Rebuttal Report

Simon Jackman

April 17, 2017

1 Introduction

In this rebuttal report, I respond to criticisms made by Professor James G. Gimpel, Professor M.V. Hood III, and Sean P. Trende in their respective expert reports. In brief, defendants' experts appear to agree that North Carolina's current congressional districting plan is an exceptionally severe and durable partisan gerrymander. They present no compelling evidence that any neutral factor justifies the partisan asymmetry inherent in the plan. They also raise objections to the efficiency gap that are invalid and improperly conflate the efficiency gap with plaintiffs' broader three-part test for partisan gerrymandering.

Before turning to the specific points raised by defendants' experts, I address them more generally by presenting data on two other measures of partisan asymmetry: partisan bias and the mean-median difference. These metrics are widely used by social scientists, and they confirm what is clear from my efficiency gap analysis: that North Carolina's current congressional districting plan is indeed an outlier compared to the distribution of congressional plans from 1972 to the present.

2 Other Measures of Partisan Asymmetry

In my original report, I discussed two measures of partisan asymmetry: the efficiency gap and partisan bias (pp. 14-18, 61-62). The efficiency gap, again, is the difference between the parties' respective wasted votes in an election, divided by the total number of votes cast in the election (where a vote is wasted if it is cast for a losing candidate or for a winning candidate but in excess of the 50%-plus-one threshold needed for victory). Partisan bias, in contrast, is the difference between the shares of *seats* that the major parties would win if they each received the same share (typically 50%) of the statewide *vote*. (*LULAC v. Perry* (2006), pp. 420, 466; Grofman & King 2007, pp. 6-13.) For example, if Democrats would win 55% of a plan's districts if they received 50% of the statewide vote (leaving 45% of the districts to be won by Republicans), then the plan would have a pro-Democratic bias of 5%.

All three of defendants' experts criticize the efficiency gap and Trende also criticizes partisan bias (paras. 52, 148). I therefore consider it necessary to show that my conclusions about North Carolina's current congressional plan, which were based primarily on the efficiency gap, are robust to the use of other measures of partisan asymmetry. In addition to partisan bias, I address the mean-median difference, that is, the difference between a party's *mean* vote share and *median* vote share across all of the districts in a plan. All three of these metrics point in exactly the same direction in this case.

Beginning with partisan bias, its calculation is relatively straightforward. An analyst first obtains district-by-district electoral results as well as the statewide vote share for each party. Next, the analyst *shifts* the observed vote share in each district by the same amount (a "uniform swing"): the amount necessary to simulate a tied statewide election. The analyst then tallies how many districts each party would have won and lost in this hypothetical election. The difference between the parties' seat shares and an even split of the seats in the hypothetical election is an estimate of the partisan bias of the underlying districting plan. For instance, if Republicans won 47% of the statewide vote, then the observed vote share in each district would be increased by 3% to simulate a tied election. Partisan bias would be determined by comparing the parties' seat shares after this uniform swing was carried out. (Grofman & King 2007; Gelman & King 1994; King & Browning 1987.)

When a statewide election is competitive (say, closer than 55% to 45%), partisan bias and the efficiency gap tend to exhibit similar values. This is because, under these conditions, the uniform swing that must be conducted to compute partisan bias is relatively small, so there is not much opportunity for the measure to diverge from the efficiency gap (Stephanopoulos & McGhee 2015, p. 856; McGhee 2014, p. 69). Since North Carolina has generally had competitive congressional elections over the last half-century, we would expect its partisan bias and efficiency gap trends to be comparable. As Figure 1 illustrates, this is indeed the case. From 1972 to 2016, the metrics mostly rise and fall in tandem—somewhat less so in the 1970s and 1980s, when Democrats often won more than 55% of the statewide vote, and somewhat more so in recent years, when elections have been closer to parity. Notably, both partisan bias and the efficiency gap swing precipitously in a pro-Republican direction between 2010 and 2012, the first election held under North Carolina's 2011 congressional plan. Both partisan bias and the efficiency gap also exhibit consistent (and enormous) pro-Republican values in 2012, 2014, and 2016.

