regression-to-the-mean" pattern, with a positive regression slope of less than one (as indeed we should, given that the first-election *EG* on the horizontal axis contributes to the average plotted on the vertical axis). But the relationship here is especially strong. The variation in plan-average efficiency gaps explained by this regression is quite large, about 73%; after taking into account the uncertainty in the *EG* scores (stemming from the imputation procedures used for uncontested districts; see my initial report) a 95% confidence interval on the variance explained measure ranges from 67% to 74% (the uncertainty has the consequence of tending to make the regression fit slightly less well). That is, even given the uncertainty that accompanies *EG* measures due to uncontestedness, the relationship between first-election *EG* and plan-average *EG* is quite strong.

In particular, at the threshold values of +/- 0.07 there is very little doubt as to the plan-average value of the efficiency gap. The historical relationship between first-election *EG* and plan-average *EG* shown in Figure 7 indicates that a first-election *EG* of -.07 is typically associated with a plan-average *EG* of about -0.053 (95% CI -0.111 to 0.004); the probability that the resulting, expected plan-average *EG* is negative is 96.5%. Conditional on a first-election *EG* of .07 we typically see a plan-average *EG* of about 0.037 (95% CI -0.021 to 0.093); the probability that the resulting, expected plan-average *EG* is positive is 89.8%. This constitutes additional, powerful evidence that (a) first-election *EG* estimates are predictive with respect to the *EG* estimates that will be observed over the life of the plan; and (b) the threshold values of +/- 0.07 are conservative, generating high-confidence predictions as to the behavior of the district plan in successive elections.

In the particular case of Wisconsin in 2012—the first election under the plan in question—I estimated the efficiency gap to be -0.133 (95% CI -0.146 to -0.121). The analysis of historical data discussed above—and graphed in Figure 7—indicates that the plan-average *EG* for this plan will be -0.095 (95% CI -0.152 to -0.032)[1], a quite large value by historical standards, placing the current Wisconsin district plan among the five to ten most disadvantageous district plans for Democrats in the data available for analysis. The probability that the Wisconsin plan—if left undisturbed—will turn out to have a positive, pro-Democratic, average efficiency gap is for all practical purposes zero (less than 0.1%).

---

1 It is also worth stressing that the confidence interval is computed so as to take into account uncertainty from all known sources: in the underlying efficiency gap scores themselves, the fact that the 2012 *EG* scores for Wisconsin are large by historical standards, and in the regression relationship between first-election *EG* and plan-average *EG*.

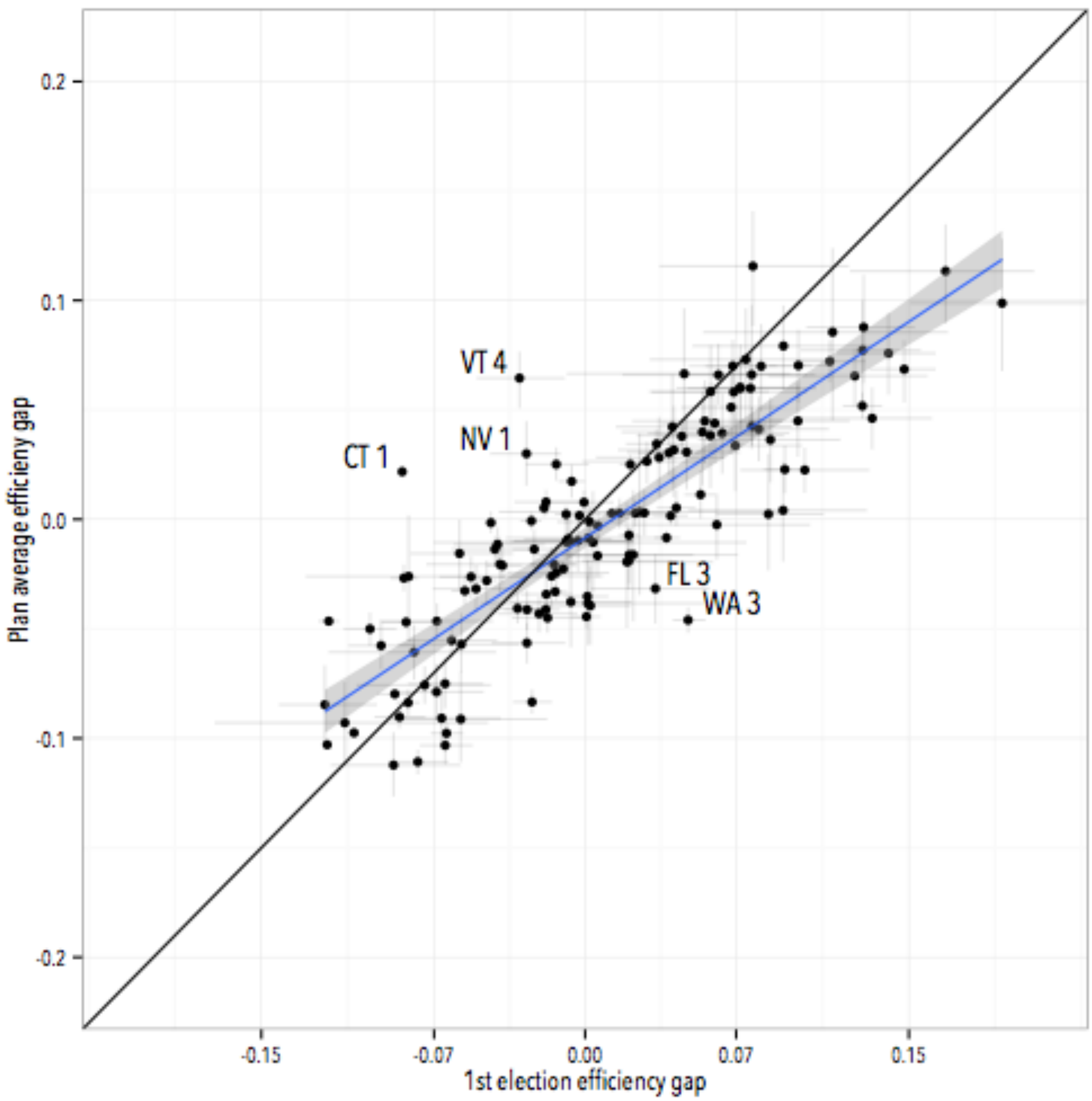


Figure 7: Scatterplot of first-election efficiency gap scores (horizontal axis) and plan-average efficiency gap scores (vertical axis). The diagonal black line is a 45-degree line; the data would lie on this line if first-election efficiency gaps coincided with plan-average efficiency gaps. The solid blue line is a linear regression with slope .64 (95% CI 0.57 to 0.72); the shaded region around the blue line is a 95% confidence interval for the regression line. Vertical and horizontal lines extending from each data point cover 95% confidence intervals in either direction, summarizing the uncertainty in both first-election *EG* and plan-average *EG*, stemming from imputations for uncontested districts. Outliers are labeled (state, plan). Analysis restricted to plans with at least three elections (1972-2010), omitting plans adopted after the 2010 Census. The first-election *EG* for the current Wisconsin plan is -0.133 (95% CI -0.146 to -0.121).

5 Party control as an explanation for change in the efficiency gap

Both Trende and Goedert point out that, on average, state house plans have exhibited pro-Republican efficiency gaps in recent years (Trende, paragraphs 129-30; Goedert p. 19). They then argue that this pro-Republican mean is attributable to a natural pro-Republican political geography in many states. However, as I found in my initial report, the *overall* efficiency gap average, over the entire 1972-2014 period, is very close to zero (Jackman Report, p. 35, 45, 57). There is thus no sign of a natural pro-Republican advantage in the dataset as a whole, nor any evidence (despite Trende and Goedert's unsupported assertions to the contrary) that states' political geography is changing in ways that favor Republicans.

In fact, the one historical change that *is* undeniable is the trend toward unified Republican control over redistricting. As Figure 8 displays, only about 10% of all state house plans were designed by Republicans in full control of the state government in the 1990s, compared to about 30% by Democrats in full control and about 60% by another institution (divided government, a commission, or a court). But in the 2000s, Republicans were fully responsible for slightly *more* plans than were Democrats (about 20% versus about 15%). And in the 2010s, the partisan gap jumped again, to about *40%* of plans designed entirely by Republicans, versus less than 20% designed entirely by Democrats.

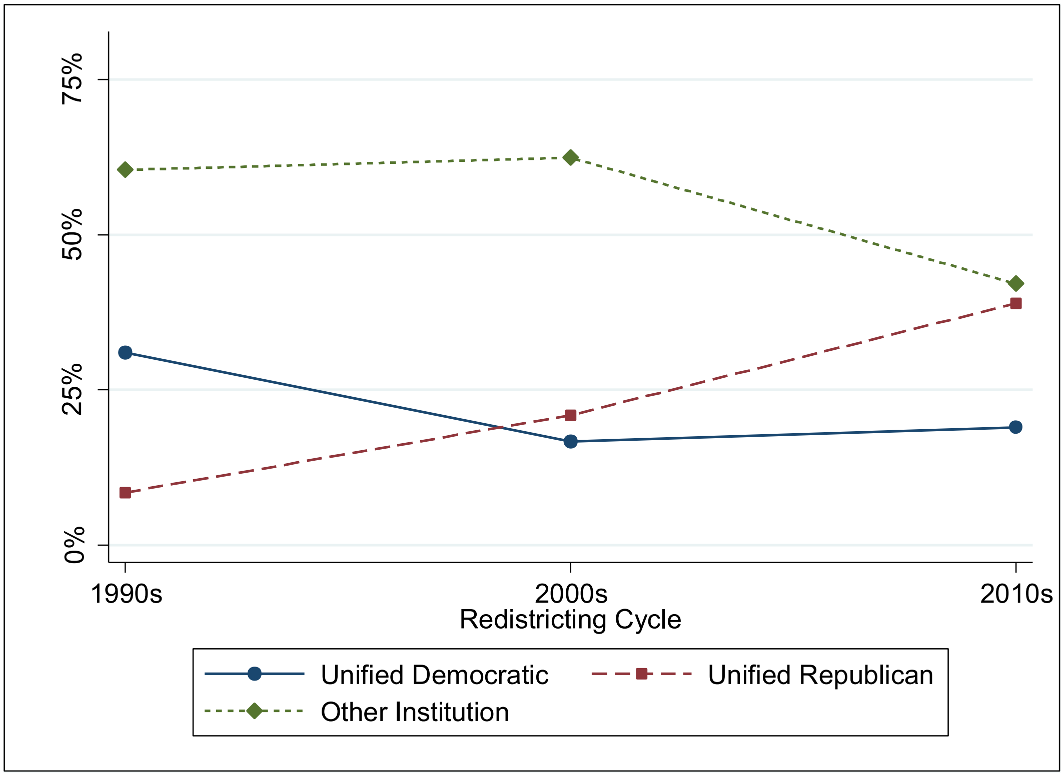


Figure 8: Share of all state house plans, by cycle, designed by Democrats in unified control of state government, by Republicans in unified control of state government, or by another institution (divided state government, commission, or court).

To determine the impact of this change in party control on the change in the efficiency gap over the last generation, I carry out three regressions, one for the 1990 redistricting cycle, one for the 2000 cycle, and one for the 2010 cycle. In each case, state house plans' efficiency gaps are the dependent variable, and unified Democratic control over redistricting and unified Republican control over redistricting are the independent variables. (The omitted category is any other institution responsible for redistricting, such as divided government, a court, or a commission.) Figure 9 then displays the *actual* average efficiency gap for each cycle, as well as the *predicted* average efficiency gap if the distribution of party control over redistricting had remained unchanged since the 1990s.

As is evident from the chart, state house plans' average efficiency gap in the 2000 cycle would have been substantially less pro-Republican (by about 0.5 percentage points) had Republicans not gained control of more state governments in this cycle relative to the 1990s. And in the current cycle, *all* of the efficiency gap's movement in a Republican direction would have been erased had the distribution of party control over redistricting not changed since the 1990s. That is, if the same distribution of party control had existed in this cycle as in the 1990s, state house plans' average efficiency gap would have been very close to zero, not over 3% in a Republican direction. Accordingly, it is the change in party control that appears to account for essentially all of the pro-Republican trend in the efficiency gap over the past two decades—and not, as claimed by Trende and Goedert, a dramatic alteration of the country's political geography.

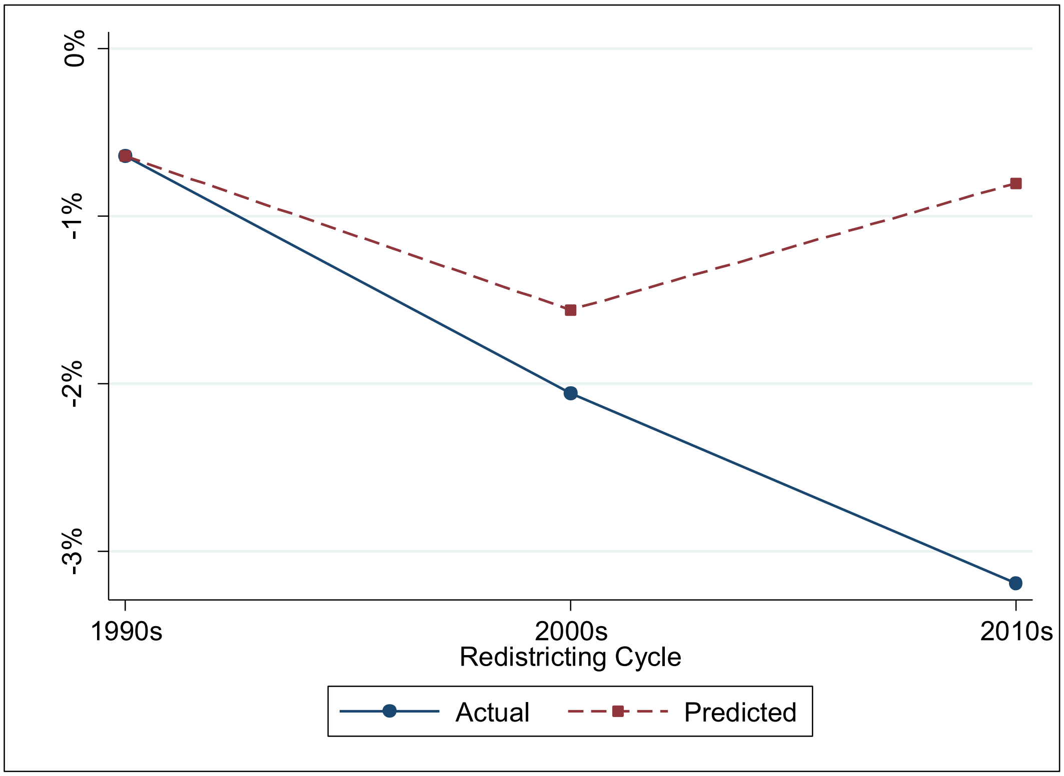


Figure 9: Actual and predicted values of state house plans' average efficiency gaps by cycle. Predicted values calculated assuming that the 1990s distribution of party control over redistricting remained constant in subsequent cycles.

6 Response to the Chen and Rodden map simulations

Both Trende and Goedert cite a recent article by Chen and Rodden (2013) that purports to find, based on simulations of hypothetical district maps, that random redistricting would benefit Republicans because of their more efficient spatial allocation (Trende, paragraphs 89, 126; Goedert, pp. 13, 18, 21). While I respect Chen and Rodden's contribution, there are several issues with their work that make it inapplicable here.

First, Chen and Rodden do not even attempt to simulate *lawful* plans. Rather, they simulate plans "using only the traditional districting criteria of equal apportionment and geographic contiguity and compactness" (Chen and Rodden, 248). They do not take into account Section 2 of the Voting Rights Act, which often requires majority-minority districts to be constructed. They also do not take into account Section 5 of the VRA, which until 2013 meant that existing majority-minority districts could not be eliminated in certain states. And they do not take into account state-level criteria such as respect for political subdivisions and respect for communities of interest, which are in effect in a majority of states (NCSL 2010, pp. 125-27).

Second, Chen and Rodden only use *presidential* election results in their analysis, but these outcomes may diverge from *state legislative* election results due to voter roll-off as well as voter preferences that vary by election level. As Stephanopoulos and McGhee have noted, "If certain voters consistently support Republicans at the presidential level and Democrats at the legislative level, then presidential data may produce more pro-Republican estimates than legislative data" (Stephanopoulos & McGhee, 870). In fact, this is exactly what seems to be occurring; at the congressional level, efficiency gaps are about 6% more Republican when they are calculating using presidential data than when they are computed on the basis of congressional election results.

Third, Chen and Rodden's simulated maps do not constitute a representative sample of the entire plan solution space. Their simulation algorithm has "no theoretical justification," is "best described as ad-hoc," and is not "designed to yield a representative sample of redistricting plans" (Fifield et al. 2015, pp. 2-3; Altman & McDonald 2010, p. 108). The explanation for this lack of representativeness is highly technical and involves the details of the particular simulation approach adopted by Chen and Rodden. But its implication is clear: that no conclusions can yet be drawn about the partisan consequences of randomly drawn maps.

Lastly, Chen and Rodden's results are directly contradicted by Fryer and Holden, who also simulated contiguous, compact, and equipopulous districts for multiple states. Unlike Chen and Rodden, Fryer and Holden found that, "[u]nder maximally compact districting, measures of Bias are slightly *smaller* in all states except [one]" (Fryer & Holden 2011, p. 514). Fryer and Holden also found that "[i]n terms of responsiveness . . . there are large and statistically significant" *increases* in all states, sometimes on the order of a fivefold rise (p. 514). Their analysis thus leads to the opposite inference from Chen and Rodden's: that randomly drawn contiguous and compact districts favor *neither* party and substantially boost electoral responsiveness.

7 Trende's analysis of particular plans

Trende devotes a large portion of his report (paragraphs 106-31) to analyzing the efficiency gaps of particular state legislative and congressional plans. He first examines a set of seventeen state legislative plans that had efficiency gaps favoring the same party over their entire lifespans, arguing that not all of these plans were gerrymanders (paragraphs 106-14). He then cites a series of congressional plans, some of which he claims had large efficiency gaps despite not being gerrymanders, and others of which allegedly had small efficiency gaps despite being gerrymanders (paragraphs 115-24). All of this analysis is riddled with conceptual and methodological errors that, in my judgment, renders it unreliable and unhelpful to the court.

Beginning with the set of seventeen state legislative plans that had efficiency gaps of the same sign throughout their lifespans, Trende asserts that they "would be included in the definition of a gerrymander," and are a "list of gerrymandered states" (paragraphs 109-10). But neither plaintiffs nor I argue that these plans should have been held unconstitutional. That is, neither plaintiffs nor I argue that these plans were designed with partisan intent (the first element of plaintiffs' proposed test), that their initial efficiency gaps exceeded a reasonable threshold (the second element), or that their efficiency gaps could have been avoided (the third element). To the contrary, I simply included these plans in my report to illuminate historical cases in which the efficiency gap's direction did not change over the course of a decade. I never stated or implied that these plans should have been deemed unlawful.

However, if we focus on the plans among the seventeen that likely *would* have failed plaintiffs' proposed test (at least the first two elements), we see that both the test and the efficiency gap perform exceptionally well. Five of the seventeen plans featured unified control by a single party over redistricting (from which, like Goedert (2014) and Goedert (2015), we can infer partisan intent) as well as an initial efficiency gap above 7% (the threshold I recommended in my initial report): Florida in the 1970s, Florida in the 2000s, Michigan in the 2000s, New York in the 1970s, and Ohio in the 2000s. Assuming that these plans' large efficiency gaps were avoidable (a granular inquiry that cannot be carried out here), it would have been quite reasonable for all of these maps to attract heightened judicial scrutiny. In particular:

- Florida's plan in the 1970s was designed exclusively by Democrats, opened with a 9.9% pro-Democratic efficiency gap, averaged a 7.0% pro-Democratic efficiency gap over its lifespan, and never once favored Republicans.

- Florida's plan in the 2000s was designed exclusively by Republicans, opened with a 8.9% pro-Republican efficiency gap, averaged a 11.2% pro-Republican efficiency gap over its lifespan, and never once favored Democrats.
- Michigan's plan in the 2000s was designed exclusively by Republicans, opened with a 12.0% pro-Republican efficiency gap, averaged a 10.3% pro-Republican efficiency gap over its lifespan, and never once favored Democrats.
- New York's plan in the 1970s was designed exclusively by Republicans, opened with a 10.7% pro-Republican efficiency gap, averaged a 9.7% pro-Republican efficiency gap over its lifespan, and never once favored Democrats.
- Ohio's plan in the 2000s was designed exclusively by Republicans, opened with a 8.6% pro-Republican efficiency gap, averaged a 9.0% pro-Republican efficiency gap over its lifespan, and never once favored Democrats.

Accordingly, we see that if my report's set of seventeen plans is analyzed properly, the opposite conclusion emerges from the one advocated by Trende. Only a subset of the seventeen plans likely would have failed plaintiffs' proposed test. But *every member* of this subset turns out to have been an exceptionally severe and durable gerrymander, featuring a very large and consistent efficiency gap over its lifespan. These are *precisely* the historical cases in which judicial intervention may have been advisable.

After commenting on these seventeen state legislative plans, Trende discusses a series of *congressional* plans, all from the 2000 and 2010 redistricting cycles. These congressional plans are entirely irrelevant to this case, which deals only with state legislative redistricting. Neither in their complaint nor in their subsequent filings do plaintiffs ever argue that their approach should be applied to congressional plans. And neither Mayer nor I provide any empirical analysis of congressional plans. In my initial report, in particular, I examined state legislative plans from 1972 to the present, but no congressional plans at all.

This state legislative focus has two explanations. First, and more importantly, each congressional delegation is *not* a legislative chamber in its own right, but rather a portion (often a very small portion) of the U.S. House of Representatives. Methods applicable to entire chambers cannot simply be transferred wholesale to delegations that make up only fractions of Congress. Second, most congressional delegations have many fewer seats than most state houses. The efficiency gap becomes lumpier when there are fewer seats, because each seat accounts for a larger proportion of the seat total, and the efficiency gap thus shifts more as each seat changes hands. This lumpiness is entirely avoided when state legislative plans, which typically have dozens or even hundreds of districts, are at issue.

For these reasons, Stephanopoulos and McGhee make two adjustments when analyzing congressional plans in their work on the efficiency gap. First, they convert the efficiency gap from percentage points to *seats* by multiplying the raw efficiency gap by each state's number of congressional districts. As they explain their method, "What matters in congressional plans is their impact on the total number of *seats* held by each party at the national level. Conversely, state houses are self-contained bodies of varying sizes, for which *seat shares* reveal the scale of parties' advantages and enable temporal and spatial comparability" (Stephanopoulos & McGhee, 869). Second, they only calculate efficiency gaps for states with at least eight congressional districts. Efficiency gaps are lumpier for states with fewer than eight districts, and additionally, congressional "redistricting in smaller states has only a minor influence on the national balance of power" (Stephanopoulos & McGhee, 868).

In his report, Trende fails to make either of these necessary adjustments when examining congressional plans. That is, he does not convert the efficiency gap from percentage points to seats, and he calculates the efficiency gap for small congressional delegations with fewer than eight seats. There is no authority in the literature for his methodological choices, and he is unable to cite any. And his flawed methods have serious substantive consequences that render his results entirely untrustworthy.

Take Trende's failure to convert the efficiency gap from percentage points to House seats. He claims that Alabama's congressional plan had an efficiency gap of -12.5% in 2002, that Arizona's congressional plan had an efficiency gap of 16% in 2012, that Colorado's congressional plan had an efficiency gap of -9% in 2002 and -10% in 2012, that Illinois's congressional plan had an efficiency gap of -9% in 2002, and that Iowa's congressional plan had an efficiency gap of -20% in 2002—all above my suggested 7% threshold for state legislative plans (paragraphs 115-16, 118-19, 121-22). But when converted to seats, *all* of these efficiency gaps become quite small, lower in all cases than the two-seat threshold proposed in the literature for congressional plans (Stephanopoulos & McGhee, 887-88). Specifically, using Trende's own calculations—which, as I discuss below, are incorrect in any event—Alabama had an efficiency gap of -0.9 seats in 2002, Arizona had an efficiency gap of 1.4 seats in 2012, Colorado had an efficiency gap of -0.6 seats in 2002 and -0.7 seats in 2012, Illinois had an efficiency gap of -1.7 seats in 2002, and Iowa had an efficiency gap of -1.0 seats in 2002. *None* of these scores are high enough to rise to presumptive unlawfulness under the literature's suggested two-seat threshold, meaning that we come to exactly the *opposite* conclusion as Trende after making the necessary adjustment.

Next take Trende's consideration of Alabama's congressional plan in 2002 (which had seven districts), Iowa's congressional plan in 2002 (five districts), and Colorado's congressional plans in 2002 and 2012 (seven districts each) (paragraphs 115-16, 119, 122). All four of these plans have fewer than eight districts, and so, based on the literature, should not be included in any efficiency gap analysis because of the measure's lumpiness when applied to so few seats. Trende nowhere acknowledges this limitation, and indeed appears unaware of its existence.

Moreover, Trende's study of congressional plans is marred by two further flaws, one conceptual and the other methodological. The conceptual defect is that, as in his earlier discussion of state legislative plans, he assumes that a large efficiency gap is all that is necessary to render a plan unconstitutional. He writes that efficiency gaps of -12.5%, -9%, -9%, -20%, and 16% "would invite court scrutiny as a Republican gerrymander" or "would invite court scrutiny as a Democratic gerrymander" (paragraphs 115, 116, 118, 119, 121, 122). But again, this is not plaintiffs' proposed test. A large efficiency gap is only a single prong of the test, and does not result in a verdict of unconstitutionality unless it is paired with a finding of partisan intent *and* a finding that it could have been avoided. Trende entirely overlooks these other elements.

The methodological defect is that whenever there were uncontested congressional races, Trende simply *substituted* presidential election results for the missing congressional results. As he put it in his deposition, he "used presidential results" and "imputed those results to the congressional races" whenever the races were uncontested (Trende deposition, p. 83). This is an exceptionally crude method that is guaranteed to produce errors, both because there is voter roll-off from the presidential to the congressional level and because voters may have different presidential and congressional preferences. Of course, presidential results can be used as the *inputs* to a regression model that *predicts* the outcomes of uncontested congressional races. Indeed, this is the preferred approach in the literature, and the approach I employed in my initial report. But presidential results cannot simply be plugged in without any adjustment, and no competent social scientist would have done so.

Accordingly, in my judgment, Trende's examination of particular state legislative and congressional plans is unreliable and entitled to no weight by the court. The state legislative analysis ignores the actual elements of plaintiffs' proposed test, and would have led to the opposite conclusion if these elements had been taken into account. Likewise, the congressional analysis ignores the test's prongs, fails to convert the efficiency gap from percentage points to seats, improperly considers states with small House delegations, improperly substitutes presidential election results whenever congressional results are missing—and deals with federal elections that simply are not part of this case.

Dated December 21, 2015

/s/ Simon Jackman

Simon Jackman, PhD

Department of Political Science

Stanford University

References

Altman, Micah and McDonald, Michael 2010, "The Promise and Perils of Computers in Redistricting." 2010. *Duke Journal of Constitutional Law & Public Policy* 5:69-111.

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

Fifield, Benjamin, Higgins, Michael, Imai, Kosuke. Tarr, Alexander. 2015. "A New Automated Redistricting Simulator Using Markov Chain Monte Carlo." *Working Paper*, available at http://imai.princeton.edu/research/files/redist.pdf.

Goedert, Nicholas. 2014. "Gerrymandering or Geography?: How Democrats Won the Popular Vote but Lost the Congress in 2012." *Research & Politics* 1(1): 2053168014528683.

Goedert, Nicholas. 2015. "The Case of the Disappearing Bias: A 2014 Update to the "Gerrymandering or Geography Debate." Forthcoming in *Research and Politics*, November 2015.

National Conference of State Legislators. September 29, 2009. *Redistricting Law 2010*.

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

Expert Report of Professor Nicholas Goedert in *Whitford v. Nichol*. December 2, 2015. "Use of Efficiency Gap in Analyzing Partisan Gerrymandering, Report for State of Wisconsin, *Whitford v. Nichol*.

Expert Report of Professor Simon Jackman in *Whitford v. Nichol*. July 7, 2015. "Assessing the Current Wisconsin State Legislative Districting Plan."

Expert Report of Professor Ken Mayer in *Whitford v. Nichol*. July 3, 2015. "Analysis of the Efficiency Gaps of Wisconsin's Current Legislative District Plan and Plaintiffs' Demonstration Plan"

Declaration of Sean P. Trende in *Whitford v. Nichol*. December 2, 2015.

Deposition of Sean P. Trende in *Whitford v Nichol*. December 14, 2015.

Case

Vieth v. Jubilerer, 541 U.S. 267 (2004).

Legislative Studies Quarterly

Comparative Legislative Research Center
The University of Iowa

ERIC MCGHEE
Public Policy Institute of California

Measuring Partisan Bias in Single-Member District Electoral Systems

In recent decades, the literature has coalesced around either symmetry or responsiveness as measures of partisan bias in single-member district systems. I argue neither accurately captures the traditional idea of an "efficient" gerrymander, where one party claims more seats without more votes. I suggest a better measure of efficiency and then use this new measure to reconsider a classic study of partisan gerrymandering. Contrary to the original study findings, I show that the effects of party control on bias are small and decay rapidly, suggesting that redistricting is at best a blunt tool for promoting partisan interests.

The literature on single-member district (SMD) elections has long been concerned with the problem of partisan efficiency. Only a bare plurality of votes is required to win any seat, so additional votes are "wasted," and a clever party can manipulate this fact to its advantage.[1] The more a party manipulates wasted votes, the more seats it gains at the opposing party's expense. Moreover, a party can gain the benefits of efficiency even if no deliberate manipulation is involved. Partisan disparities in wasted votes can occur for any number of reasons, including minority-representation requirements and the distribution of partisan supporters in geographic space through demographic and socioeconomic change.

In recent decades, the literature has coalesced around the twin concepts of *symmetry* and *responsiveness* as the best means of measuring this sort of partisan advantage. Symmetry hinges on uniform treatment—a system is asymmetrical if one party receives a larger share of seats than the other party for the same share of votes—while responsiveness captures the overall competitiveness of the system, as measured by the number of seats that change hands for a given shift in the aggregate vote. These measures are potentially useful because they are easy to calculate with simple election returns, and they promise consistency of meaning in a way that simplifies comparisons across time and space.

DOI: 10.1111/lsq.12033

In this article, I suggest that symmetry and responsiveness, while certainly of value for some purposes and concepts, are not logically related to the seat gains from efficiency that concern many scholars and analysts. I suggest a more direct measure of efficiency and use data from state legislative elections to demonstrate that it is a more appropriate means of representing this particular concept. The new measure is also more flexible. Analysts can use it to evaluate electoral systems where one party typically dominates others—a condition advocates of symmetry studiously avoid—and it actually converges on symmetry as a party's aggregate vote share approaches 0.5. Thus, the new measure can be considered a more general extension of the symmetry measure itself—one that preserves some of symmetry's normative meaning for uncompetitive SMD systems.

The value of this new measure is largely unrelated to the specific strategy a party takes to maximize its seat share. Several important works have suggested caveats to the seat-maximization strategy to account for the uncertainty of partisan support and other factors (Friedman and Holden 2008; Owen and Grofman 1988). But the end goal of all of these strategies is to maximize expected seats given expected votes, so the measure of efficiency I offer here will still be of considerable utility.

I then use this new measure of partisan bias to reconsider our understanding of redistricting in the United States. In a classic study, Gelman and King (1994) argued that party control of redistricting has substantial effects on bias that endure through the redistricting cycle but that the disruptive nonpartisan effects of redistricting are strong enough to make the process normatively beneficial all the same. I show that these findings are artifacts of symmetry; when the new measure is used instead, redistricting has only small and transient effects on bias and is hardly more disruptive than the natural back and forth of electoral competition. This casts doubt on the utility of redistricting as a means of achieving an efficient seat-share advantage and thus on part of the rationale of judicial intervention in partisan gerrymandering cases.

Efficiency, Symmetry, and Responsiveness

Scholarship on redistricting has long understood that a particular plan can be biased toward one party based on the efficient distribution of partisan voters across districts. The candidate who receives one vote more than 50% (in a two-party contest) is the winner, so any vote beyond that threshold is wasted in the sense that it does not contribute to victory. All the votes cast for a losing candidate are also wasted because they do not support a winning effort. In either case, a party could avoid wasting the votes by moving them to a different district where they could be part of a minimally winning coalition. The party with greater efficiency—fewer wasted votes—will generally win more seats by smaller margins than the opposition. Along these lines, experts frequently speak of "cracking"—dividing a partisan community between two or more districts when they could be a majority in a single district—and "packing"—filling a district with more supporters of the opposing party than needed to win the seat. Both aim to make supporters of the disadvantaged party either a slim minority in each district or the overwhelming majority. This is the essence of partisan efficiency.

Some version of efficiency is typically the core concept of interest in the literature on redistricting. Most of this literature focuses on deliberate efforts to advantage one party or the other through gerrymandering (e.g., Cain 1985; Engstrom 2006; Owen and Grofman 1988). It is easy to see why the effort to maximize seat share this way might be troubling to normative democratic theory, since it offers a party a means of increasing its margin of control over policy without winning more votes from the public. However, the importance of efficiency need not be limited to partisan gerrymandering. A substantial literature ponders the unintended partisan consequences of a variety of phenomena with no overt partisan origin or objective. These include redistricting to boost minority representation (Brace, Grofman, and Handley 1987; Schotts 2001), differential rates of turnout across districts (Campbell 1996; Grofman, Koetzle, and Brunell 1997), and simple differences in residential patterns for supporters of each party (Chen and Rodden 2011). In all these examples, efficiency is believed to reflect a more fundamental partisan disparity. It is a question of lost potential: a biased system wastes more votes on one party and so restricts the number of seats that party might otherwise win.

While efficiency is the concept of greatest interest, the most common measure of partisan advantage in the literature is *symmetry*. Symmetry is achieved when each party receives the same share of seats for identical shares of votes. The idea is easiest to understand at an aggregate vote share of 0.5—a party that receives half the vote ought to receive half the seats—but a similar logic can apply across the "seats-votes curve" that traces out how seat shares change as vote shares rise and fall. For example, if a party receives a vote share of 0.55 and a seat share of 0.60, the opposing party should also expect to receive a seat share of 0.60 if it were to receive a vote share of 0.55. An unbiased system means that for V share of the votes a party should receive S share of the seats, and this should be true for all parties and vote percentages (Niemi and Deegan 1978).

Symmetry has never been strictly compared to efficiency. In place of such an assessment, the literature has taken one of three different approaches. The first simply defines partisan bias as asymmetry in the seats-votes curve. Some of the earliest expositions of partisan symmetry were most likely to take this approach. For example, Soper and Rydon (1958) equate the median vote (i.e., 50%) with what they call the "effective" vote—a term that is now typically reserved specifically for partisan efficiency. Brookes (1959) also starts the derivations of the measures he proposes under the presumption that symmetry is the concept of interest. And while this approach started early, there are more recent examples of authors who adopt symmetry without much further explanation (Altman 2002; Gelman and King 1994b; Jackman 1994; Niemi and Deegan 1978). The implication is that the measure is facially valid.

The second approach recognizes the existence of partisan efficiency, but either argues that symmetry is a superior measure or ceases to mention efficiency after some cursory introduction. The principal example here is Grofman and King, who recognize that "Journalistic accounts of partisan gerrymandering often describe it as a process of packing one's opponents into as few districts as possible and seeking to win the remaining districts by the barest of margins" (2007, 13), but who then dismiss this notion as flawed and define a partisan gerrymander in terms other than seat gain.

The most common perspective simply equates the two measures, either implicitly or explicitly. This approach dates back to Tufte's (1973) seminal analysis of the relationship between votes and seats, which at first seems to ignore efficiency and advocate symmetry, but then emphasizes differential constituency size as a possible explanation for the shifts in symmetry he finds. (Constituency size is a matter of efficiency because the party that wins districts with smaller numbers of voters claims more seats with fewer votes.) Many important studies since then have also used symmetry as a measure but described some form of efficiency as the concept of interest (e.g., Campagna and Grofman 1990; Cox and Katz 2002; Engstrom 2006; Erikson 1972; Gilligan and Matsusaka 1999; Grofman, Koetzle, and Brunell 1997; Kastellec, Gelman, and Chandler 2008).

While symmetry is the most common measure of bias, a few scholars have suggested that *responsiveness* can also serve as an effective measure of an efficient seat-maximizing gerrymander (Cox and Katz 1999). Responsiveness records how many seats change hands as vote shares rise and fall, so it can be thought of as the slope of the seats-votes curve across a range of vote shares. Usually this range is between 10 and 30 points wide and is centered upon either the actual election result or a vote share of 0.5. Responsiveness is greater than 1 when seat share rises faster than vote share and less than 1 when the opposite is true. Because responsiveness will be higher when more seats are clustered close to the 0.5 threshold, it will be higher when there are more competitive seats. A party drawing an aggressive seat-maximizing gerrymander will strive to win its seats by as narrow a margin as possible, so such a gerrymander might plausibly increase the overall responsiveness of the system. Some scholars have also suggested measuring party advantage as the deviation from an ideal level of responsiveness, usually defined as proportionality (i.e., responsiveness = 1.0), but such approaches have never caught on in the literature on SMD systems.[2]

Two aspects of symmetry and responsiveness deserve special attention. First, they are almost always counterfactual calculations. For any given election, there is only one actual seats-votes pair. To identify the seats-votes curve—and with it, partisan symmetry and responsiveness—the vote share in every seat must be shifted a uniform amount. For each new hypothetical vote share, the new seat share is recorded based on the number of seats that changed hands as a result of the exercise. Since each district's vote share is shifted a uniform amount, the hypothetical assumes that the order of the district outcomes would not change if overall party performance were suddenly equal (Grofman and King 2007).[3] There are some partial exceptions to this counterfactual—in particular, the work of Gelman and King (1994b) allows for a certain amount of randomness around this vote shift, in the understanding that no counterfactual's consequences can ever be known with certainty—but the core logic of the measure is the same.

Second, both measures reflect the shape of the seats-votes curve and not any particular point on it, so they are relatively immune to shifts in party performance. For instance, a partisan tide that moves all seats a uniform amount alters only where the actual election result falls on the curve; it cannot change the shape of the curve itself, and so it cannot have much if any impact on symmetry or responsiveness. Even many shifts in relative district position—which do change the curve—will not affect symmetry or responsiveness unless they fall within range of the counterfactual. By necessity, then, both symmetry and responsiveness presume bias is a strongly enduring feature of a redistricting plan, something that should not and usually will not respond to the dynamics of ordinary partisan competition.

There is a portion of the literature that largely or entirely disregards symmetry and responsiveness and adopts a more direct measure of efficiency instead. Though the exact method varies somewhat, these studies all compare outcomes before the redistricting to what they would have been for the same election under the new plan, and the intent is always to isolate the gain in seat share induced by the redistricting alone (Born 1985; Cain 1985; Engstrom 2006; Kousser 1996). This approach is probably closest to the concept of efficiency, but it also imposes significant costs on the analyst. It only applies to before-and-after comparisons, and it is most effective and valid where it is possible to aggregate political data from before the redistricting into the new district lines. In many circumstances, either the relevant data are unavailable or there is no redistricting intervention to evaluate in the first place. Thus, it remains attractive to have a measure that is meaningful across a broad range of circumstances and which can be calculated with easily available data. Both symmetry and responsiveness strive to fill that need.

In short, much of the literature seeks to measure efficiency while assuming that either symmetry or responsiveness is an adequate measure of the concept. There is, of course, no reason why an analyst *must* measure efficiency. One can certainly have a broader normative debate about the value of efficiency as a measure of partisan fairness compared to either symmetry or responsiveness. (I will return to this question again in the conclusion.) But many studies actually assume that either symmetry or responsiveness, or both, accurately measure efficiency. This is a far stronger, and fundamentally empirical, claim. Given the dependence of so many studies on this core idea, it is surprising that no systematic evaluation of it has ever been made. Should we expect these measures to capture efficiency, or are the concepts distinct?

Understanding Efficiency

To evaluate symmetry and responsiveness as measures of efficiency, we need some standard for comparing them. What does it mean for a plan to be efficient? First, a party that strives to maximize efficiency wants to win more seats, not protect the ones it already holds. A redistricting party might certainly draw safe seats to preserve gains it has recently made, especially if it feels those gains are unlikely to be surpassed before the next redistricting. Moreover, scholars have rightly noted that an effective gerrymander will avoid making majority party seats so competitive that the gerrymander comes undone in the face of contrary partisan tides (Friedman and Holden 2008; Owen and Grofman 1988). However, neither of these ideas represents the core concept of efficiency as described above and as employed in the literature, so I set them aside for now. I will take up these points again in the discussion.

Second, efficient bias is not just about winning more seats—it is about winning them with the same number of supporters: that is, “the effect of redistricting on the allocation of seats between the parties *given their average district votes*" (Gelman and King 1994a, 550, emphasis added). Thus, it assumes a fixed number of party voters who are then redistributed among the districts. It is an advantage that stems from where a party's supporters live and how the district lines have been drawn around them. The less reliable a party's supporters, the more difficult it is to make an efficient gerrymander "stick" (Friedman and Holden 2008).

These two components—higher seat share and a constant vote share—suggest the following proposition:

Proposition 1: If a party's seat share grows without any change in its vote share, then the efficiency of the system also shifts in that party's favor.

It is important to note that Proposition 1 is definitional: it identifies a condition under which a party's advantage in efficiency *must* be changing, since it is the very nature of efficiency to extract more seats from the same level of partisan support. Proposition 1 also implies that an appropriate measure of efficiency should capture the shift in seat share described. This suggests a second proposition:

Proposition 2: A valid measure of efficiency will suggest a monotonic increase (decrease) in the advantage for one party if that party gains (loses) seat share without any corresponding increase (decrease) in its vote share.

Proposition 2 is a necessary condition for a valid measure of efficiency: any measure that does not satisfy the proposition may be appropriate for some other concept of interest, but it is not, ipso facto, a measure of efficient bias. Thus, to cast doubt on symmetry and responsiveness as measures of efficiency, we need only establish that the measures can, under certain circumstances, violate Proposition 2. I turn now to some theoretical examples that produce just such an unintended result.

First consider symmetry under the case when there are only two parties, a government (G) and an opposition (O), who fight to control 10 seats. Party G's share of the vote in any given district is V_i^G, its actual (i.e., prior to counterfactual) aggregate share of the total vote across all districts is V_G, and its actual share of all seats is S_G. G receives a majority of the vote, so $V_G > 0.5$. As with any two-party system, 0.5 is the win-loss threshold: any seat with an actual government vote share above this threshold (i.e., $V_i^G > 0.5$) will be held by G, while any seat below it will be held by the opposition O. Panel 1 of Figure 1 presents one possible configuration of these values across the continuum of vote shares, with G holding six of 10 seats ($S_G = 0.6$).

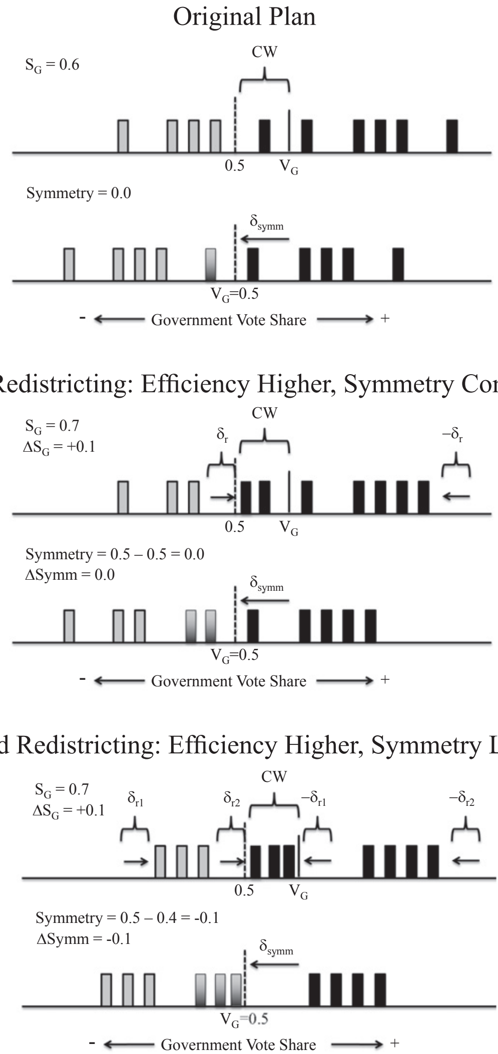
Since calculating symmetry requires shifting the vote shares of all districts and recording the number of seats that change hands, there exists a range of vote shares within which any seat will change hands as a consequence of this counterfactual, and outside of which no seat will change hands. Let us refer to this range of votes as the *counterfactual window*, or CW. Symmetry typically involves shifting V_G to 0.5, so the CW in this case is the range of vote shares defined by $0.5 < V_i^G < V_G$. Any seat falling in that range belongs to party G under the actual election results but will belong to party O under the counterfactual scenario. Panel 1 of Figure 1 demonstrates this effect: G only retains the seats that fall above V_G, which in this case means "losing" one seat. Moreover, in this example the system is symmetrical, since after the counterfactual each party holds half the seats for half the votes.

Now consider what happens under the following redistricting plans where G improves its efficiency by increasing its seat share without any change in vote share:[4]

1) *Figure 1, Panel 2: A seat held by Party O moves into the CW (Party G efficiency gain)*. Under this scenario, presented as Panel 2 of Figure 1, the opposition party loses a seat to the government as G transfers some of its supporters from its safest seat to the most marginal seat held by O. (This is, in fact, a typical ploy for a partisan gerrymander.) Efficiency increases for G, as it gains a seat without gaining any additional votes. Yet, using symmetry, the seat it gains is lost again under the counterfactual, leaving no trace of the change.
2) *Figure 1, Panel 3: A government and an opposition seat are both moved into the CW (Party G efficiency gain)*. The government increases its efficiency by gaining a seat without any gain in vote share, but under the counterfactual, it loses this seat and an additional one besides. This leads to the conclusion that the opposition has gained from the redistricting, when the government is the real beneficiary.

If symmetry is sometimes a poor measure of efficiency, responsiveness may have even greater limitations. Symmetry is a count of the number of seats above and below a particular threshold (the aggregate vote share, in the case of the most common measure of symmetry). Responsiveness, by contrast, is effectively a count of the number of

FIGURE 1
Redistricting and Symmetry in Three Hypothetical Plans



Note: Each pair of figures corresponds to a scenario described in the text. The top figure is the actual election outcome and the bottom figure the outcome under the symmetry counterfactual. δ_{symm} is the shift in votes required for the symmetry counterfactual. δ_r is the shift in votes for particular districts from redistricting. ΔS_G is the change in the government's share of seats relative to the baseline redistricting plan in the first panel, and ΔSymm is the change in symmetry relative to the same baseline plan. All other values are defined in the text.

districts in the CW. If that number increases, responsiveness will also increase. Specifically, it is defined as:

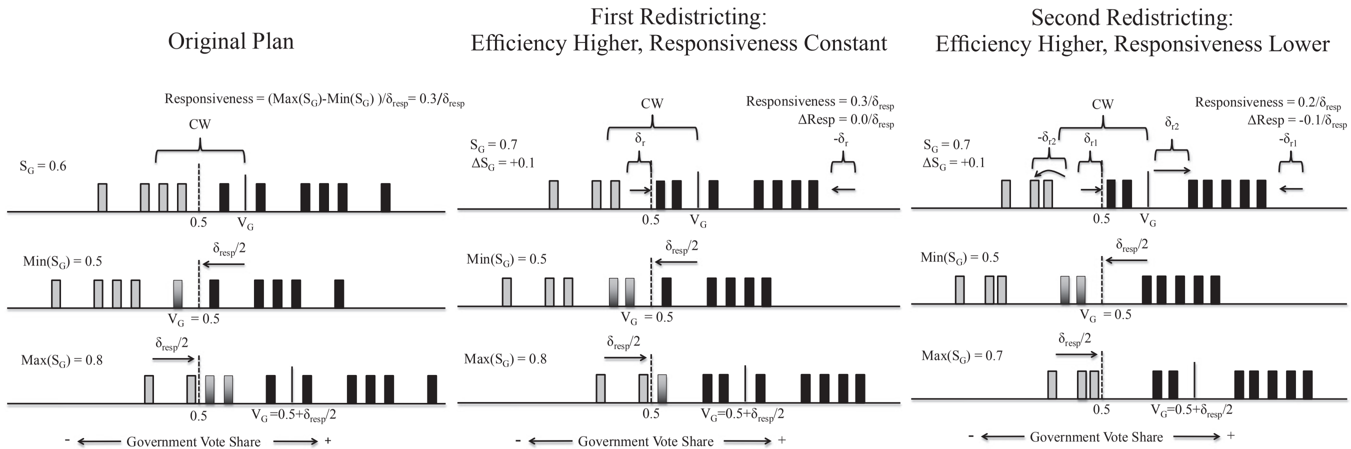
$$Responsiveness = (Max(S_G) - Min(S_G))/\delta_{resp}. \tag{1}$$

δ_{resp} is the range of vote shares considered for the calculation. This range is somewhat arbitrary, and analysts have used values ranging from as high as +/- 15 percentage points off the actual aggregate vote share to as small as virtually nothing (specifically, the derivative of the seats-votes curve at the point of the actual election outcome). $Max(S_G)$ and $Min(S_G)$ are the maximum and minimum seat shares for the government in that range of vote shares. The first panel in Figure 2 presents a base plan altered slightly from the one found in Figure 1, with a responsiveness of $0.3/\delta_{resp}$. From this baseline, the following two scenarios produce the results similar to the ones in Figure 1:

1) *Figure 2, Panel 2: A seat held by Party O crosses the win-loss threshold but remains in the CW, while a seat held by Party G becomes more competitive.* Party G's relative efficiency increases, but there is no change in responsiveness because the number of seats in the CW does not change.
2) *Figure 2, Panel 3: The redistricting plan in the first scenario also moves an opposition seat out of the CW while making a government seat less competitive*. The first scenario already improved efficiency for the government without changing responsiveness; these additional changes also decrease the responsiveness of the plan as well. The result is an improvement in efficiency for the government but a decline in responsiveness.

While these examples demonstrate specific violations of Proposition 2, there are more conceptual reasons to think responsiveness does not make sense as a measure of efficiency. It was designed to measure competition, not partisan advantage, and so it has no partisan valence. A redistricting plan that improves efficiency for either party may also improve responsiveness by increasing the number of competitive seats for that party. That means responsiveness cannot distinguish between a plan that favors one party from a plan that favors the other. Moreover, even a classic partisan gerrymander might not improve responsiveness, since the redistricting party could be expected to make minority seats uncompetitive at the same time that it makes its own seats competitive (Cain 1985). If these changes balance each other out and leave the same

FIGURE 2
Redistricting and Responsiveness in Three Hypothetical Plans



Note: Each pair of figures corresponds to a scenario described in the text. The top figure is the actual election outcome and the bottom figure the outcome under the responsiveness counterfactual. δ_{resp} is the shift in votes required for the responsiveness counterfactual. δ_r is the shift in votes for particular districts, as induced by various alternative redistricting plans. ΔS_G is the change in the government's share of seats relative to the baseline redistricting plan in the first panel and ΔResp is the change in responsiveness relative to the same baseline plan. All other values are defined in the text.

number of seats in the CW, there will be no change to responsiveness at all.

It is important to emphasize that these examples have been specifically selected to demonstrate a violation of Proposition 2. As such, they show only that symmetry and responsiveness are distinct from efficiency. They do not, and cannot, tell us whether symmetry and responsiveness are *correlated* with efficiency, such that they might serve as reasonable proxies for the concept. I address that related notion in the next section.

Testing Symmetry and Responsiveness

The arguments above present evidence that symmetry and responsiveness may be poor measures of seat gains from redistricting. Yet the sort of scenarios I have identified as problematic may be rare in the real world. How well do each of these measures correlate with the seat share they must reflect in order to be valid measures of efficiency? Do symmetry and responsiveness correspond better to real redistricting plans than they appear to in the theoretical discussion above?

To answer this question, I turn to a data set of elections to lower state legislative chambers from 1970 to 2003 (Carsey et al. 2008). The data are the standard source for studies of legislative elections, and past efforts at validation have found them to be in excellent condition (Gelman and King 1994a). Because the claims I have presented here only apply to single-member district elections, I drop all multimember districts from the analysis. I then use JudgeIt (Gelman, King, and Thomas 2008) to calculate estimated vote shares and seat shares, as well as symmetry and responsiveness, for every state in every election year.[5]

Consistent with Proposition 2, I assume that an appropriate measure of efficiency should reflect changes in seat share when vote share is controlled: in other words, it should identify additional seat share above and beyond what is predicted by vote share. To test this idea, I regress seat share separately on both symmetry and responsiveness, controlling for aggregate vote share. Since symmetry is calculated at an aggregate vote share of 0.5, it becomes a more accurate proxy for efficiency as the actual vote share approaches that value. Thus, I also present a second model for each measure where I interact the predictors with an indicator for elections in which one party received more than 55% of the two-party vote. (This distinction splits the data roughly in half, with 219 (44%) state-election pairs falling in the uncompetitive category.) If the arguments I have made are correct, symmetry ought to predict seat share

TABLE 1
Explaining Seat Share in State Legislative Races with Three Measures of Partisan Advantage, 1970–2003

	Symmetry		Responsiveness	
	(1)	(2)	(1)	(2)
Partisan Advantage Measure	0.246***	0.731***	0.010**	0.000
	(0.036)	(0.050)	(0.004)	(0.004)
Vote Share	2.038***	1.797***	2.359***	2.700***
	(0.054)	(0.116)	(0.040)	(0.117)
Uncompetitive = 1	—	0.015**	—	−0.012*
		(0.006)		(0.006)
Advantage × Uncompetitive	—	−0.799***	—	0.031***
		(0.064)		(0.008)
Vote Share × Uncompetitive	—	0.581***	—	−0.254*
		(0.129)		(0.128)
Intercept	−0.010***	0.046	0.000	0.013**
Adjusted R^2	0.898	0.924	0.891	0.895
Root MSE	0.050	0.043	0.051	0.050
N	501	501	501	501

Note: Cell entries are ordinary least-squares coefficients, and estimation was run in R 2.14.0. The dependent variable in each case was two-party seat share. The measure of partisan advantage used in each model is identified in the column heading. Vote share, seat share, and both measures of partisan advantage are mean deviated.
$^*p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

poorly, but better in the competitive range of outcomes. That means the interaction model should produce a modest main effect and a strongly negative interaction. Responsiveness, on the other hand, ought to perform poorly in both competitive and uncompetitive elections.

The results can be found in Table 1 and largely conform to expectations. In the first model with no interactions, responsiveness has almost no predictive power (0.010), while symmetry's relationship to seat share is weak (0.246). When the model is split into competitive and uncompetitive elections, the results for responsiveness are unchanged—a zero coefficient becomes 0.031 in uncompetitive elections—but the models for symmetry now suggest an important distinction. Symmetry's predictive power is reasonably good in the competitive range (0.731) but becomes effectively zero and even slightly negative (−0.068) where one party is stronger than the other.

Thus, it appears the concerns raised about symmetry and responsiveness extend beyond hypotheticals and apply to a broad range of actual election results. Responsiveness has particular trouble, but symmetry also struggles outside the relatively narrow range of elections contemplated by the symmetry counterfactual.

An Alternative Measure of Partisan Efficiency

It is possible to imagine a simple measure of efficiency that avoids many of the problems of symmetry and responsiveness and does not require any counterfactual at all. If a party gains seats without any increase in vote share, it ought to be making better use of the votes it wins by claiming its seats with smaller majorities. This suggests the following measure, which we might call *relative wasted votes (ω)* and which is defined as follows for a system with two parties:

$$\omega = (W_O - W_G) \Big/ \sum v_i^t. \quad (2)$$

W_O and W_G are the total surplus votes (in excess of the number required to win) plus the total lost votes (cast for the party in races that party lost) for the opposition and government parties, respectively. W_O and W_G are more formally described as:

$$W_O = \sum Sur_O + \sum Lost_O, \quad (3)$$

$$W_G = \sum Sur_G + \sum Lost_G. \quad (4)$$

Where Sur_G and Sur_O are the surplus votes, and $Lost_G$ and $Lost_O$ the lost votes, for the government and opposition parties. The difference between these numbers should give a sense of relative advantage in wasted votes. This number is then divided by the total number of votes cast in all districts to standardize the value for comparison across electoral systems. In the special case where there are only two parties and all districts are equal in population—a set of constraints virtually universal in the research on symmetry and responsiveness[6]—Equation (2) reduces to an even simpler form:

$$\omega = S_{marg} - 2 * V_{marg}. \quad (5)$$

Where

$$S_{marg} = S_G - 0.5, \tag{6}$$

$$V_{marg} = V_G - 0.5. \tag{7}$$

The derivation of Equation (5) can be found in Appendix B. In words, the equation states that ω is equal to G's excess seat share over 0.5, minus twice its excess vote share over 0.5. Thus, the balance of wasted votes in Equation (5) will be equal whenever the majority's margin in seats is twice its margin in votes.

Equation (5) is consistent with Proposition 2. When vote share is constant as Proposition 2 requires, V_{marg} is also constant, and ω can only be changed by altering the seat share. Furthermore, when the seat share changes without any change in vote share, ω will change by the same amount, and this rate of change is constant across all values of V_{marg}.[7] Thus, for an SMD system with the characteristics identified, ω is an exact measure for the changes in seat share contemplated by Proposition 2. Indeed, if ω is inserted into the basic model in Table 1, it perfectly predicts seat share with a coefficient of 1.0 for relative wasted votes and 2.0 for vote share.[8]

It is not essential to measure relative wasted votes with the precise functional form in Equation (2). But since any alternative must still reliably identify when Party G has more wasted votes than Party O, the number of options is limited. The most obvious alternative is to use the ratio of wasted votes (i.e., W_O/W_G) rather than the difference, since the ratio still identifies the key condition of equivalence between the parties (in this case, when the ratio is equal to 1.0). But the ratio is a nonlinear function, and so it adds needless complexity without improving on ω in other respects. In fact, *even using the ratio, wasted votes are equal when the majority party's margin in seats is twice its margin in votes*. The proof for this claim can be found in Appendix B. It suggests that, far from a quirk of ω, a "twice vote margin" rule of thumb for the equivalence of wasted votes may be a fundamental feature of this sort of SMD system.

In addition to its other advantages, ω is also flexible. It can be applied to any redistricting system, regardless whether the parties are competitive with one another or not. By contrast, analysts are cautioned to avoid using symmetry unless the actual aggregate seat or vote share is close to 0.5 (Gelman and King 1994a; Grofman and King 2007).

Moreover, ω actually collapses on symmetry when the aggregate vote share is 0.5: if that value is substituted into Equation (5), the V_{marg} term disappears and only the difference between seat share and 0.5 remains. If an aggregate vote share of 0.5 is a substantively interesting value, then it can be used with the confidence that the results will be identical to symmetry; otherwise, the actual outcome or some more realistic counterfactual may be used instead.

ω is not without limitations in extreme cases. When the majority holds all the seats (i.e., $S_{marg} = 1$), any additional votes will actually reduce the party's efficiency advantage and even put it at a disadvantage for any vote share above 0.75. This property of ω is strange and arguably violates a sense that a measure of party advantage should always increase as a party grows stronger. Strictly speaking, however, it does not violate Proposition 2, which concerns changes in seat share for a constant vote share, not vice versa. Indeed, once a party wins all available seats, additional votes are wasted by definition: they do not contribute to additional victories. More important, the levels of partisan dominance required to produce this result are highly unusual. In the legislative elections data, no party has ever won all the seats or even won more than 75% of the vote. By contrast, majorities between 55 and 75% of the vote—which cause serious problems for symmetry but raise no issues for relative wasted votes—are quite common and comprise close to half the data set (44%).

Equation 5 also demonstrates that vote share, seat share, and efficiency are linked: the only way to change seat share for a constant vote share is to alter the distribution of wasted votes between the parties. Thus, ω is both a tally of all the cracking and packing decisions in a given SMD system and a composite measure of aggregate vote and seat share. Far from standing above the political fray, efficiency is and ought to be sensitive to changes in vote share. Symmetry and responsiveness assume that such changes in vote share—especially uniform shifts—are irrelevant for partisan bias. This is a design decision that reflects nothing inherent in the concept of efficiency and can in fact misrepresent efficiency in at least some circumstances, as the earlier discussion makes clear. Even the simple notion that a partisan gerrymander might come "undone" in the face of contrary partisan tides becomes highly unlikely if symmetry or responsiveness is the measure of bias. Unless the tide is stronger in some seats than others—and thereby alters the shape of the seats-votes curve—these measures will consider the change to be irrelevant for evaluating the plan's bias. The measures will thereby suggest stability in efficiency where none may in fact exist.

Partisan Control and Partisan Bias

Thus far I have presented theoretical and empirical evidence that both symmetry and responsiveness are poor measures of efficiency and offered an alternative to replace them. In this section, I apply this new measure to a key dispute in the literature and demonstrate how it alters our understanding of the relevant dynamics.

A long-standing argument in the redistricting literature concerns whether control of the redistricting process matters for the party that possesses it. Most politicians and journalists, and many political scientists, assume that a party with unfettered control will successfully exploit its position to maximize its seat share, and several empirical studies have presented evidence for precisely this result (Engstrom 2006; McDonald 2004). At the same time, many scholars have questioned the link between party control and partisan outcomes. A party in control wants to maximize its seat share, but its incumbents also want safe seats that will protect their political careers. These goals are in tension, since safe seats necessarily waste votes that could be distributed more efficiently elsewhere (Cain 1984). Even if the party seeks to maximize seats, it is not always clear that the plan will produce the results the party expects (Born 1985; Cain 1985; Campagna and Grofman 1990; Niemi and Jackman 1991). The more fluid the party support in the electorate, the less reliable the gerrymander.

Gelman and King summarize the debate this way: "one side holds that gerrymanderers. . .maximize only their party's seat advantage and have a large and lasting effect, while the other argues that whatever gerrymanderers maximize, they have only a small and transitory effect" (1994a, 543). Gelman and King (GK) offer a truce between these two perspectives, arguing that each is partly right. The "large and lasting effect" group is correct that the party in control of redistricting shifts bias in its favor, and that "the effect is substantial and fades only very gradually over the following 10 years" (543); the "small and transitory effect" group is right in the sense that a partisan redistricting plan still upsets established political relationships enough to leave the total level of bias lower than before.

To reach these conclusions, GK develop a parsimonious model where partisan bias is regressed on its lag, an indicator for redistricting years, the interaction between these two variables, an indicator for the party in control of the redistricting process (−1 Republican, 0 bipartisan/nonpartisan, 1 Democratic), and a set of state fixed effects. Formally, the model can be written as follows for state *i* in year *t*:

$$Y_{it} = \alpha Y_{i,t-1} + \beta R_{it} + \gamma (Y_{i,t-1} * R_{it}) + \delta P_{it} + s_i + \varepsilon_{it}, \quad (8)$$

where $Y_{i,t-1}$ is the lagged dependent variable, R_{it} is the redistricting dummy, P_{it} is party control, s_i is the set of state fixed effects, ε_{it} is the error term, and α, β, γ, and δ are coefficients to be estimated. α measures the endurance of a plan's bias through time, γ shows how well redistricting disrupts this stability, and δ captures the tendency for a party to skew redistricting in its favor, independent of these other factors. When they estimate this equation using symmetry, GK find a positive δ and a strong and positive α, which says that partisan gerrymandering is real and persists through the life of a redistricting plan. But they also find a strongly negative γ, meaning redistricting so thoroughly "resets" the system as to decrease bias overall.

The main limitation of this analysis as it pertains to efficiency is that it uses symmetry as the measure of bias. As we have seen, symmetry discounts shifts in aggregate vote share by design and focuses instead on the relative distribution of votes across seats as the phenomenon of interest. This makes symmetry more stable within redistricting cycles, where uniform shifts in vote share are more common, but also less stable between cycles, where changes in the distribution prevail. In short, structural features of symmetry incline the measure toward finding the very pattern reported by Gelman and King, whether this pattern reflects actual seat gains or not.

In Table 2, I have updated GK's symmetry model with legislative data through 2003. I use their original coding of party control for the 1972 and 1982 redistricting cycles and then two additional sources for the 1992 and 2002 redistricting years that have been added to the analysis.[9] GK run their model on only 16 states, after omitting those where one party did not win either a vote or a seat majority at any point in the study period.[10] If we follow this approach and limit the analysis to the same 16 states, we confirm GK's basic result: a strong lagged effect ($\alpha = 0.756$), a modest party control effect ($\delta = 0.017$), and a strong interaction ($\gamma = -0.316$). The interaction here suggests that redistricting disrupts almost half the legacy effect from one plan to the next. Yet when the same model is run with relative wasted votes, a different story emerges. A similar party-control effect is still visible ($\delta = 0.020$), but the lagged effect is about a third the size ($\alpha = 0.256$).[11] Furthermore, the interaction term is small and statistically insignificant (though still negative), suggesting that there is little about redistricting that is especially disruptive.

The remaining columns in Table 2 expand the model to include all the states in the sample, thus introducing a comparatively uncompetitive

TABLE 2
Effects of Redistricting on Symmetry and Relative Wasted Votes in State Legislative Elections, 1970–2003

	Competitive States		All States	
	Symmetry	ω	Symmetry	ω
Lagged symmetry/ω (α)	0.752***	0.256***	0.684***	0.227***
	(0.054)	(0.076)	(0.055)	(0.060)
Redistricting year (β)	–0.018*	–0.006	–0.016**	–0.003
	(0.007)	(0.007)	(0.006)	(0.005)
Redistricting year × lagged symmetry/ω (γ)	–0.318**	–0.091	–0.184*	–0.026
	(0.010)	(0.116)	(0.081)	(0.081)
Partisan control (δ)	0.017*	0.020*	0.020**	0.014**
	(0.008)	(0.008)	(0.006)	(0.006)
(State fixed effects)				
Intercept	0.011	0.019	0.009	0.019
Adjusted R^2	0.701	0.375	0.749	0.433
Root MSE	0.044	0.044	0.045	0.041
N	244	244	463	463

Note: Cell entries are ordinary least squares coefficients. Models include state fixed effects and heteroskedastic-consistent standard errors, as estimated in R 2.14.0. Dependent variable is either symmetry or relative wasted votes (ω), depending on the column. "Competitive states" include the 16 identified as such in Gelman and King (1994a).
$^{\#}p < 0.10$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

set of election outcomes into the analysis. For relative wasted votes, the results are broadly similar, although the coefficient on partisan control is somewhat smaller than in the original specification. For symmetry, the additional states actually increase the impact of party control very slightly, while cutting the interaction coefficient almost in half. However, the basic distinction between the two measures remains intact: symmetry suggests a larger effect from redistricting and more persistence through time.

It is also worth noting the significant difference between the two equations in terms of variance explained. The adjusted R^2 for the symmetry model (0.749) is significantly larger than that of the relative wasted-votes model (0.433). Some of this can be explained by the fact that the symmetry variable has a higher variance to begin with, but almost half this difference can be explained by the greater importance of the lag term in the symmetry model. Removing it makes very little difference to the variance explained of the relative wasted-votes model (new $R^2 = 0.403$), but it makes a substantial difference to the variance explained of the symmetry model (new $R^2 = 0.581$). Thus, the greater instability of relative wasted votes, despite its stronger logical connection to efficiency and seat share, remains one of its most important distinctions from symmetry.

Florida's lower house offers a good case study of this fading effect. The redistricting process was controlled by the Democrats in 1991, and they drew a plan that largely appeared to preserve the status quo, albeit marginally improving the party's relative wasted-votes score from 0.03 to 0.04. But the party's seat share dropped through the remainder of the decade, and by the time of the next redistricting cycle, the Republicans were firmly in control with 64% of the seats. Relative wasted votes largely tracked this mid-decade shift, showing a Democratic advantage as high as 0.09 in 1994 that flipped to a Republican advantage of 0.12 in 2000. By contrast, symmetry suggested very little change over the same period: it remained within a narrow range that never gave either party better than a 0.02 advantage.

In sum, relative wasted votes paints a different picture of redistricting than symmetry. Both measures agree that parties seek to use redistricting to shift bias in their favor and that they are successful in these efforts, at least initially. But where symmetry suggests these efforts largely endure until the next redistricting shock, relative wasted votes implies the long-term consequences of a plan are unknowable. Indeed, the estimates suggest the effect of relative wasted votes decays about as rapidly in a single election as the effect of symmetry does in four elections. Drawing new lines adds little to this inherent uncertainty, perhaps because the uncertainty is so substantial in the first place.

Because it translates more directly into seat share, relative wasted votes also gives us better insight into the magnitude of the party-control effect. The coefficients in Table 2 suggest about a 3-point difference in seat share between states where Republicans control redistricting and states where Democrats hold the levers of power. This translates to roughly two to four seats for the lower chambers of most states. While this is a real effect that might have practical political implications, it is not hard to see how it might fade over the course of a redistricting cycle.

On balance, then, these findings support the "small and transitory" perspective on redistricting in contrast to the conclusions one would reach using symmetry. They also cast doubt on any normative benefits that might stem from the supposed power of redistricting to shatter old biases even as it creates new ones. Instead, the normal give and take of elections is probably disruptive enough to overwhelm such benefits.

Discussion

The findings presented here offer reasons to abandon symmetry and responsiveness as measures of partisan efficiency and adopt relative wasted votes (ω). The new measure offers two practical advantages. First, it is more accurate at measuring a core concept of interest in partisan gerrymandering, which is whether one party gains seats through redistricting alone. The seat-gain concept is best evaluated at the actual level of support for each party, not at a counterfactual level the system may or may not reach at any point in the near future. Second, the new measure is available for the entire range of aggregate vote shares a party might expect to receive. Symmetry, by contrast, must be restricted to electoral systems where party control is seriously at issue because the counterfactual is otherwise too unrealistic (Grofman and King 2007).

The new measure also sharpens the link between efficiency and seat share. In the case of two parties and equipopulous districts, there is no distinction between the concepts: a party that gains seat share through redistricting necessarily improves its relative efficiency. Indeed, information about relative efficiency is contained in the seats-votes curve itself: as Equation (5) makes clear, when responsiveness is greater than 2.0, a party's relative efficiency improves as its vote share rises; when it is less than 2.0, a higher vote share has the opposite effect. The literature has always assumed that any level of responsiveness is just as "fair" as any other, but the analysis here casts some doubt on that assumption, at least as concerns wasted votes.

These theoretical insights lead to the key empirical finding of the article. Because efficiency is a function of both vote share and seat share, it is sensitive to changes in party performance. Thus, the effects of partisan gerrymanders, though real, are easily undone. The partisan legacy of the last plan is usually gone by the following redistricting, and redistricting itself is not as disruptive a force as symmetry would suggest. This raises important questions about the utility of court intervention in partisan gerrymandering cases. If partisan gerrymandering has a large, enduring effect, then it might be important for the courts to insert themselves into the process to prevent the advantages a party might otherwise obtain. But if the effects of gerrymandering, though real, are small and ephemeral, court intervention in such an inherently political process might do more harm than good.

One might object that bias ought to capture something immutable about a redistricting plan, or at least something independent of the back and forth of partisan tides, and that symmetry should be preferred for that reason. However, the concept of interest here—partisan efficiency—

requires identifiable and reliable partisan voters. If a large number of voters can change their minds—whether idiosyncratically or as part of a broad partisan tide—their support cannot be counted on and the expected gains from redistricting will never materialize. A measure should not force stability in such circumstances, or it risks ignoring some of the very properties it is meant to test.

Gerrymandering can certainly embrace a wider range of goals than strict efficiency, and it is worth considering some here. Owen and Grofman (1988) incorporate uncertainty into their model of gerrymandering, showing that a smart redistricting party will avoid cutting its margins too close in the districts it wins for fear of losing those seats in future partisan tides and limiting its expected efficiency gains. Friedman and Holden (2008) have extended this logic to situations where gradations of support in the electorate can be identified, demonstrating that under those conditions, it is not sensible to create a handful of districts that the opposing party is guaranteed to win.[12] Though these models do refine the goals of a classic gerrymander, their implications for the discussion here should not be overstated. Models that account for uncertainty do not discount efficient seat maximization as a goal; rather, they offer different tactics for achieving that end, and they emphasize that what matters most is the expected value of efficiency across a number of elections. Since maximizing seat share for a given share of the vote is still the goal, relative wasted votes ought to be a good measure of the concept. Indeed, I have argued at length that relative wasted votes is better than symmetry at capturing the consequences of this sort of uncertainty.

Another possible goal comes from Yoshinaka and Murphy (2009), who note that a gerrymandering party will often destabilize the support base for opposing-party incumbents, even if it gains no additional seats as a result. This is an important point and a key secondary goal of a gerrymander. Nonetheless, in making this point, Yoshinaka and Murphy do not dismiss efficiency as an objective of partisan gerrymandering. On the contrary, they explicitly test for it, albeit with different measures than the one I have offered here. It is conceivable that they might have conducted that portion of their study using relative wasted votes instead.

Finally, one could certainly argue that symmetry offers a distinct and valuable notion of fairness—that equal parties should be treated equally—which should not be discarded lightly. Because an asymmetric system can allow a party to maintain its chamber majority without winning a majority of the vote, it raises special normative concerns that will likely always be relevant to discussions of redistricting. That said, even if symmetry remains a valuable normative concept, it may be more fruitful to treat it as a special case of relative wasted votes, since the two measures collapse on each other at $V = 0.5$. Rather than calculate relative wasted votes and symmetry separately, an analyst could calculate relative wasted votes across a range of substantively meaningful values. If it seems possible for the minority party to claim half the votes in a future election, then that outcome could be part of the range of values that is explored, and symmetry would therefore be part of the discussion. If not, then it need not be included. Again, the point is not to dismiss symmetry as a normative concept but to suggest that an important aspect of it is already captured in the broader notion of wasted votes.

Regardless, if the goal is to measure efficiency, then the twin concepts of symmetry and responsiveness look to be inappropriate for many SMD electoral systems. The measure I have offered here—relative wasted votes—is arguably a more valid and flexible measure of the concept of efficiency in most cases, and it subsumes symmetry at a point in the seats-votes curve where symmetry's counterfactual is often evaluated. Most important, using relative wasted votes in place of symmetry offers us a different understanding of substantive questions about redistricting. If the traditional measures are used to evaluate efficiency, it should only be with great caution and significant caveats, because doing so could in many cases lead analysts to the wrong conclusions.

Eric McGhee <mcghee@ppic.org> *is a Research Fellow at the Public Policy Institute of California, 500 Washington Street, Suite 600, San Francisco, CA 94111.*

APPENDIX A

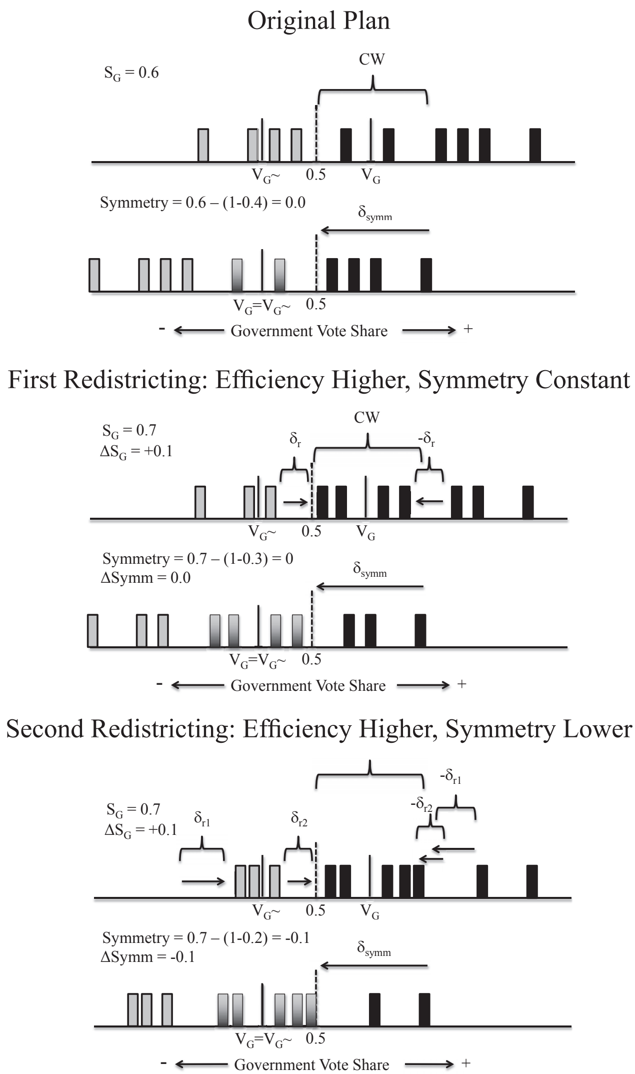
Alternative Measure of Symmetry and Its Consistency with Proposition 2

Symmetry is sometimes calculated as the difference between the seat share that the government receives and the share the opposition would receive for the same share of the vote. If S_G is the share of seats received by the government for an aggregate vote share V_G, let $V_{\tilde{G}} = 1 - V_G$ and $S_{\tilde{G}}$ be the share of seats the opposition would receive for $V_{\tilde{G}}$. This alternative measure of symmetry is simply the difference between the two seat shares: $S_G - S_{\tilde{G}}$. Negative values favor the opposition, while positive values favor the government. A possible configuration for these values is in the first panel of Figure A1.

Consider now two scenarios whose results for this version of symmetry mimic the results for the version of symmetry presented in the main text:

1) *Panel 2, Figure A1: Both a seat held by Party G and a seat held by Party O move into the CW, but only the seat held by Party O crosses the win-loss threshold.* Party

FIGURE A1
An Alternative Measure of Symmetry in Three Hypothetical Plans



Note: Each pair of figures corresponds to a scenario described in the text. The top figure is the actual election outcome and the bottom figure the outcome under the symmetry counterfactual. δ_{symm} is the shift in votes required for the symmetry counterfactual. δ_r is the shift in votes for particular districts from redistricting. ΔS_G is the change in the government's share of seats relative to the baseline redistricting plan in the first panel, and ΔSymm is the change in symmetry relative to the same baseline plan. All other values are defined in the text.

G gains a seat and so improves its relative efficiency, but there is no change in symmetry.

2) *Panel 3, Figure A1: One uncompetitive seat held by Party O becomes more competitive, one seat held by O changes hands, and two Party G seats move into the CW.* Party G increases its relative efficiency, but symmetry suggests Party O improves its situation instead.

APPENDIX B

Deriving and Exploring Equation 5

Here I offer the derivation of Equation 5 in the text. When there are only two parties and each district has exactly the same number of voters, proportions can be substituted for raw votes in all of the formulas. In this case, the total vote in each district becomes equal to 1.0, so,

$$\sum_1^{d_t} V_i^t = d_t \tag{B1}$$

where d_t is simply the total number of districts in the electoral system. Moreover, surplus and lost votes become:

$$Sur_O = \sum_1^{d_O} (0.5 - V_i^G) = 0.5 * d_O - \sum_1^{d_O} V_i^G \tag{B2}$$

$$Sur_G = \sum_1^{d_G} (V_i^G - 0.5) = \sum_1^{d_G} V_i^G - 0.5 * d_G \tag{B3}$$

$$Lost_O = \sum_1^{d_G} (1 - V_i^G) = d_G - \sum_1^{d_G} V_i^G \tag{B4}$$

$$Lost_G = \sum_1^{d_O} V_i^G \tag{B5}$$

where d_O is the total number of districts won by the opposition party, and d_G is the same for the government. Wasted votes then become:

$$W_O = 0.5 * d_O - \sum_1^{d_O} V_i^G + d_G - \sum_1^{d_G} V_i^G \tag{B6}$$

$$W_G = \sum_1^{d_G} V_i^G - 0.5 * d_G + \sum_1^{d_O} V_i^G \tag{B7}$$

Substituting into Equation (1), the formula for ω becomes:

$$\begin{aligned}\omega &= (W_O - W_G)\Big/\sum_1^{d_t} V_i^t \\ &= \Big\{\Big(0.5 * d_O - \sum_1^{d_O} V_i^G + d_G - \sum_1^{d_G} V_i^G\Big) - \Big(\sum_1^{d_G} V_i^G - 0.5 * d_G + \sum_1^{d_O} V_i^G\Big)\Big\}\Big/ d_t \\ &= \Big\{0.5 * (d_O + d_G) - 2 * \Big(\sum_1^{d_O} V_i^G - \sum_1^{d_G} V_i^G\Big) + d_G\Big\}\Big/ d_t\end{aligned} \quad \text{(B8)}$$

Because there are only two parties, and because total votes equals total seats, we can make three additional substitutions:

$$d_t = d_O + d_G \quad \text{(B9)}$$

$$S_G = \frac{d_G}{d_t} \quad \text{(B10)}$$

$$V_G = \Big(\sum_1^{d_O} V_i^G + \sum_1^{d_G} V_i^G\Big)\Big/ d_t \quad \text{(B11)}$$

Equation B9 states that the government and opposition districts comprise the total universe of seats. Equation B10 is a simple identity: the government's share of seats is equal to the number of seats it holds divided by the total number of seats. Equation B11 states that the government's average vote share across all districts is equal to the sum of its shares in seats it holds plus its shares in seats it does not hold, divided by the total number of seats. Rearranging and substituting into Equation (B8), we obtain:

$$\omega = 0.5 - 2 * V_G + S_G$$

If we define $V_{marg} = V_G - 0.5$ and $S_{marg} = S_G - 0.5$, then this rearranges into:

$$\begin{aligned}\omega &= 0.5 - 2 * (V_{marg} + 0.5) + S_G \\ &= 0.5 - 2 * V_{marg} - 1 + S_G \\ &= S_G - 0.5 - 2 * V_{marg} \\ &= S_{marg} - 2 * V_{marg}\end{aligned} \quad \text{(B12)}$$

Equation (B12) is exactly equal to Equation (5) in the text, and it suggests that wasted votes are always equal when one party holds a margin in seats twice as large as its margin in votes (i.e., when the expression is equal to zero, indicating no difference in wasted votes).

As also noted in the text, this core result holds if ω is expressed as the ratio of wasted opposition votes to wasted government votes, rather than the difference. The precise values of ω would change with such a measure, but wasted votes will still be equal when the margin in seats is twice the margin in votes. For the proof, first imagine a measure ω* such that

$$\omega^* = \left(W_O \Big/ \sum_1^{d_t} V_i^t\right) \Big/ \left(W_G \Big/ \sum_1^{d_t} V_i^t\right)$$

From Equations (B6), (B7), (B9), and (B10), and from the fact that $\frac{d_O}{d_t} = S_O = 1 - S_G$, it follows that

$$\begin{aligned}\omega^* &= \frac{\left(0.5 * d_O - \sum_1^{d_O} V_i^G + d_G - \sum_1^{d_G} V_i^G\right) \Big/ d_t}{\left(\sum_1^{d_G} V_i^G - 0.5 * d_G + \sum_1^{d_O} V_i^G\right) \Big/ d_t} \\ &= \frac{0.5 - 0.5 * S_G - V_G + S_G}{V_G - 0.5 * S_G} = \frac{0.5 - V_G + 0.5 * S_G}{V_G - 0.5 * S_G}\end{aligned} \quad \text{(B13)}$$

Substituting $V_{marg} + 0.5$ for V_G and $S_{marg} + 0.5$ for S_G and reducing yields

$$\omega^* = \frac{0.5 * S_{marg} + 0.25 - V_{marg}}{V_{marg} + 0.5 - 0.5 * S_{marg} - 0.25} \quad \text{(B14)}$$

Unlike Equation (5), we must set Equation (B14) equal to 1.0 to identify the point when wasted votes are equal for each party. This leads to the following

$$\frac{0.5 * S_{marg} + 0.25 - V_{marg}}{V_{marg} + 0.5 - 0.5 * S_{marg} - 0.25} = 1$$

$$0.5 * S_{marg} + 0.25 - V_{marg} = V_{marg} + 0.5 - 0.5 * S_{marg} - 0.25$$

$$S_{marg} = 2 * V_{marg}$$

While the ratio approach does confirm the basic finding that seat margin is twice the vote margin when wasted votes are equal, outside this point of equivalence the equation is far more awkward to use. The partial derivative of Equation (B13) with respect to S_G (which identifies how ω* changes under the key condition that seat-share changes without a change in vote share) is $\frac{1}{4 * (V_G - 0.5 * S_G)^2}$, making it both nonlinear and dependent on V_G. The same derivative for Equation (5) is a constant of 1.0, meaning a one-unit change in seat share always leads to a one-unit change in ω. Equation (5) is therefore a far cleaner approach that communicates the same information about which party is advantaged while also offering substantively useful information about the consequences of that advantage for seat share.

NOTES

The author would like to thank Bruce Cain, Benjamin Highton, Vladimir Kogan, and Nicholas Stephanopolous for comments on earlier drafts.

1. An SMD system need not always require a plurality. A runoff system, for instance, would ultimately require a clear majority to win. But the logic of an efficient seat-maximizing gerrymander is the same regardless.

2. See, e.g., Taagepera and Shugart (1989; chaps. 10 and 14) for a relevant discussion of such measures.

3. Strictly speaking, neither symmetry nor responsiveness need be completely hypothetical. In place of a uniform vote shift, some scholars have regressed actual seat share on actual vote share across several elections and then used this model's prediction at 0.5 as a measure of symmetry and its coefficient as a measure of responsiveness. This approach has its limitations (see, e.g., Grofman and King 2007), but it does have the advantage of rooting the estimates in data as much as possible. Nonetheless, the two approaches are more similar than different. If we assume that the only source of change in aggregate vote share is a uniform partisan tide—that is, if the distribution of votes across seats remains constant from one election to the next—then the two approaches are identical. If the distribution is not constant in this way, then it is not clear whether this approach to symmetry is measuring anything related to redistricting at all. Moreover, the approach still requires a hypothetical extrapolation if the vote share in a given electoral system never crosses 0.5 during the study period.

4. The logic of these two scenarios applies when symmetry is calculated as the difference between seat share and vote share at an aggregate vote share of 0.5. Slightly different logic applies if symmetry is instead a comparison between the seat share one party holds and the seat share the opposing party *would* hold *if* it held the same share of votes. With that version of symmetry, the examples above do not always violate Proposition 2. However, it is easy to imagine slightly different examples where this second version of symmetry does violate Proposition 2, so it is no less problematic as a functional measure of efficiency. These examples are described in greater detail in Appendix A.