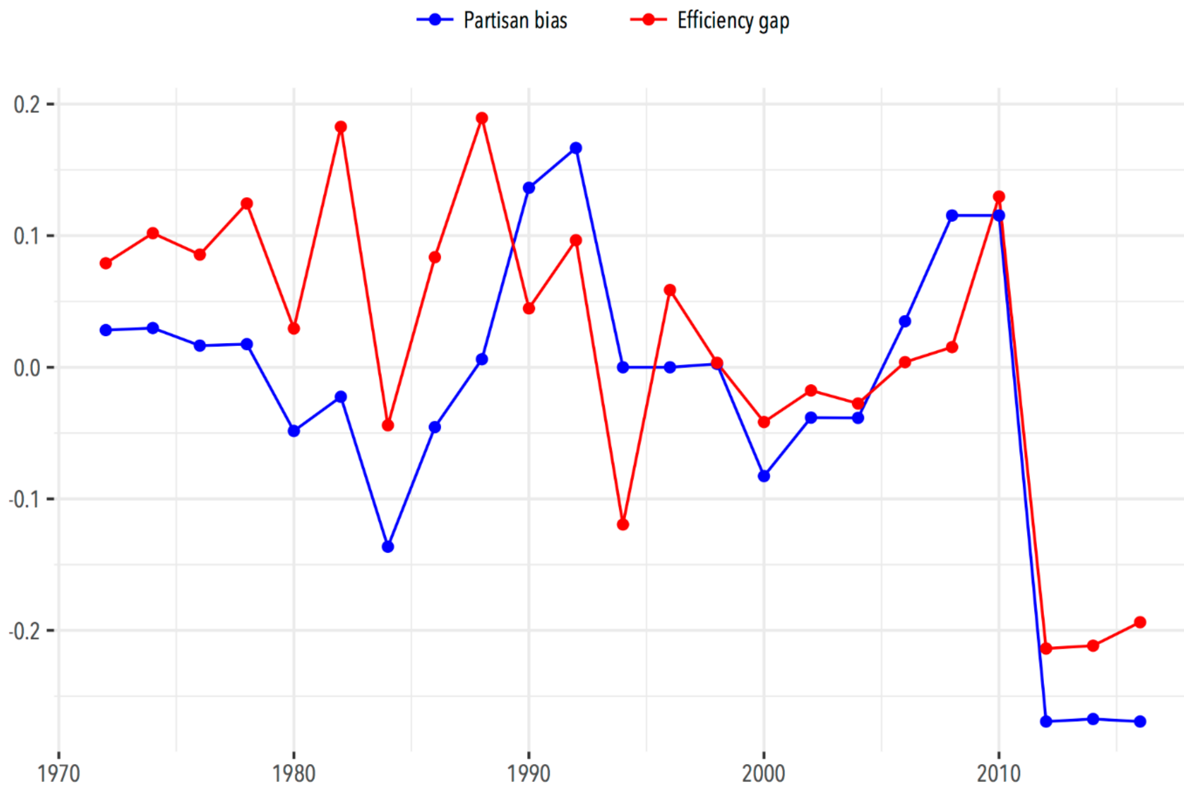

Figure 1: Efficiency gap and partisan bias for North Carolina Congressional elections, 1972-2016.

To further highlight the partisan biases of North Carolina's congressional plans over the last three elections, Figure 2 is a histogram showing the partisan biases for all 283 elections in my database that were closer statewide than 55% to 45%. It is clear that North Carolina's 2011 plan and 2016 plan are true outliers. Indeed, their partisan biases of about -27% (in all three elections) are the *second-largest* on record, roughly *three standard deviations* from the historical mean. This is powerful corroborative evidence indicating that there is nothing idiosyncratic about the conclusions I reached based on the efficiency gap. Partisan bias tells exactly the same story.

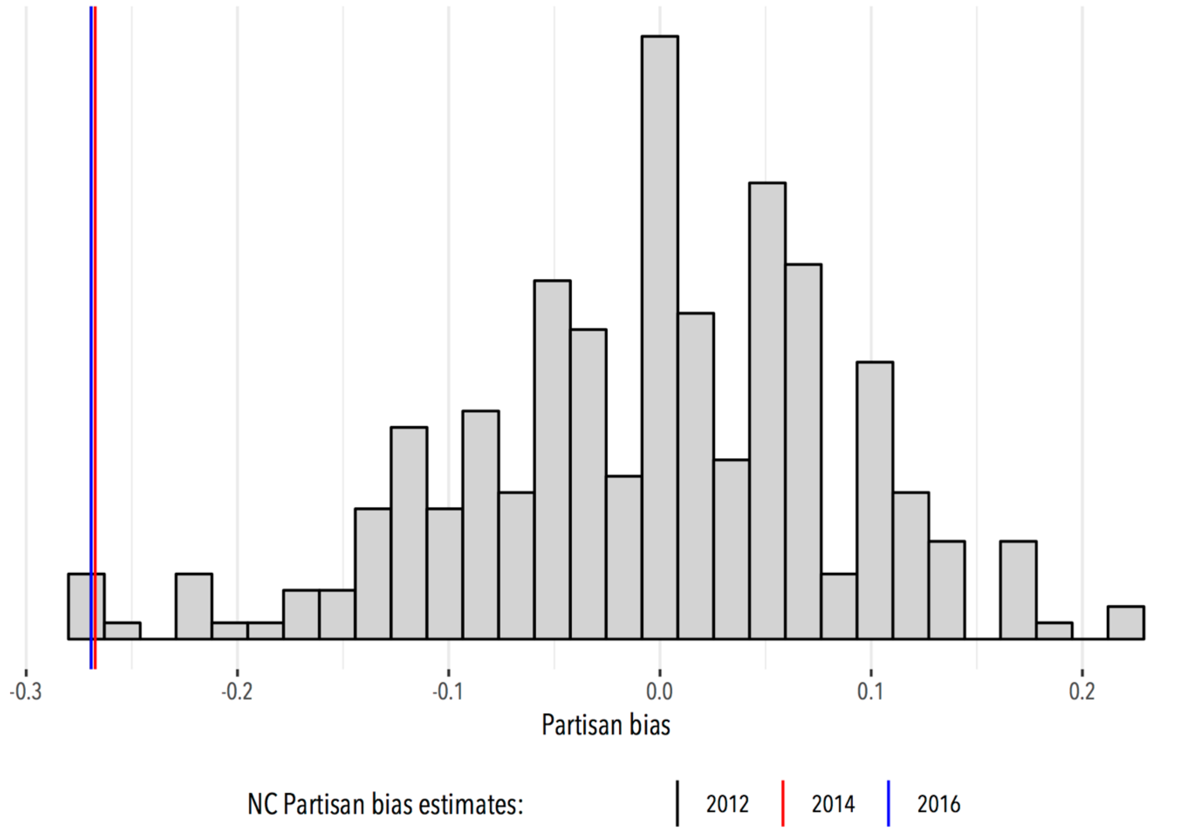

Figure 2: Histogram of partisan bias in 282 Congressional elections closer statewide than 55% to 45%, 1972-2016. The three vertical lines indicate North Carolina's scores in 2012, 2014, and 2016.