5. JudgeIt requires a model of elections to improve estimates and reduce error: I use *incumbency* (–1 Republican, 0 open, 1 Democratic), *party control* (–1 Republican, 1 Democratic), and *uncontestedness* (–1 uncontested Republican, 0 contested, 1 uncontested Democrat) as predictors, plus lagged vote share in years where no redistricting has occurred. I apply JudgeIt's default setting for uncontested races, which assigns uncontested Republicans a vote share of 0.25 and uncontested Democrats a vote share of 0.75. In redistricting years, I presume that a seat is held by the party coded as an incumbent in the data and otherwise by the party that held the same seat number before the redistricting. JudgeIt pools error across several elections to improve estimates, so I have also dropped states in years where there is not enough data for JudgeIt to function properly. After these purges, the data include 38 states, with 501 state-election year pairs.

The states are Alabama, Alaska, California, Colorado, Connecticut, Delaware, Florida, Georgia, Hawaii, Illinois, Indiana, Iowa, Kansas, Kentucky, Massachusetts, Maine, Michigan, Minnesota, Missouri, Montana, Mississippi, North Carolina, Nevada, New York, New Mexico, Ohio, Oklahoma, Oregon, Pennsylvania, Rhode Island, South Carolina, Tennessee, Texas, Utah, Virginia, Wisconsin, West Virginia, and Wyoming.

6. This necessarily assumes away differences in efficiency due to turnout. Ignoring turnout differences in this way is legally mandated for redistricting in the United States, but turnout variation is still a worthy topic of study. In fact, future research could use Equation (2) instead of Equation (5) to explore the subject. For more on turnout effects in SMD elections, see Campbell (1996) and Grofman, Koetzle, and Brunell (1997).

7. More formally, the partial derivative of ω with respect to S_{marg} is a constant of 1.0.

8. This can be seen by solving Equation (5) for seat share, which produces the equivalence $\omega + 2 * V_G - 0.5 = S_G$. I also reran the models from Table 1 with relative wasted votes, and the results confirmed these expectations.

9. For the 1992 redistricting cycle, I rely on a report from the National Conference of State Legislatures (1989) from just prior to the redistricting that spells out the process for drawing legislative plans in each state. I then use information about legislative and gubernatorial partisanship from the redistricting year to assess partisan control. For the 2002 redistricting cycle, I use the coding provided by McDonald (2004).

10. The 16 states are California, Colorado, Connecticut, Delaware, Florida, Iowa, Kansas, Michigan, Montana, Nevada, New York, Ohio, Pennsylvania, Tennessee, Utah, and Wisconsin. Though Gelman and King (1994a) refer to this subset as competitive, many of these states had mean vote shares well beyond the competitive range at some point in the study period. Thus, for many of these states, the symmetry hypothetical is probably unrealistic. Likewise, since the data here extend Gelman and King's original time series by about 14 years, many more states fall into the competitive category as they define it, further blurring the distinction.

11. The standardized versions of these lagged coefficients are virtually identical.

12. However, in the case where a gerrymanderer can only measure support as a binary concept, the Friedman and Holden (2008) model collapses on the Owen and Grofman (1988) model, and some degree of packing becomes optimal.

REFERENCES

Altman, Micah. 2002. "A Bayesian Approach to Detecting Electoral Manipulation." *Political Geography* 21: 39–48.

Born, Richard. 1985. "Partisan Intentions and Election Day Realities in the Congressional Redistricting Process." *The American Political Science Review* 79 (2): 305–19.

Brace, Kimball, Bernard Grofman, and Lisa Handley. 1987. "Does Redistricting Aimed to Help Blacks Necessarily Help Republicans?" *Journal of Politics* 49 (1): 169–85.

Brookes, Ralph Herbert. 1959. "Electoral Distortion in New Zealand." *The Australian Journal of Politics and History* 5: 218–23.

Cain, Bruce E. 1984. *The Reapportionment Puzzle*. Berkeley: University of California Press.

Cain, Bruce E. 1985. "Assessing the Partisan Effects of Redistricting." *The American Political Science Review* 79 (2): 320–33.

Campagna, Janet, and Bernard Grofman. 1990. "Party Control and Partisan Bias in 1980s Congressional Redistricting." *Journal of Politics* 52 (4): 1242–57.

Campbell, James E. 1996. *Cheap Seats: The Democratic Party's Advantage in U.S. House Elections*. Columbus: The Ohio State University Press.

Carsey, Thomas M., William D. Berry, Richard G. Niemi, Lynda W. Powell, and James M. Snyder Jr. 2008. "State Legislative Election Returns, 1967–2003." ICPSR #21480.

Chen, Jowei, and Jonathan Rodden. 2011. "Using Legislative Districting Simulations to Measure Electoral Bias in Legislatures." Presented at the annual meeting of the Midwest Political Science Association, Chicago.

Cox, Gary W., and Jonathan N. Katz. 1999. "The Reapportionment Revolution and Bias in U. S. Congressional Elections." *American Journal of Political Science* 43 (3): 812–41.

Cox, Gary W., and Jonathan N. Katz. 2002. *Elbridge Gerry's Salamander: The Electoral Consequences of the Reapportionment Revolution*. Cambridge: Cambridge University Press.

Engstrom, Erik J. 2006. "Stacking the States, Stacking the House: The Partisan Consequences of Congressional Redistricting in the 19th Century." *American Political Science Review* 100 (03): 419–27.

Erikson, Robert S. 1972. "Malapportionment, Gerrymandering, and Party Fortunes in Congressional Elections." *American Political Science Review* 66 (4): 1234–45.

Friedman, John N., and Richard T. Holden. 2008. "Optimal Gerrymandering: Sometimes Pack, But Never Crack." *American Economic Review* 98 (1): 113–44.

Gelman, Andrew, and Gary King. 1994a. "Enhancing Democracy Through Legislative Redistricting." *The American Political Science Review* 88 (3): 541–59.

Gelman, Andrew, and Gary King. 1994b. "A Unified Method of Evaluating Electoral Systems and Redistricting Plans." *American Journal of Political Science* 38 (2): 514–54.

Gelman, Andrew, Gary King, and Andrew C. Thomas. 2008. "JudgeIt II: A Program for Evaluating Electoral Systems and Redistricting Plans."

Gilligan, Thomas W., and John G. Matsusaka. 1999. "Structural Constraints on Partisan Bias Under the Efficient Gerrymander." *Public Choice* 100: 65–84.

Grofman, Bernard, and Gary King. 2007. "The Future of Partisan Symmetry as a Judicial Test for Partisan Gerrymandering after LULAC v. Perry." *Election Law Journal* 6 (1): 2–35.

Grofman, Bernard, William Koetzle, and Thomas Brunell. 1997. "An Integrated Perspective on the Three Potential Sources of Partisan Bias: Malapportionment, Turnout Differences, and the Geographic Distribution of Party Vote Shares." *Electoral Studies* 16 (4): 457–70.

Jackman, Simon. 1994. "Measuring Electoral Bias: Australia, 1949–93." *British Journal of Political Science* 24 (3): 319–57.

Kastellec, Jonathan P., Andrew Gelman, and Jamie P. Chandler. 2008. "The Playing Field Shifts: Predicting the Seats-Votes Curve in the 2008 U.S. House Elections." *PS: Political Science & Politics* 41 (4): 729–32.

Kousser, J. Morgan. 1996. "Estimating the Partisan Consequences of Redistricting Plans-Simply." *Legislative Studies Quarterly* 21 (4): 521–41.

McDonald, Michael P. 2004. "A Comparative Analysis of Redistricting Institutions in the United States, 2001–02." *State Politics & Policy Quarterly* 4 (4): 371–95.

National Conference of State Legislature Reapportionment Task Force. 1989. "Redistricting Provisions: 50 State Profiles." Washington, DC: National Conference of State Legislatures.

Niemi, Richard G., and John Deegan, Jr. 1978. "A Theory of Political Districting." *The American Political Science Review* 72 (4): 1304–23.

Niemi, Richard G., and Simon Jackman. 1991. "Bias and Responsiveness in State Legislative Districting." *Legislative Studies Quarterly* 16 (2): 183–202.

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

Schotts, Kenneth W. 2001. "The Effect of Majority-Minority Mandates on Partisan Gerrymandering." *American Journal of Political Science* 45 (1): 120–35.

Soper, C.S., and Joan Rydon. 1958. "Under-Representation and Electoral Participation." *The Australian Journal of Politics and History* 4 (1): 94–106.

Taagepera, Rein, and Matthew Soberg Shugart. 1989. *Seats and Votes: The Effects and Determinants of Electoral Systems*. New Haven, CT: Yale University Press.

Tufte, Edward R. 1973. "The Relationship Between Seats and Votes in Two-Party Systems." *The American Political Science Review* 67 (2): 540–54.

Yoshinaka, Antoine, and Chad Murphy. 2009. "Partisan Gerrymandering and Population Instability: Completing the Redistricting Puzzle." *Political Geography* 28: 451–62.

EXHIBIT A

DR. JOWEI CHEN'S ANALYSIS OF WISCONSIN'S ACT 43

I was able to apply the same computer simulation methodology and statistical tests developed in my published article cited by the Defendants to Wisconsin's Act 43.[1] I employed substantially the same computer simulation methodology to produce and analyze a large number of state assembly plans drawn using traditional districting principles. I sought to answer the following three questions:

1) What level of electoral bias emerges from a non-partisan process that draws Wisconsin's assembly districts by following the traditional districting principles of equal apportionment, preserving communities of interest (county and municipal boundaries), respecting the Voting Rights Act, and maximizing compactness?

2) How likely is such a non-partisan process to produce a state assembly districting plan with minimal electoral bias, as measured by the plan's efficiency gap?

3) How likely is such a non-partisan process to produce a districting plan with an efficiency gap similar to that of the enacted Act 43?

Executive Summary

The results of my simulation analysis, as outlined below, demonstrate that a non-partisan districting process following traditional districting principles generally produces a state assembly plan with minimal bias. In fact, 144 of the 200 random districting plans produced by the non-gerrymandered computer simulation process exhibit an efficiency gap of within 3% of zero, indicating no substantial favoring of either Democrats or Republicans.

The remaining fifty-six simulated plans exhibit an efficiency gap between -5.8% and -3.0%. Because a negative efficiency gap indicates electoral bias in favor of Republicans, these

[1] Those methodologies are described more fully in the article previously cited by Defendants and provided with their papers: Jowei Chen & Jonathan Rodden, *Unintentional Gerrymandering: Political Geography and Electoral Bias in Legislatures,* Quarterly Journal of Political Science, Vol. 8, No. 3: 239-269 (2013).

results suggest that Wisconsin's natural political geography, combined with a non-partisan process following traditional districting principles, could plausibly produce a plan with a modest amount of Republican-favoring electoral bias.

These levels of natural electoral bias pale in comparison to the much more extreme electoral bias exhibited by the Act 43 plan. The Act 43 plan exhibits a Republican-favoring efficiency gap several times that of most simulated plans, and over twice as large as even the most biased of the 200 plans produced by the non-partisan computer simulation process. In sum, statistically speaking, it is extremely unlikely that a neutral districting process, using traditional factors, would have produced a plan exhibiting electoral bias as significant as that of Act 43.

Discussion

I begin with an explanation of the logic of the districting simulation approach, followed by an overview of the simulation technique. Then, I present the results of the simulations and illustrate how the Act 43 plan is a statistical outlier.

The Logic of Redistricting Simulations

When political representation is based on winner-take-all districts, asymmetries between votes and seats can emerge merely because one party's supporters are more clustered in space than those of the other party. When this happens, the party with a more concentrated support base achieves a smaller seat share because it racks up large numbers of "surplus" votes in the districts it wins, while falling just short of the winning threshold in many of the districts it loses. This can happen quite naturally in cities due to such factors as racial segregation, housing and labor markets, transportation infrastructure, and residential sorting by income and lifestyle.

By generating a large number of randomly drawn districting plans, optimizing traditional districting criteria, the computer simulation process demonstrates the range of districting plans that would likely emerge from a neutral process. Courts and litigants can then draw inferences by comparing the partisanship of enacted plans against this range of simulated plans.

In my published academic research on legislative districting, partisan and racial gerrymandering, and electoral bias, I have developed computer simulation programming techniques that allow me to randomly produce a large number of alternative districting plans in any given state or county using precincts as building blocks. Most importantly, these computer simulations can be programmed to optimize districts with respect to any specified traditional districting goal while ignoring partisan considerations.

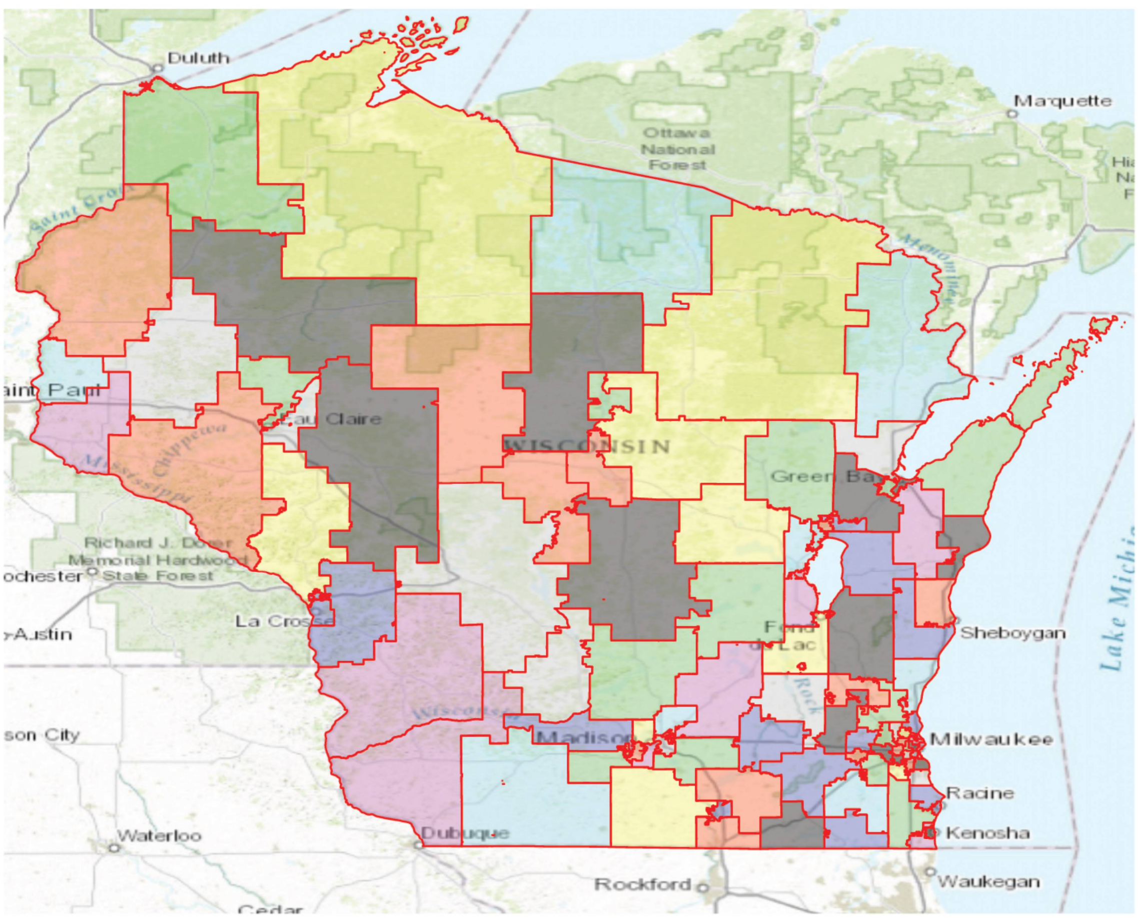
I use this simulation approach to analyze Wisconsin's Act 43 plan. First, I analyze the Legislature's districting plans and identify areas in which these enacted plans deviate significantly from equally populated districts. To analyze the Legislature's motivations for these population deviations, I use computer simulations to randomly generate two-hundred districting plans that optimize four criteria: equal apportionment within 1% of ideal district population, preservation of municipal boundaries, preservation of county boundaries, and maximization of geographic compactness. Additionally, to comply with the Voting Rights Act, I preserve the one majority-Hispanic and six majority-Black districts that were drawn in the Act 43 plan.

I then compare the computer-generated, non-partisan districting plans to the Act 43 plan using various measures of partisanship and electoral bias. The results show that computer-simulated districting plans produced by a non-partisan process preserve significantly more county and municipal boundaries than the Act 43 plan. More importantly, the simulated plans always produce significantly less electoral bias than Act 43, often resulting in a partisan efficiency gap of close to 0%. Thus, Wisconsin's Act 43 plan creates a level of electoral bias falling completely outside the range of likely outcomes under a non-partisan districting process that creates equally populated districts while maximizing compactness and preserving county and municipal boundaries.

Figure 1 illustrates an example of a computer simulated plan producing ninety-nine Assembly districts in Wisconsin. The plan displayed in Figure 1 is one of 200 complete districting plans produced by the simulation process and analyzed in this brief.

FIGURE 1

Example of a Simulated Districting Plan (Simulation #22)
20 Counties and 1,842 Municipalities Preserved Intact
Average Reock Compactness: 0.46
Efficiency Gap: +0.38% (702,384 wasted Rep. votes, 812,091 wasted Dem. votes)
41 Republican−Leaning Districts, 58 Democratic−Leaning Districts



In simulating Assembly districting plans for Wisconsin, the computer algorithm follows the following traditional districting criteria:

1) ***Equal Apportionment:*** Wisconsin's 2010 Census population was 5,686,986, so each of the 99 Assembly districts has an ideal population of 57,444.3. The computer simulation algorithm is designed to draw ninety-nine districts so that every district is within 1% of the ideal district population. As a result of this criterion, every computer simulated district produced for this analysis contains a population ranging from 56,871 to 58,017.

2) ***County Boundaries***: Wisconsin contains seventy-two counties, and Act 43 preserves fourteen of these counties intact while splitting each of the remaining fifty-eight counties into two or more Assembly districts. The left column of Table 1 lists the fourteen counties that Act 43 preserves intact.

All of the computer simulated plans preserve intact a significantly higher number of county boundaries. As Figure 2 illustrates, each of the simulated plans preserves from 18 to 25 counties intact. Though the precise set of intact counties differs from one simulated plan to the next, there are eighteen counties that are always preserved intact in 100% of the simulated plans. These counties are the first eighteen counties listed in the right column of Table 1. An additional twenty counties are preserved intact in some, but not all, of the simulated plans. These twenty additional counties are also listed in the right column of Table 1, along with the frequency with which each county is preserved intact in the simulated plans.

TABLE 1: Counties Preserved Intact in Enacted and Simulated Districting Plans

Counties Preserved Intact in Act 43 Assembly Plan	Counties Preserved Intact in Computer-Simulated Ninety-Nine -District Plans
Ashland	Ashland (100% of Simulated Plans)
Bayfield	Bayfield (100% of Simulated Plans)
Crawford	Buffalo (100% of Simulated Plans)
Door	Burnett (100% of Simulated Plans)
Florence	Crawford (100% of Simulated Plans)
Grant	Door (100% of Simulated Plans)
Iron	Florence (100% of Simulated Plans)
Kewaunee	Forest (100% of Simulated Plans)
Lincoln	Grant (100% of Simulated Plans)
Menominee	Iron (100% of Simulated Plans)
Pepin	Kewaunee (100% of Simulated Plans)
Price	Lincoln (100% of Simulated Plans)
Rusk	Marquette (100% of Simulated Plans)
Taylor	Menominee (100% of Simulated Plans)
	Pepin (100% of Simulated Plans)
	Price (100% of Simulated Plans)
	Rusk (100% of Simulated Plans)
	Taylor (100% of Simulated Plans)
	Marinette (58% of Simulated Plans)
	Douglas (51% of Simulated Plans)
	Vilas (38% of Simulated Plans)
	Lafayette (21% of Simulated Plans)
	Oneida (16% of Simulated Plans)
	Green Lake (15% of Simulated Plans)
	Langlade (12% of Simulated Plans)
	Richland (10% of Simulated Plans)
	Trempealeau (5% of Simulated Plans)
	Polk (4% of Simulated Plans)
	Vernon (3% of Simulated Plans)
	Sawyer (3% of Simulated Plans)
	Dunn (3% of Simulated Plans)
	Waushara (2% of Simulated Plans)
	Pierce (2% of Simulated Plans)
	Barron (2% of Simulated Plans)
	Iowa (1% of Simulated Plans)
	Green (1% of Simulated Plans)
	Clark (1% of Simulated Plans)
	Adams (1% of Simulated Plans)

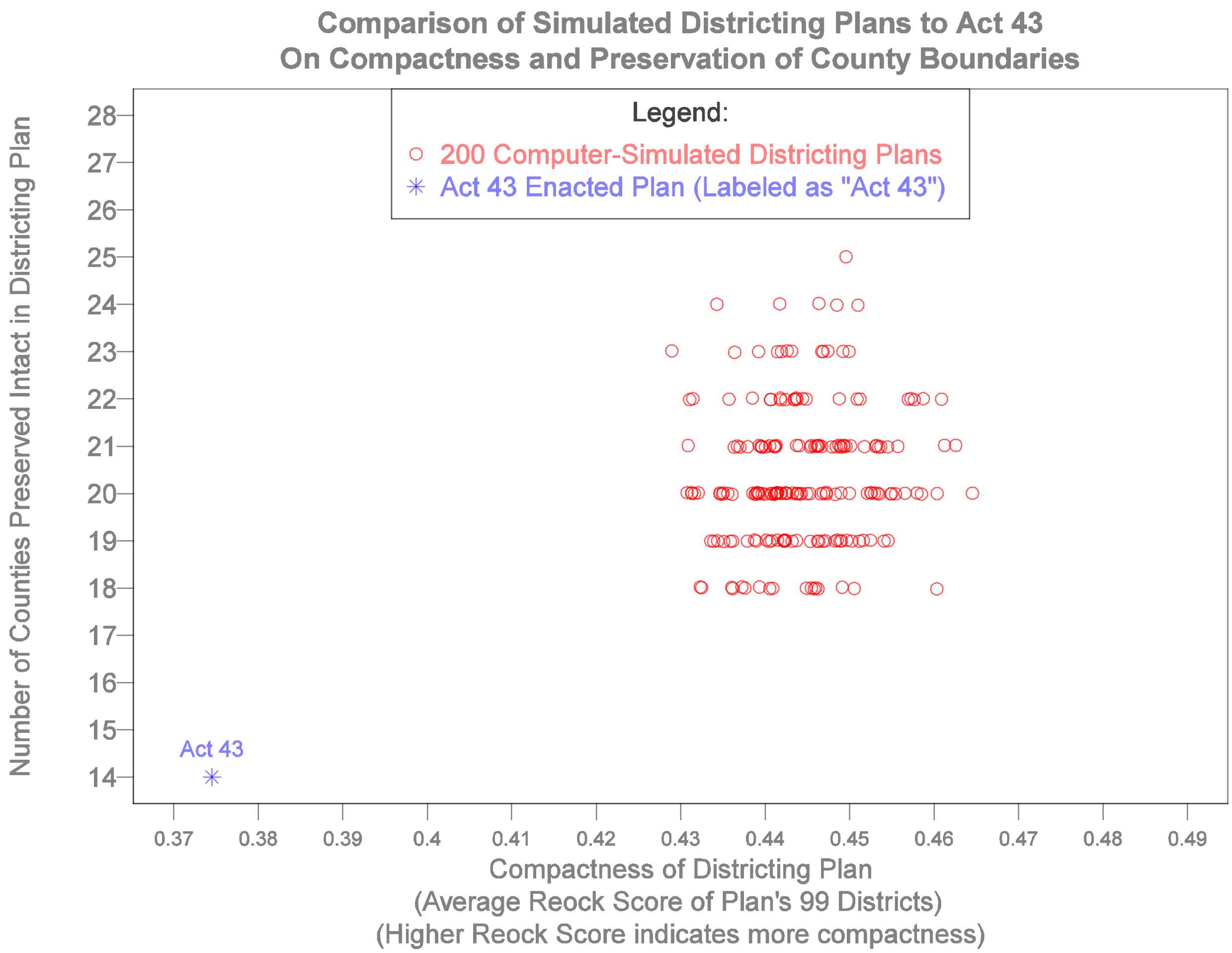
3) ***Municipalities Boundaries***: Wisconsin contains a total of 1,896 municipalities, which include cities, towns, and villages. For purposes of counting municipal splits, I treated each Census Minor Civil Division (MCD) as a separate municipality, even if two MCD's have the same name.

Act 43 preserves intact the boundaries of 1,825 municipalities. All of the computer simulated plans preserve intact a significantly higher number of municipalities. The number of municipalities preserved intact in the simulations ranges from 1,837 to 1,853.

4) ***Geographic Compactness***: Beyond preserving county and municipal boundaries, the simulation algorithm prioritizes the drawing of geographically compact districts. Compactness is quantifiably measured by a Reock score for each district in any given plan. The Reock score of a district is calculated by first drawing the smallest possible bounding circle that completely encloses the district's borders; hence, the bounding circle will always be at least as large as the district itself. The Reock score is then calculated as the ratio of the district's area to the area of the bounding circle. Therefore, the Reock score will always be a fraction less than or equal to one, with a higher Reock score indicating a more compact district. The Reock score for an entire plan is then calculated as the average score for the ninety-nine Assembly districts within the plan.

To compare the compactness of the computer simulated plans and the Act 43 Assembly plan, the horizontal axis of Figure 2 measures the Reock score for the Act 43 Assembly plan as well as the 200 computer simulated plans. Figure 2 illustrates that plans produced by the partisan-neutral computer simulation process are always significantly more compact than the Act 43 Assembly plan. While the Act 43 plan has a Reock score of 0.37, the 200 computer simulated plans exhibit Reock scores ranging from 0.43 to 0.46, indicating that 100% of the simulated plans are substantially more compact than the plan enacted by the Wisconsin Legislature.

FIGURE 2



5) ***The Voting Rights Act:*** Act 43 produces one majority-Hispanic district (Assembly District 8) and six majority-African-American districts (Assembly Districts 10, 11, 12, 16, 17, and 18). To comply with the Voting Rights Act, the computer simulated plans preserve each of these seven majority-minority districts exactly as they were drawn in the Act 43 plan. In other words, these seven districts from Act 43 appear in each of my computer-simulated plans exactly as they were drawn by the Wisconsin Legislature with no modifications.

Simulation Results

The following describes the simulation results and inferences about Act 43.

Efficiency Gap: To calculate the efficiency gap of Act 43 and of each simulated plan, I first calculate the partisanship of each simulated district and each Act 43 Assembly district by calculating Republican Mitt Romney's share of the two-party presidential vote in November 2012 within each district. Using Mitt Romney and Barack Obama votes as a simple measure of district partisanship, I then calculate the districting plan's efficiency gap using the method outlined in *Partisan Gerrymandering and the Efficiency Gap*.[2] Districts are classified as Republican victories if Romney votes exceeded Obama votes in November 2012 and as Democratic victories if Obama garnered more votes than Romney. For each party, I then calculate the total sum of surplus votes in districts the party won and lost votes in districts where the party lost. The efficiency gap is then calculated as total wasted Republican votes minus total wasted Democratic votes, divided by the total number of two-party votes cast statewide.

Figure 3 illustrates the efficiency gap of the two-hundred simulated Assembly districting plans produced using the traditional districting criteria described in the previous section, and of Act 43. Each red circle in Figure 3 represents a complete simulated districting plan, with its efficiency gap measured along the horizontal axis. The vertical axis measures the total number of counties preserved intact by the plan, a number that, as noted above, ranges from eighteen to twenty-five counties for each simulation.

[2] Nicholas O. Stephanopoulos & Eric M. McGhee, *Partisan Gerrymandering and the Efficiency Gap,* 82 University of Chicago Law Review 831 (2015).

FIGURE 3

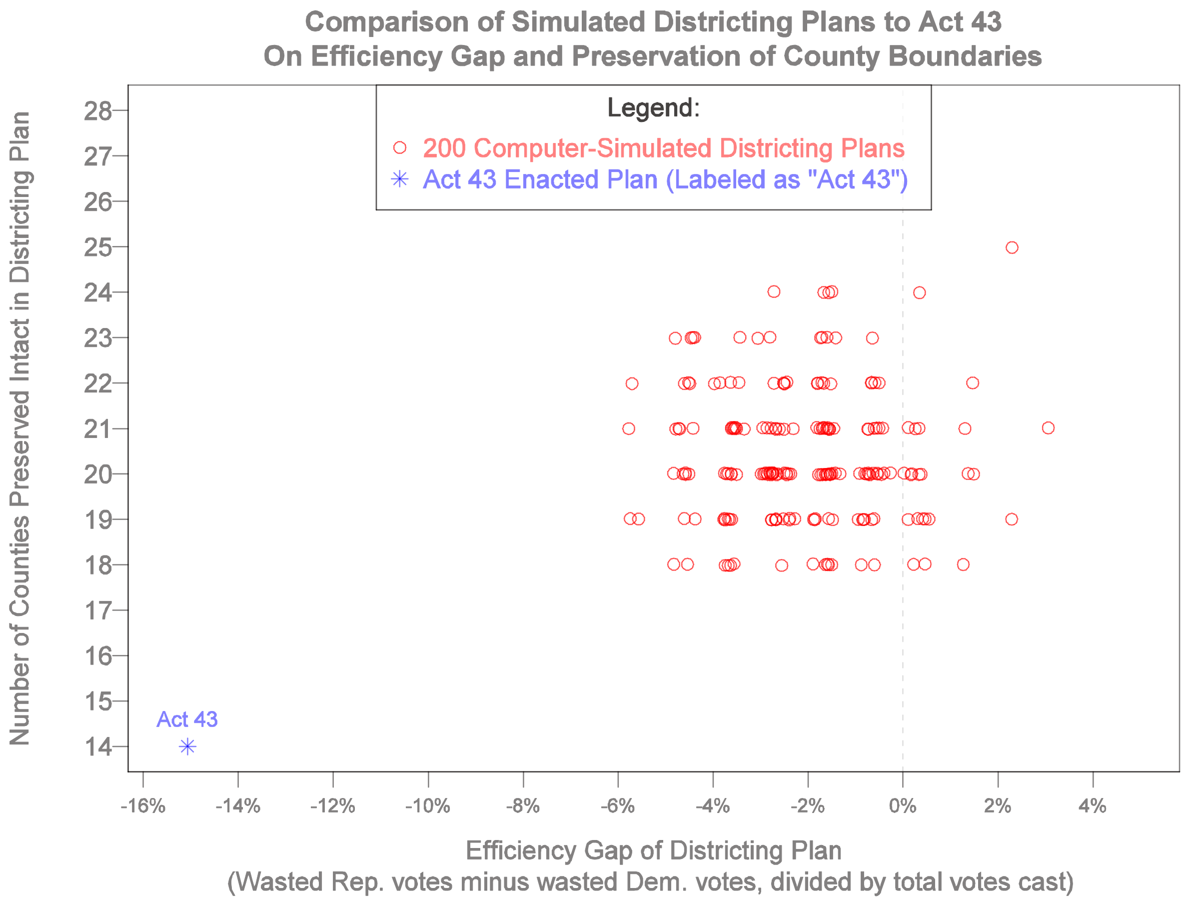


Figure 3 reveals that the simulated districting plans are reasonably neutral with respect to electoral bias. About 72% of the simulated plans exhibit an efficiency gap within 3% of zero, indicating de minimis electoral bias in favor of either party. In fact, 23% of the simulations produce an efficiency gap between -1.0% and +1.0%. These patterns illustrate that a non-partisan districting process following traditional criteria very commonly produces a neutral Assembly plan in Wisconsin with minimal electoral bias.

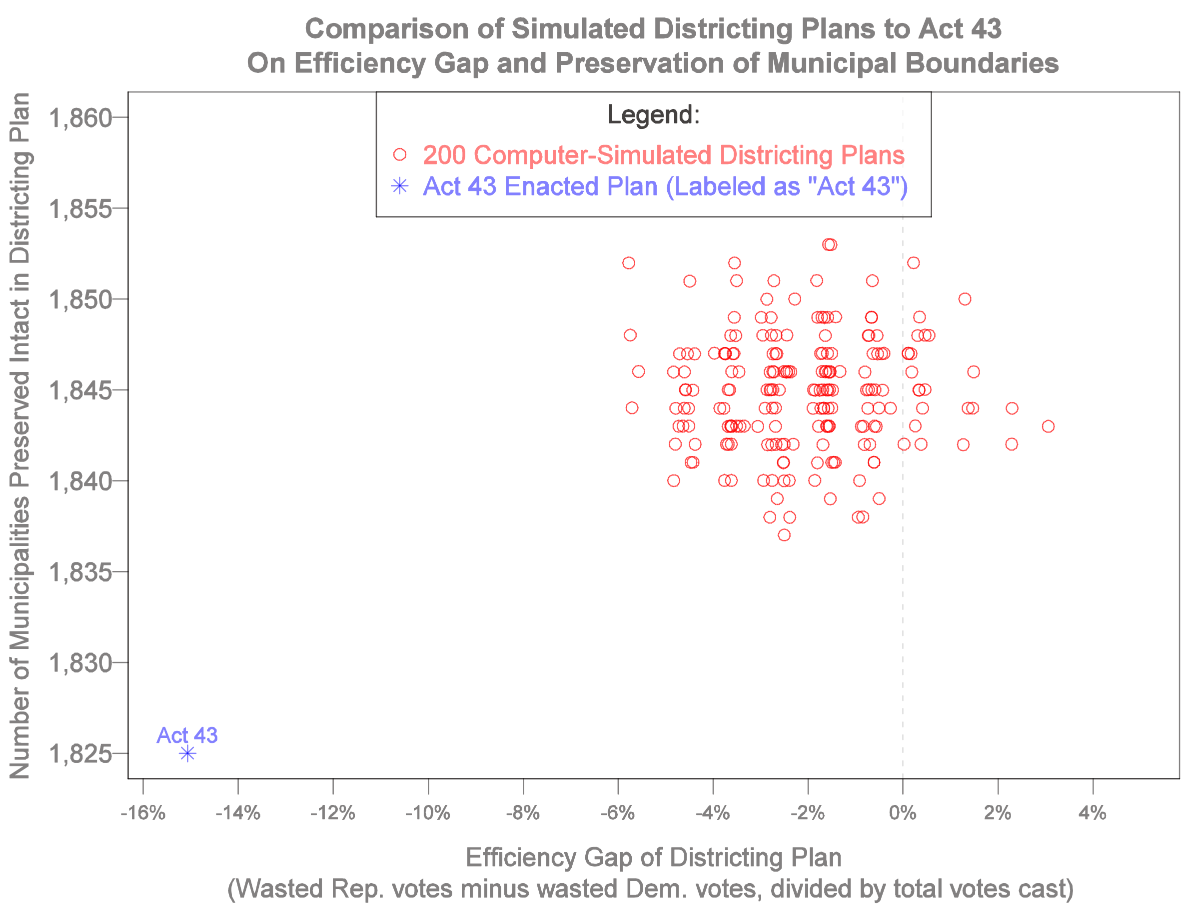
It is important to note that the simulations produce plans with both positive and negative efficiency gaps. Although the efficiency gap of every simulated plan is relatively small in magnitude, 90% of plans exhibit a negative efficiency gap, indicating slightly more wasted Democratic votes than wasted Republican votes. But 10% of the plans exhibit a positive efficiency gap, reflecting more wasted Republican votes. Hence, it is not extraordinary for Wisconsin's political geography, combined with traditional redistricting criteria, to naturally produce a districting plan that somewhat favors Republicans.

The blue star in the lower left corner of Figure 3 represents the Assembly plan enacted by Act 43. This blue star depicts the enacted plan's efficiency gap of -15.1%, reflecting significantly more wasted Democratic votes than wasted Republican votes. Thus, the level of electoral bias in the Act 43 Assembly plan is not only entirely outside of the range produced by the simulated plans, the enacted plan's efficiency gap is well over twice as biased as the most biased of the two-hundred simulated plans. The improbable nature of the Act 43 efficiency gap allows us to conclude with high statistical certainty that neutral, non-partisan districting criteria, combined with Wisconsin's natural political geography, would not have produced a districting plan as electorally skewed as the Act 43 Assembly plan.

Figure 3 additionally illustrates that the Act 43 plan preserves intact far fewer counties than would have been reasonably possible under a neutral process prioritizing traditional districting criteria. The Act 43 plan keeps intact only fourteen of Wisconsin's seventy-two counties. Meanwhile, each of the simulated plans preserves eighteen to twenty-five counties fully intact. Figure 3 suggests a possible connection between Act 43 plan's creation of an extreme efficiency gap and the plan's splitting up of far more counties than what could have been reasonably expected under a partisan-neutral districting process.

Figure 4 illustrates the same patterns regarding the splitting of municipal boundaries. As before, the horizontal axis of Figure 4 measures the efficiency gap of the simulated plans and the Act 43 Assembly plan. The vertical axis in Figure 4 measures the number of municipalities kept intact within each plan. Figure 4 illustrates that Act 43 is a statistical outlier not only in terms of its large, Republican-favoring efficiency gap, but also in its splitting of far more municipalities than any of the simulated plans.

FIGURE 4



As an additional measure of the partisanship of each plan, Figure 5 reports the number of Republican-leaning districts—defined as districts in which Romney voters outnumbered Obama voters in November 2012—in each plan. The horizontal axis in Figure 5 measures the number of

Republican districts (out of the ninety-nine Assembly districts) created by each simulated plan and by the Act 43 Assembly plan. The vertical axis measures the number of counties preserved intact by each plan. As before, red circles denote the two-hundred computer simulated plans, while the blue star represents the Act 43 plan.

FIGURE 5

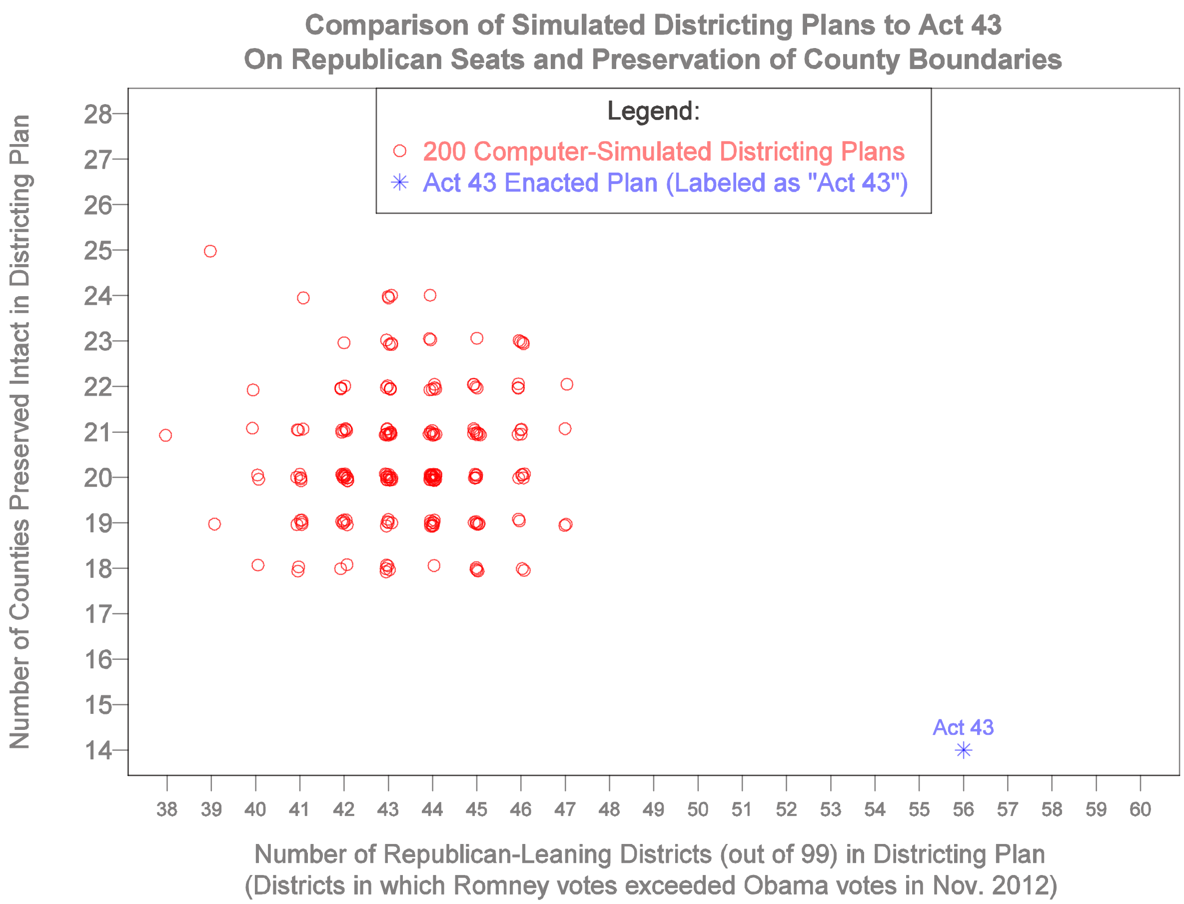


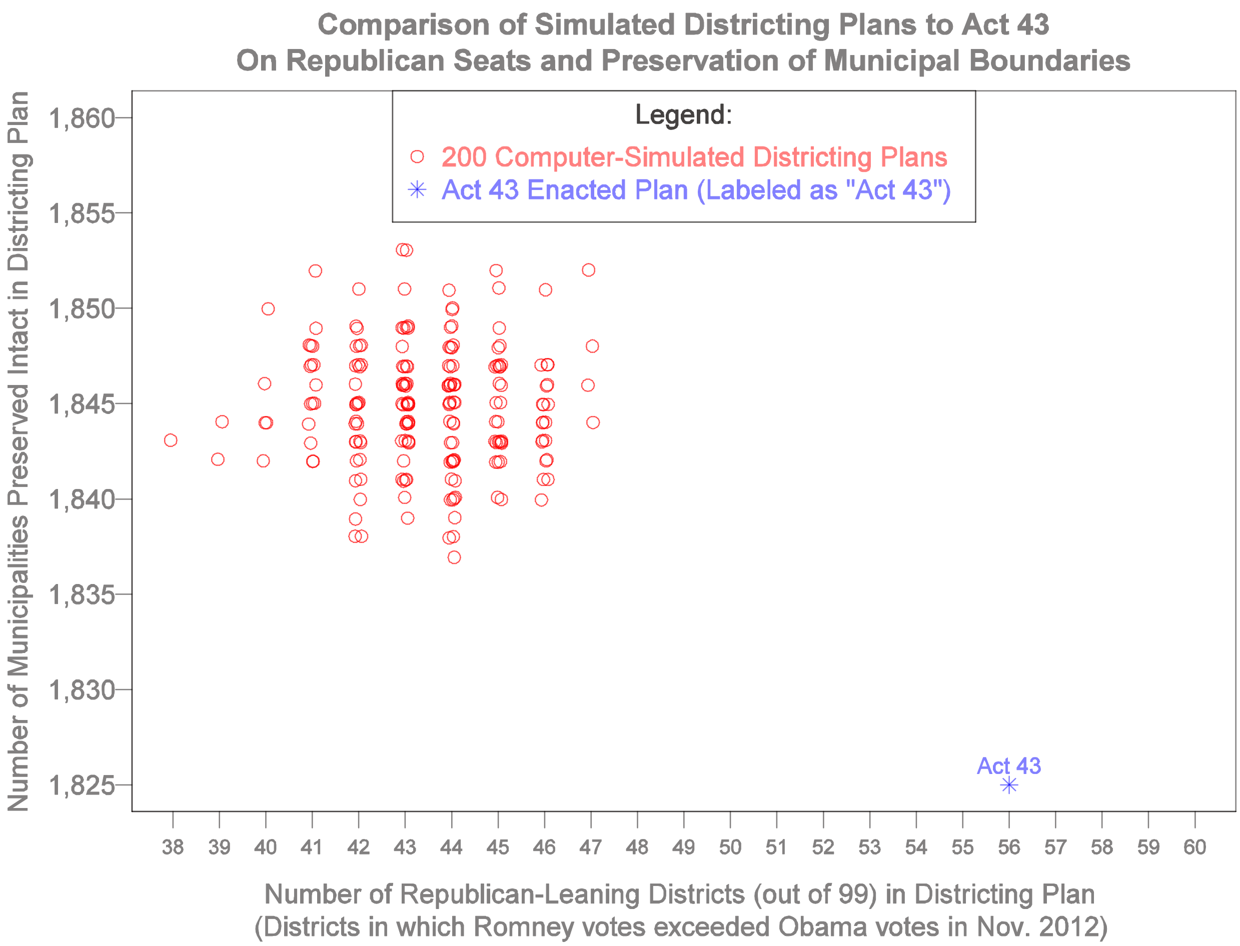
Figure 5 illustrates the contrast between the simulated plans and the Act 43 plan in terms of their partisan division of Assembly seats. In the simulated plans (drawn in a non-partisan manner respecting traditional districting criteria), between thirty-eight and forty-seven districts contain more Republican than Democratic voters. This range translates to a 38.4% to 47.4%

Republican share of the ninety-nine total Assembly districts, a range consistent with and reflective of the Republicans' statewide 46.5% share of the November 2012 presidential vote.