My discussion to this point has only considered competitive elections closer statewide than 55% to 45%. In *uncompetitive* settings, however, partisan bias becomes less reliable and, in my opinion, should not be used. If an actual election is uncompetitive, then the amount of uniform swing required to construct the counterfactual of a tied election is large, politically implausible and unrealistic, if not whimsical. Consider trying to predict what would happen if Massachusetts or Utah suddenly became tossup states. For precisely this reason, even advocates of partisan bias recommend applying the measure only to competitive statewide elections (Grofman & King 2007, p. 19; Gelman & King 1994, p. 545).

The two charts below illustrate the unreliability of partisan bias in uncompetitive settings. Figure 3 plots the difference between the efficiency gap and partisan bias versus the Democratic share of the statewide vote in congressional elections from 1972 to 2016. The data points resemble a bowtie, tightly bunched when elections are competitive but fanning out in all directions when they are uncompetitive (see also Stephanopoulos & McGhee, p. 858).

Figure 4 indicates how the efficiency gap and partisan bias are related in competitive (closer than 55% to 45%) and uncompetitive (further apart than 55% to 45%) congressional elections from 1972 to 2016. In competitive elections, the measures are highly correlated (r = 0.77) and cluster closely around the best fit line. But in uncompetitive elections, the metrics are only modestly correlated (r = 0.29) and diverge much more from the best fit line. I therefore recommend that partisan bias be used as a robustness check only when statewide elections are relatively close.

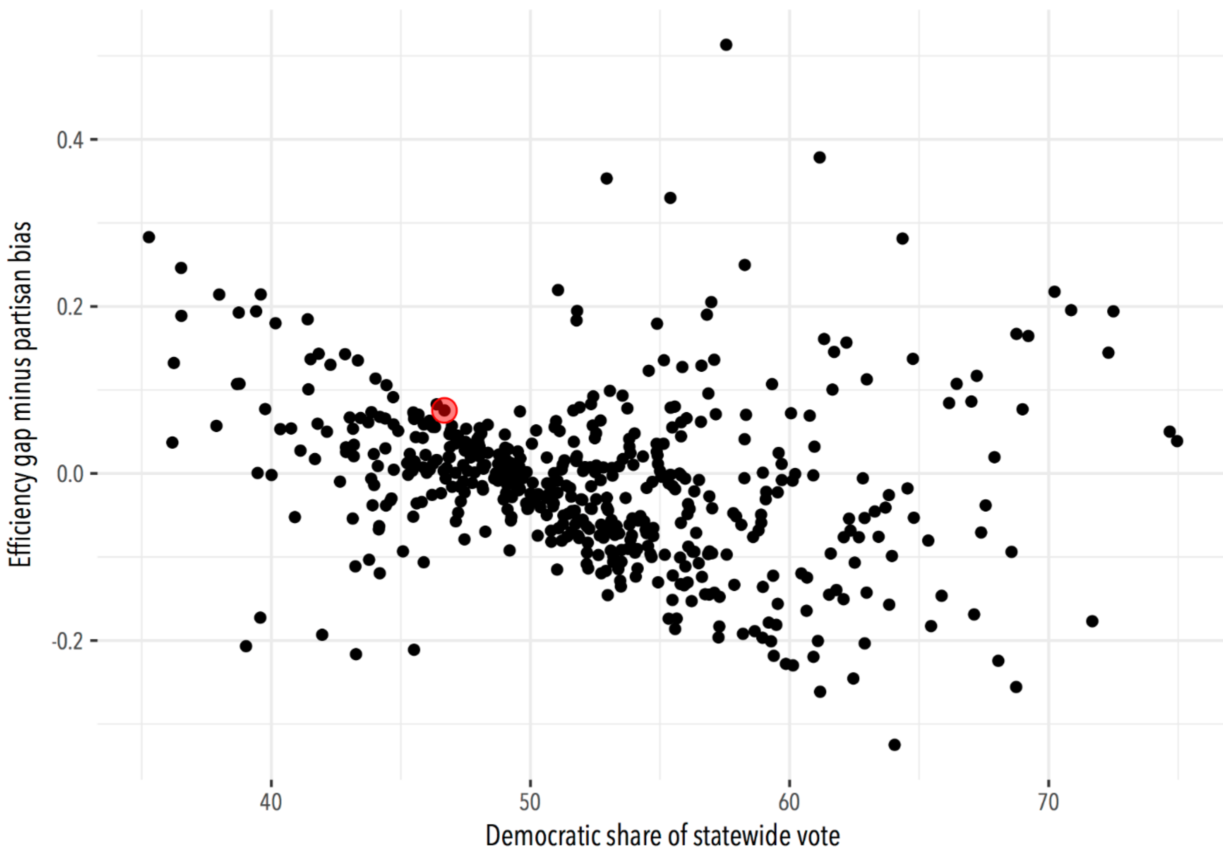

Figure 3: Efficiency gap minus partisan bias versus the Democratic share of the statewide vote, Congressional elections, 1972-2016. North Carolina's 2016 plan is highlighted in red.