Yet the Act 43 plan creates a total of fifty-six Republican-leaning Assembly districts, as measured by 2012 presidential vote share. This total is far outside of the range of partisan outcomes observed in the simulations, indicating that the Act 43 plan was the product of an intentional effort to craft more Republican-leaning districts than was possible under a partisan-neutral map-drawing process following traditional districting criteria. As before, the fact that Act 43 preserved intact far fewer counties than any of the simulated plans suggests that the Act 43 Assembly plan had to violate the traditional districting principle of respecting county boundaries in order to achieve fifty-six Republican-leaning districts, an extremely improbable outcome.

Figure 6 illustrates the same pattern regarding the splitting of municipal boundaries. As in Figure 5, the horizontal axis of Figure 6 measures the number of Republican districts (out of ninety-nine) created by each simulated plan and by the Act 43 Assembly plan. But the vertical axis in Figure 6 measures the number of municipalities kept intact within each plan. This Figure illustrates that the Act 43 plan's statistically extreme creation of fifty-six Republican districts came at the expense of preserving far fewer municipalities intact than were reasonably possible under the partisan-neutral process followed by the computer simulations.

FIGURE 6

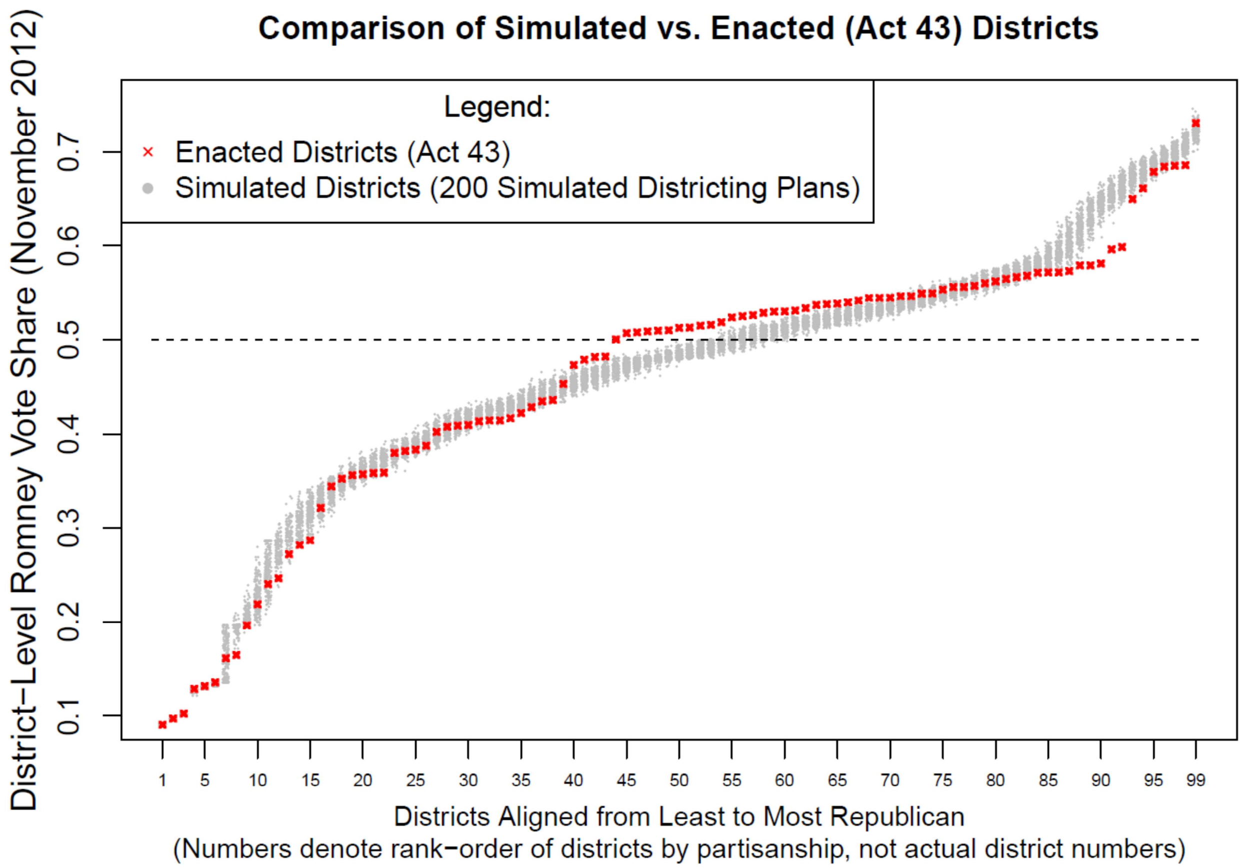


How did Act 43 create such a statistically improbable Assembly plan in terms of its partisan division of seats? Figure 7 provides suggestive evidence. Figure 7 displays the partisanship, measured by the Romney share of the November 2012 vote, of every single district in all simulated districting plans and the enacted Act 43 plan. The vertical axis measures each district's partisanship, with gray dots representing simulated districts and red stars representing the ninety-nine Assembly districts created under Act 43.

Figure 7 contains a total of ninety-nine columns. For each simulated plan and for the Act 43 plan, the ninety-nine districts are aligned from left to right by partisanship. In other words, the left-most red star represents the most Democratic-leaning Act 43 district (Assembly District 16, in which Romney won 9.0% of the presidential vote), while the right-most red star represents the most Republican-leaning Act 43 district (Assembly District 99, in which Romney won 73.1% of the presidential vote). The gray dots representing districts for each simulated plan are similarly aligned by partisanship across the ninety-nine columns in Figure 7.

Overall, Figure 7 allows comparison of the enacted and the simulated districting plans with respect to their distribution of partisanship across districts. Most strikingly, Figure 7 illustrates how Act 43 created its unusually large sum of fifty-six Republican-leaning districts. As illustrated in the middle portion of Figure 7, Act 43 created eleven Republican-leaning districts that would instead have been Democratic-leaning districts when drawn by the partisan-neutral simulation process. This creation is evidenced by the noticeable divergence of the red stars away from the entire range of gray circles in the middle portion of Figure 7. In order to convert these Democratic-leaning districts into Republican-leaning districts, the Act 43 plan appears to have pulled Republican voters away from what would otherwise have been more heavily Republican districts, as illustrated in the far right portion of Figure 7. In the right-most fifteen columns in Figure 7, the red stars often fall under the entire range of gray circles, showing that Act 43 unpacked some Republican voters from these safe Republican districts and placed more Republican votes into what would otherwise have been slightly Democratic districts, tipping them into Republican-leaning districts.

FIGURE 7



Conclusion

Using computer simulations to generate a large baseline sample of legally valid districting plans under a partisan-neutral map-drawing process following traditional districting criteria, we find that drawing a minimally biased Assembly map is reasonably possible. The results show that the non-partisan simulation process successfully produces valid districting plans with a neutral efficiency gap with striking frequency.

Furthermore, we are able to discover not merely the ways in which the enacted Act 43 plan deviates from traditional districting criteria, but also the partisan consequences of such deviations. Act 43 not only created an extremely biased Assembly plan with an efficiency gap far outside of any gap observed in 200 simulations, the enacted plan achieved this partisan outcome at the expense of traditional districting principles, splitting apart far more counties and municipalities than were necessary.

Respectfully submitted,

Jowei Chen

March 17, 2016

Assessing the Current Wisconsin State Legislative Districting Plan

Simon Jackman

July 7, 2015

EXHIBIT 3

Contents

1 Introduction

My name is Simon Jackman. I am currently a Professor of Political Science at Stanford University, and, by courtesy, a Professor of Statistics. I joined the Stanford faculty in 1996. I teach classes on American politics and statistical methods in the social sciences.

I have been asked by counsel representing the plaintiffs in this lawsuit (the "Plaintiffs") to analyze relevant data and provide expert opinions in the case titled above. More specifically, I have been asked

- to determine if the current Wisconsin legislative districting plan constitutes a partisan gerrymander;
- to explain a summary measure of a districting plan known as "the efficiency gap" (Stephanopolous and McGhee, 2015), what it measures, how it is calculated, and to assess how well it measures partisan gerrymandering;
- to compare the efficiency gap to extant summary measures of districting plans such as partisan bias;
- to analyze data from state legislative elections in recent decades, so as to assess the properties of the efficiency gap and to identify plans with high values of the efficiency gap;
- to suggest a threshold or other measure that can be used to determine if a districting plan is an extreme partisan gerrymander;
- to describe how the efficiency gap for the Wisconsin districting plan compares to the values of the efficiency gap observed in recent decades elsewhere in the United States;
- to describe where the efficiency gap for the current Wisconsin districting plan lies in comparison with the threshold for determining if a districting plan constitutes an extreme partisan gerrymander.

My opinions are based on the knowledge I have amassed over my education, training and experience, and follow from statistical analysis of the following data:

- a large, canonical data set on candidacies and results in state legislative elections, 1967 to the present available from the Inter-University Consortium for Political and Social Research (ICPSR study number 34297); I use a release of the data updated through 2014, maintained by Karl Klarner (Indiana State University and Harvard University).
- presidential election returns, 2000-2012, aggregated to state legislative districts.

2 Qualifications, Publications and Compensation

My Ph.D. is in Political Science, from the University of Rochester, where my graduate training included courses in econometrics and statistics. My curriculum vitae is attached to this report.

All publications that I have authored and published in the past ten years appear in my curriculum vitae. Those publications include peer-reviewed journals such as: *The Journal of Politics*, *Electoral Studies*, *The American Journal of Political Science*, *Legislative Studies Quarterly*, *Election Law Journal*, *Public Opinion Quarterly*, *Journal of Elections, Public Opinion and Parties*, and *PS: Political Science and Politics*.

I have published on properties of electoral systems and election administration in *Legislative Studies Quarterly*, the *Australian Journal of Political Science*, the *British Journal of Political Science*, and the *Democratic Audit of Australia*. I am a Fellow of the Society for Political Methodology and a member of the American Academy of Arts and Sciences.

I am being compensated at a rate of $250 per hour.

3 Summary

1. **Partisan gerrymandering and wasted votes.** In two-party, single-member district electoral systems, a partisan gerrymander operates by effectively "wasting" more votes cast for one party than for the other. Wasted votes are votes for a party in excess of what the party needed to win a given district or votes cast for a party in districts that the party doesn't win. Differences

in wasted vote rates between political parties measure the extent of partisan gerrymandering.

2. **The efficiency gap** (*EG*) is a relative, wasted vote measure, the ratio of one party's wasted vote rate to the other party's wasted vote rate. *EG* can be computed directly from a given election's results, without recourse to extensive statistical modeling or assumptions about counter-factual or hypothetical election outcomes, unlike other extant measures of the fairness of an electoral system (e.g., partisan bias).

3. The efficiency gap is an "excess seats" measure, reflecting the nature of a partisan gerrymander. An efficiency gap in favor one party sees it wasting fewer votes than its opponent, thus translating its votes across the jurisdiction into seats more efficiently than its opponent. This results in the party winning more seats than we'd expect given its vote share (V) and if wasted vote rates were the same between the parties. $EG = 0$ corresponds to no efficiency gap between the parties, or no partisan difference in wasted vote rates. In this analysis (but without loss of generality) *EG* is normed such that negative *EG* values indicate higher wasted vote rates for Democrats relative to Republicans, and $EG > 0$ the converse.

4. A districting plan in which *EG* is consistently observed to be positive is evidence that the plan embodies a pro-Democratic gerrymander; the magnitudes of the *EG* measures speak to the severity of the gerrymander. Conversely, a districting plan with consistently negative values of the efficiency gap is consistent with the plan embodying a pro-Republican gerrymander.

5. **Performance of the efficiency gap in 786 state legislative elections.** My analysis of 786 state legislative elections (1972-2014) examines properties of the efficiency gap. *EG* is estimated with some uncertainty in the presence of uncontested districts (and uncontested districts are quite prevalent in state legislative elections), but this source of uncertainty is small relative to differences in the *EG* across states and across districting plans.

6. **Stability of the efficiency gap.** *EG* is stable in pairs of temporally adjacent elections held under the same districting plan. In 580 pairs of consecutive

EG measures, the probability that each *EG* measure has the same sign is 74%. In 141 districting plans with three or more elections, 35% have a better than 95% probability of *EG* being negative or positive for the entire duration of the plan; in about half of the districting plans the probability that *EG* doesn't change sign is above 75%.

7. **Recent decades show more pro-Republican gerrymandering, as measured by the efficiency gap.** Efficiency gap measures in recent decades show a pronounced shift in a negative direction, indicative of an increased prevalence of districting plans favoring Republicans. Among the 10 most pro-Democratic *EG* measures in my analysis, *none* were recorded after 2000.

8. **The current Wisconsin state legislative districting plan** (the "Current Wisconsin Plan"). In Wisconsin in 2012, the average Democratic share of district-level, two-party vote (V) is estimated to be 51.4% ($\pm$0.6, the uncertainty stemming from imputations for uncontested seats); recall that Obama won 53.5% of the two-party presidential vote in Wisconsin in 2012. Yet Democrats won only 39 seats in the 99 seat legislature (S = 39.4%), making Wisconsin one of 7 states in 2012 where we estimate $V > 50\%$ but $S < 50\%$. In Wisconsin in 2014, V is estimated to be 48.0% ($\pm$0.8) and Democrats won 36 of 99 seats (S = 36.4%).

9. Accordingly, Wisconsin's *EG* measures in 2012 and 2014 are large and negative: -.13 and -.10 (to two digits of precision). The 2012 estimate is the largest *EG* estimate in Wisconsin over the 42 year period spanned by this analysis (1972-2014).

10. Among 79 *EG* measures generated from state legislative elections after the 2010 round of redistricting, Wisconsin's *EG* scores rank 9th (2012, 95% CI 4 to 13) and 18th (2014, 95% CI 14 to 21). Among 786 *EG* measures in the 1972-2014 analysis, the magnitude of Wisconsin's 2012 *EG* measure is surpassed by only 27 (3.4%) other cases.

11. Analysis of efficiency gaps measures in the post-1990 era indicates that conditional on the magnitude of the Wisconsin 2012 efficiency gap (the first election under the Current Wisconsin Plan), there is a 100% probability

that *all subsequent elections* held under that plan will also have efficiency gaps disadvantageous to Democrats.

12. **The Current Wisconsin Plan presents overwhelming evidence of being a pro-Republican gerrymander**. In the entire set of 786 state legislative elections and their accompanying *EG* measures, there are *no precedents* prior to this cycle in which a districting plan generates an initial two-election sequence of *EG* scores that are each as large as those observed in *WI*.

13. The Current Wisconsin Plan is generating *EG* measures that make it *extremely likely* that it has a systematic, historically large and enduring, pro-Republican advantage in the translation of votes into seats in Wisconsin's state legislative elections.

14. **An actionable threshold based on the efficiency gap.** Historical analysis of the relationship between the first *EG* measure we observe under a new districting plan and the subsequent *EG* measures lets us assess the extent to which that first *EG* estimate is a *reliable* indicators of a *durable* and hence *systematic* feature of the plan. In turn, this let us assess the *confidence* associated with a range of possible *actionable EG thresholds*.

15. My analysis suggests that *EG* greater than .07 in absolute value be used as an actionable threshold. Relatively few plans produce a first election with an *EG* measure in excess of this threshold, and of those that do, the historical analysis suggests that most go on to produce a sequence of *EG* estimates indicative of systematic, partisan advantage consistent with the first election *EG* estimates, At the 0.07 threshold, 95% of plans would be either (a) undisturbed by the courts, or (b) struck down because we are sufficiently confident that the plan, if left undisturbed, would go on to produce a one-sided sequence of *EG* estimates, consistent with the plan being a partisan gerrymander. In short, our "confidence level" in the 0.07 threshold is 95%.

16. **The Current Wisconsin Plan is generating estimates of the efficiency gap far in excess of this proposed, actionable threshold**. In 2012 elections to the Wisconsin state legislature, the efficiency gap is estimated to be -.13; in

2014, the efficiency gap is estimated to be -.10. Both measures are separately well beyond the conservative .07 threshold suggested by the analysis of efficiency gap measures observed from 1972 to the present.

A vivid, graphical summary of my analysis appears in Figure 1, showing the average value of the efficiency gap in 206 districting plans, spanning 41 states and 786 state legislative elections from 1972 to 2014. The Current Wisconsin Plan has been in place for two elections (2012 and 2014), with an average efficiency gap of -.115. Details on the interpretation and calculation of the efficiency gap come later in my report, but for now note that negative values of the efficiency gap indicate a districting plan favoring Republicans, while positive values indicate a plan favoring Democrats. Note that *only four other districting plans have lower average efficiency gap scores than the Current Wisconsin Plan*, and these are also from the post-2010 round of redistricting. That is, Wisconsin's current plan is generating the 5th lowest average efficiency gap observed in over 200 other districting plans used in state legislative elections throughout the United States over the last 40 years. The analysis I report here documents why the efficiency gap is a valid and reliable measure of partian gerrymandering and why are confident that the current Wisconsin plan exceeds even a conservative definition of partisan gerrymandering.

4 Redistricting plans

A districting plan is an exercise in map drawing, partitioning a jurisdiction into districts, typically required to be contiguous, mutually exclusive and exhaustive regions, and — at least in the contemporary United States — of approximately the same population size. In a single-member, simple plurality (SMSP) electoral system, the highest vote getter in each district is declared the winner of the election. Partisan gerrymandering is the process of drawing districts that favor one party, typically by creating a set of districts that help the party win an excess of seats (districts) relative to its jurisdiction-wide level of support.

What might constitute evidence of partisan gerrymandering? One indication might be a series of elections conducted under the same districting plan in which a party's seat share (S) is unusually large (or small) relative to its vote share (V).

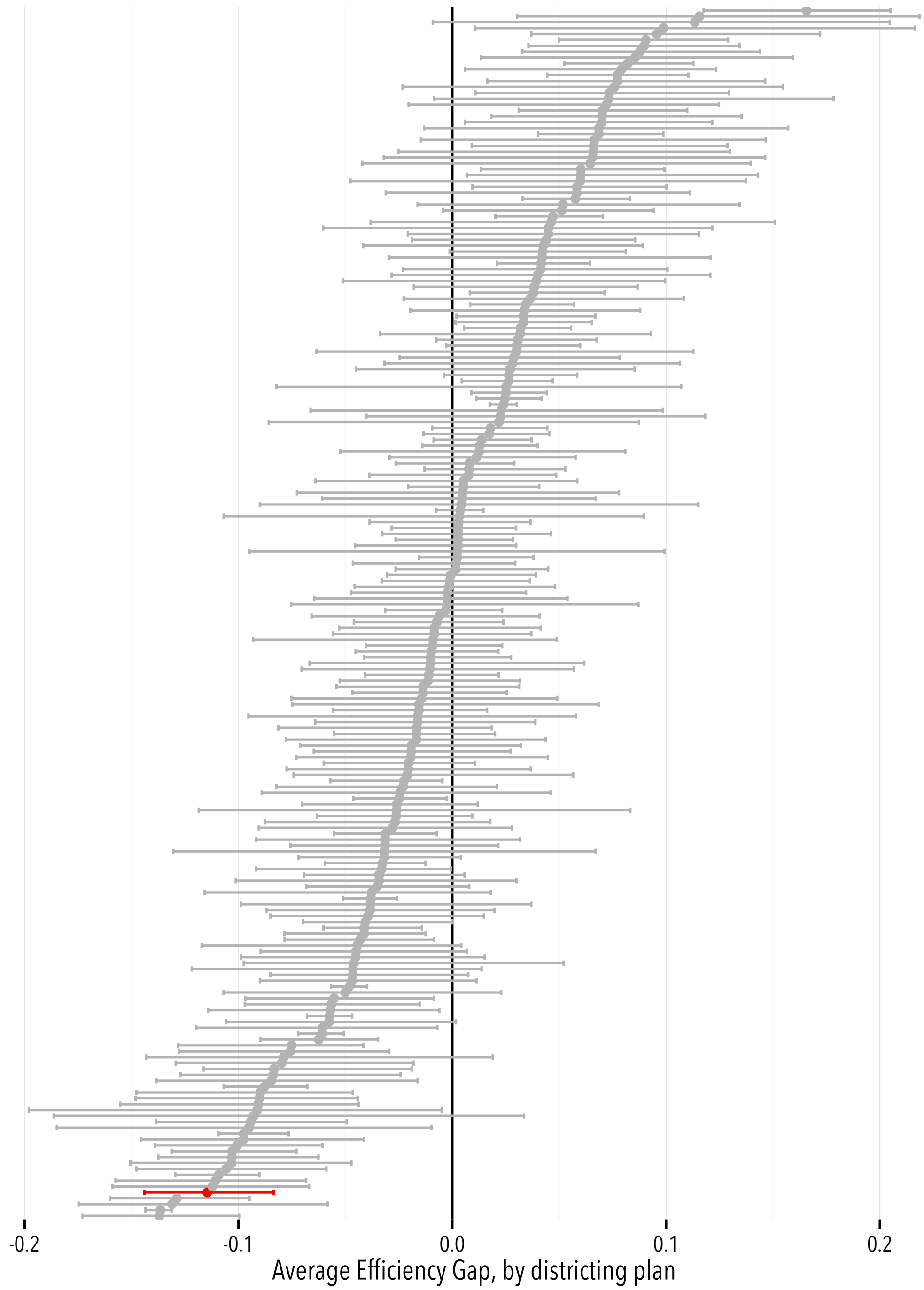


Figure 1: Average efficiency gap score, 206 districting plans, 1972-2014. Plans have been sorted from low average *EG* scores to high. Horizontal lines cover 95% confidence intervals. Negative efficiency gap scores are plans that disadvantage Democrats; positive efficiency gap scores favor Democrats. The Current Wisconsin Plan is shown in red. See also Figure 36.

There may be elections where a party wins a majority of seats (and control of the jurisdiction's legislature) despite not winning a majority of votes: $S > .5$ while $V < .5$ and vice-versa. In fact, there are numerous instances of mismatches between the party winning the statewide vote and the party controlling the state legislature in recent decades. I estimate that since 1972 there have been 63 cases of Democrats winning a majority of the vote in state legislative elections, while not winning a majority of the seats, and 23 cases of the reverse phenomenon, where Democrats won a majority of the seats with less than 50% of the statewide, two-party vote.

Geographic clustering of partisans is typically a prerequisite for partisan gerrymandering. This is nothing other than partisan "packing": a gerrymandered districting plan creates a relatively small number of districts that have unusually large proportions of partisans from party *B*. The geographic concentration of party *B* partisans might make creating these districts a straightforward task. In other districts in the jurisdiction, party *B* supporters never (or seldom) constitute a majority (or a plurality), making those districts "safe" for party *A*. This districting plan helps ensure party *A* wins a majority of seats even though party *B* has a majority of support across the jurisdiction, or at the very least, the districting plan helps ensures that party *A*'s seat share exceeds its vote share in any given election.

It is conventional in political science to say that such a plan allows party *A* to "more efficiently" translate its votes into seats, relative to the way the plan translates party *B*'s votes into seats. This nomenclature is telling, as we will see when we consider the *efficiency gap* measure, below.

Assessing the partisan fairness of a districting plan is fundamentally about measuring a party's excess (or deficit) in its seat share relative to its vote share. The efficiency gap is such a summary measure. To assess the properties of the efficiency gap, I first review some core concepts in the analysis of districting plans: vote shares, seat shares, and the relationship between the two quantities in single-member districts.

4.1 Seats-Votes Curves

Electoral systems translate parties' vote shares (V) into seat shares (S). Both V and S are proportions. Plotting the two quantities V and S against one another yields the "seats-votes" curve, a staple in the analysis of electoral systems and districting plans. Two seats-votes curves are shown in Figure 2, one showing a non-linear relationship between seats and votes typical of single-member district systems,[1] the other showing a linear relationship between seats and votes observed under proportional representation systems.

In pure proportional representation (PR) voting systems, seats-votes curves are 45 degree lines by design, crossing the $(V, S) = (.5, .5)$ point: i.e., under PR, $S = V$ and a party that wins 50% of the vote will be allocated 50% of the seats. Absent a deterministic allocation rule like pure PR, seats-votes curves are most usefully thought of in probabilistic terms, due to the fact that there are many possible configurations of district-specific outcomes corresponding to a given jurisdiction-wide V, and hence uncertainty — represented by a probability *distribution* — over possible values of S given V.

In single-member, simple plurality (SMSP) systems, we often see non-linear, "S"-shaped seats-votes curves. With an approximately symmetric mix of districts (in terms of partisan leanings), large changes in seat shares (S) can result from relatively small changes in votes shares (V) at the middle of the distribution of district types. This presumes a districting plan such that both parties have a small number of "strongholds," with extremely large changes in vote shares needed to threaten these districts, and so the seats-votes curve tends to "flatten out" as jurisdiction-wide vote share (V) takes on relatively large or small values. Other shapes are possible too: e.g., bipartisan, incumbent-protection plans generate seats-votes curves that are largely flat for most values of V, save for the constraint that the curve run through the points $(V, S) = (0, 0)$ and $(1, 1)$; i.e., relatively large movements in V generates relatively little change in seats shares.

[1]The curve labeled "Cube Law" in Figure 2 is generated assuming that $S/(1-S) = [V/(1-V)]^3$, an approximation for the lack of proportionality we observe in single-member district systems, though hardly a "law."

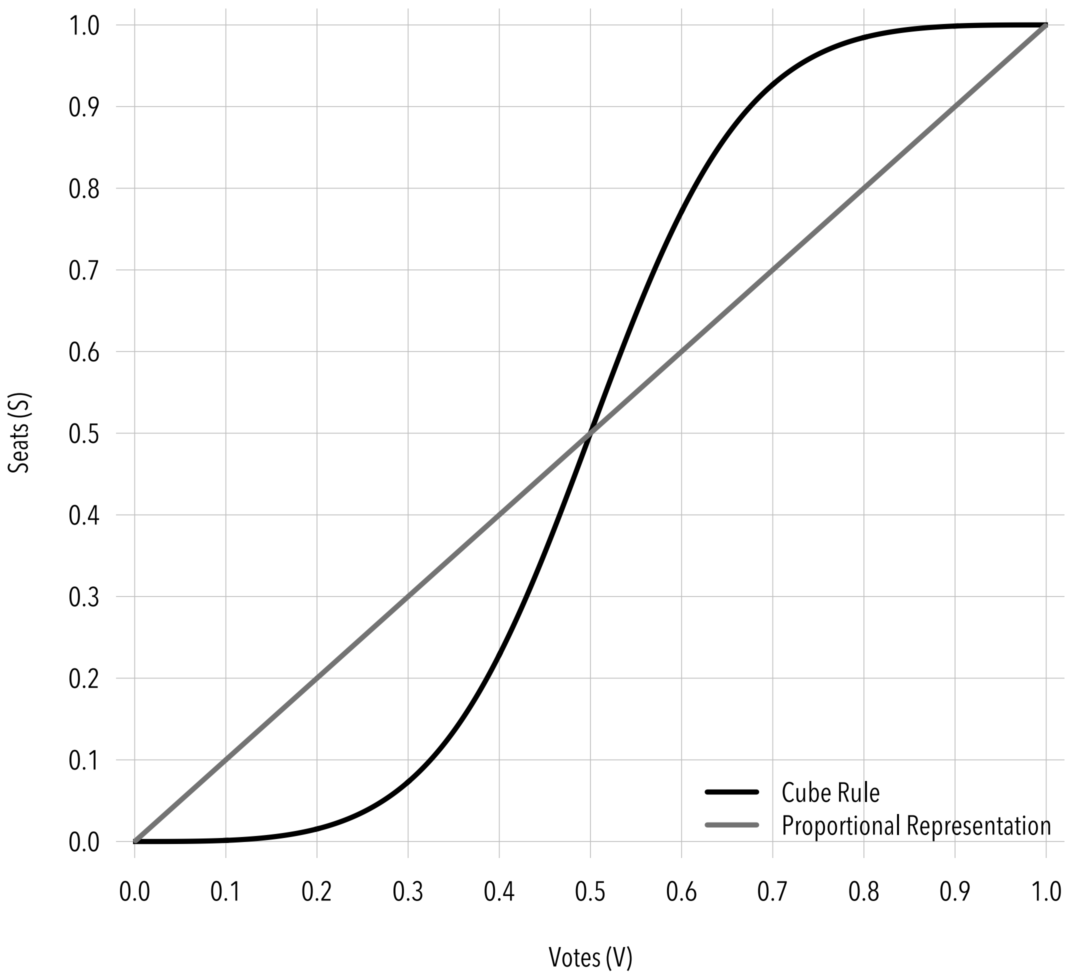


Figure 2: Two Theoretical Seats-Votes Curves

5 Partisan bias

Both of the hypothetical seats-votes curves in Figure 2 run through the "50-50" point, where $V = .5$ and $S = .5$. An interesting empirical question is whether *actual* seats-votes curves run through this point, or more generally, whether the seats-votes curve is symmetric about $V = .5$. Formally, symmetry of the seats-vote curve is the condition that $E(S|V) = 1 - E(S|1 - V)$, where E is the expectation operator, averaging over the uncertainty with respect to S given V. The vertical offset from the $(.5, .5)$ point for a seats-votes curve is known as *partisan bias*: the extent to which a party's expected seat share lies above or below 50%, conditional on that party winning 50% of the jurisdiction-wide vote.

Figure 3 shows three seats-votes curves, with the graph clipped to the region $V \in [.4, 6.]$ and $S \in [.4, .6]$ so as to emphasize the nature of partisan bias. The blue, positive bias curve "lifts" the seats-votes curve; it crosses $S = .5$ with $V < .5$ and passes through the upper-left quadrant of the graph. That is, with positive bias, a party can win a majority of the seats with *less* then a majority of the jurisdiction-wide or average vote; equivalently, if the party wins $V = .5$, it can expect to win *more* than 50% of the seats. Conversely, with negative bias, the opposite phenomenon occurs: the party can't expect to win a majority of the seats until it wins more than a majority of the jurisdiction-wide or average vote.

5.1 Multi-year method

With data from multiple elections under the same district plan, partisan bias can be estimated by fitting a seats-votes curve to the observed seat and vote shares, typically via a simple statistical technique such as linear regression; this approach has a long and distinguished lineage in both political science and statistics (e.g., Edgeworth, 1898; Kendall and Stuart, 1950; Tufte, 1973). Niemi and Fett (1986) referred to this method of estimating the partisan bias of an electoral system as the "multi-year" method, reflecting the fact that the underlying data comes from a sequence of elections.

This approach is of limited utility when assessing a new or proposed districting plan. More generally, it is of no great help to insist that a sequence of elections must be conducted under a redistricting plan before the plan can be properly assessed. Indeed, few plans stay intact long enough to permit reliable analysis in

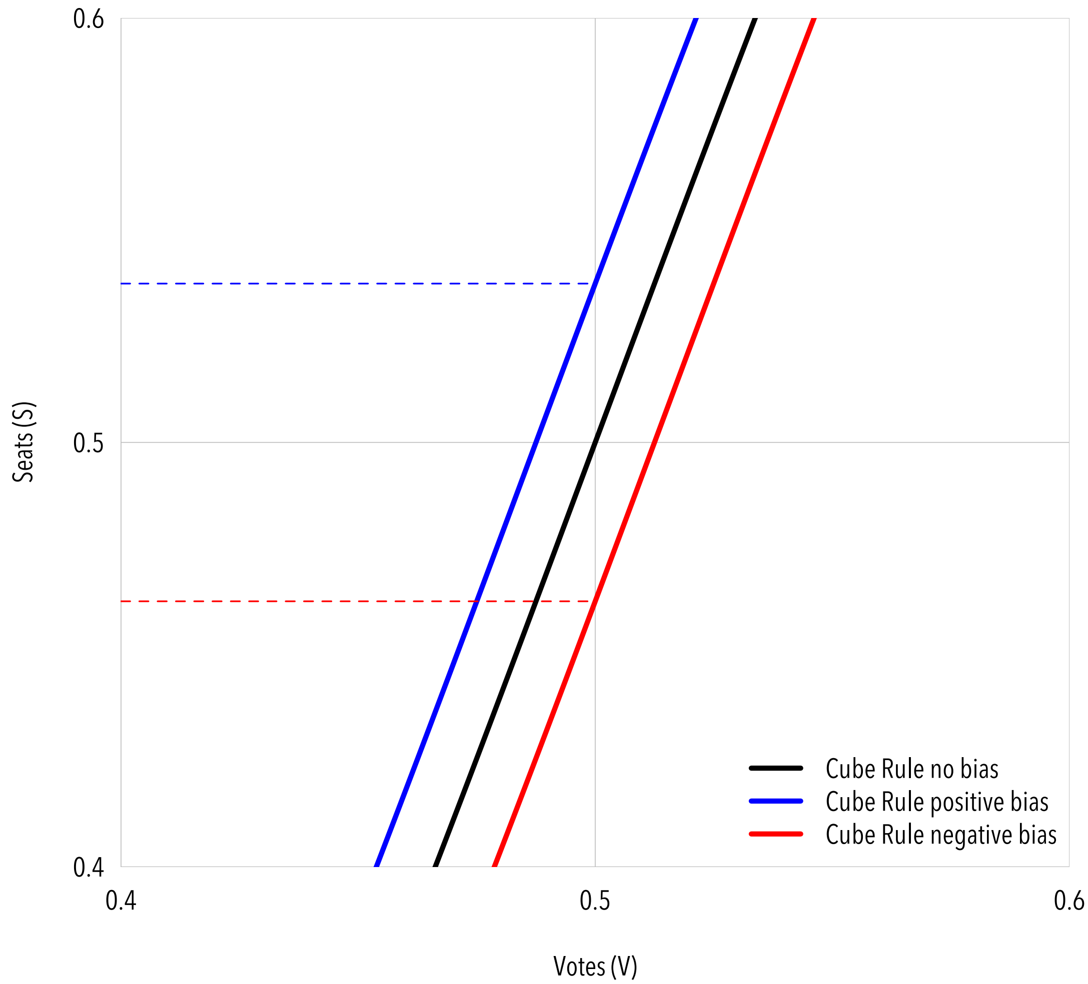


Figure 3: Theoretical seats-votes curves, with different levels of partisan bias. This graph is "zoomed in" on the region $V \in [.4, .6]$ and $S \in [.4, .6]$; the seats-votes "curves" are approximately linear in this region.

this way. State-level plans in the United States might generate as many five elections between decennial censuses. Accordingly, many uses of the "multi-year" method pool multiple plans and/or across jurisdictions, so as to estimate average partisan bias. For instance, Niemi and Jackman (1991) estimated average levels of partisan bias in state legislative districting plans, collecting data spanning multiple decades and multiple states, and grouping districting plans by the partisanship of the plan's authors (e.g., plans drawn under Republican control, Democratic control, mixed, or independent).

Assessing the properties of a districting plan after a tiny number of elections — or *no* elections — requires some assumptions and/or modeling. A single election yields just a single (V, S) data point, through which no unique seats-vote curve can be fitted and so partisan bias can't be estimated without further assumptions. Absent *any* actual elections under the plan, we might examine votes from a previous election, say, with precinct level results re-aggregated to the new districts.

5.2 Uniform swing

One approach—dating back to Sir David Butler's (1974) pioneering work on British elections—is the uniform partisan swing approach. Let $\mathbf{v} = (v_1, \dots, v_n)'$ be the set of vote shares for party A observed in an election with n districts. Party A wins seat i if $v_i > .5$, assuming just two parties (or defining v as the share of two-party vote); i.e., $s_i = 1$ if $v_i > .5$) and otherwise $s_i = 0$. Party A's seat share is $S = \frac{1}{n}\sum_{i=1}^{n} s_i$. V is the jurisdiction-wide vote share for party A, and if each district had the same number of voters $V = \bar{v} = \frac{1}{n}\sum_{i=1}^{n} v_i$, the average of the district-level v_i. Districts are never *exactly* equal sized, in which case we can define V as follows: let t_i be the number of voters in district i, and $V = \sum_{i=1}^{n} t_i v_i / \sum_{i=1}^{n} t_i$.

The uniform swing approach perturbs the observed district-level results $\mathbf{v}$ by a constant factor δ, corresponding to a hypothetical amount of *uniform swing* across all districts. For a given δ, let $v_i^* = v_i + \delta$ which in turn generates $V^* = V + \delta$ and an implied seat share S^*. Now let δ vary over a grid of values ranging from $-V$ to $1 - V$; then V^* varies from 0 to 1 and a corresponding value of S^* can also be computed at every grid point. The resulting set of (V^*, S^*) points are then plotted to form a seats-vote curve (actually, a step function). Partisan bias is simply "read off" this set of results, computed as $S^*|(V^* = .5) - .5$.

There is an elegant simplicity to this approach, taking an observed set of district-level vote shares $\mathbf{v}$ and shifting them by the constant δ. The observed distribution of district level vote shares observed in a given election is presumed to hold under *any* election we might observe under the redistricting plan, save for the shift given by the uniform swing term δ.

5.3 Critiques of partisan bias

Among political scientists, the uniform swing approach was criticized for its determinism. Swings are never exactly uniform across districts. There are many permutations of observed vote shares that generate a statewide vote share of 50% other than simply shifting observed district-level results by a constant factor. A less deterministic approach to assessing partisan bias was developed over a series of papers by Gary King and Andrew Gelman in the early 1990s (e.g., Gelman and King, 1990). This approach fits a statistical model to district-level vote shares — and, optionally, utilizing available predictors of district-level vote shares — to model the way particular districts might exhibit bigger or smaller swings than a given level of state-wide swing. Perhaps one way to think about the approach is that it is "approximate" uniform swing, with statistical models fit to historical election results to predict and bound variation around a state-wide average swing. The result is a seats-vote curve and an estimate of partisan bias that comes equipped with uncertainty measures, reflecting uncertainty in the way that individual districts might plausibly deviate from the state-wide average swing yet still produce a state-wide average vote of 50%.

The King and Gelman model-based simulation approaches remain the most sophisticated methods of generating seats-votes curves, extrapolating from as little as one election to estimate a seats-votes curve and hence an estimate of partisan bias. Despite the technical sophistication with which we can estimate partisan bias, legal debate has centered on a more fundamental issue, the *hypothetical* character of partisan bias itself. Recall that partisan bias is defined as "seats in excess of 50% *had the jurisdiction-wide vote split 50-50.*" The premise that $V = .5$ is the problem, since this will almost always be a counter-factual or hypothetical scenario. The further V is away from .5 in a given election, the counter-factual we must contemplate (when assessing the partisan bias of a districting plan) becomes all the more speculative.

In no small measure this is a marketing failure, of sorts. Partisan bias (at least under the uniform swing assumption) is essentially a measure of skew or asymmetry in *actual* vote shares. Partisan bias garners great rhetorical and normative appeal by directing attention to what happens at $V = .5$; it seems only "fair" that if a party wins 50% or more of the vote it should expect to win a majority of the districts.

Yet this distracts us from the fact that *asymmetry* in the distribution of vote shares across districts is the key, operative feature of a districting plan, and the extent to which it advantages one party or the other. Critically, we need not make appeals to counter-factual, hypothetical elections in order to assess this asymmetry.

6 The Efficiency Gap

The efficiency gap (*EG*) is also an asymmetry measure, as we see below. But unlike partisan bias, the interpretation of the efficiency gap is *not* explicitly tied to any counter-factual election outcome. In this way, the efficiency gap provides a way to assess districting plans that is free of the criticisms that have stymied the partisan bias measure.

Stephanopoulos and McGhee (2015) derive the *EG* measure with the concept of wasted votes. A party only needs v_i = 50% + 1 of the votes to win district i. Anything more are votes that could have been deployed in other districts. Conversely, votes in districts where the party doesn't win are "wasted," from the perspective of generating seats: any districts with $v_i < .5$ generate no seats.

Wasted votes get at the core of what partisan gerrymandering is, and how it operates. A gerrymander against party A creates a relatively small number of districts that "lock up" a lot of its votes ("packing" with $v_i > .5$) and a larger number of districts that disperse votes through districts won by party B ("cracking" with $v_i < .5$). To be sure, both parties are wasting votes. But partisan advantage ensues when one party is wasting fewer votes than the other, or, equivalently, more efficiently translating votes into seats. Note also how the efficiency gap measure is also closely tied to asymmetry in the distribution of v_i.

Some notation will help make the point more clearly. If $v_i > .5$ then party A wins the district and $s_i = 1$; otherwise $s_i = 0$. The efficiency gap is defined by McGhee (2014, 68) as "relative wasted votes" or

$$EG = \frac{W_B}{n} - \frac{W_A}{n}$$

where

$$W_A = \sum_{i=1}^{n} s_i(v_i - .5) + (1 - s_i)v_i$$

is the sum of wasted vote proportions for party A and

$$W_B = \sum_{i=1}^{n} (1 - s_i)(.5 - v_i) + s_i(1 - v_i)$$

is the sum of wasted vote proportions for party B and n is the number of districts in the jurisdiction. If $EG > 0$ then party B is wasting more votes than A, or A is translating votes into seats more efficiently than B; if $EG < 0$ then the converse, party A is wasting more votes than B and B is translating votes into seats more efficiently than A.

6.1 The efficiency gap when districts are of equal size

Under the assumption of equally sized districts McGhee (2014, 80) re-expresses the efficiency gap as:

$$EG = S - .5 - 2(V - .5) \tag{1}$$

recalling that $S = n^{-1} \sum_{i=1}^{n} s_i$ is the proportion of seats won by party A and $V = n^{-1} \sum_{i=1}^{n} v_i$ is the proportion of votes won by party A.

The assumption of equally-sized districts is especially helpful for the analysis reported below, since the calculation of EG in a given election then reduces to using the jurisdiction-level quantities S and V as in equation 1. For the analysis of historical election results reported below, it isn't possible to obtain measures of district populations, meaning that we really have no option other than to rely on the jurisdiction-level quantities S and V when estimating the EG.

I operationalize V as the average (over districts) of the Democratic share of the two-party vote, in seats won by either a Democratic or Republican candidate;

this set of seats includes uncontested seats, where I will use imputation procedures to estimate two-party vote share. If districts are of equal size (and ignoring seats won by independents and minor party candidates) then this average over districts will correspond to the Democratic share of the state-wide, two-party vote.

6.2 The seats-vote curve when the efficiency gap is zero

This simple expression for the efficiency gap implies that *if the efficiency gap is zero*, we obtain a particular type of seats-votes curve, shown in Figure 4:

1. the seats-votes curve runs through the 50-50 point. If the jurisdiction wide vote is split 50-50 between party A and party B then with an efficiency gap of zero, $S = .5$.

2. conditional on $V = .5$ (an even split of the vote), the efficiency gap is the same as partisan bias: $V = .5 \iff EG = S - .5$, the seat share for party A in excess of 50%. That is, the efficiency gap reduces to partisan bias *under the counter-factual scenario* $V = .5$ that the partisan bias measure requires us to contemplate. On the other hand, the efficiency gap is not premised on that counter-factual holding, or any other counter-factual for that matter; the efficiency gap summarizes the distribution of observed district-level vote shares v_i.

3. the seats-votes curve is linear through the 50-50 point with a slope of 2. That is, with $EG = 0$, $S = 2V - .5$. Or, with a zero efficiency gap, each additional percentage point of vote share for party A generates *two* additional percentage points of seat share. A zero efficiency gap does not imply proportional representation (a seats-votes that is simply a 45 degree line).

4. a party winning 25% or less of the jurisdiction-wide vote should win zero seats under a plan with a zero efficiency gap; a party winning 75% or more of the jurisdiction-wide vote should win all of the seats under a plan with a zero efficiency gap. This is a consequence of the "2-to-1" seats/vote ratio and the symmetry implied by a zero efficiency gap. A party that wins an extremely low share of the vote ($V < .25$) can only be winning any seats if it enjoys an efficiency advantage over its opponent.

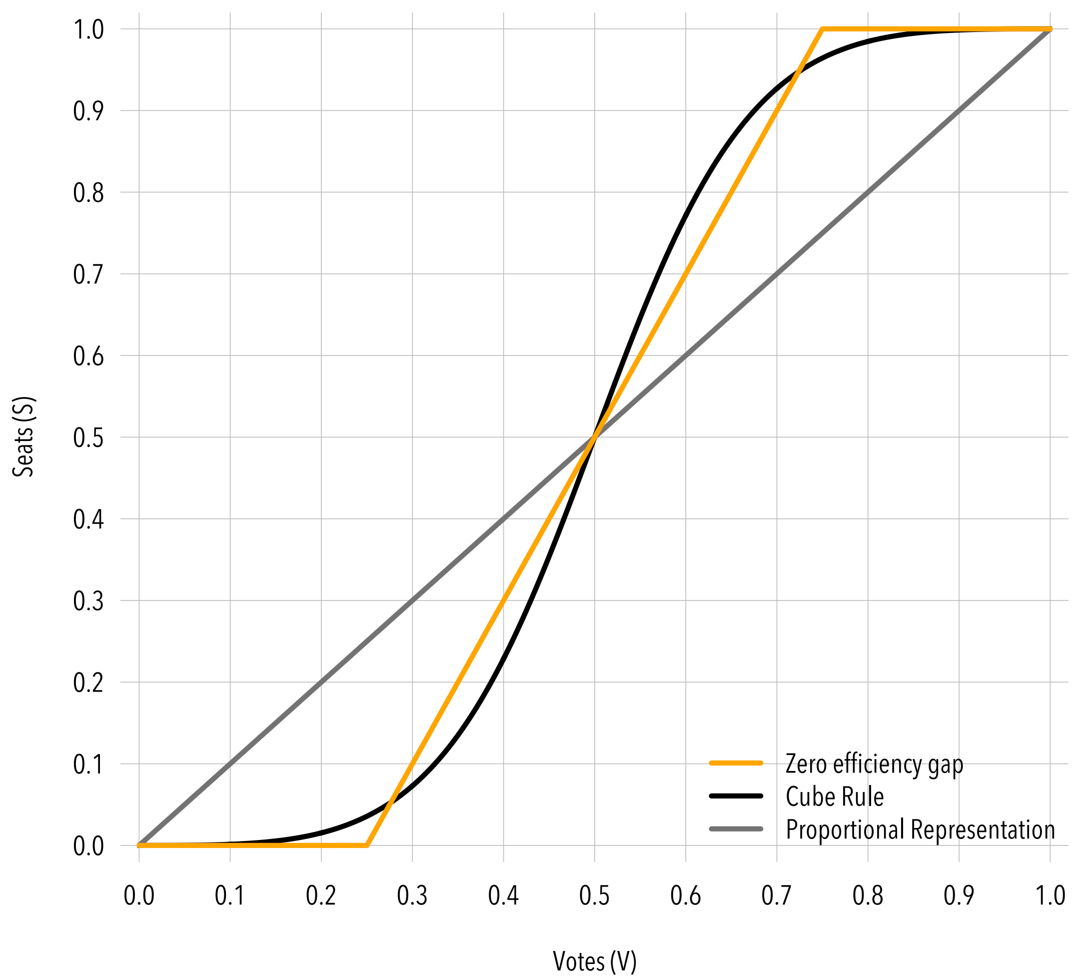


Figure 4: Theoretical seats-votes curves. The $EG = 0$ curve implies that (a) a party winning less than $V = .25$ jurisdiction-wide should not win any seats; (b) symmetrically, a party winning more than $V = .75$ jurisdistion-wide should win all the seats; and (c) the relationship between seat shares S and vote shares V over the interval $V \in [.25, .75]$ is a linear function with slope two (i.e., for every one percentage point gain in vote share, seat share should go up by two percentage points).

Moreover, the efficiency gap is trivial to compute once we have V and S for a given election. We don't need a sequence of elections under a plan in order to compute EG, nor do we need to anchor ourselves to a counter-factual scenario such as $V = .5$ as we do when computing partisan bias. For any given observed V, the hypothesis of zero efficiency gap tells us what level of S to expect.

6.3 The efficiency gap as an excess seats measure

In this sense the efficiency gap can be interpreted even more simply as an "excess seats" measure. Recall that $EG = 0 \iff S = 2V - .5$. In a given election we observe $EG = S - .5 - 2(V - .5)$. The efficiency gap can be computed by noting how far the observed S lies above or below the orange line in Figure 4.

A positive EG means "excess" seats for party A relative to a zero efficiency gap standard given the observed V in that election; conversely, a negative EG mean a deficit in seats for party A relative to a zero efficiency gap standard given the observed V.

7 State legislative elections, 1972-2014

We estimate the efficiency gap in state legislative elections over a large set of states and districting plans, covering the period 1972 to 2014. We begin the analysis in 1972 for two primary reasons: (a) state legislative election returns are harder to acquire prior to the mid-1960s, and not part of the large, canonical data collection we rely on (see below); and (b) districting plans and sequences of elections from 1972 onwards can be reasonably considered to be from the post-malapportionment era.

For each election we recover an estimate of the efficiency gap based on the election results actually observed in that election. To do this, I compute two quantities for each election:

1. V, the statewide share of the two-party vote for Democratic candidates, formed by averaging the district-level election results v_i (the Democratic share of the two-party vote in district i) in seats won by major party candidates, including uncontested seats, and

2. *S*, the Democratic share of seats won by major parties.

Recall that these quantities are the inputs required when computing the efficiency gap (equation 1).

The analysis that follows relies on a data set widely used in political science and freely available from the Inter-University Consortium for Political and Social Research (ICPSR study number 34297). The release of the data I utilize covers state legislative election results from 1967 to 2014, updated by Karl Klarner (Indiana State University and Harvard University). I subset the original data set to general election results since 1972 in states whose lower houses are elected via single-member districts, or where single-member districts are the norm. Multi-member districts "with positions" are treated as if they are single-member districts.

Figure 5 provides a graphical depiction of the elections that satisfy the selection criteria described above.

- Arizona, Idaho, Louisiana, Maryland, Nebraska, New Hampshire, New Jersey, North Dakota and South Dakota all drop out of the analysis entirely, because of exceedingly high rates of uncontested races, using multi-member districts, non-partisan elections, or the use of a run-off system (Louisiana).

- Alaska, Hawaii, Illinois, Indiana, Kentucky, Maine, Minnesota, Montana, North Carolina, Vermont, Virginia, West Virginia and Wyoming do not supply data over the entire 1972-2014 span; this is sometimes due to earlier elections being subject to exceedingly high rates of uncontestedness, the use of multi-member districts or non-partisan elections.

- Alabama and Mississippi have four-year terms in their lower houses, contributing data at only half the rate of the vast bulk of states with two-year legislative terms.

- Twenty-three states supply data every two years from 1972 to 2014, including Michigan and Wisconsin.

- Data is more abundant in recent decades. For the period 2000 to 2014, 41 states contribute data to the analysis at two or four year intervals.

In summary, the data available for analysis span 83,269 district-level state legislative contests, from 786 elections across 41 states.

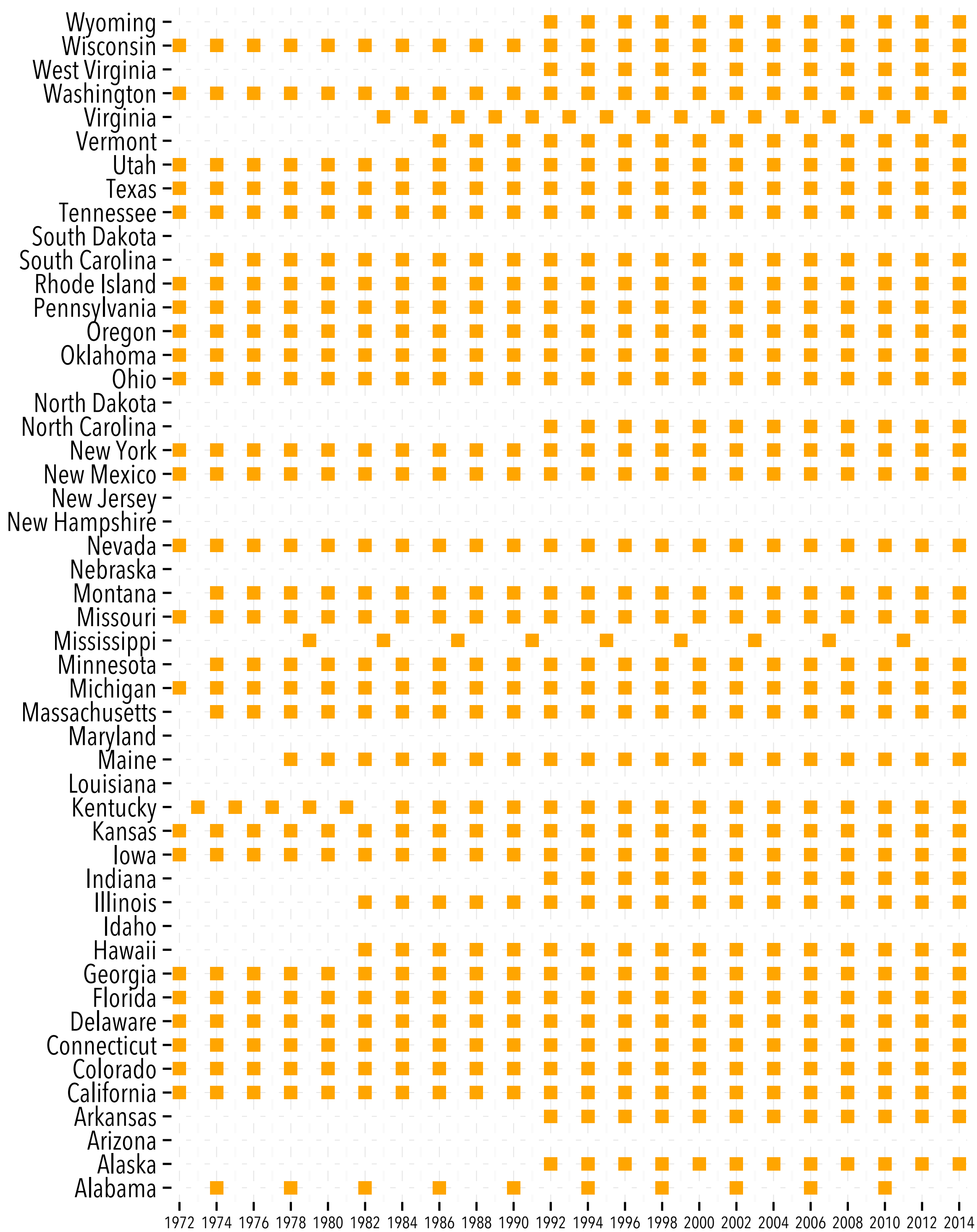


Figure 5: 786 state legislative elections available for analysis, 1972-2014, by state.

7.1 Grouping elections into redistricting plans

Districting plans remain in place for sequences of elections. An important component of my analysis involves tracking the efficiency gap across a series of elections held under the same districting plan. A key question is how much variation in the *EG* do we observe *within* districting plans, versus variation in the *EG* *between* districting plans.

To the extent that the *EG* is a feature of a districting plan per se, we should observe a small amount of within-plan variation relative to between plan variation. To perform this analysis we must group sequences of elections within states by the districting plan in place at the time.

Stephanopolous and McGhee (2015) provide a unique identifier for the districting plan in place for each state legislative election, for which I adopt here.

Figure 6 displays how the elections available for analysis group by districting plan. Districts are typically redrawn after each decennial census; the first election conducted under new district boundaries is often the "2" election (1982, 1992, etc). Occasionally we see just one election under a plan: examples include Alabama 1982, California, Hawaii 1982, Tennessee 1982, Ohio 1992, South Carolina 1992, North Carolina 2002, and South Carolina 2002.

Alaska, Kentucky, Pennsylvania and Texas held just one election under their respective districting plans adopted after the 2010 Census. In each of those states a different plan was in place for 2014 state legislative elections. Alabama's state legislature has a four year term and we observe only the 2014 election under its post-2010 plan. The last election from Mississippi was in 2011 and was held under the plan in place for its 2003 and 2007 elections.

7.2 Uncontested races

Uncontested races are common in state legislative elections, and are even the norm in some states. For 38.7% of the district-level results in this analysis, it isn't possible to directly compute a two-party vote share (v_i), either because the seat was uncontested or not contested by both a Democratic and Republican candidate, or (in a tiny handful of cases) the data are missing.

In some states, for some elections, the proportion of uncontested races is so high that we drop the election from the analysis. As noted earlier, examples

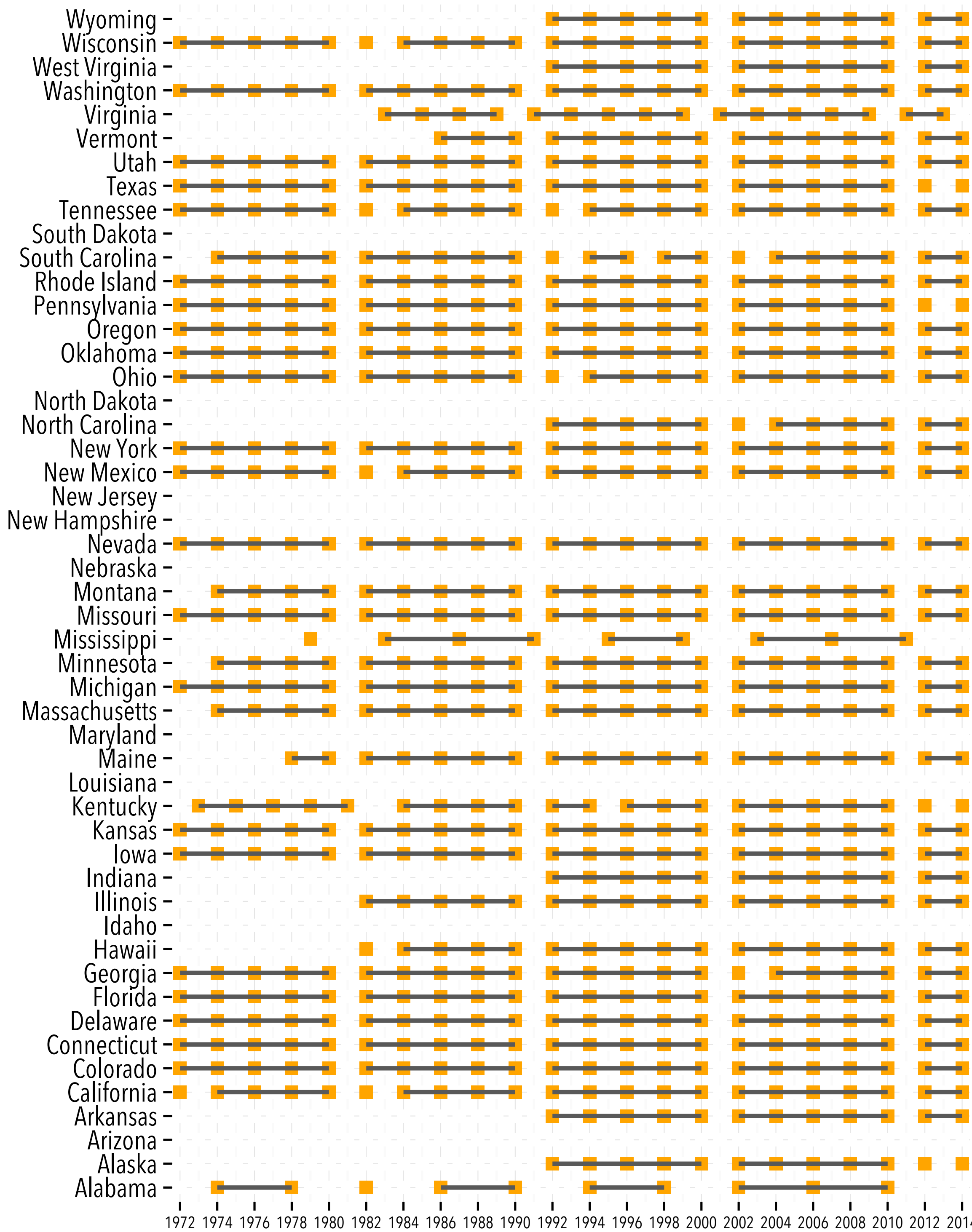


Figure 6: 786 state legislative elections available for analysis, 1972-2014, by state, grouped by districting plan (horizontal line).

include Arkansas elections prior to 1992 and South Carolina in 1972.

Even with these elections dropped from the analysis, the extent of uncontestedness in the remaining set of state legislative election results is too large to be ignored. Of the remaining elections, 31% have missing two-party results in at least half of the districts.

A graphical summary of the prevalence of uncontested districts appears in Figure 7, showing the percentage of districts without Democratic and Republican vote counts, by election and by state. Uncontested races are the norm in a number of Southern states: e.g., Georgia, South Carolina, Mississippi, Arkansas, Texas, Alabama, Virginia, Kentucky and Tennessee record rates of uncontestedness that seldom, if ever, drop below 50% for the period covered by this analysis. Wyoming also records a high proportion of districts that do not have Democratic versus Republican contests. States that lean Democratic also have high levels of uncontestedness too: see Rhode Island, Massachusetts, Illinois and, in recent decades, Pennsylvania.

Michigan and Minnesota are among the states with the lowest levels of uncontested districts in their state legislative elections. Over the set of 786 state legislative elections we examine, there are just *three* instances of elections with Democrats and Republicans running candidates in every district: Michigan supplies two of these cases (2014 and 1996) and Minnesota the other (2008).

8 Imputations for Uncontested Races

Stephanopolous and McGhee (2015) note the prevalence of uncontested races and report using a statistical model to impute vote shares to uncontested districts. They write:

> We strongly discourage analysts from either dropping uncontested races from the computation or treating them as if they produced unanimous support for a party. The former approach eliminates important information about a plan, while the latter assumes that coerced votes accurately reflect political support.

I concur with this advice, utilizing an imputation strategy for uncontested districts with *two* distinct statistical models, predicting Democratic, two-party

Percent single-member districts without D and R candidates/vote counts, by state & election

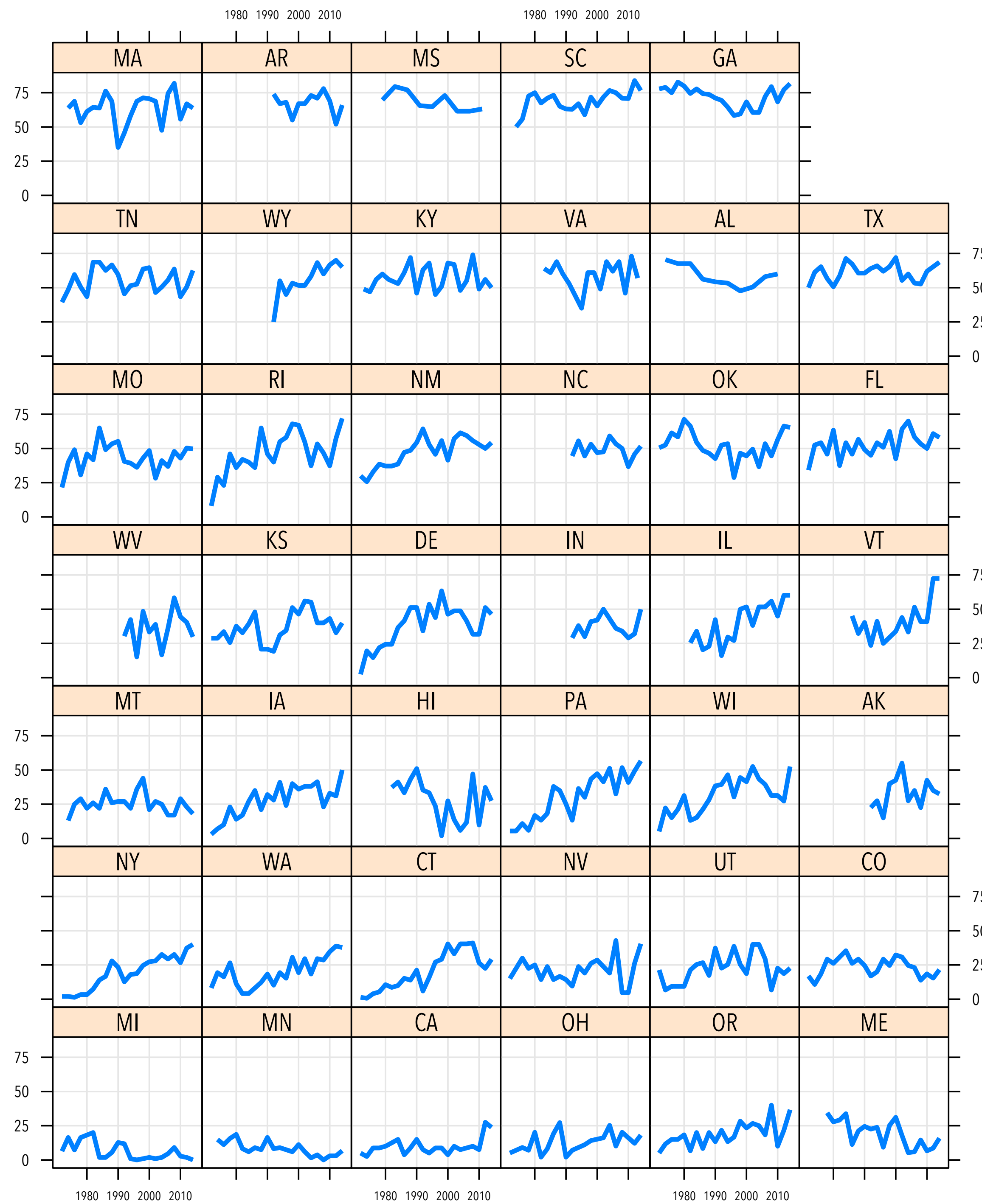


Figure 7: Percentage of districts missing two-party vote shares, by election, in 786 state legislative elections, 1972-2014. Missing data is almost always due to districts being uncontested by both major parties.

vote share in state legislative districts (v_i).

8.1 Imputation model 1: presidential vote shares

The first imputation model relies on presidential election returns reported at the level of state legislative districts. Presidential election returns are excellent predictors of state legislative election outcomes and observed even when state legislative elections are uncontested. I fit a series of linear regressions of v_i on the Democratic share of the two-party vote for president in district i, as recorded in the most temporally-proximate presidential election for which data is available and for which the current election's districting plan was in place; separate slopes and intercepts are estimated depending on the incumbency status of district i (Democratic, Open/Other, Republican).

The model also embodies the following assumptions in generating imputations for unobserved vote shares in uncontested districts. In districts where a Republican incumbent ran unopposed, we assume that the Democratic share of the two-party vote would have been less than 50%; conversely, where Democratic incumbents ran unopposed, we assume that the Democratic share of the vote would have been greater than 50%.

In most states the analysis predicts 2014 and 2012 state legislative election results v_i using 2012 presidential vote shares; 2006, 2008 and 2010 v_i is regressed on 2008 presidential vote shares, and so on. Some care is needed matching state and presidential election results in states that hold their state legislative elections in odd-numbered years, or where redistricting intervenes. In a small number of cases, presidential election returns are not available, or are recorded with district identifiers that can't be matched in the state legislative elections data. We lack data on presidential election results by state legislative district prior to 2000, so 1992 is the earliest election with which we can match state legislative election results to presidential election results at the district level.

The imputation model generally fits well. Across the 447 elections, the median r^2 statistic is 0.82. The cases fitting less well include Vermont in 2012 ($r^2 = 0.29$), with relatively few contested seats and multi-member districts with positions.

We examine the performance of the imputation model in a series of graphs, below, for six sets of elections: Wisconsin in 2012 and 2014, Michigan in 2014

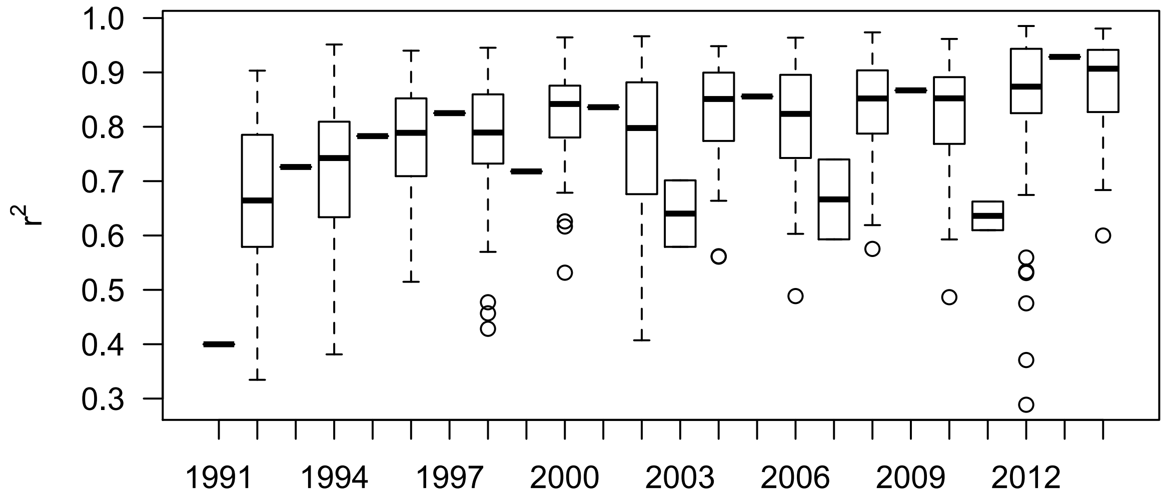


Figure 8: Distribution of r^2 statistics, regressions of Democratic share of two-party vote in state legislative election outcomes on Democratic share of the two-party for president.

(with no uncontested districts), South Carolina in 2012 (with the highest proportion of uncontested seats in the 2012 data), Virginia in 2013 and Wyoming in 2012 (the latter two generating extremely large, negative values of the efficiency gap). Vertical lines indicate 95% confidence intervals around imputed values for the Democratic share of the two-party vote in state legislative elections (vertical axis). Separate slopes and intercepts are fit for each incumbency type. Note also that the imputed data almost always lie on the regression lines.

Imputations for uncontested districts are accompanied by uncertainty. Although the imputation models generally fit well, like any realistic model they provides less than a perfect fit to the data. Note too that in any given election, there is only a finite amount of data and hence a limit to the precision with which we can make inferences about unobserved vote shares based on the relationship between observed vote shares and presidential vote shares.

Uncertainty in the imputations for v in uncontested districts generates uncertainty in "downstream" quantities of interest such as statewide Democratic vote share V and the efficiency gap measure EG. This is key, given the fact that uncontestedness is so pervasive in these data. We want any conclusions about the efficiency gap's properties or inferences about particular levels of the efficiency gap to reflect the uncertainty resulting from imputing vote shares in uncontested districts.

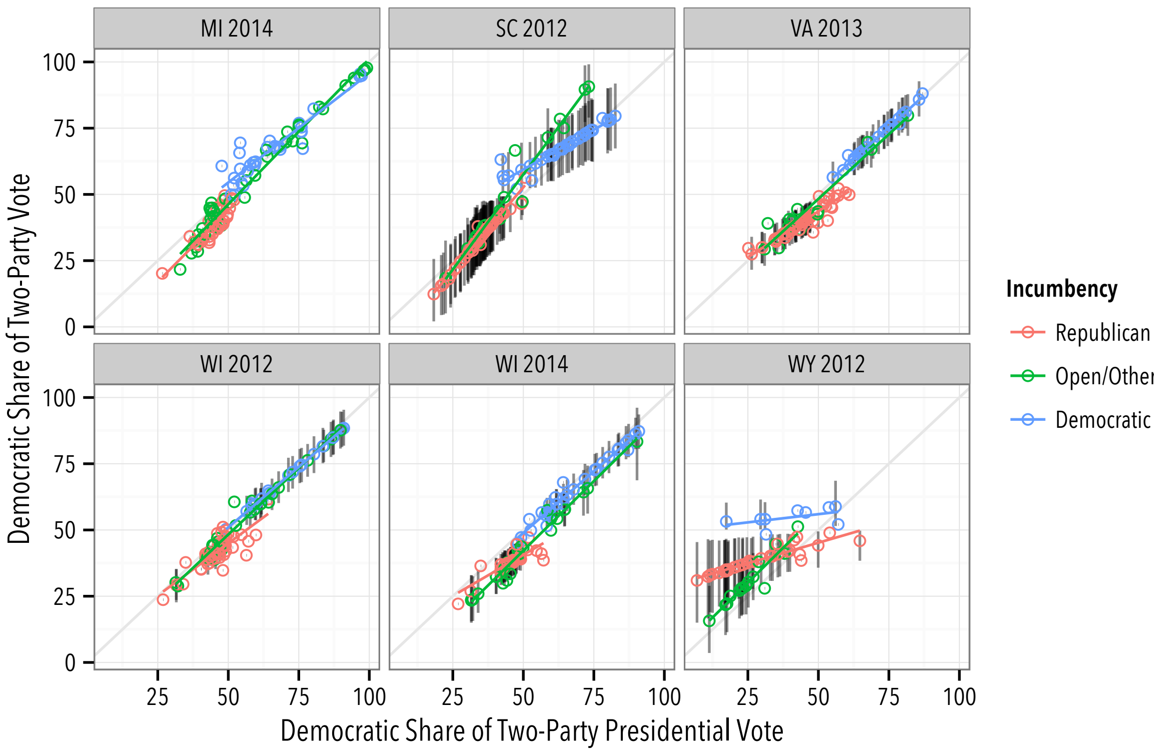


Figure 9: Regression model for imputing unobserved vote shares in 6 selected elections. Vertical lines indicate 95% confidence intervals around imputed values for the Democratic share of the two-party vote in state legislative elections (vertical axis). Separate slopes and intercepts are fit for each incumbency type. Note also that the imputed data almost always lie on the regression lines.

8.2 Imputation model 2

We rely on imputations based on presidential election returns when they are available. But presidential vote isn't always available at the level of state legislative districts (not before 1992, in this analysis). To handle these cases, we rely on a second imputation procedure, one that models sequences of election results observed under a redistricting plan, interpolating unobserved Democratic vote shares given (1) previous and future results for a given district; (2) statewide swing in a given state election; and (3) change in the incumbency status of a given district. This model also embodies the assumption that unobserved vote shares would nonetheless be consistent with what we *did* observe in a given seat: where a Democrat wins in an uncontested district, any imputation for v in that district must lie above 50%, and where a Republican wins an uncontested district, any imputation for v must lie below 50%.

8.3 Combining the two sets of imputations

We now have two sets of imputations for uncontested districts: (1) using presidential vote as a basis for imputation, where available (447 state legislative elections from 1992 to 2014); and (2) the imputation model that relies on the trajectory of district results over the history of a districting plan, including incumbency and estimates of swing, which supplies imputations for uncontested districts in all years.

When there are no uncontested districts, obviously the two imputations must agree, for the trivial reason that are no imputations to perform. As the number of uncontested districts rises, the imputations from the two models have room to diverge. Where the two sets of imputations are available for a given election (elections where presidential vote shares by state legislative districts are available) we generally see a high level of agreement between the two methods.

The two sets of imputations for V correlate at .99. With only a few exceptions (see Figure 10), the discrepancies are generally small relative to the uncertainty in the imputations themselves. As the proportion of districts with missing data increases, clearly the scope for divergence between the two models increases.

To re-iterate, we prefer the imputations from "Model 1" based on the regressions utilizing presidential vote shares in state legislative districts, and use them whenever available (i.e., for most states in the analysis, the period 1992-2014). We only rely on "Model 2" when presidential vote shares are not available. We model the difference between the two sets of imputations, adjusting the "Model 2" imputations of *V* to better match what we have obtained from "Model 1", had the necessary presidential vote shares by state legislative district been available.

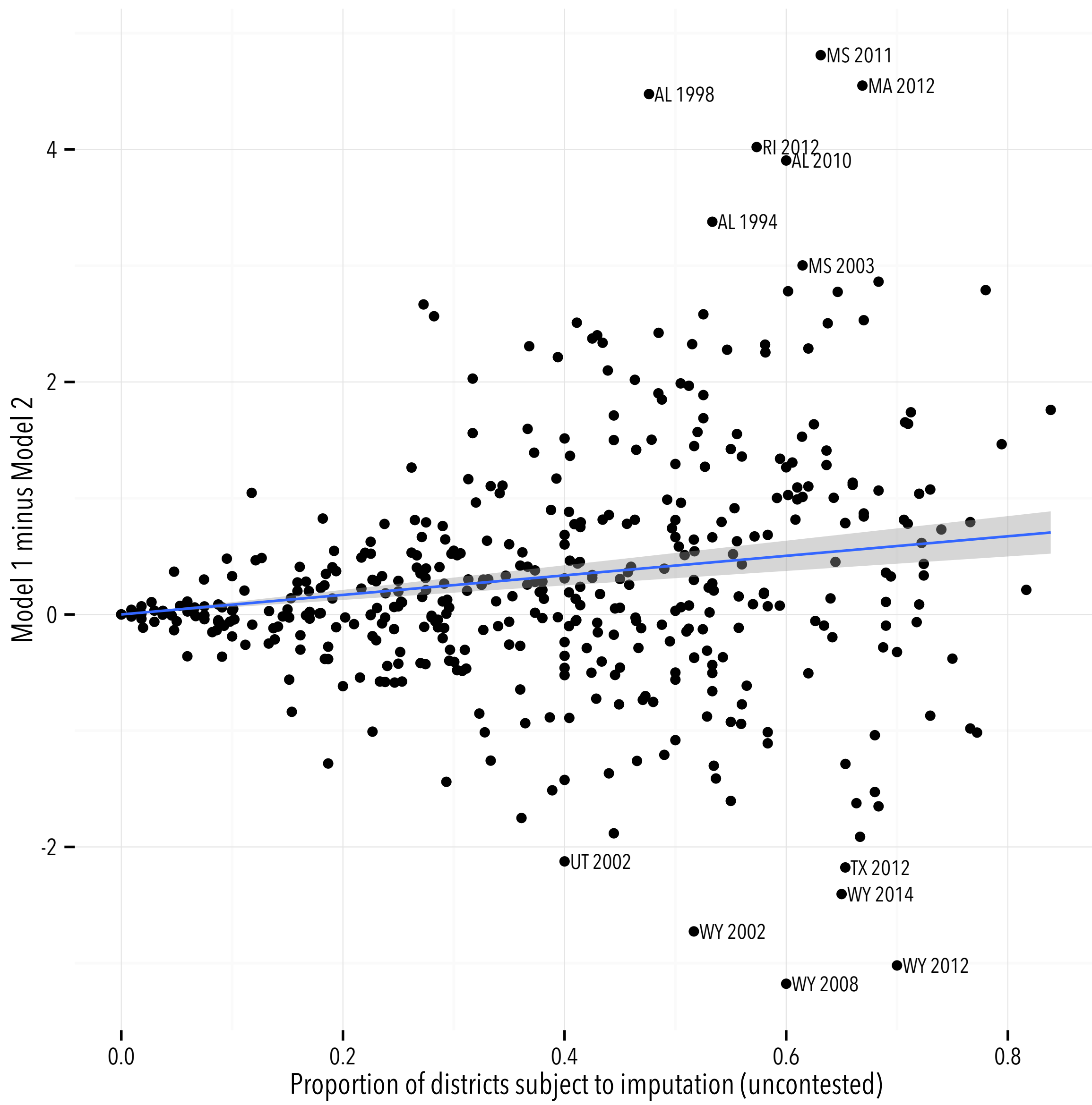


Figure 10: Difference between imputations for V by proportion of uncontested seats. The fitted regression line is constrained to respect the constraint that the imputations must coincide when there are no uncontested seats.

8.4 Seat and vote shares in 786 state legislative elections

After imputations for missing data, each election generates a seats-votes (V, S) pair. In Figure 11 we plot *all* of the V and S combinations over the 786 state elections in the analysis. We also overlay the seats-vote curve corresponding to an efficiency gap of zero. This provides us with a crude, visual sense of how often we see large departures from the zero EG benchmark.

The horizontal lines around each plotted point show the uncertainty associated with each estimate of V (statewide, Democratic, two-party vote share), given the imputations made for uncontested and missing district-level vote shares. Uncontested seats do not generate uncertainty with respect to the party winning the seat, and so the resulting uncertainty is with respect to vote shares, on the horizontal axis in Figure 11.

The efficiency gap in each election is the vertical displacement of each plotted (V, S) point from the orange, zero-efficiency gap line in Figure 11. Uncertainty as to the horizontal co-ordinate V (due to imputations for uncontested races) generates uncertainty in determining how far each point lies above or below the orange, zero efficiency gap benchmark.

9 The efficiency gap, by state and election

We now turn to the centerpiece of the analysis: assessing variation in the efficiency gap across districting plans.

We have 786 efficiency gap measures in 41 states, spanning 43 election years. These are computed by substituting each state election's estimate of V and the corresponding, observed seat share S into equation 1.

Figure 12 shows the efficiency gap estimates for each state election, grouped by state and ordered by year; vertical lines indicate 95% credible intervals arising from the fact that the imputation model for uncontested seats induces uncertainty in V and any quantity depending on V such as EG (recall equation 1). In many cases the uncertainty in EG stemming from imputation for uncontested seats is small relative to variation in EG both between and within districting plans.

We observe considerable variation in the EG estimates across states and elections. Some highlights:

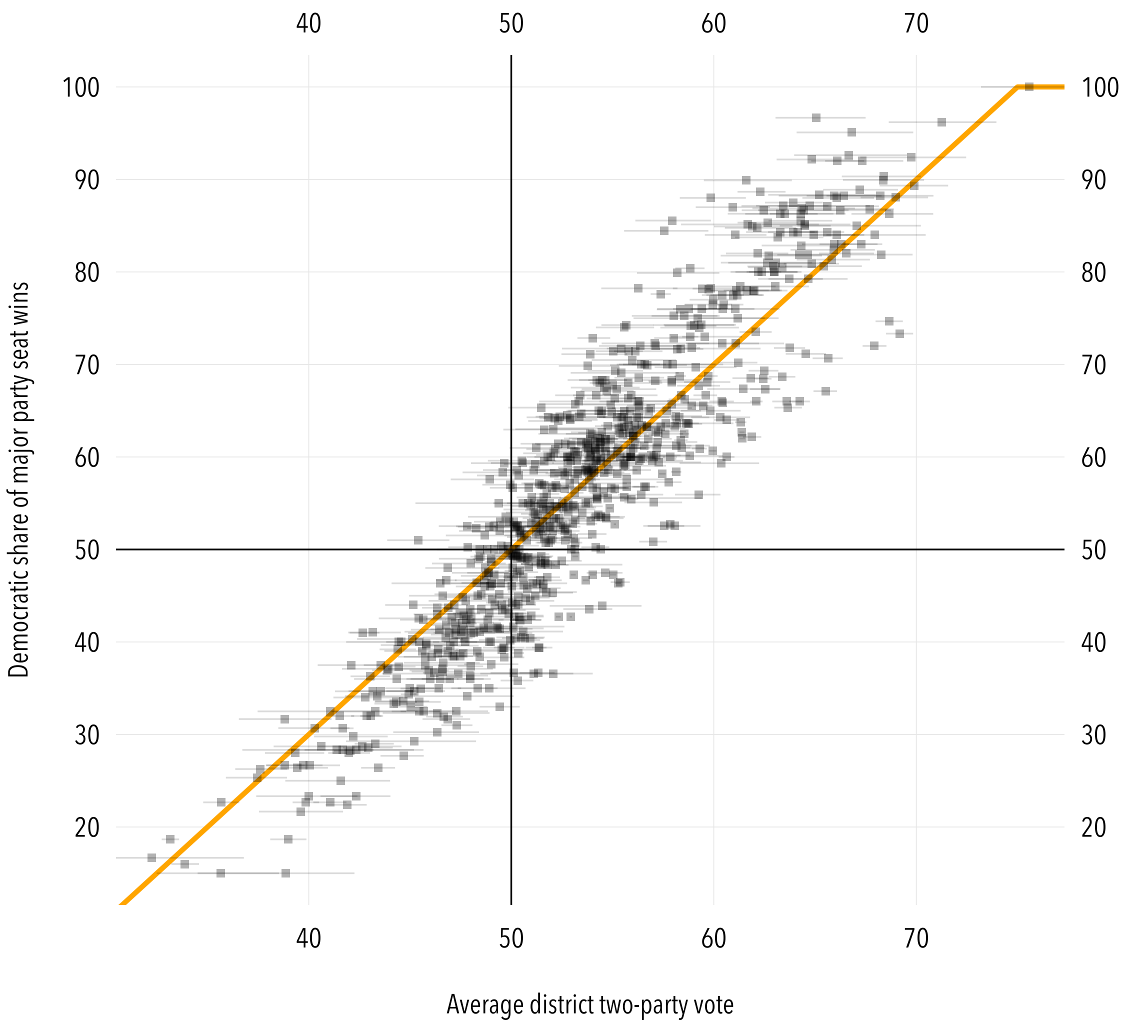


Figure 11: Democratic seat shares (S) and vote shares (V) in 786 state legislative elections, 1972-2014, in 41 states. Seat shares are defined with respect to single-member districts won by either a Republican or a Democratic candidate, including uncontested districts. Vote shares are defined as the average of district-level, Democratic share of the two-party vote, in the same set of districts used in defining seat shares. Horizontal lines indicate 95% credible intervals with respect to V, due to uncertainty arising from imputations for district-level vote shares in uncontested seats. The orange line shows the seats-votes relationship we expect if the efficiency gap were zero. Elections below the orange line have $EG < 0$ (Democratic disadvantage); points above the orange line have $EG > 0$ (Democratic advantage).