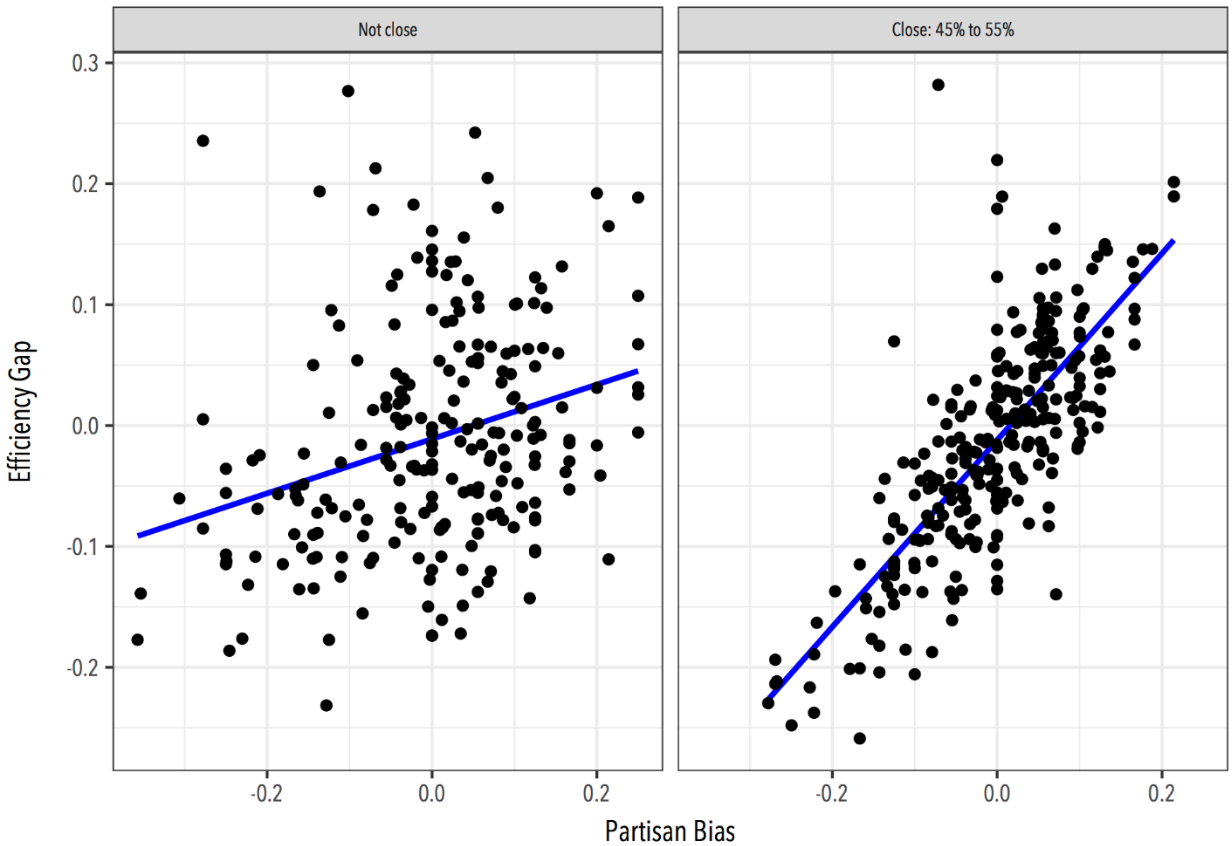

Figure 4: Efficiency gap versus partisan bias, Congressional elections, 1972-2016, competitive elections (closer than 55% to 45%) and uncompetitive elections.

While partisan bias and the efficiency gap are the most established measures of partisan asymmetry, scholars have recently advanced another metric: the mean-median difference (Wang 2016; McDonald & Best 2015). This measure is simply the difference between a party's mean vote share and median vote share across all of the districts in a given jurisdiction. The intuition is that when the mean and the median diverge significantly, the distribution of district-level vote shares is skewed in favor of one party and against its opponent—consistent with the classic gerrymandering techniques of "packing" partisans into a relatively small number of districts and/or "cracking" partisans among a larger number of districts. Conversely, when the mean and the median are close, the distribution of district-level vote shares is more symmetric.

But unlike partisan bias and the efficiency gap, the mean-median difference is denominated in units of *vote share* rather than *seat share*. While measuring the skew of the district-level vote shares, the metric ignores a critical feature of this distribution: how many district-level vote shares lie above or below the 50% point, the point where a seat changes hands. That is, the mean-median measure ignores which party actually *wins* each district, as this is immaterial to the calculation of the mean and the median. The mean-median difference also has an arithmetical relationship with partisan bias. It is partisan bias divided by the slope of a plan's seat-vote curve at the point of a tied election. As this slope is usually close to two, the magnitude of the mean-median difference is usually about half that of partisan bias (McDonald & Best, p. 315).

Because they are so closely connected as a matter of arithmetic, the mean-median difference and partisan bias are highly correlated in both competitive ($r = 0.78$) and uncompetitive ($r = 0.77$) elections. Since the efficiency gap and partisan bias are closely related in competitive elections, so too is the mean-median difference and the efficiency gap. Specifically, the mean-median difference is moderately correlated with the efficiency gap in competitive elections ($r = 0.60$) but only weakly correlated with the efficiency gap in uncompetitive elections ($r = 0.19$). Both because the mean-median difference is so similar to partisan bias, and because its validity as a measure of gerrymandering is reduced by its exclusive focus on votes rather than seats, I recommend using it only as a secondary robustness check in competitive settings.

North Carolina had competitive congressional elections in 2012, 2014 and 2016. For this reason, it is worthwhile to note the mean-median differences in North Carolina in these years. In 2012, the mean Democratic vote share across North Carolina's thirteen districts was 50.9% and the median Democratic vote share was 43.2%, resulting in a pro-Republican differential of 7.7%. In 2014, the mean Democratic vote share was 46.2% and the median Democratic vote share was 39.0%, for a pro-Republican differential of 7.2%. And in 2016, the mean Democratic vote share was 46.7% and the median Democratic vote share was 41.6%, for a pro-Republican differential of 5.1%. These are very large mean-median differences—North Carolina's average mean-median difference from 1972 to 2016 was just 1.0%—further confirming the severity of the partisan asymmetry in the 2011 and 2016 North Carolina districting plans.

3 Gimpel: Efficiency Gap Calculations

I turn next to defendants' experts, beginning with Gimpel. Gimpel first appears to concede that North Carolina's current congressional plan is skewed in a pro-Republican direction. He writes that it "is obvious to visual inspection and a few minutes of data analysis" that the plan "show[s] a Republican advantage" (p. 2). He also states that it "is not really worth disputing" that "the 2016 map adopted by the North Carolina legislature is the result of a partisan gerrymander" (p. 7). I agree with these comments.

Next, Gimpel attempts to calculate the efficiency gap for North Carolina's 2001 plan, 2011 plan, and 2016 plan. However, he uses an incorrect definition for the efficiency gap: "% popular vote - % seats" (p. 13). The efficiency gap plainly is *not* equivalent to simple disproportionality, that is, the difference between a party's vote share and seat share in an election. This error has serious consequences for Gimpel's computations. Table 1 shows the correct efficiency gaps for North Carolina's 2001 plan, 2011 plan, and 2016 plan, all using the same 2004-2010 data that Gimpel employed (p. 10).

District	2001 Dem	2001 Rep	2001 Wasted Dem	2001 Wasted Rep	2011 Dem	2011 Rep	2011 Wasted Dem	2011 Wasted Rep	2016 Dem	2016 Rep	2016 Wasted Dem	2016 Wasted Rep
1	0.654	0.346	0.154	0.346	0.717	0.283	0.217	0.283	0.680	0.320	0.180	0.320
2	0.531	0.469	0.031	0.469	0.445	0.555	0.445	0.055	0.448	0.552	0.448	0.052
3	0.420	0.580	0.420	0.080	0.470	0.530	0.470	0.030	0.466	0.534	0.466	0.034
4	0.586	0.414	0.086	0.414	0.691	0.309	0.191	0.309	0.618	0.382	0.118	0.382
5	0.402	0.598	0.402	0.098	0.428	0.572	0.428	0.072	0.453	0.547	0.453	0.047
6	0.375	0.625	0.375	0.125	0.447	0.553	0.447	0.053	0.470	0.530	0.470	0.030
7	0.518	0.482	0.018	0.482	0.459	0.541	0.459	0.041	0.484	0.516	0.484	0.016
8	0.535	0.465	0.035	0.465	0.453	0.547	0.453	0.047	0.463	0.537	0.463	0.037
9	0.419	0.581	0.419	0.081	0.418	0.582	0.418	0.082	0.463	0.537	0.463	0.037
10	0.390	0.610	0.390	0.110	0.443	0.557	0.443	0.057	0.440	0.560	0.440	0.060
11	0.491	0.509	0.491	0.009	0.439	0.561	0.439	0.061	0.450	0.550	0.450	0.050
12	0.681	0.319	0.181	0.319	0.757	0.243	0.257	0.243	0.640	0.360	0.140	0.360
13	0.577	0.423	0.077	0.423	0.445	0.555	0.445	0.055	0.481	0.519	0.481	0.019
Total			3.079	3.421			5.112	1.388			5.056	1.444
Efficiency Gap			2.6%				-28.6%				-27.8%	

Table 1: Correct efficiency gap calculations for North Carolina's 2001 plan, 2011 plan, and 2016 plan, using Gimpel's 2004-2010 data.

According to the correct calculations, the 2001 plan actually has a small pro-Democratic efficiency gap of 2.6%, not 14.8% (p. 11). The 2011 plan actually has an enormous pro-Republican efficiency gap of 28.6%, not 16.3% (p. 12). And the 2016 plan actually has an enormous pro-Republican efficiency gap of 27.8%, not 7.6% (p. 14). Thus when Gimpel's errors are remedied, his data strongly confirms my own findings:

namely, that the 2001 plan was close to perfectly symmetric while the 2011 plan and the 2016 plan are extraordinarily tilted toward Republicans.[1]

4 Gimpel: Partisan Fairness Versus Descriptive Representation

Gimpel devotes a significant portion of his report to the argument that there is a tradeoff between partisan fairness and descriptive representation for African Americans (pp. 18-21). He claims that the drafters of North Carolina's current congressional plan had to "uphold the value of descriptive representation" by leaving unchanged "large parts of Districts 1 and 12," thus causing "the remaining districts [to] exhibit a decidedly more Republican tilt" (p. 18). Gimpel purports to support this argument by showing that if the voters of Districts 1 and 12 (under the 2011 plan) were *entirely omitted* from a new plan, that map would favor Republicans (pp. 18-20).

There are several problems with this line of analysis. First, the drafters of North Carolina's current congressional plan explicitly declined to take race into account when designing the plan's districts. Their "adopted criteria" included a statement that "[d]ata identifying the race of individuals or voters *shall not be used* in the construction or consideration of districts" (Compl. Ex. A, p. 1). Gimpel seems unaware of this important element in the drafting history of the North Carolina plan.