Efficiency gap, by state and year

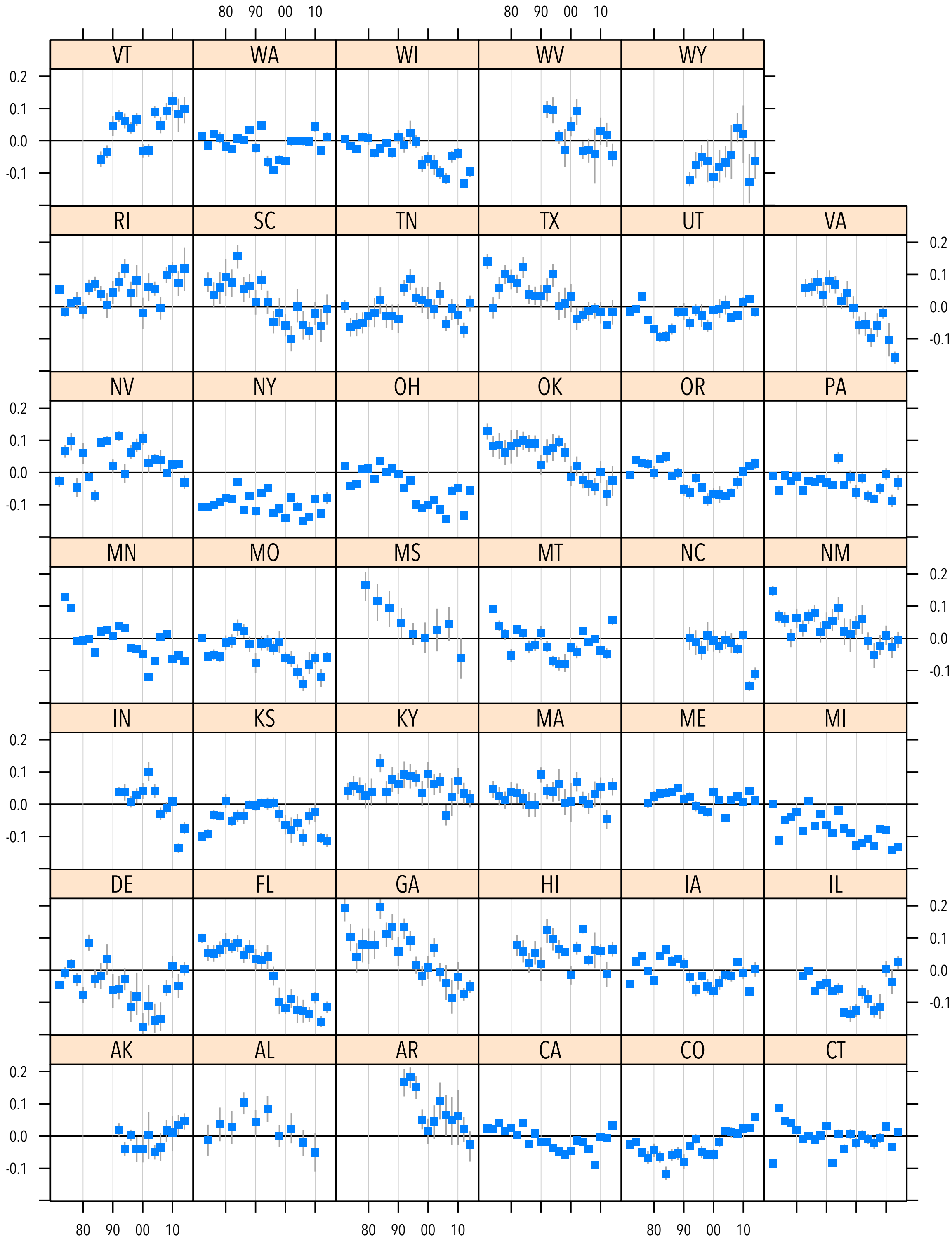


Figure 12: Efficiency gap estimates in 786 state legislative elections, 1972-2014. Vertical lines cover 95% credible intervals.

1. estimates of *EG* range from −0.18 to 0.20 with an average value of −0.005.

2. The lowest value, −0.18 is from Delaware in 2000. There were 19 uncontested seats in the election to the 41 seat state legislature. Democrats won 15 seats ($S = 15/41 = 36.6\%$). I estimate V to be 52.1%. Via equation 1, this generates $EG = -0.18$. Considerable uncertainty accompanies this estimate, given the large number of uncontested seats. The 95% credible interval for V is $\pm$ 2.03 percentage points, and the 95% credible interval for the accompanying *EG* estimate is $\pm$ 0.04.

3. The highest value of *EG* is 0.20 is from Georgia in 1984. There were 140 uncontested seats in the election to the 180 seat state legislature. Democrats won 154 seats ($S = 154/180 = 85.6\%$). I estimate V to be 57.9%. Again, using equation 1, this generates $EG = 0.2$. Considerable uncertainty also accompanies this estimate, given the large number of uncontested seats. The 95% credible interval for V is $\pm$ 1.89 percentage points, and the 95% credible interval for the accompanying *EG* estimate is $\pm$ 0.04. Figure 13 contrasts the seats and votes recorded in Georgia against those for the entire data set, putting Georgia's large *EG* estimates in context.

4. New York has the lowest median *EG* estimates, ranging from -.15 (2006) to -.028 (1984). Statewide V ranges from 53.7% to 69.2%, but Democrats only win 70 (1972) to 112 (2012) seats in the 150 seat state legislature, so S ranges from .47 to .75, considerably below that we'd expect to see given the vote shares recorded by Democrats if the efficiency gap were zero. See Figure 15.

5. Arkansas has the highest median *EG* score by state, .10; see Figure 14.

6. Connecticut has the median, within-state median *EG* score of approximately zero; Figure 16 shows Connecticut's seats and votes have generally stayed close to the $EG = 0$ benchmark.

7. Michigan has the third lowest median *EG* scores by state, surpassed only by New York and Wyoming. Michigan's *EG* scores range from -.14 (2012) to .01 (1984). V ranges from 50.3% to 60.6%, a figure we estimate confidently given low and occasionally even zero levels of uncontested districts

in Michigan state legislative elections. Yet *S* ranges from 42.7% (Democrats won 47 out of 110 seats in 2002, 2010 and 2014) to 63.6% (Democrats won 70 out of 110 seats in 1978). See Figure 17.

8. Wisconsin's *EG* estimates range from -.14 (2012) to .02 (1994). Although the *EG* estimates for WI are not very large relative to other states in other years, Wisconsin has recorded an unbroken run of negative *EG* estimates from 1998 to 2014 and records two very large estimates of the efficiency gap in elections held under its current plan: -.13 (2012) and -.10 (2014). In short, Democrats are underperforming in state legislative elections in Wisconsin, winning fewer seats than a zero efficiency gap benchmark would imply, given, their statewide level of support. See Figure 18.

9.1 Are efficiency gap estimates statistically significant?

Recall that $EG < 0$ means that Democrats are disadvantaged, with relatively more wasted votes than Republicans; conversely $EG > 0$ means that Democrats are the beneficiaries of an efficiency gap, in that Democrats have fewer wasted votes than Republicans. But *EG* does vary from election to election, even with the same districting plan in place and *EG* is almost always not measured perfectly, but is estimated with imputations for uncontested seats.

In Figure 19 we plot the imprecision of each efficiency gap estimate (the half-width of its 95% credible interval) against the estimated *EG* value itself. Points lying inside the cones have *EG* estimates that are small relative to their credible intervals, such that we would not distinguish them from zero at conventional levels of statistical significance. Not all *EG* estimates can be distinguished from zero at conventional levels of statistical significance, nor should they. But many estimates of the *EG* *are* unambiguously non-zero. Critically, the two most recent Wisconsin *EG* estimates (-.13 in 2012, -.10 in 2014) are clearly non-negative, lying far away from the "cone of ambiguity" shown in Figure 19; the 95% credible interval for the 2012 estimates runs from -.146 to -.121 and from -.113 to -.081 for the 2014 estimate.

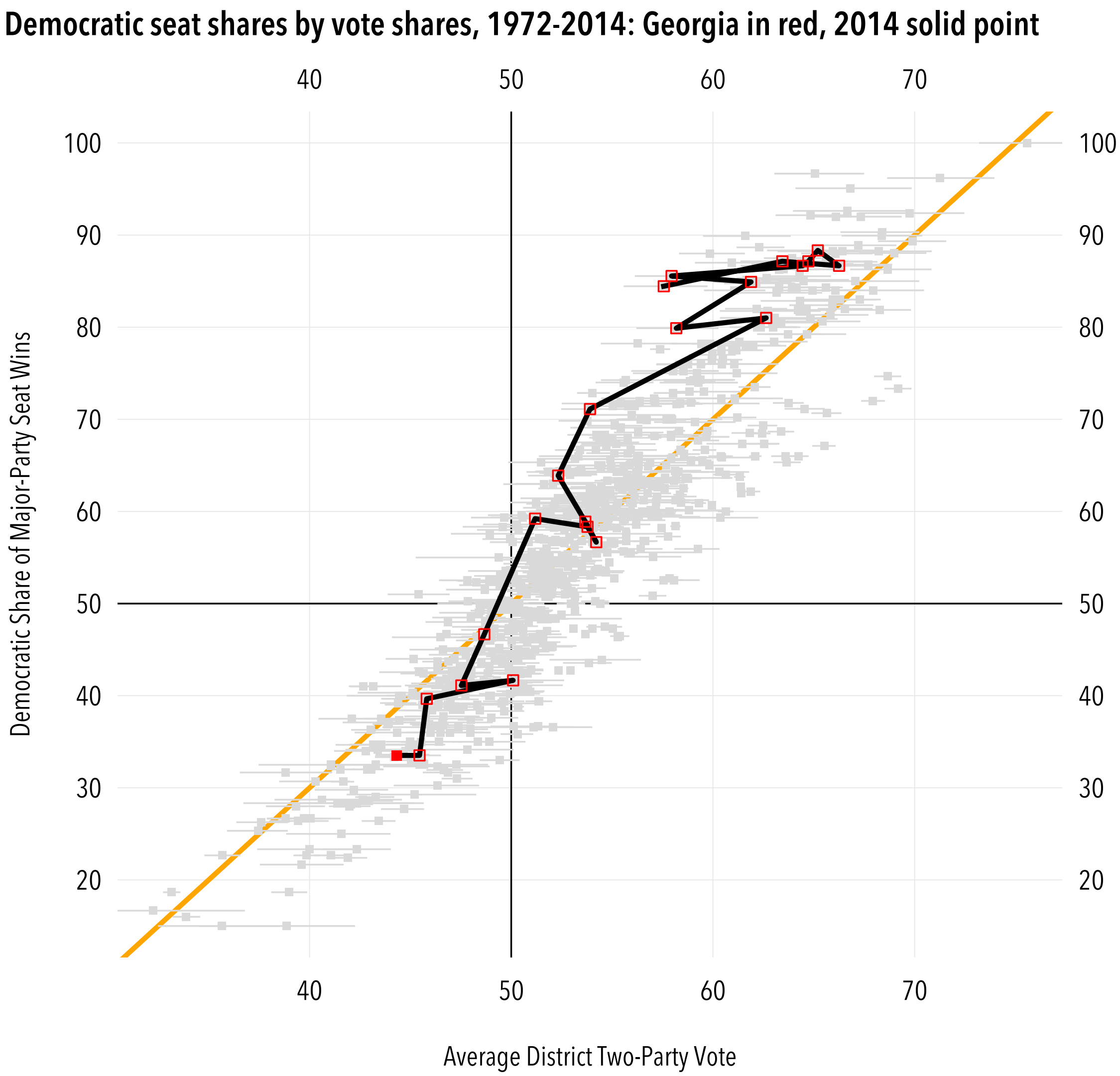


Figure 13: Georgia, Democratic seat share and average district two-party vote share, 1972-2014. Orange line shows the seats-votes curve if the efficiency gap were zero; the efficiency gap in any election is the vertical distance from the corresponding data point to the orange line. Gray points indicate elections from other states and elections (1972-2014). Horizontal lines cover a 95% credible interval for Democratic average district two-party vote share, given imputations in uncontested districts.

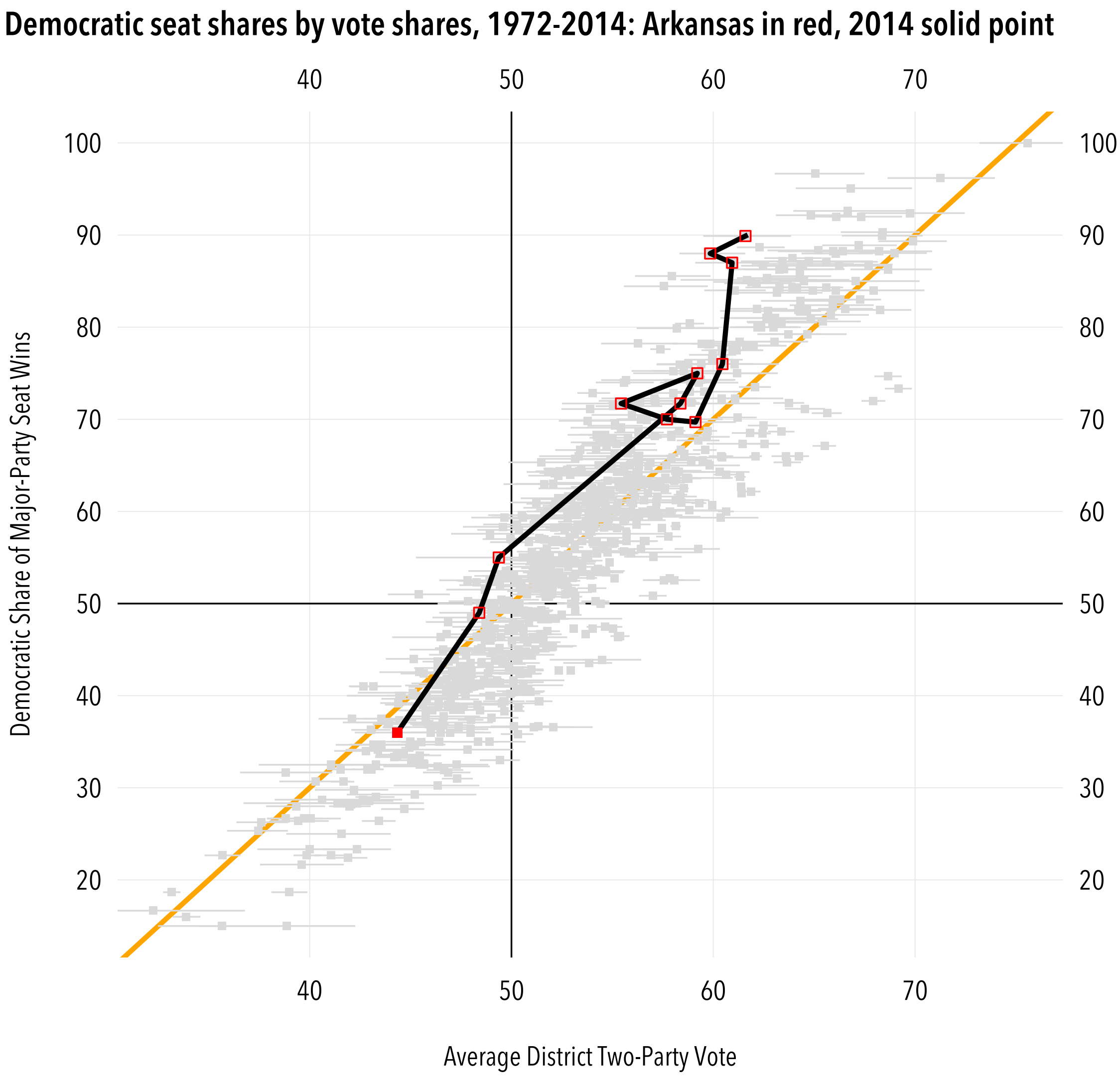


Figure 14: Arkansas, Democratic seat share and average district two-party vote share, 1992-2014. Orange line shows the seats-votes curve if the efficiency gap were zero; the efficiency gap in any election is the vertical distance from the corresponding data point to the orange line. Gray points indicate elections from other states and elections (1972-2014). Horizontal lines cover a 95% credible interval for Democratic average district two-party vote share, given imputations in uncontested districts.

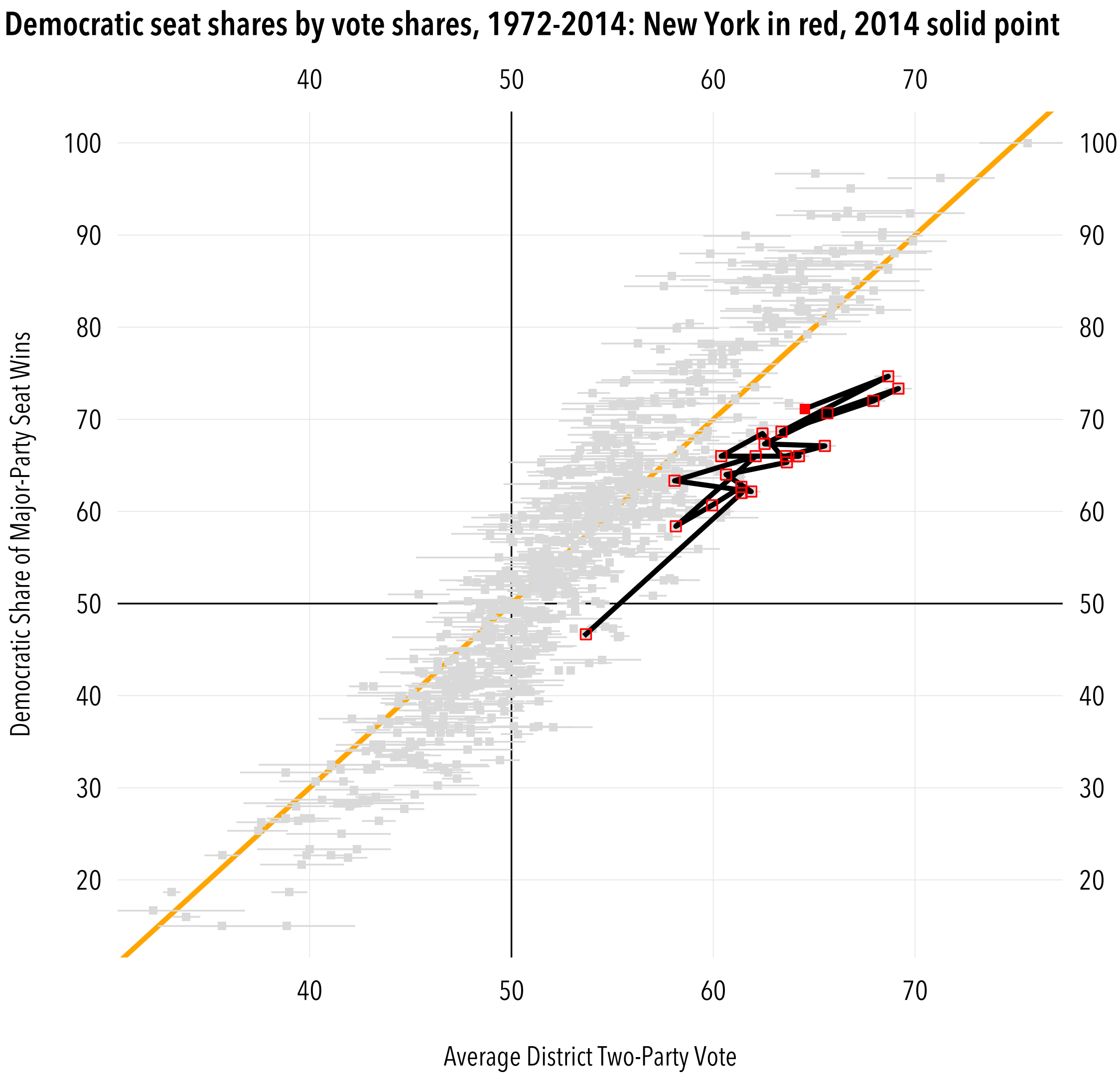


Figure 15: New York, Democratic seat share and average district two-party vote share, 1972-2014. Orange line shows the seats-votes curve if the efficiency gap were zero; the efficiency gap in any election is the vertical distance from the corresponding data point to the orange line. Gray points indicate elections from other states and elections (1972-2014). Horizontal lines cover a 95% credible interval for Democratic average district two-party vote share, given imputations in uncontested districts.

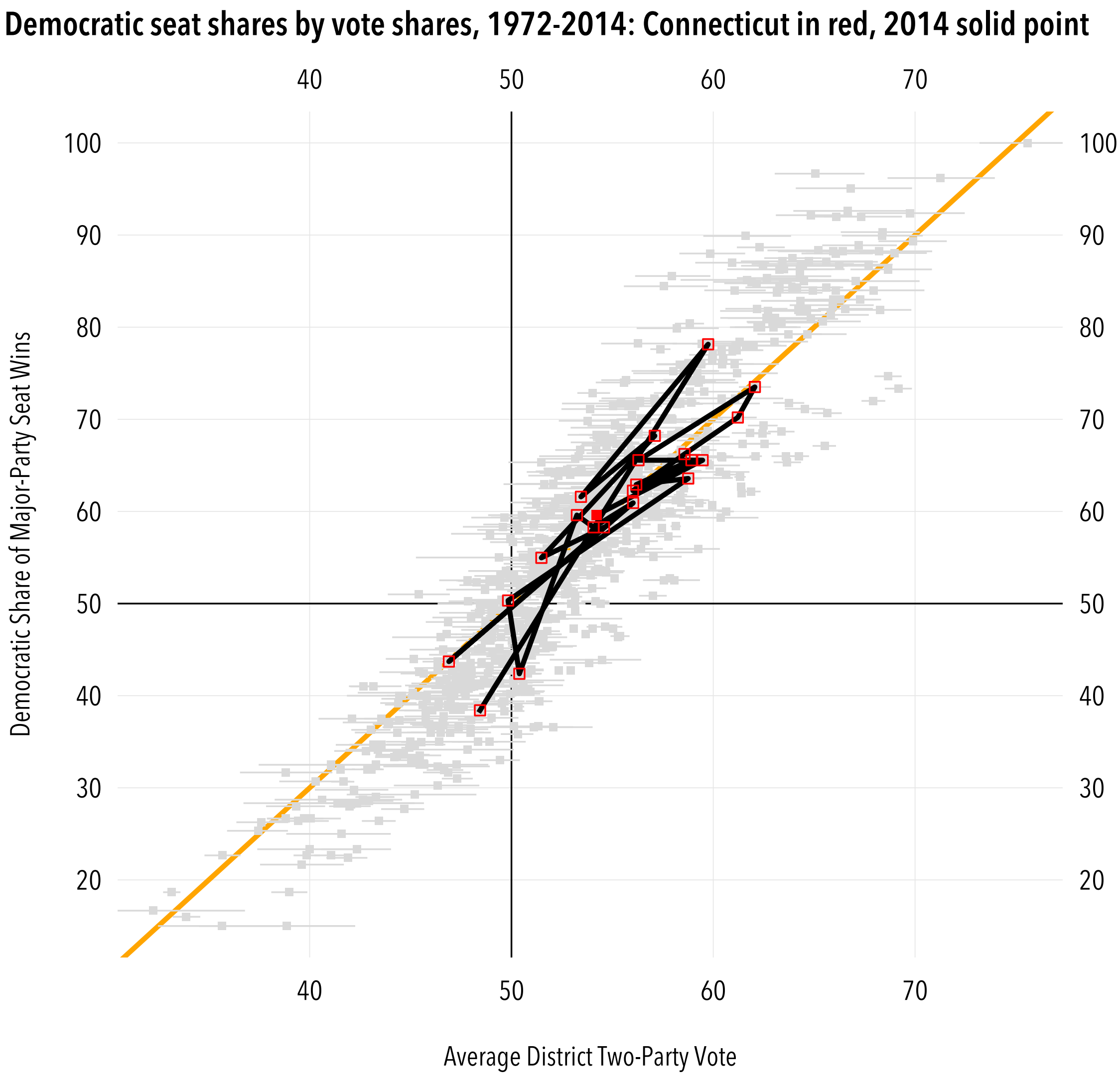


Figure 16: Connecticut, Democratic seat share and average district two-party vote share, 1972-2014. Orange line shows the seats-votes curve if the efficiency gap were zero; the efficiency gap in any election is the vertical distance from the corresponding data point to the orange line. Gray points indicate elections from other states and elections (1972-2014). Horizontal lines cover a 95% credible interval for Democratic average district two-party vote share, given imputations in uncontested districts.

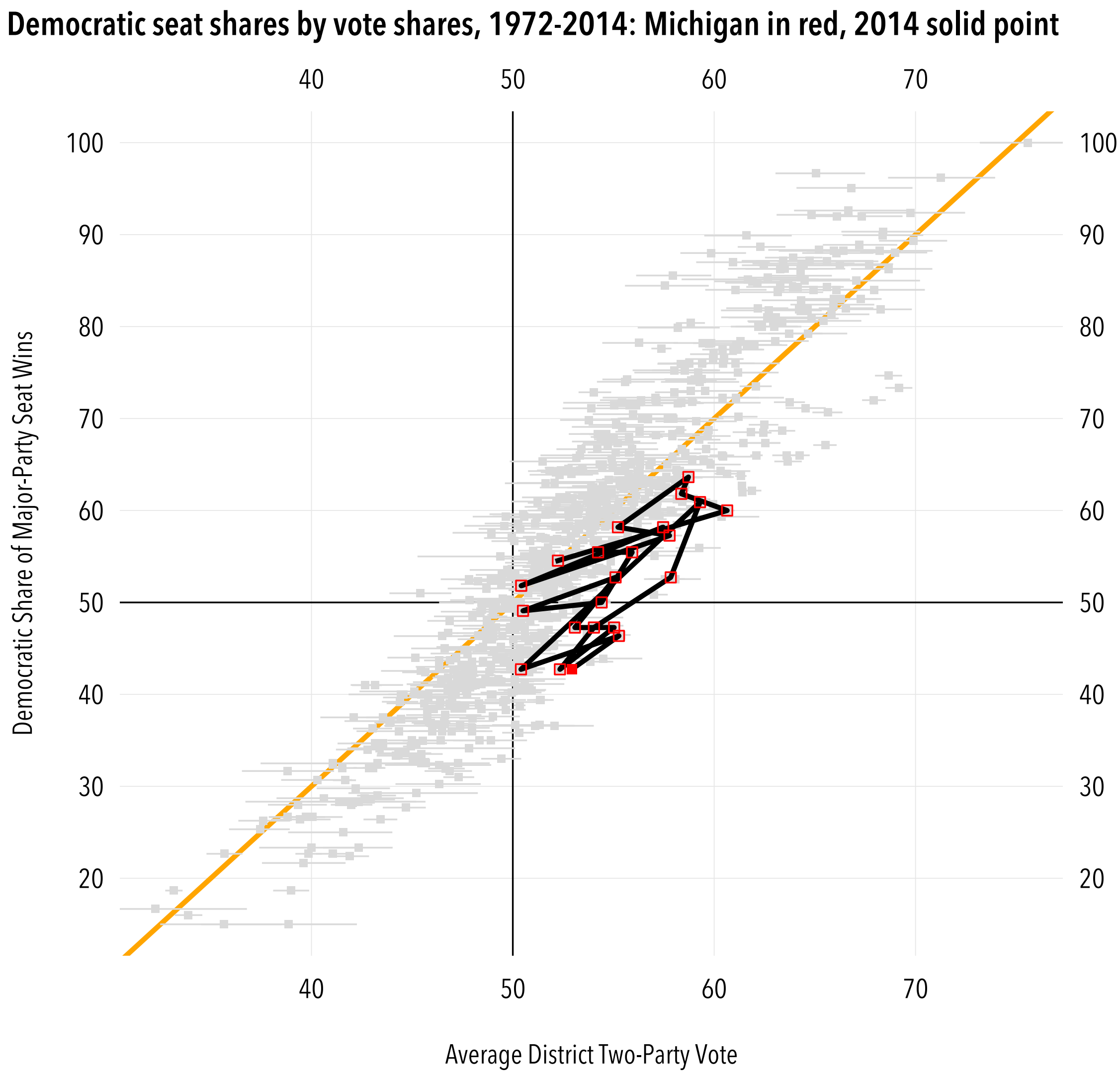


Figure 17: Michigan, Democratic seat share and average district two-party vote share, 1972-2014. Orange line shows the seats-votes curve if the efficiency gap were zero; the efficiency gap in any election is the vertical distance from the corresponding data point to the orange line. Gray points indicate elections from other states and elections (1972-2014). Horizontal lines cover a 95% credible interval for Democratic average district two-party vote share, given imputations in uncontested districts.

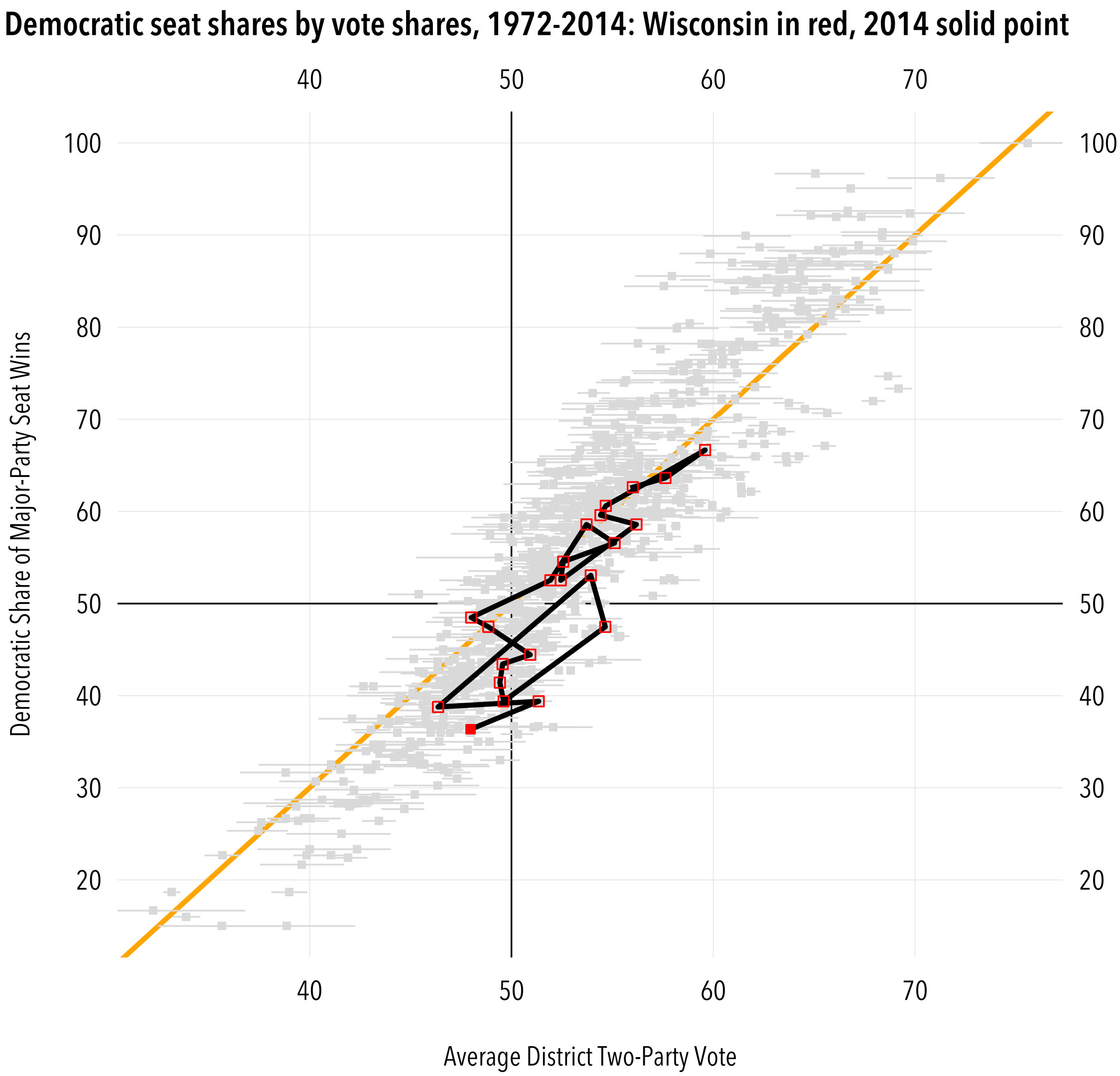


Figure 18: Wisconsin, Democratic seat share and average district two-party vote share, 1972-2014. Orange line shows the seats-votes curve if the efficiency gap were zero; the efficiency gap in any election is the vertical distance from the corresponding data point to the orange line. Gray points indicate elections from other states and elections (1972-2014). Horizontal lines cover a 95% credible interval for Democratic average district two-party vote share, given imputations in uncontested districts.

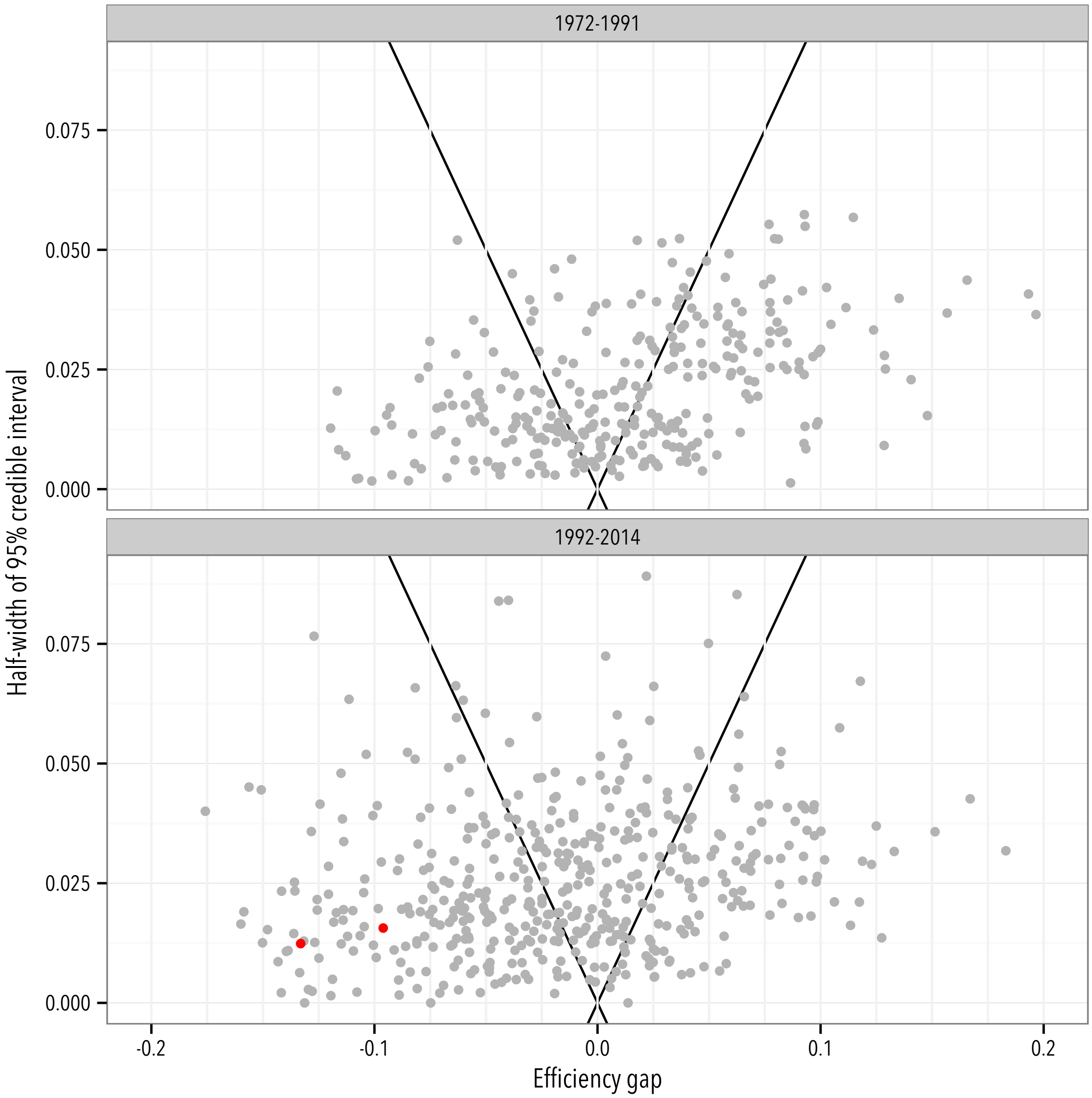


Figure 19: Uncertainty in the efficiency gap, against the *EG* estimate itself. The vertical axis is the half-width of the 95% credible interval for each *EG* estimate (plotted against the horizontal axis); points lying inside the cones have *EG* estimates that are small relative to their credible intervals, such that we would not distinguish them from zero at conventional levels of statistical significance. *EG* estimates from Wisconsin in 2012 and 2014 are shown as red points in the lower panel. Note the greater prevalence of large, negative and precisely estimated *EG* measures in recent decades.

9.2 Over-time change in the efficiency gap

Are large values of the efficiency gap less likely to be observed in recent decades? This is relevant to any discussion of a standard by which to assess redistricting plans. If recent decades have generally seen smaller values of the efficiency gap relative to past decades, then this might be informative as to how we should assess contemporary districting plans and their corresponding values of the *EG*.

Figure 20 plots *EG* estimates over time, overlaying estimates of the smoothed, weighted quantiles (25th, 50th and 75th) of the *EG* measures (the weights capture the uncertainty accompanying each estimate of the *EG*). The distribution of *EG* measures in the 1970s and 1980s appeared to slightly favor Democrats; about two-thirds of all *EG* measures in this period were positive. The distribution of *EG* measures trends in a pro-Republican direction through the 1990s, such that by the 2000s, *EG* measures were more likely to be negative (Republican efficiency advantage over Democrats); see Figure 21.

There is some evidence that the 2010 round of redistricting has generated an increase in the magnitude of the efficiency gap in state legislative elections. For most of the period under study, there seems to be no distinct trend in the magnitudes of the efficiency gap over time; see Figure 22. The median, absolute value of the efficiency gap has stayed around 0.04 over much of the period spanned by this analysis; elections since 2010 are producing higher levels of *EG* in magnitude.

It is also interesting to note that the estimate of the 75th percentile of the distribution of *EG* magnitudes jumps markedly after 2010, suggesting that districting plans enacted after the 2010 census are systematically more gerrymandered than in previous decades. Of the almost 800 *EG* estimates in the analysis, spanning 42 years of elections, the largest, negative estimates (an efficiency gap disadvantaging Democrats) are more likely to be recorded in the short series of elections after 2010. These include Alabama in 2014 (-.18), Florida in 2012 (-.16), Virginia in 2013 (-.16), North Carolina in 2012 (-.15) and Michigan in 2012 (-.14); these five elections are among the 10 least favorable to Democrats we observe in the entire set of elections. Among the 10 most pro-Democratic *EG* scores, *none* were recorded after 2000. The most favorable election to Democrats in terms of *EG* since 2010 is the 2014 election in Rhode Island ($EG = .12$), which is only the 20th largest (pro-Democratic) *EG* in the entire analysis.

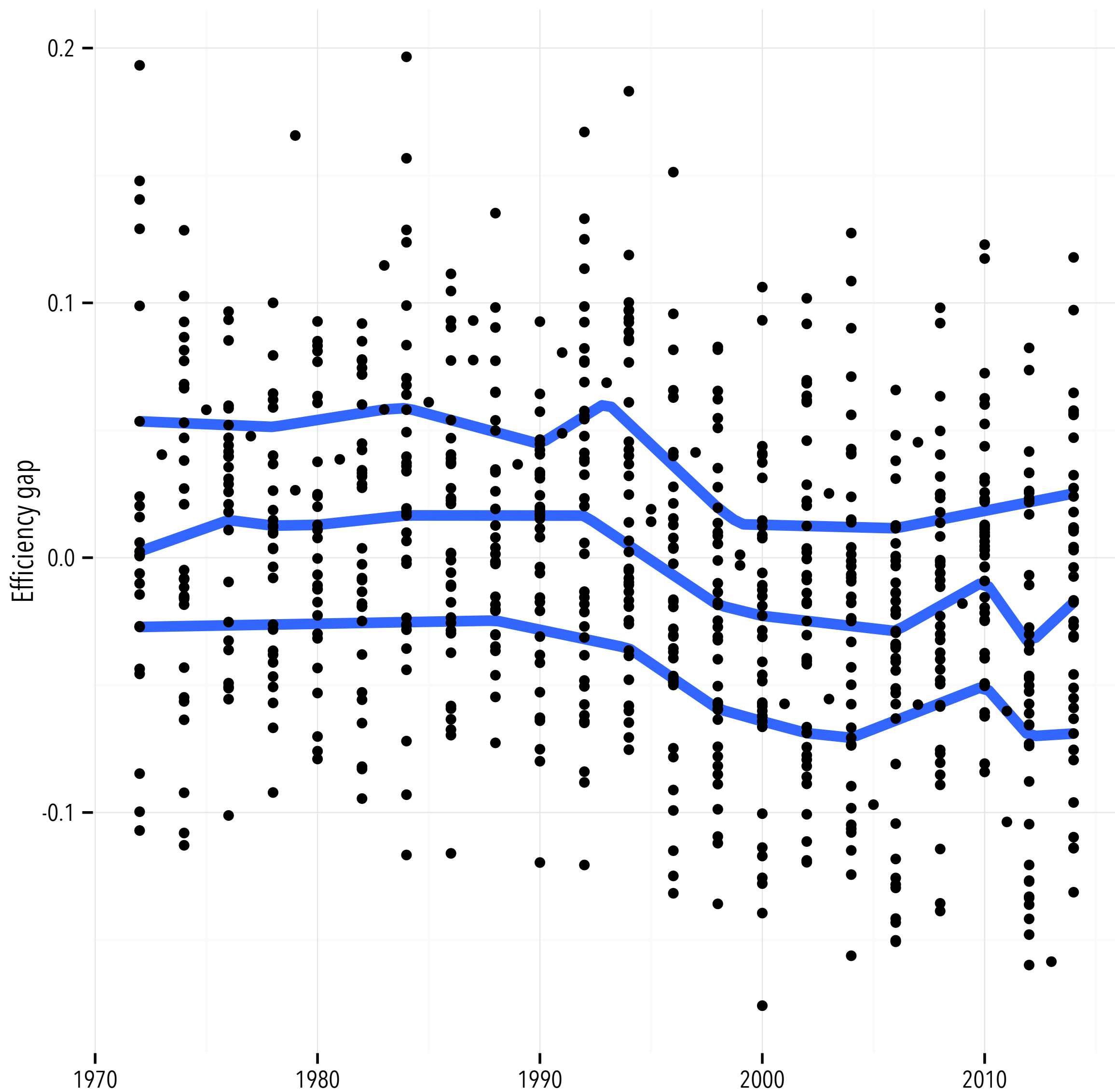


Figure 20: Efficiency gap estimates, over time. The lines are smoothed estimates of the 25th, 50th and 75th quantiles of the efficiency gap measures, weighted by the precision of each *EG* measure.