Second, the relationship between partisan fairness and descriptive representation cannot be examined by crudely excluding most African Americans from the study. They are voters who must be taken into account by any district plan. Obviously, if all of the voters in two heavily *Republican* districts were removed from the analysis, North Carolina would appear to tilt sharply in a *pro-Democratic* direction. Neither this fact, nor Gimpel's converse finding, establishes anything of interest.

Third, to investigate how partisan fairness and descriptive representation actually are related, I created Figures 5 and 6. They are scatter plots for all congressional elections in my database from 1972 to 2016, with the efficiency gap on the y-axis and the share of U.S. House seats held by African American members and Latino members, respectively,

[1] Gimpel also appears to have erred in his compactness, county split, and population deviation calculations. Gimpel states that the 2011 plan had an average compactness of 0.08 and that the 2016 plan had an average compactness of 0.26 (pp. 12, 14). These are different figures from those reported by Hood (p. 22). Similarly, Gimpel states that the 2011 plan split 50 counties (p. 12), while Hood found that it split 40 (p. 23). And Gimpel states that the 2016 plan has a population deviation of 0.15 (p. 14), while Hood asserts that it achieves perfect population equality (p. 21).

on the *x*-axis. It is apparent from a visual inspection that there is almost no correlation between partisan fairness (i.e., the efficiency gap) and descriptive representation (i.e., the share of African American or Latino members). In both cases, the loess curve is mostly flat, indicating that a state's efficiency gap neither rises nor falls substantially as the proportion of minority House members in that state's delegation increases. This is a much more thorough analysis than any attempted by Gimpel, and it contradicts his claim about there being a tradeoff between partisan fairness and descriptive representation.

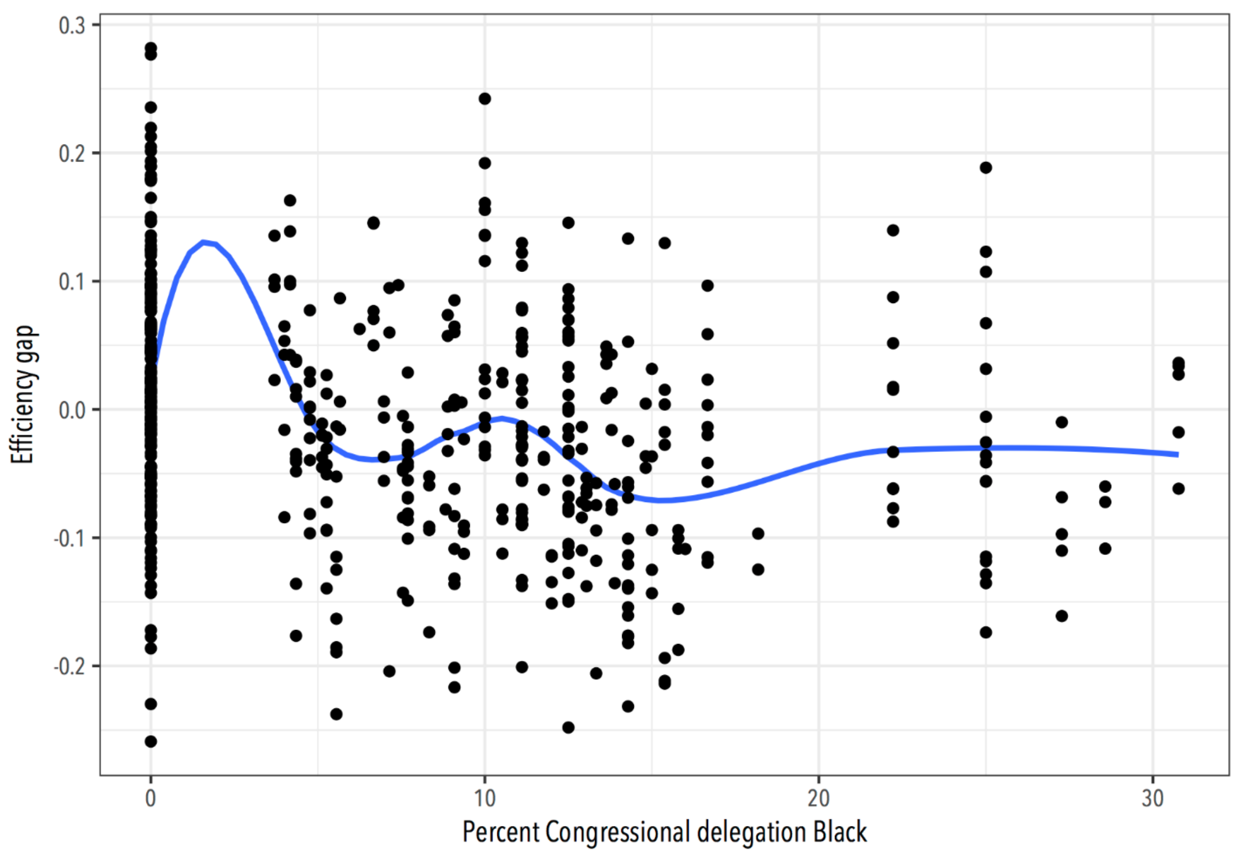

Figure 5: Efficiency gap versus proportion of Congressional seats held by African American members, Congressional elections by state and year, 1972-2016. The blue line is a loess curve summarizing the relationship between the two variables.

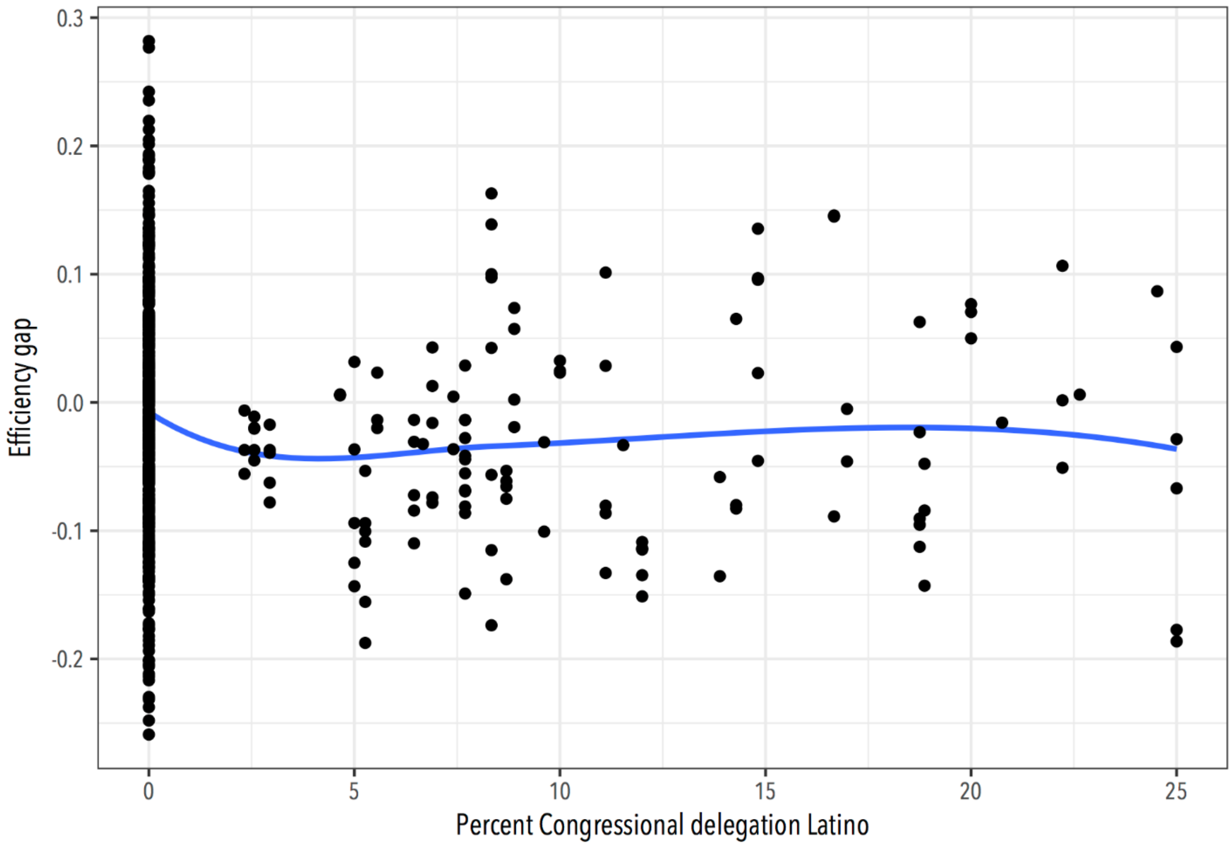

Figure 6: Efficiency gap versus proportion of Congressional seats held by Latino members, Congressional elections by state and year, 1972-2016. The blue line is a loess curve summarizing the relationship between the two variables.

5 Hood: North Carolina's Partisan Balance

Hood argues that "Republicans have a political edge in North Carolina" by highlighting the rising share of Republican seats in the state legislature (p. 3). But there is a glaring problem with assessing a state's partisan balance by looking solely at the legislative seats held by each party: those seats could be won not by appealing to voters but rather by gerrymandering the legislative maps.

Figure 7 shows that the Republican "political edge" identified by Hood is indeed the product of gerrymandering, not the will of the electorate. Like Hood's chart (p. 4), Figure 7 plots the share of Democratic state house seats from 1992 to 2014. But unlike Hood's chart, Figure 7 also plots the Democratic share of the statewide vote in state house elections over this period. Clearly, Democratic seat share and Democratic vote share moved in tandem from 1992 to 2010, gently rising for the most part but plummeting in the Republican wave election of 2010. Equally clearly, Democratic seat share and Democratic vote share diverged widely in 2012 and 2014, to a greater extent than at any previous point in modern North Carolina history. Democratic vote share hovered around 50%, while Democratic seat share fell below *40%*. The more valid conclusion to be drawn from Hood's data thus is not that North Carolina is a Republican state—but rather that North Carolina is a state currently *gerrymandered* by Republicans, at both the congressional and state legislative levels.