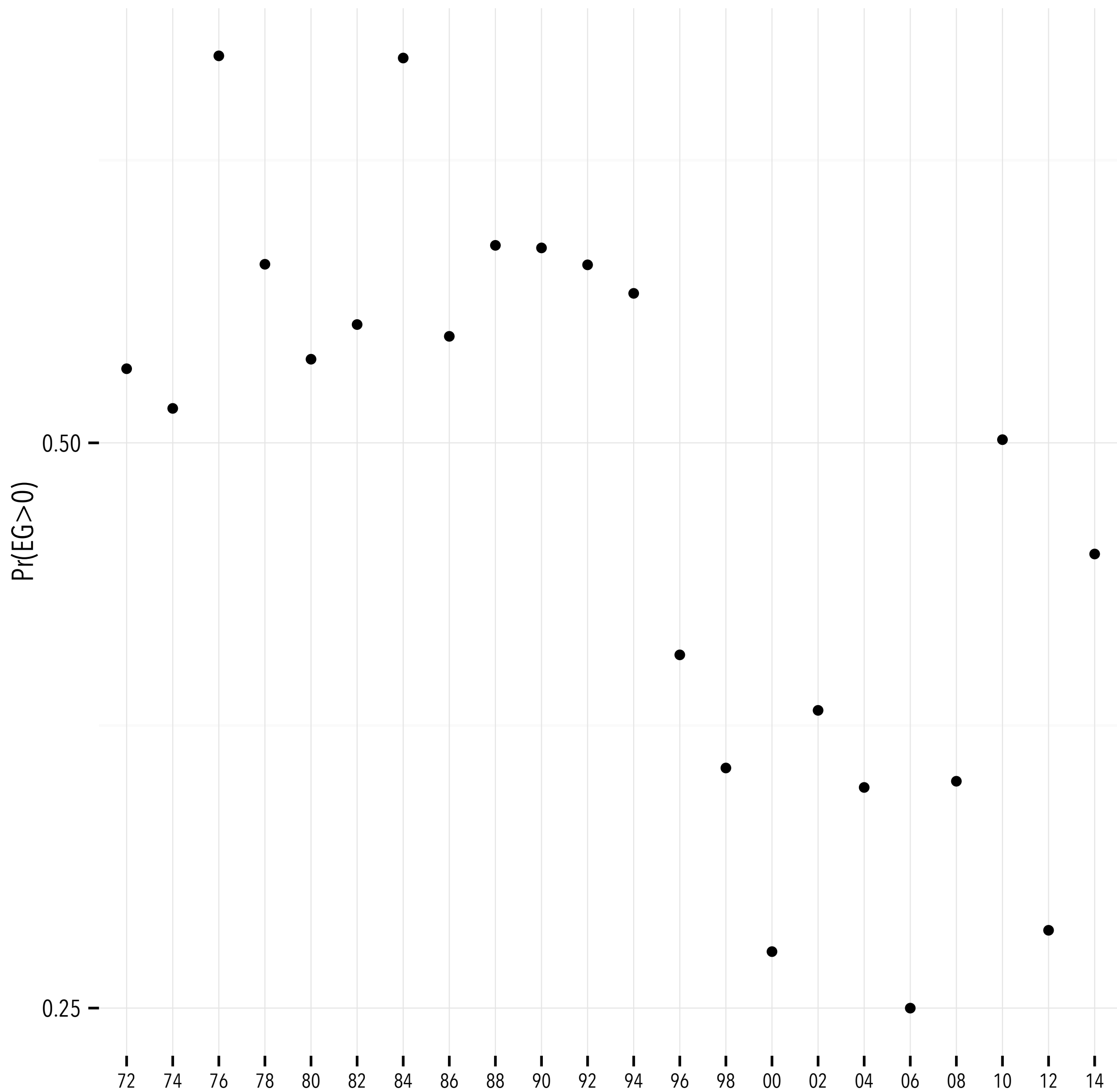


Figure 21: Proportion of efficiency gap measures that are positive, by two year intervals.

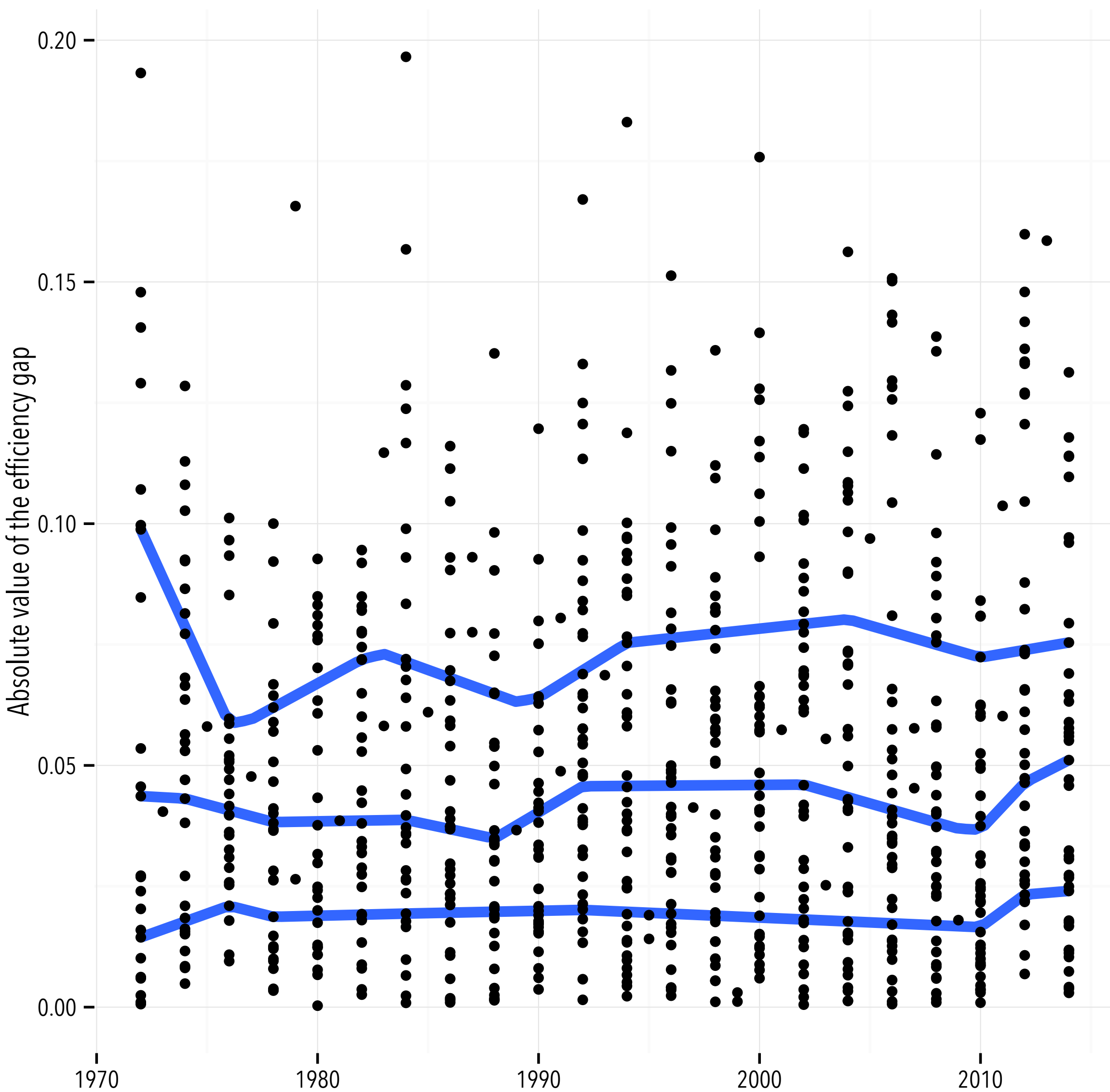


Figure 22: Absolute value of efficiency gap measures, over time. The lines are smoothed estimates of the 25th, 50th and 75th quantiles of the absolute value of the efficiency gap measure, weighted by the precision of each *EG* measure.

9.3 Within-plan variation in the efficiency gap

The efficiency gap is measured at each election, with a given districting plan typically generating up to five elections and hence five efficiency gap measures. Efficiency gap measures will change from election to election as the distribution of district-level vote shares varies over elections. Some of this variation is to be expected. Even with the same districting plan in place, districts will display "demographic drift," gradually changing the political complexion of those districts. Incumbents lose, retire or die in office; sometimes incumbents face major opposition, sometimes they don't. Variation in turnout — most prominently, from on-year to off-year — will also cause the distribution of vote shares to vary from election to election, even with the districting plan unchanged. All these election-specific factors will contribute to election-to-election variation in the efficiency gap.

Precisely because we expect a reasonable degree of election-to-election variation in the efficiency gap, we assess the magnitude of this "within-plan" variability in the measure. If a plan is a partisan gerrymander — with a systematic advantage for one party over the other — then the "between-plan" variation in *EG* should be relatively large relative to the "within-plan" variation in *EG*.

About 76% of the variation in the *EG* estimates is between-plan variation. The *EG* measure does vary election-to-election, but there is a moderate to strong "plan-specific" component to variation in the *EG* scores. We conclude that the efficiency gap *is* measuring an enduring feature of a districting plan.

We examine some particular districting plans. The 786 elections in this analysis span 150 districting plans. For plans with more than one election, we compute the standard deviation of the sequence of election-specific *EG* measures observed under the plan. These standard deviations range from .011 (Kentucky's plan in place for just two elections in 1992 and 1994, or Indiana's plan 1992-2000) to .079 (Delaware's plan between 2002 and 2010).

A highly variable plan: Deleware 2002-2010. Figure 23 shows the seats, votes and *EG* estimates produced under the Delaware 2002-2010 plan. This is among the most variable plans we observe with respect to the *EG* measure. An efficiency gap running against the Democrats for 2002, 2004 and 2006 (the latter election saw Democrats win only 18 seats out of 41 with 54.5% of the state wide vote) falls to a small gap in 2008 ($V = 0.584, S = 25/41 = .61, EG = -0.058$) and

Delaware ends the decade with a positive efficiency gap in 2010. The Democratic district-average two-party vote share fell to $V = 0.561$ in 2010, but translated into $S = 26/41 = .63, EG = 0.012$.

A plan with moderate variability in the EG. The median, within-plan standard deviation of the *EG* is about .03. This roughly corresponds to the within-plan standard deviation of the *EG* observed under the plan in place for five Wisconsin state legislative elections 1992-2000, presented in Figure 24. This was a plan that generated relatively small values of *EG* that alternated sign over the life of the plan: negative in 1992, positive in 1994 and 1996, and negative in 1998 and 2000.

A low variance plan, Indiana 1992-2000. See Figure 25. The *EG* measures recorded under this plan are all relatively small and positive, ranging from 0.008 to 0.041 and correspond to an interesting period in Indiana state politics. Democrats won 55 of the 100 seats in the Indiana state house in the 1992 election with what I estimate to be just over 50% of the district-average vote (29 of 100 seats were uncontested). Democratic vote share fell to about 45% in the 1994 election (38 uncontested seats), and Democrats lost control of the legislature. The 1996 election resulted in a 50-50 split in the legislature. Democrats won legislative majorities in the 1998 and 2000 elections, while the last election might have been won by Democrats with just less than 50% of the district-vote; I estimate $V = 0.495 \pm .012$ and $EG = 0.041$.

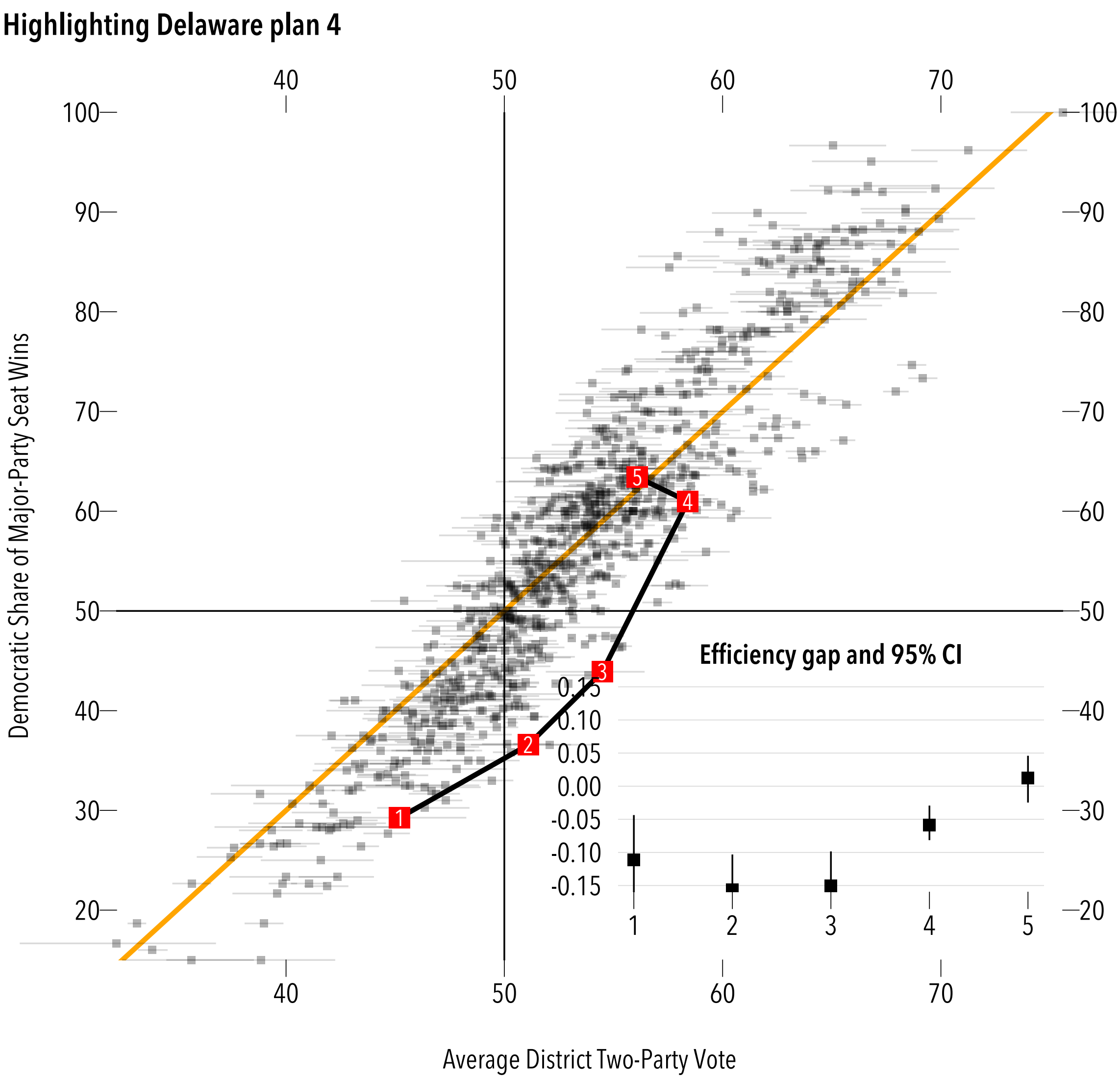


Figure 23: Seats, votes and the efficiency gap recorded under the Delaware plan, 2002-2010. Orange line shows the seats-votes curve if the efficiency gap were zero; the efficiency gap in any election is the vertical distance from the corresponding data point to the orange line. Gray points indicate elections from other states and elections (1972-2014). Horizontal lines cover a 95% credible interval for Democratic average district two-party vote share, given imputations in uncontested districts. The inset in the lower right shows the sequence of efficiency gap measures recorded under the plan; vertical lines are 95% credible intervals.

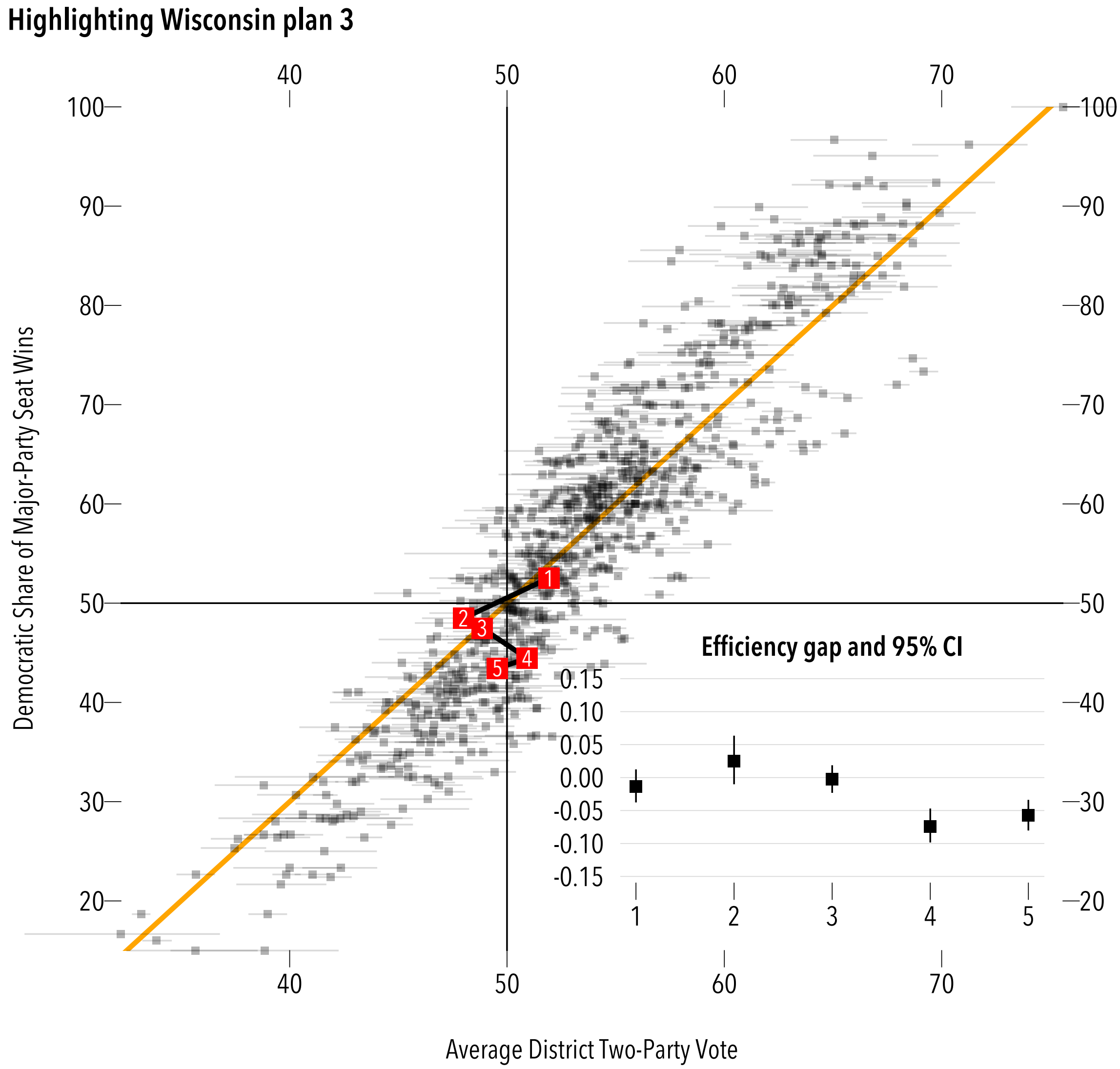


Figure 24: Seats, votes and the efficiency gap recorded under the Wisconsin plan, 1992-2000. Orange line shows the seats-votes curve if the efficiency gap were zero; the efficiency gap in any election is the vertical distance from the corresponding data point to the orange line. Gray points indicate elections from other states and elections (1972-2014). Horizontal lines cover a 95% credible interval for Democratic average district two-party vote share, given imputations in uncontested districts. The inset in the lower right shows the sequence of efficiency gap measures recorded under the plan; vertical lines are 95% credible intervals.

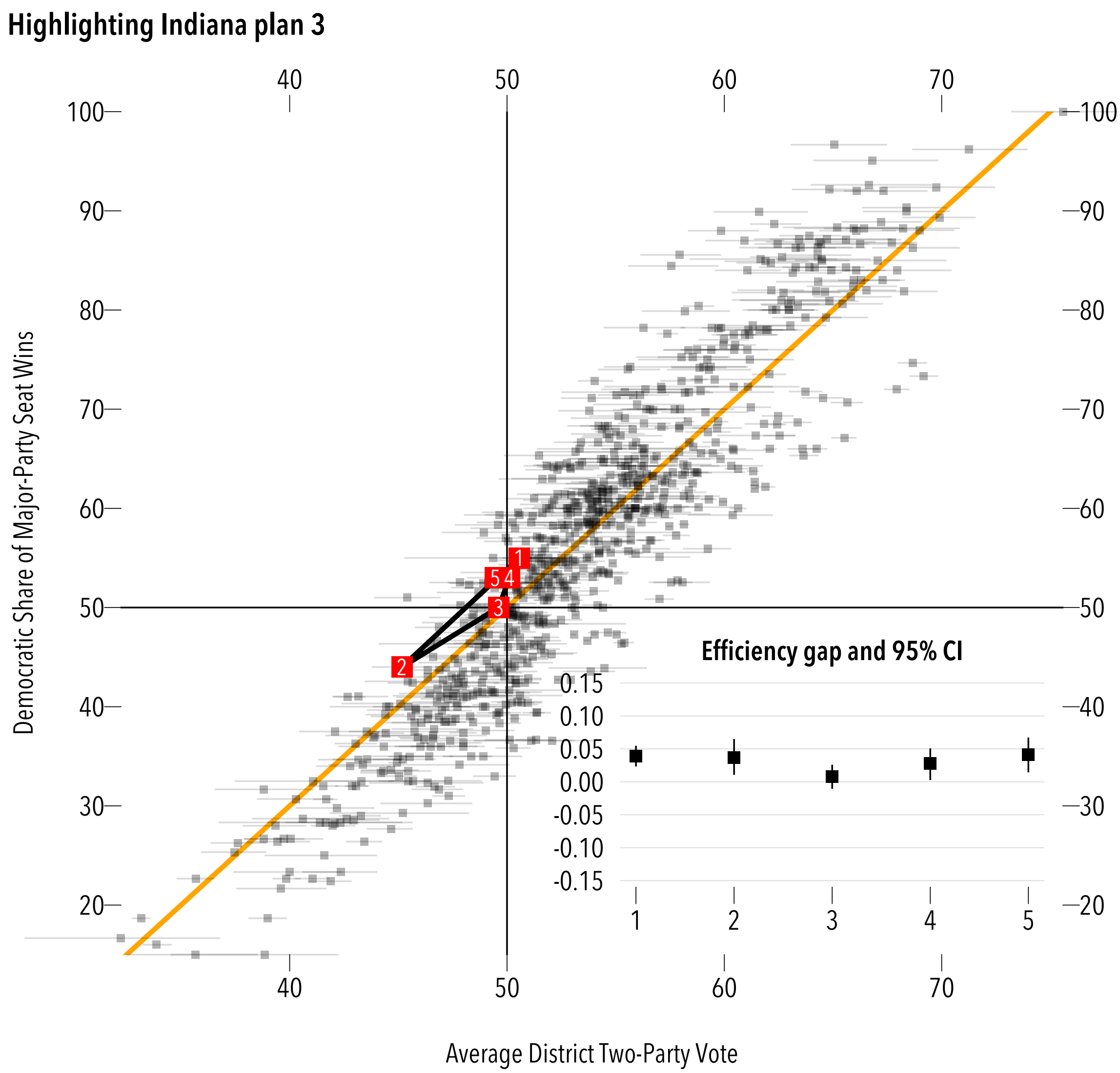


Figure 25: Seats, votes and the efficiency gap recorded under the Indiana plan, 1992-2000. Orange line shows the seats-votes curve if the efficiency gap were zero; the efficiency gap in any election is the vertical distance from the corresponding data point to the orange line. Gray points indicate elections from other states and elections (1972-2014). Horizontal lines cover a 95% credible interval for Democratic average district two-party vote share, given imputations in uncontested districts. The inset in the lower right shows the sequence of efficiency gap measures recorded under the plan; vertical lines are 95% credible intervals.

9.4 How often does the efficiency gap change sign?

Having observed a particular value of *EG*, how confident are we that:

- the *EG* measure is distinguishable from zero at conventional levels of statistical significance? That is, how sure are we as to the sign of any particular *EG* estimate? We addressed this question in section 9.1.
- it will be followed by one or more estimates of *EG* that are of the same sign?
- over the life of a districting plan, *EG* remains on one side of zero or the other?

The latter two questions are key. It is especially important that we assess the *durability* of the sign of the *EG* measure under a districting plan, if we seek to assert that a districting plan is a partisan gerrymander. We will see that *magnitude* and *durability* of the efficiency gap go together: *large* values of the efficiency gap don't seem to be capricious, but likely to be repeated over the life of a districting plan, consistent with partisan disadvantage being a systematic feature of the plan.

We begin this part of the analysis by considering temporally adjacent *pairs* of *EG* estimates. Can we be confident that these have the same sign? In general, yes. Of the full set of 786 elections for which we compute an efficiency gap estimate, 580 are temporally adjacent, within state and districting plan. Figure 26 shows that we usually see efficiency gap measures with the same sign; this probability exceeds 90% for almost half of the temporally adjacent pairs of efficiency gap measures. Averaged over all pairs, this "same sign" probability is 74%. While the efficiency gap does vary election to election, these fluctuations are not so large that the *sign* of the efficiency gap is likely to change election to election.

What about over the life of an entire redistricting plan? How likely is it that the efficiency gap retains the same sign over, say, three to five elections in a given state, taking into account election-to-election variation *and* uncertainty arising from the imputation procedures used for uncontested districts?

We have 141 plans that supply three or more elections with estimate of the efficiency gap. Of these, 17 plans are *utterly unambiguous* with respect to the sign of the efficiency gap estimates recorded over the life of the plan: for each of these plans we estimate the probability that the *EG* has the same sign over the life of the plan to be 100%. These plans are listed below in Table 1.

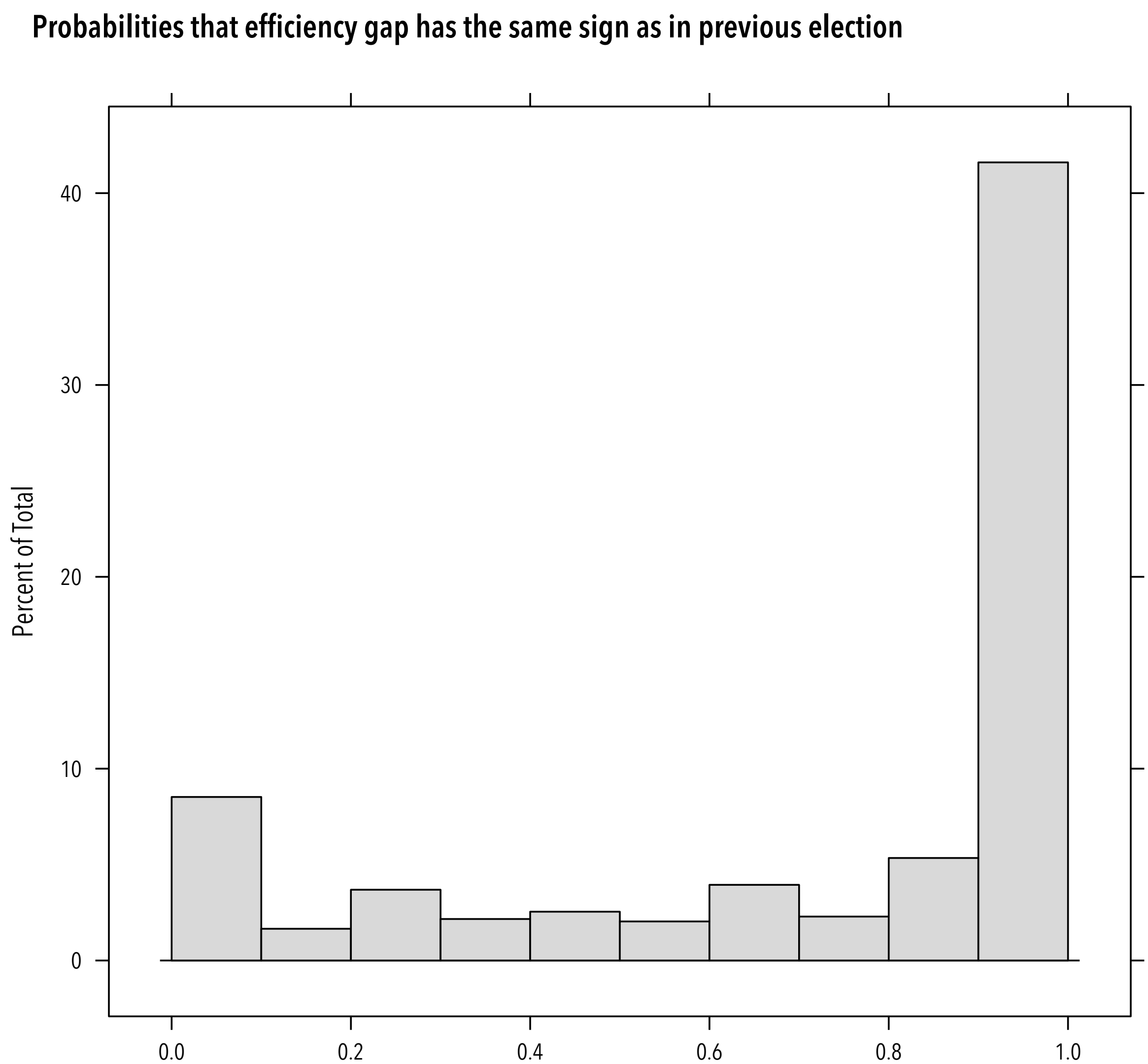


Figure 26: Stability in 580 successive pairs of efficiency gap measures

State	Plan	Start	End	EG avg	EG min	EG max
Florida	4	2002	2010	-0.112	-0.136	-0.084
New York	4	2002	2010	-0.111	-0.150	-0.078
Illinois	3	1992	2000	-0.103	-0.136	-0.058
Michigan	4	2002	2010	-0.103	-0.130	-0.077
New York	3	1992	2000	-0.098	-0.139	-0.048
New York	1	1972	1980	-0.097	-0.108	-0.079
Missouri	4	2002	2010	-0.091	-0.142	-0.061
Ohio	4	2002	2010	-0.090	-0.143	-0.049
New York	2	1982	1990	-0.084	-0.120	-0.028
Ohio	3	1994	2000	-0.083	-0.109	-0.025
Michigan	3	1992	2000	-0.080	-0.128	-0.019
Wisconsin	4	2002	2010	-0.076	-0.118	-0.039
Colorado	2	1982	1990	-0.075	-0.117	-0.055
Colorado	1	1972	1980	-0.041	-0.067	-0.018
California	3	1992	2000	-0.041	-0.057	-0.018
Pennsylvania	2	1982	1990	-0.033	-0.056	-0.020
Florida	1	1972	1980	0.070	0.052	0.099

Table 1: Plans with no doubt as to the sign of the efficiency gap over the life of the plan (3+ elections).

Interestingly, these plans with an utterly unambiguous history of one-sided *EG* measures are almost all plans with efficiency gaps that are disadvantagous to Democrats. Michigan's 2002-2010 plan is on this list, as is the plan in place in Wisconsin 2002-2010 (average *EG* of -.076).

We examine this probability of “3+ consecutive *EG* measures with the same sign” for all of the plans with 3 or more elections in this analysis. 35% of 141 plans with 3 or more elections have at least a 95% probability of recording plans with *EG* measures with the same sign. If we relax this threshold to 75%, then 46% of plans with 3 or more elections exhibit *EG* measures with the same sign. Again, there is a reasonable amount of within-plan movement in *EG*, but in a large proportion of plans the efficiency gap appears to be a stable attribute of the plan.

10 A threshold for the efficiency gap

We now turn to the question of what might determine a threshold for determining if the EG is a *large and enduring* characteristic of a plan. We pose the problem as follows:

> for a given threshold $EG^* > 0$, what is the probability that having observed a value of $EG \geq EG^*$ we then see $EG < 0$ in the remainder of the plan?

To answer this we compute

- if (and optionally, when) a plan has $EG \geq EG^*$;
- conditional on seeing $EG \geq EG^*$, do we also observe $EG < 0$ (a sign flip) in the same districting plan?

For $EG < 0$, the computations are reversed: conditional on seeing $EG < EG^*$, do we also see $EG > 0$ under the same plan?

Figure 27 displays two proportions, plotted against a series of potential thresholds on the horizontal axis. The two plotted proportions are:

- the proportion of plans in which we observe an EG more extreme than the specified threshold EG^* (on the horizontal axis);
- among the plans that trip the specified threshold, the proportion in which we see a EG in the same plan with a different sign to EG^*.

Plans with at least one election with $|EG| > .07$ are reasonably common: over the entire set of plans analyzed here — and again, with the uncertainty in EG estimates taken into account — there is about a 20% chance that a plan will have at least one election with $|EG| < .07$.

Observing $EG > .07$ is not a particularly informative signal with respect to the other elections in the plan. Conditional on observing an election with $EG > .07$ (an efficency gap favoring Democrats), there is an a 45% chance that *under the same plan* we will observe $EG < 0$. That is, making an inference about a plan on the basis of one election with $EG > .07$ would be quite risky. Estimates

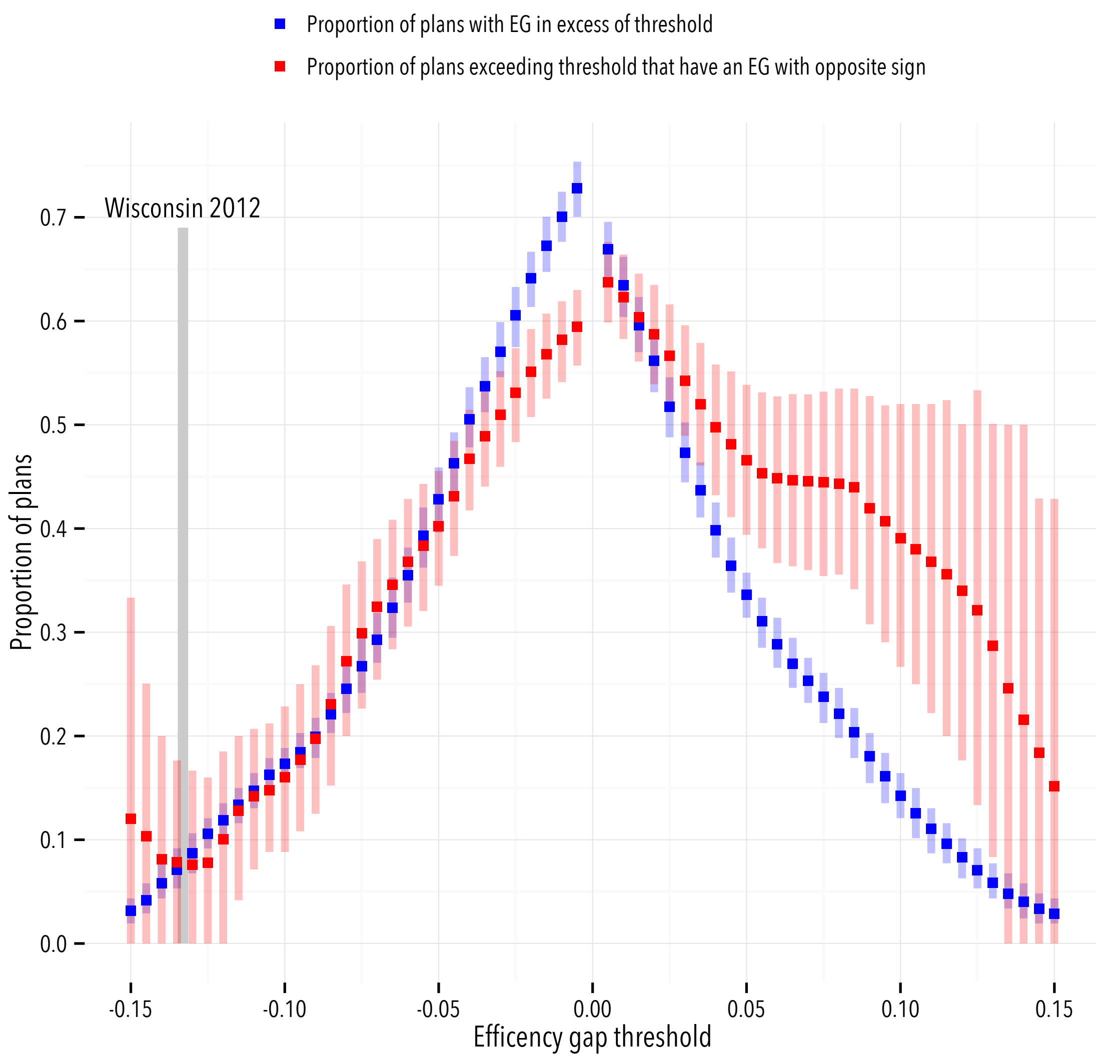


Figure 27: Proportion of plans that (a) record an efficiency gap measure at least as extreme as the value on the horizontal axis; and (b) conditional on at least one election with *EG* in excess of this threshold (not necessarily the first election), the proportion of plans where there is another election in the plan with an *EG* of the opposite sign.

of the "sign flip" rate conditional on a plan generating a relatively large, pro-Democratic *EG* estimates are quite unreliable because there are so few plans generating large, pro-Democratic *EG* estimates to begin with; note the confidence intervals on the "sign flip" rate get very wide as the data become more scarce on the right hand side of the graph.

This finding is not symmetric. The "signal" $EG < -.07$ (an efficiency gap disadvantageous to Democrats) is much more informative about other elections in the plan than the opposite signal $EG > .10$ (a pro-Democratic efficiency gap). If any single election in the plan has $EG < -.07$ then the probability that *all* elections in the plan have $EG < 0$ is about .80. That is, there is a smaller degree of within-plan volatility in plans that disadvantage Democrats. Observing a relatively low value of the *EG* such as $EG < -.07$ is much more presumptive of a systematic and enduring feature of a redistricting plan than the opposite signal $EG > .07$. Efficiency gap measures that appear to indicate a disadvantage for Democrats are thus more reliable signals about the respective districting plan than efficiency gap measures indicating an advantage for Democrats.

We repeat this previous exercise, but restricting attention to more recent elections and plans, with the results displayed in Figure 28. Again we see that plans with pro-Democratic *EG* measures are quite likely to also generate an election with $EG < 0$; and again, note that estimates of the "sign flip" rate are quite unreliable because there are so few plans generating large, pro-Democratic *EG* estimates to begin with.

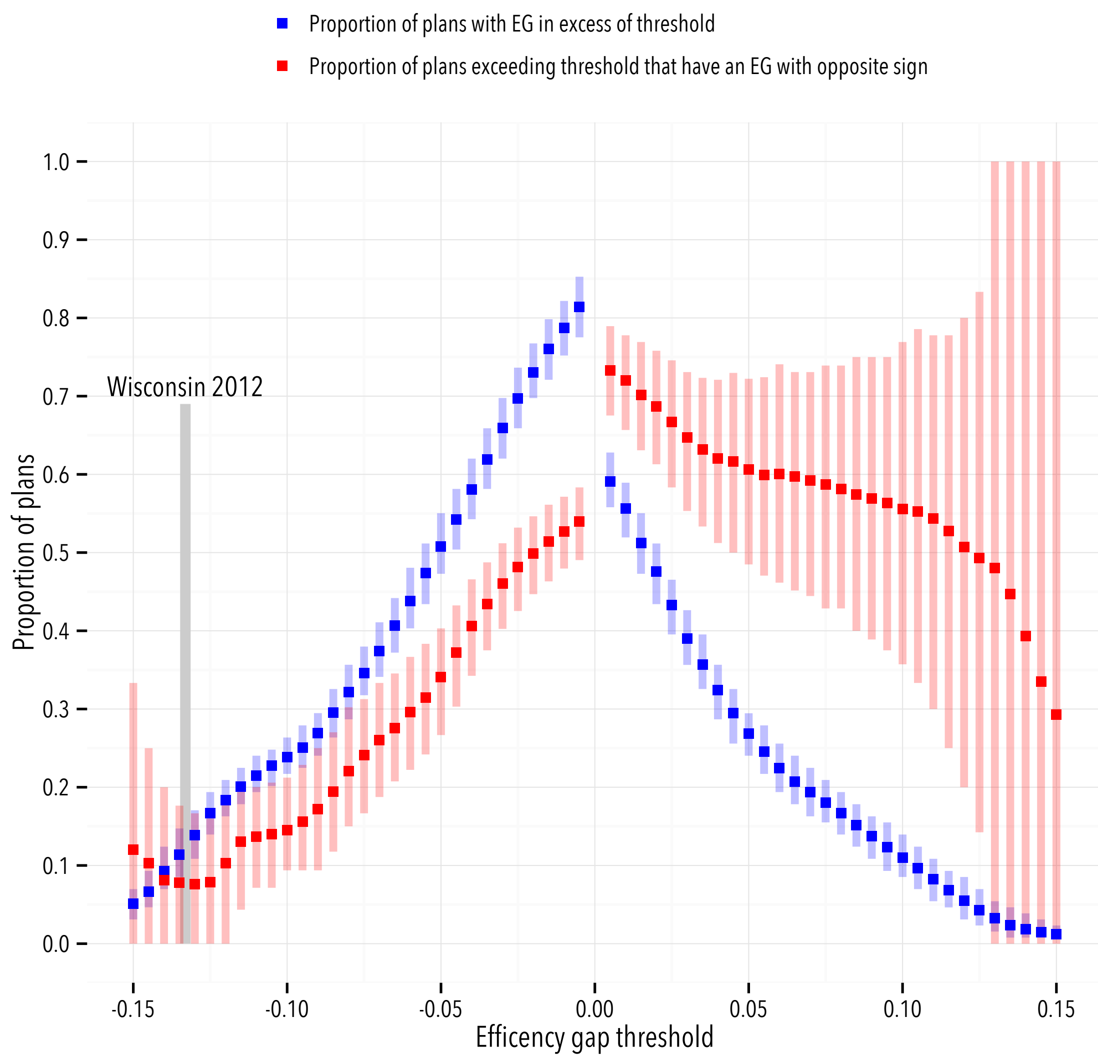


Figure 28: Proportion of plans in which (a) the efficiency gap measure is at least as extreme as the value on the horizontal axis; and (b) of these plans with at least one election with *EG* in excess of this threshold (not necessarily the first election), the proportion of plans in which there is another election in the plan with an *EG* of the *opposite* sign. Analysis of state legislative elections in 129 plans, 1991-present.

10.1 Conditioning on the first election in a districting plan

We also compute this probability of a sign flip in *EG* conditional on the magnitude of the *EG* observed with the *first* election under a districting plan. We perform this analysis twice: (1) for all elections in the data set and (2) for elections held under plans adopted in 1991 or later.

Figures 29 and 30 display the results of these analyses. First, over the full set of data (Figure 29) we observe a roughly symmetric set of *EG* scores in the first election under a plan. But we seldom see plans in the 1990s or later that commence with a large, pro-Democratic efficiency gap; the probability of a first election having $EG > .10$ is zero and the probability of a first election having $EG > .05$ (historically, not a large *EG*) is only about 11%. Negative efficiency gaps (not favoring Democrats) are much more likely under the first election in the post-1990 plans: almost 40% of plans open with $EG < -.05$ and about 20% of plans open with $EG < -.10$.

As noted earlier, pro-Democratic efficiency gaps seem much more fleeting than pro-Republican efficiency gaps. Conditional on a pro-Republican estimate of $EG > 0$ in the first election under a plan, the probability of seeing *EG* change sign over the life of the plan is almost always around 40% (1972-2014, Figure 29) or 50% (1991-present, Figure 30).

A very different conclusion holds if the first election observed under a plan indicates a sizeable efficiency gap working to disadvantage Democrats. In fact, the more negative the initial *EG* observed under a plan, the more confident we can be that we will continue to observe $EG < 0$ over the sequence of elections to follow under the plan. Conditional on a first election with $EG < -.10$, the probability of *all subsequent* efficiency gaps being negative is about 85%. Indeed, it is more likely than not that if the first election has $EG < 0$ (no matter how small), then so too will all subsequent elections (a 60% chance of this event).

Note that the Current Wisconsin Plan opens with $EG = -.13$ in the 2012 election. Analysis of efficiency gap measures in the post-1990 era (Figure 30) indicates that conditional on an *EG* measure of this size and sign, there is a 100% probability that *all subsequent elections* held under that plan will also have efficiency gaps disadvantageous to Democrats. That is, in the post-1990 era, if a plan's first election yields $EG \leq -.13$, we *never* see a subsequent election under that plan yielding a pro-Democratic efficiency gap. In short, a signal such as

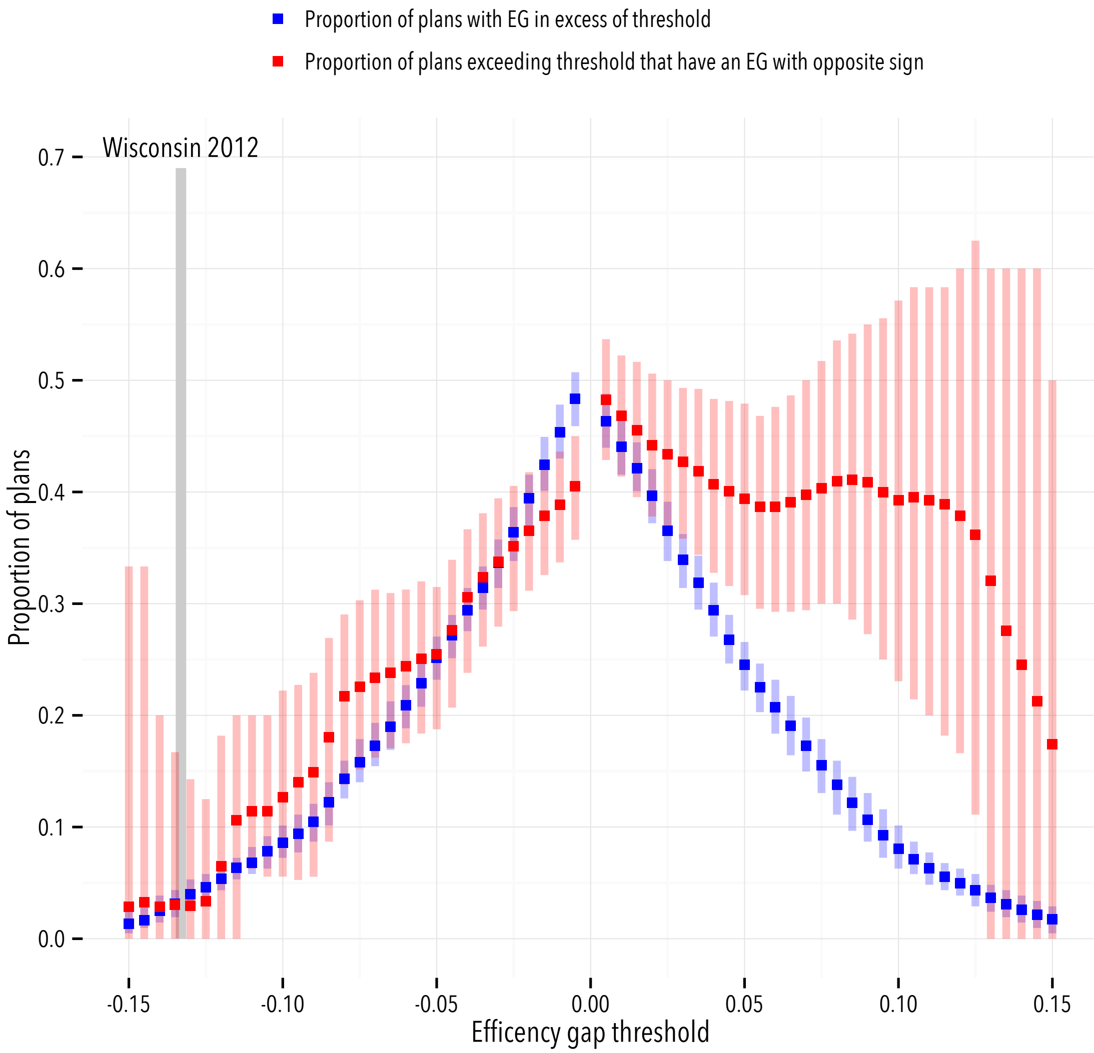


Figure 29: Proportion of plans in which the *first election* (a) has an efficiency gap measure at least as extreme as the value on the horizontal axis; and (b) conditional on the first election having an *EG* in excess of this threshold, the proportion of those plans in which a *subsequent election* has an *EG* of the *opposite* sign. Analysis of all state legislative elections in all plans with more than one election, 1972-present.

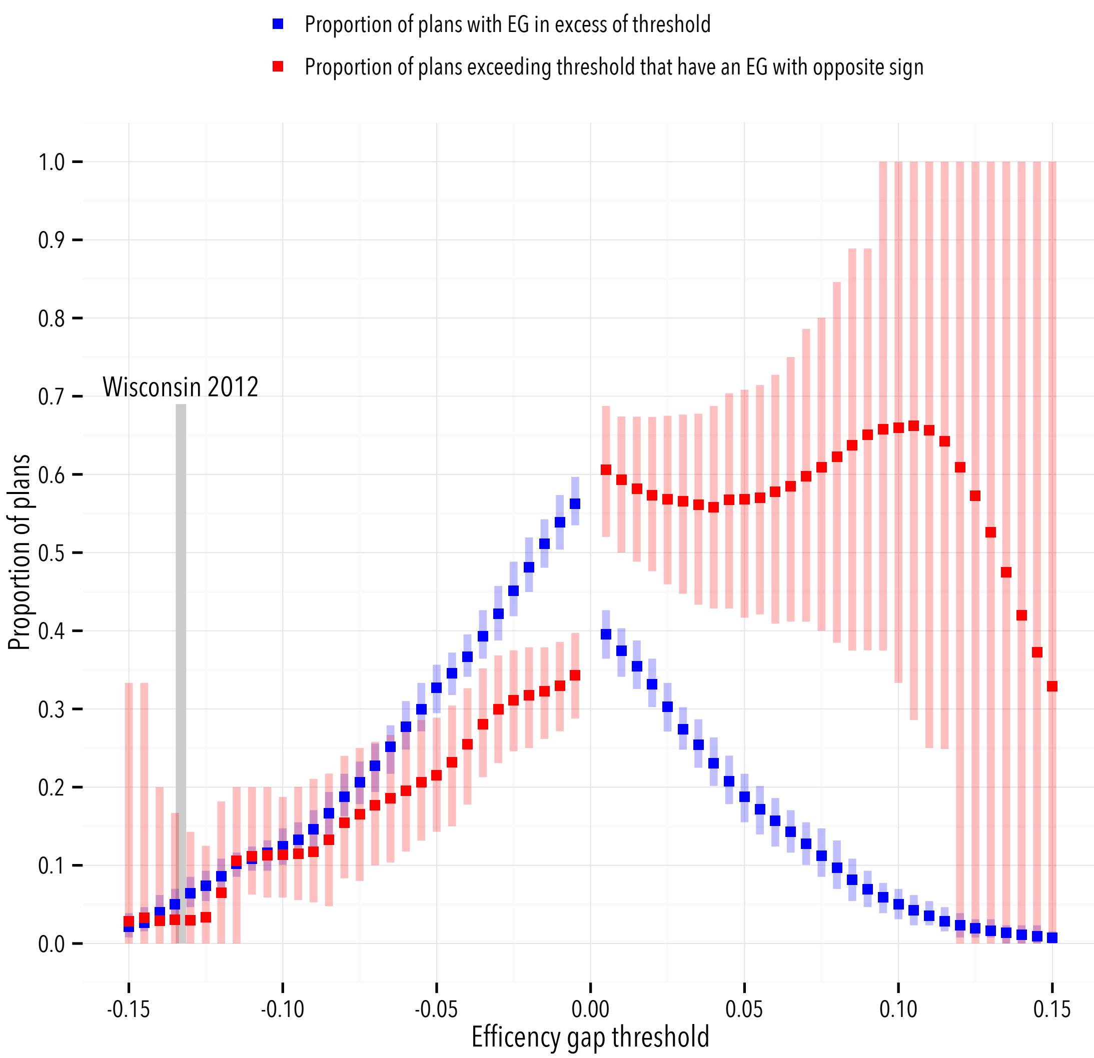


Figure 30: Proportion of plans in which the *first election* (a) has an efficiency gap measure at least as extreme as the value on the horizontal axis; (b) conditional on the first election having an *EG* in excess of this threshold, the proportion of those plans in which a *subsequent election* has an *EG* of the *opposite* sign. Analysis of state legislative elections in 129 plans, 1991-present.

$EG \leq -.13$ is extremely reliable with respect to the districting plan that generated it, at least given the post-1990 record.

10.2 Conditioning on the first two elections in a districting plan

The difficulty with conditioning on the first two elections of a districting plan is that the data start to thin out. In the entire data set there simply aren't many districting plans that equal or surpass the two, relatively large values of *EG* observed in Wisconsin in the first two elections of the current plan. Indeed, the only cases with a similar history of *EG* measures like Wisconsin's in 2012 and 2014 are contemporaneous cases: Florida, Michigan, and North Carolina in 2012 and 2014.

We relax the threshold of what counts as a similar case to encompass plans whose first two efficiency gap measures are within 75% of the magnitude of Wisconsin's 2012 and 2014 *EG* measures; we now pick up 11 roughly comparable cases, 4 of which date from earlier decades. Again, this is testament to how recent decades have seen an increase in the prevalence of larger, negative values of the efficiency gap.

For the four prior cases we plot the sequence of *EG* estimates in Figure 31. With the exception of the last election in the highly unusual Delaware sequence (among the most volatile observed in the data set; see section 9.3), the other proximate cases all go on to record efficiency gap measures that are below zero over the balance of the plan. We stress that four cases doesn't provide much basis for comparison, but this only speaks to the fact that the sequence of two large, negative values of the efficiency gap in Wisconsin in 2012 and 2014 are virtually without historical precedent. We have little guidence from the historical record as to what to expect given an opening sequence of *EG* measures like the ones observed in Wisconsin. But the little evidence we do have suggests that a stream of similarly sized, negative values of the efficiency gap are quite likely over the balance of the districting plan.

10.3 An actionable EG threshold?

We now consider a more general question: what is an actionable threshold for the efficiency gap?

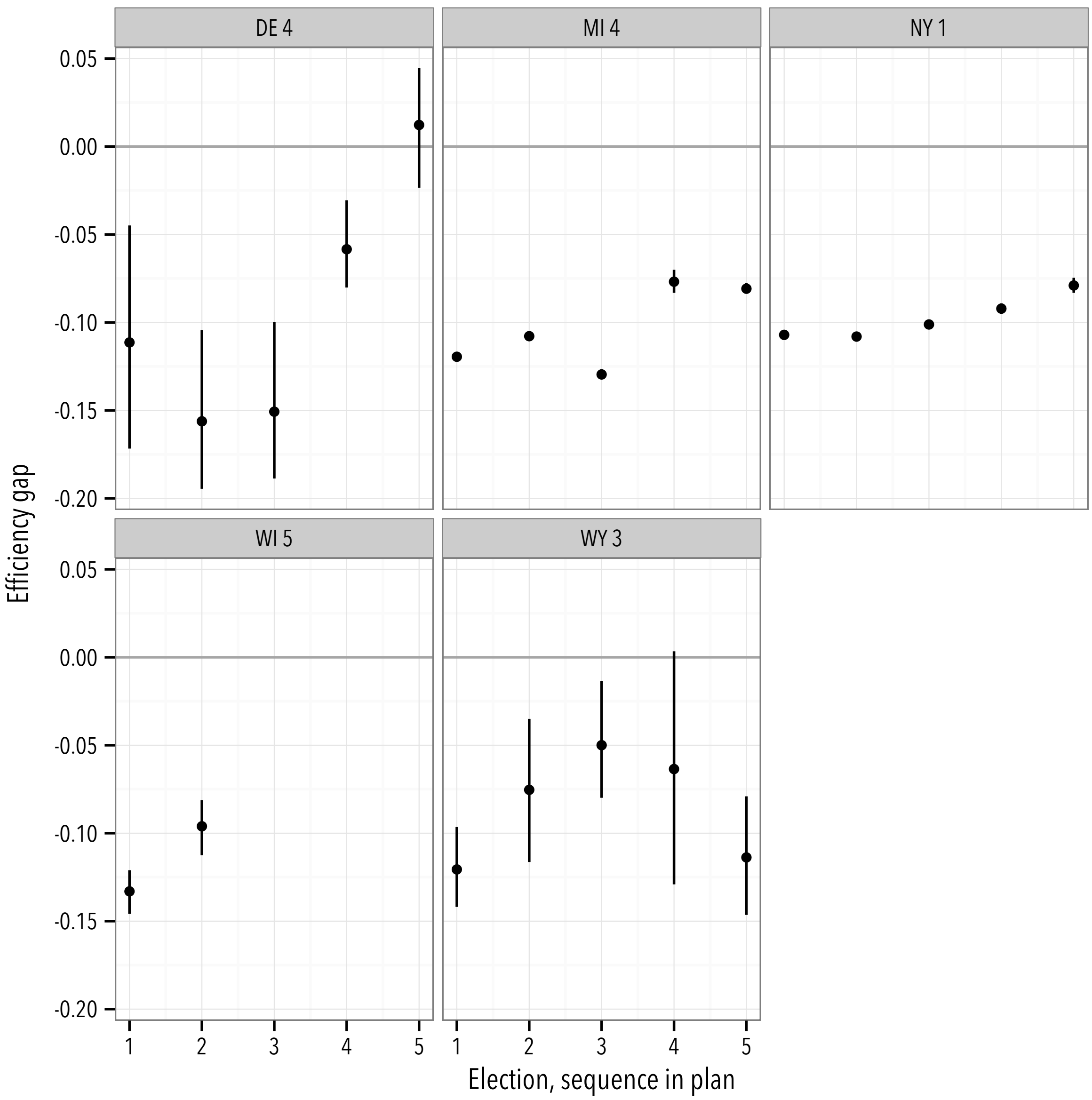


Figure 31: Sequence of *EG* estimates observed over the life of districting plans, for pre-2010 plans with first two *EG* scores within 75% of the magnitude of the *EG* scores observed in Wisconsin in 2012 and 2014.

First, recall that relatively small *EG* estimates are likely to be swamped by their estimation uncertainty, depending on the proportion of uncontested districts in the given election and the statistical procedures. In every instance though, this is an empirical question; at least in the approach I present here, each *EG* estimate I generate is accompanied with uncertainty bounds, letting us assess the *probability* that a given estimate is positive or negative. Figure 19 provides a summary of the relationship between the size of the *EG* estimate and the "statistical significance" of the estimate (in the sense that the 95% credible interval for each estimate does not overlap zero).

Second, the distribution of *EG* statistics in the 1972-2014 period is roughly symmetric around zero. Reference to this empirical distribution might also be helpful in setting actionable thresholds, and answering the question "is the *EG* measure at issues large relative to those observed in the previous 40 years of state legislative elections?" Double digit *EG* measures (-.10 or below; .10 or above) are pushing out into the extremes of the observed distribution of *EG* estimates: *EG* estimates of this magnitude are comfortably past the question of "statistical significance." Just 15% of the 786 *EG* measures generated in this analysis are below -.07; fewer than 12% are greater than .07.

We do need to be careful when making these kinds of *relative* assessments about the magnitude of the efficiency gap. If pro-Republican gerrymandering is widespread, then it will be less unusual to see a large, negative *EG* estimate, at least contemporaneously; in fact this appears to the case in the post-2010 set of elections, where the longer-term distinctiveness of the Wisconsin numbers is matched and in some cases exceeded by other states also recording unusually large, negative *EG* estimates (e.g., Florida, Michigan, Virginia and North Carolina). This speaks to the utility of the longer-term, historical analysis in both Stephanopolous and McGhee (2015) and in this report. It it is important to remember that $EG = 0$ corresponds to a partisan symmetry in wasted vote rates; we should be wary of arguments that would lead us to tolerate small to moderate levels of the efficiency gap because they appear to be the norm in some period of time, or in some set of jurisdictions.

In any litigation, much will turn on the question of *durability* in the efficiency gap, and this concern motivates much of the preceeding analysis. We cannot wait until three, four, or more elections have transpired under a plan in order to assess its properties. Courts will be asked to assess a plan based on only one *EG* estimate, or two. Analysis of the sort I provide here will be informative in these cases, assessing whether the estimate is so large that the historical record suggests that the first election's *EG* estimate is a reliable indicator as an enduring feature of the plan, and not an election-specific aberration.

10.4 Confidence in a given threshold

Figures 32 and 33 present my estimate of a "confidence rate" associated with a range of possible "actionabale thresholds" for the efficiency gap. These figures essentially re-package the information shown in Figures 29 and 30. Suppose a court rejects or amends every plan with a first election *EG* more extreme (further away from zero) than the proposed threshold shown on the horizontal axis of these graphs. A certain number of plans fail to trip this threshold, and so are upheld by the courts if they are challenged. Of those that do trip the threshold and are rejected by a court, what is our confidence that the plan, if left undisturbed, would go on to produce a sequence of *EG* measures that lie on the same side of zero as the threshold? Combining these two proportions gives us an overall confidence measure associated with a particular threshold.

This analysis points to a benchmark of about -.06 or -.07 as the actionable threshold given a first election with $EG < 0$ (Democratic disadvantage) or .08 or .09 when we observe $EG > 0$ in the first election under a redistricting plan (Democratic advantage); the asymmetry here reflects the fact that districting plans evincing apparent Democratic advantages are not as durable or as common (in recent decades) as plans presenting evidence of pro-Republican gerrymanders. At these proposed benchmarks the overall confidence rates are estimated to be 95%, with this confidence rate corresponding to a benchmark used widely in statistical decision-making in many fields of science.

Figures 32 and 33 also highlight that $EG < -.07$ or $EG > .07$ would be an extremely conservative threshold. On the pro-Democratic side, $EG > .07$ is a rare event. Districting plans unfavorable to Democrats, with $EG < -.07$ are not unusual; about 10% of post-1990 plans generate *EG* measures below -.07; the proportion of these plans that then record a sign flip is only about 10%; see Figure 30. If the presumption was that any plan with a first election showing

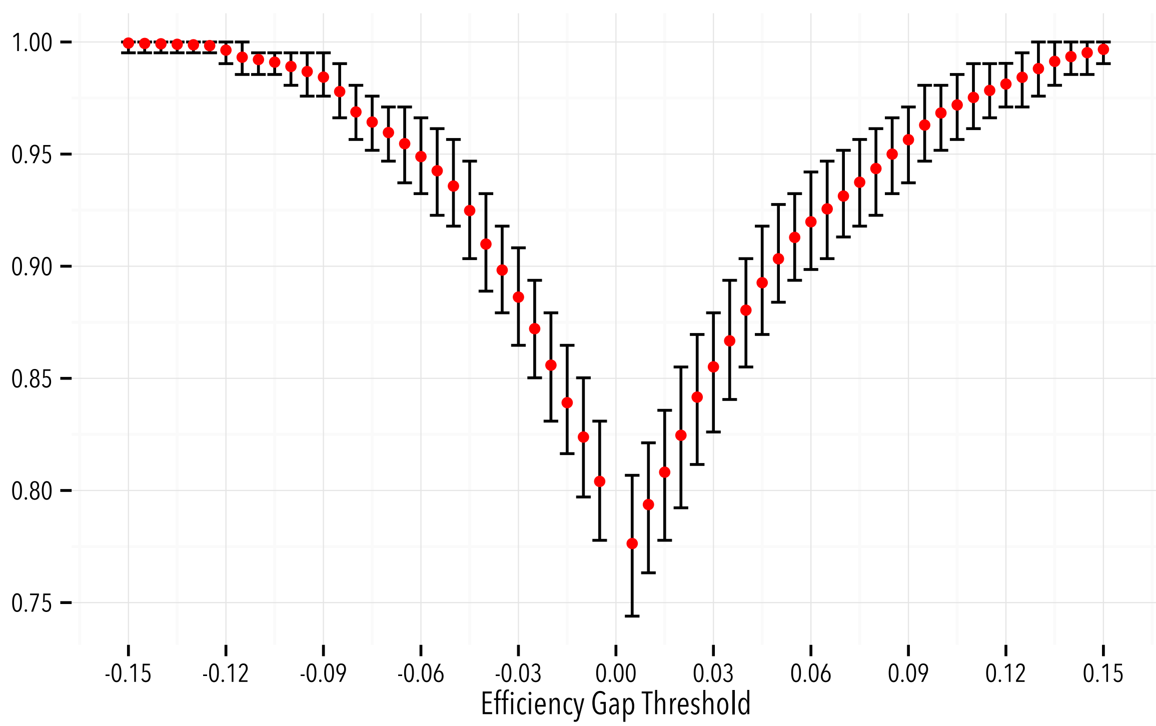


Figure 32: Proportion of plans being either (a) undisturbed or (b) if left undisturbed, would continue to produce one-sided partisan advantage (no sign change in subsequent *EG* measures), as a function of the proposed "first election," efficiency gap threshold (horizontal axis), based on analysis of all multi-election districting plans, 1972-2014. The proportion on the vertical axis is thus interpretable as the "confidence level" associated with intervention at a given first election, *EG* threshold. Vertical lines indicate 95% credible intervals.

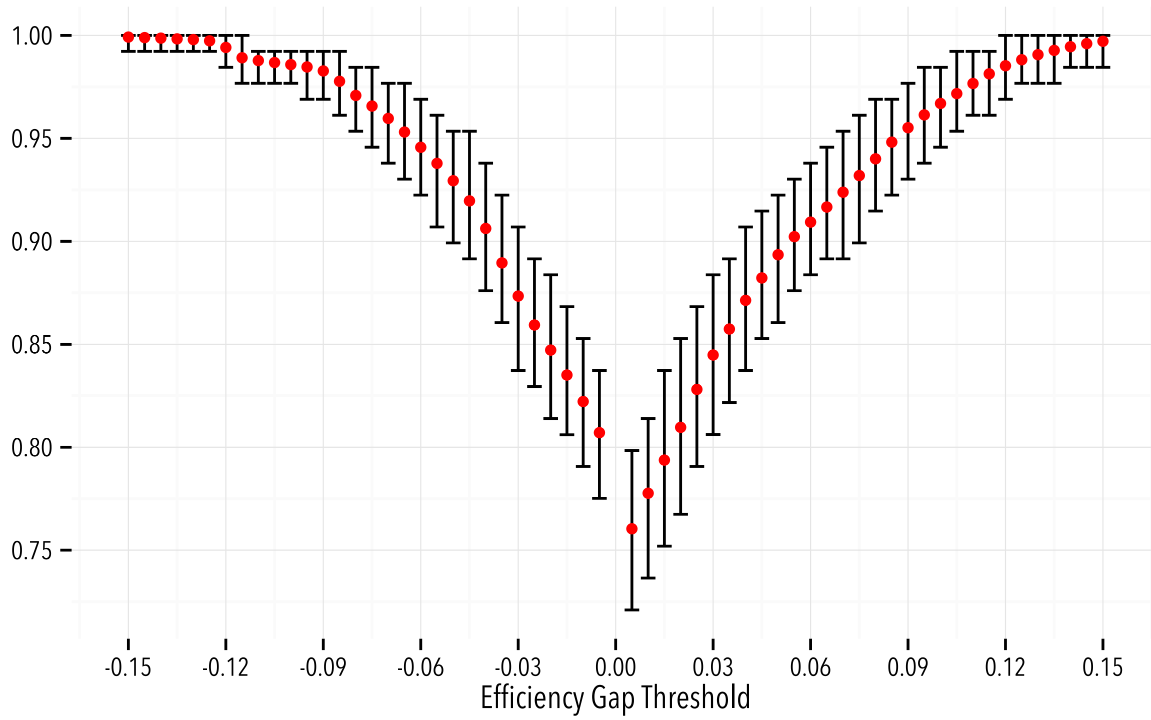


Figure 33: Proportion of plans being either (a) undisturbed or (b) if left undisturbed, would continue to produce one-sided partisan advantage (no sign change in subsequent *EG* measures), as a function of the efficiency gap threshold (horizontal axis), based on analysis of post-1990 plans and elections. The proportion on the vertical axis is thus interpretable as the "confidence level" associated with intervention at a given first election, *EG* threshold. Vertical lines indicate 95% credible intervals.

$EG < -.07$ would be rejected, then we'd be "wrong" to do so in about 10% of those cases (in the sense that if left in place, the plan would go on to produce at least one election with $EG > 0$). The total error rate in this case would be *1% of all plans*. Equivalently, 99% of all plans would be either left undisturbed or appropriately struck down or amended by a court, given the historical relationship between "first election" *EG* measures and the sequence of *EG* measures that follow.

11 Conclusion: the Wisconsin plan

Wisconsin has had two elections for its legislature under the plan currently in place, in 2012 and 2014. Both elections were subject to considerable rates of uncontestedness (27 of 99 seats in 2012 and 52 of 99 seats in 2014), but these rates are hardly unusual; Wisconsin's rates of uncontested districts in these two elections are low to moderate compared to other states. We use the relationship between state legislative election results and presidential election results in state legislative districts (and incumbency) to impute two-party vote shares in uncontested seats (see section 7.2). With a complete set of vote shares, we then compute average district-level Democratic two-party vote share (V) and note the share of seats (contested and uncontested) won by Democratic candidates (S).

In Wisconsin in 2012, and after imputations for uncontested seats, V is estimated to be 51.4% ($\pm$0.6); recall that Obama won 53.5% of the two-party presidential vote in Wisconsin in 2012. Yet Democrats won only 39 seats in the 99 seat legislature ($S = 39.4\%$), making Wisconsin one of 7 states in 2012 where we estimate $V > 50\%$ but $S < 50\%$ and where Democrats failed to win a majority of legislative seats despite $V > 50$ (the other states are Florida, Iowa, Michigan, North Carolina and Pennsylvania). In 2014, V is estimated to be 48.0% ($\pm$0.8) and Democrats won 36 of 99 seats ($S = 36.4\%$).

This provides the raw ingredients for computing the efficiency gap (EG) for these two elections (recalling equation 1). Repeating these calculations across a large set of state elections provides a basis for assessing whether the efficiency gap estimates for Wisconsin in 2012 and 2014 are noteworthy.

Wisconsin's efficiency gap measures in 2012 and 2014 are -.13 and -.10 (to two digits of precision). These negative estimates indicate the disparity between

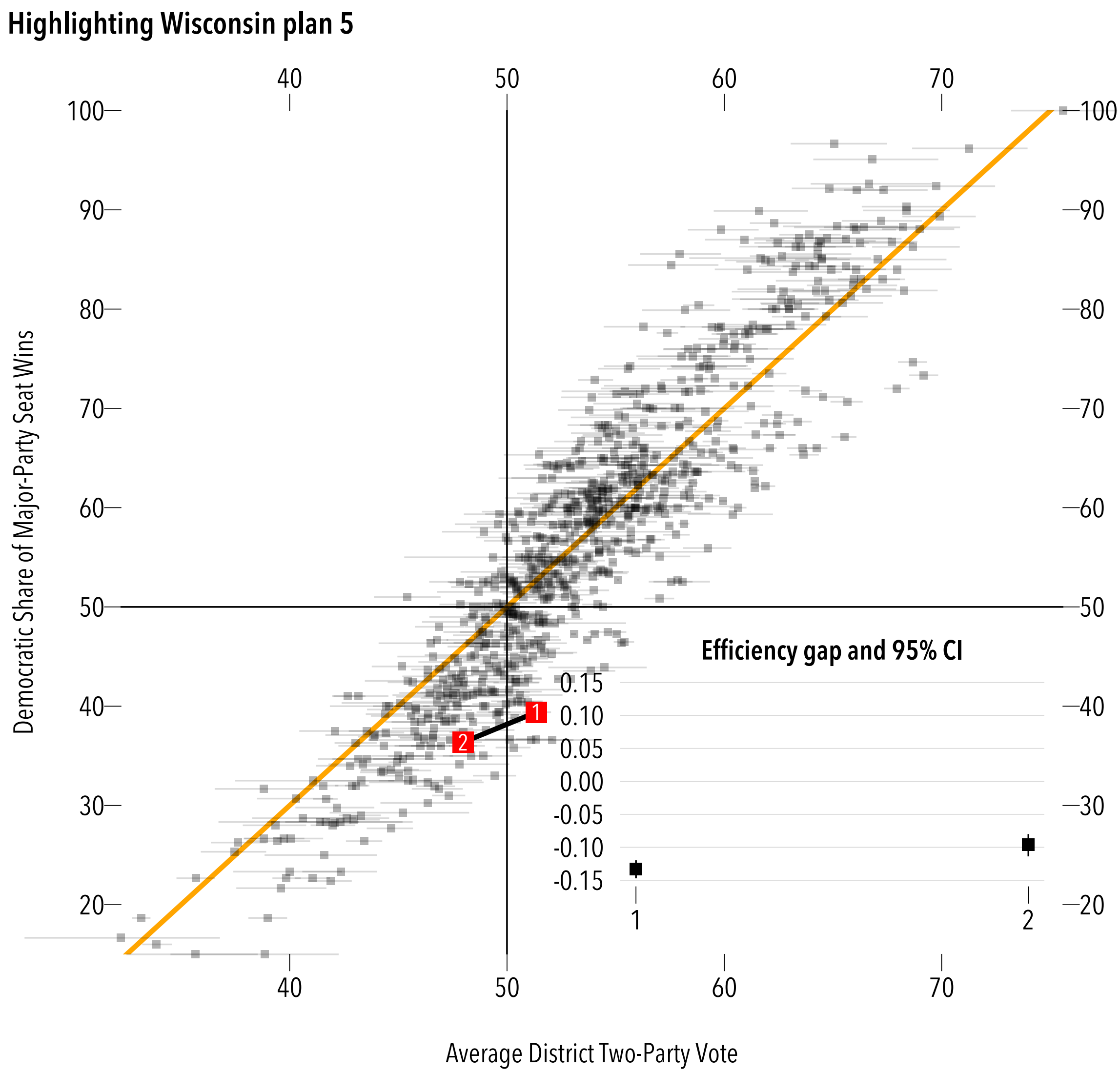


Figure 34: Seats, votes and the efficiency gap recorded under the Wisconsin plan, 2012 and 2014. Orange line shows the seats-votes curve if the efficiency gap were zero; the efficiency gap in any election is the vertical distance from the corresponding data point to the orange line. Gray points indicate elections from other states and elections (1972-2014). Horizontal lines cover a 95% credible interval for Democratic average district two-party vote share, given imputations in uncontested districts. The inset in the lower right shows the sequence of efficiency gap measures recorded under the plan; vertical lines are 95% credible intervals.

vote shares and seat shares in these elections, which in turn, is consistent with partisan gerrymandering. The negative *EG* estimates generated in 2012 and 2014 are unusual relative to Wisconsin's political history (see Figure 35). The 2012 estimate is the largest *EG* estimate in Wisconsin over the 42 year period spanned by this analysis (1972-2014); the 2014 estimate is the fourth largest (behind 2012, 2006 and 2004, although it is essentially indistinguishable from the 2004 estimate). The jump from the *EG* values being recorded towards the end of the previous districting plan in Wisconsin (2002-2010) to the 2012 and 2014 values strongly suggests that the districting plan adopted in 2011 is a driver of the change, systematically degrading the efficiency with which Democratic votes translate into Democratic seats in the Wisconsin state legislature.

Wisconsin's 2012 and 2014 *EG* estimates are also large relative to the *EG* scores being generated contemporaneously in other state legislative elections. Figure 36 shows *EG* estimates recorded under plans in place since the post-2010 census round of redistricting; the *EG* estimates are grouped by state and ordered, with Wisconsin highlighted. We have 78 *EG* scores from elections held since the last round of redistricting. Among these 79 scores, Wisconsin's *EG* scores rank eigth (2012, 95% CI 3 to 12) and seventeenth (2014, 95% CI 13 to 20).

The historical analysis reported above supports the proposition that Wisconsin's *EG* scores are likely to endure over the course of the plan. Few states ever record *EG* scores as large as those observed in Wisconsin; indeed, there is virtually no precedent for the lop-sided, two election sequence of *EG* scores generated in Wisconsin in 2012 and 2014 in the data I analyze here (1972-2014). The closest historical analogs suggest that a districting plan that generates an opening, two-election sequence of *EG* scores like those from Wisconsin will continue to do so, generating seat shares for Democrats that are well below those we would expect from a neutral plan.

The Current Wisconsin Plan is generating estimates of the efficiency gap far in excess of the proposed, actionable threshold (see section 10). In 2012 elections to the Wisconsin state legislature, the efficiency gap is estimated to be -.13; in 2014, the efficiency gap is estimated to be -.10. Both measures are separately well beyond the conservative .07 threshold suggested by the analysis of efficiency gap measures observed from 1972 to the present.

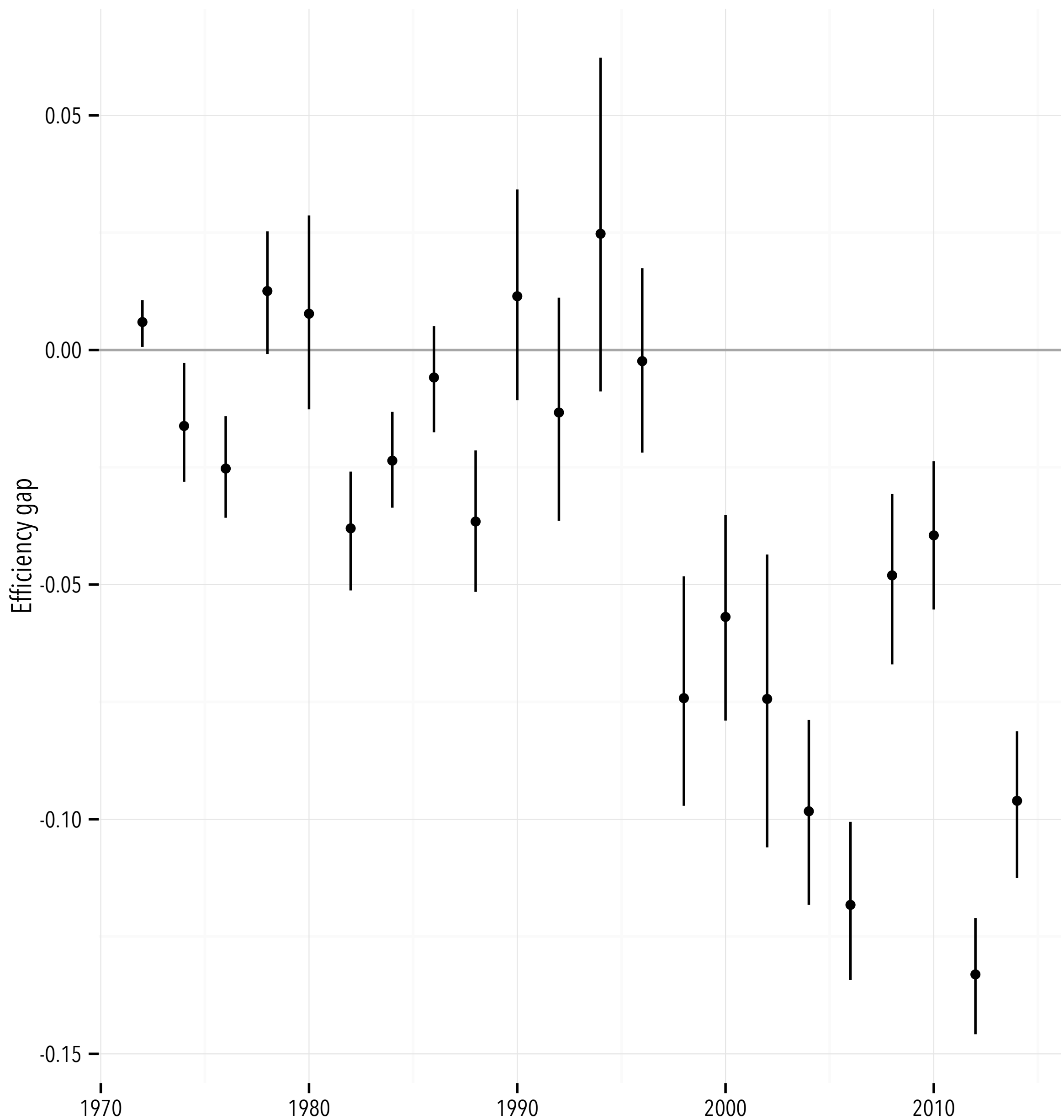


Figure 35: History of efficiency gap estimates in Wisconsin, 1972-2014. Vertical lines indicate 95% credible intervals.

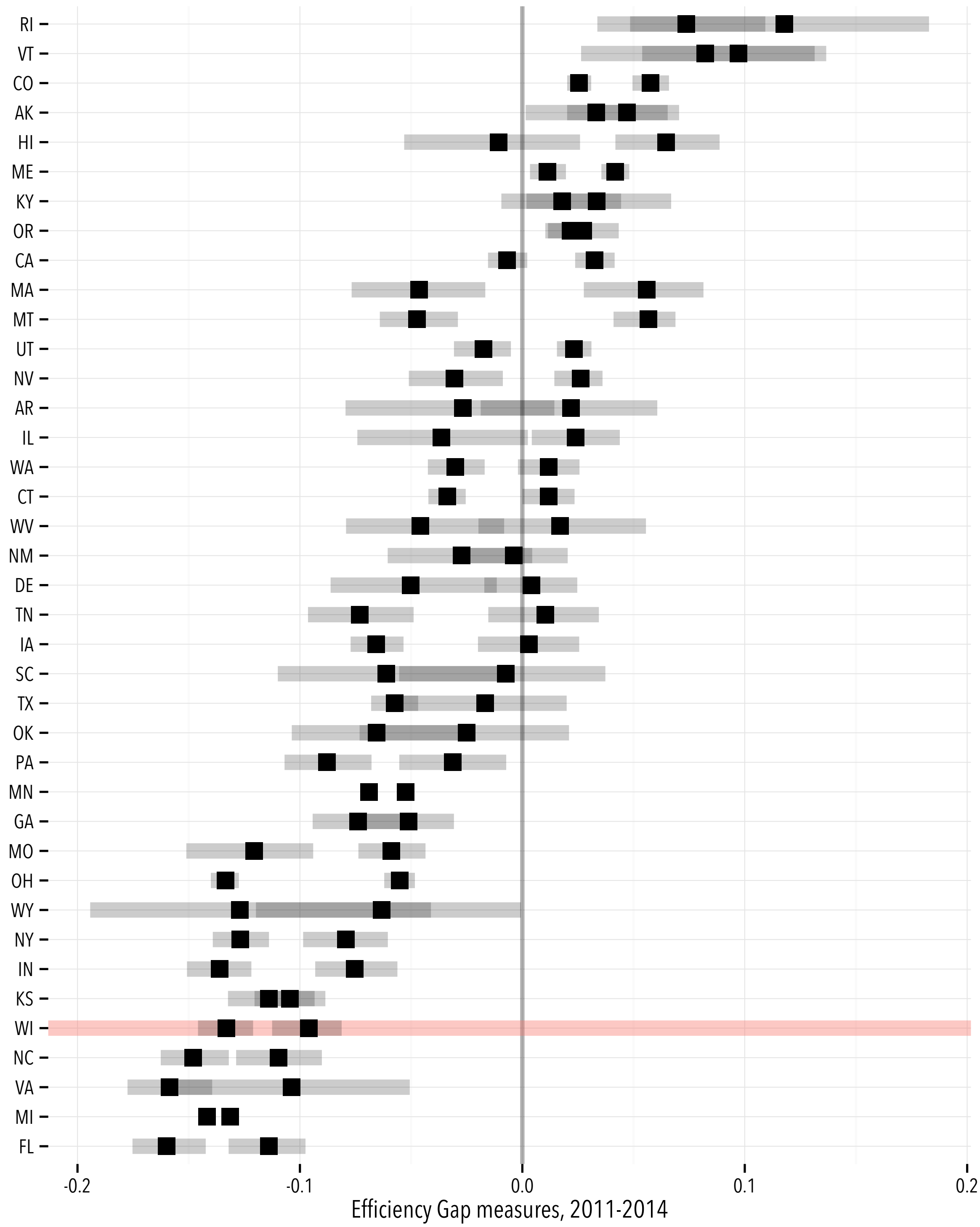


Figure 36: *EG* estimates in 2012 and 2014, grouped by state and ordered. Horizontal bars indicate 95% credible intervals.

References

Butler, David and Donald Stokes. 1974. *Political change in Britain: the evolution of electoral choice*. London: Macmillian.

Edgeworth, FY. 1898. "Applications of the Theory of Probability." *JR Statist. Soc* 51:534.

Gelman, Andrew and Gary King. 1990. "Estimating the Consequences of Electoral Redistricting." *Journal of the American Statistical Association* 85:274–82.

Kendall, M. G. and A. Stuart. 1950. "The Law of Cubic Proportion in Election Results." *The British Journal of Sociology* 1(3):183–196.

McGhee, Eric. 2014. "Measuring Partisan Bias in Single-Member District Electoral Systems." *Legislative Studies Quarterly* 39(1):55–85.

Niemi, Richard G. and Patrick Fett. 1986. "The Swing Ratio: An Explanation and an Assessment." *Legislative Studies Quarterly* 11:75–90.

Niemi, Richard G. and Simon Jackman. 1991. "Bias and Responsiveness in State Legislative Districting." *Legislative Studies Quarterly* 16:183–202.

Stephanopolous, Nicholas O. and Eric M. McGhee. 2015. "Partisan Gerrymandering and the Efficiency Gap." *University of Chicago Law Review* 82. Forthcoming.

Tufte, Edward R. 1973. "The Relationship Between Seats and Votes in Two-Party Systems." *American Political Science Review* 67:540–554.