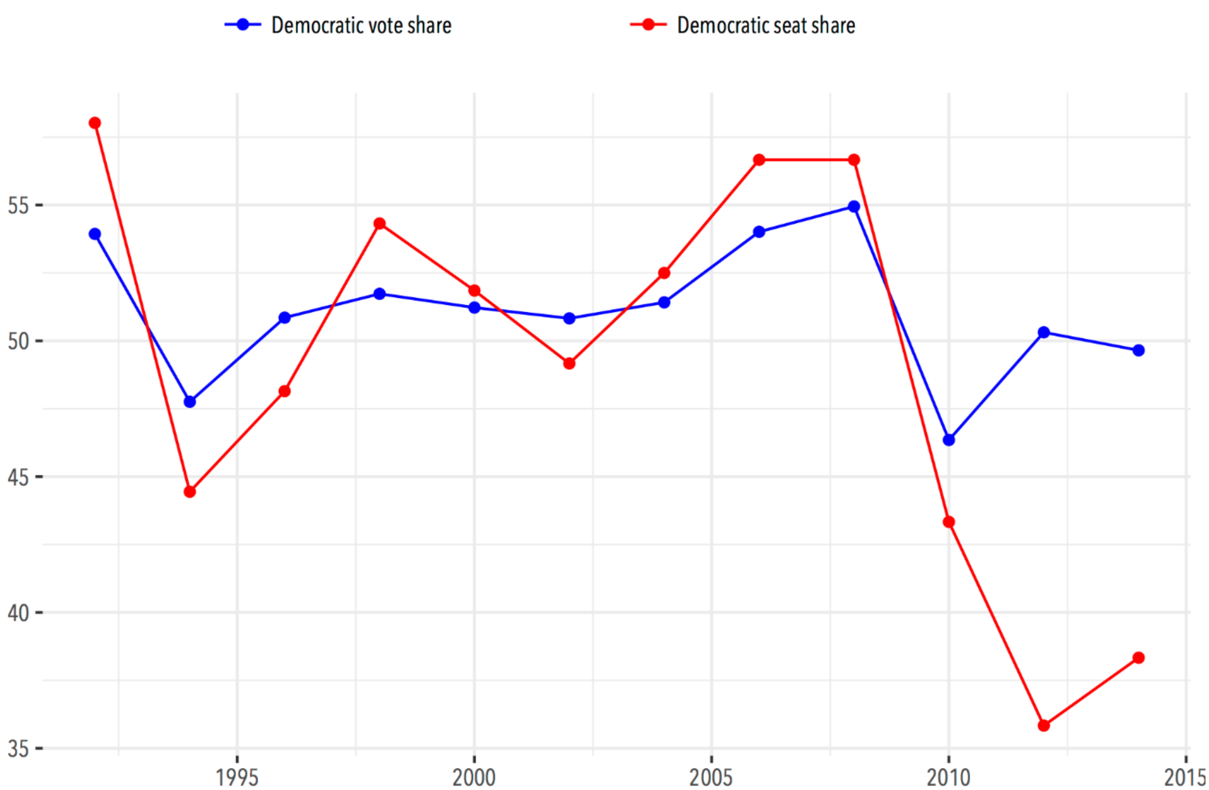

Figure 7: Democratic vote share and Democratic seat share in North Carolina General Assembly elections, 1992-2014.

6 Hood: Incumbency, Challenger Quality, and Campaign Spending

Hood suggests that Republicans' success under North Carolina's current congressional plan may be attributable to incumbency advantage, to the low quality of Democratic challengers, and to the greater campaign spending of Republican candidates (pp. 5-7). These factors obviously have *some* impact on candidates' performances—though Hood makes no effort to quantify the magnitude of any such effect. Crucially, these factors are also *endogenous* to any districting plan. In other words, who is and who isn't an incumbent, how much money a candidate is able to raise, and whether an experienced politician chooses to run, are all determined in part by the underlying competitiveness of the districts themselves. These factors thus may not be driving Republicans' success since they are themselves (at least in part) products of redistricting.

Additionally, Hood's own data shows that Republicans' edge under North Carolina's current congressional plan is *not* the result of incumbency, challenger quality, or campaign spending. He created a "partisan index" using "eleven statewide races" from 2010 to 2014 (pp. 7-8). Because these races are statewide, not district-specific, they are entirely unaffected by district-level variations in any of the factors that Hood identified. Nevertheless, using these races, North Carolina's current congressional plan looks exactly the same as when actual congressional results are considered. That is, Republicans enjoy distinct advantages in ten out of thirteen districts, even though the state as a whole leans slightly Democratic (50.2% over the eleven statewide races) (p. 25).

Notably, this is the same conclusion that Gimpel reached using a different set of statewide races from 2004 to 2010. According to Gimpel's data and analysis as well, Republicans are advantaged in ten out of thirteen districts, even though the state as a whole is marginally Democratic (50.4% over his array of statewide races) (p. 14). It is therefore clear that when incumbency, challenger quality, and campaign spending are controlled for by using statewide races rather than actual congressional results, the analysis is substantively unchanged. The 2016 plan retains its extreme pro-Republican asymmetry.

7 Hood: Efficiency Gap Analysis

Hood states in his efficiency gap analysis that the metric "increases as the number of seats won by a party increases" (p. 14). This is not necessarily true; the metric increases if a party's seat share rises at a significantly faster rate than the party's vote share. And this is more than a semantic point. The reason why seat share is an excellent predictor of the efficiency gap in North Carolina elections from 1992 to 2016 is that vote share has not varied much over this period. Indeed, the standard deviation of the two-party vote share in North Carolina's congressional elections from 1992 to 2016 is just 3.1%, indicating an impressive degree of electoral stability.

Nationally and over the entire era from 1972 to 2016, congressional elections have been substantially more volatile, with a standard deviation of 6.9% for the two-party vote share. So unsurprisingly, when I apply Hood's regression model (p. 14) to *all* elections in my database, I get a much lower *r*-squared of 0.500. In other words, while seat share may explain 83% of the variance in North Carolina's efficiency gaps from 1992 to 2016, it explains only 50% of the variance in efficiency gaps nationwide from 1972 to 2016.

Hood also engineers two hypothetical, ten-district plans that purport to resemble North Carolina in the 1970s and 1980s and North Carolina from the 1990s to the present (pp. 16-20). One thing to note about these toy examples is that Hood misstates the population shares of Group IB (it's 20%, not 25%), Group II (it's 20%, not 25%), and Group III (it's 40%, not 30%). Additionally, by showing how electoral outcomes vary as Group III's preferences shift by *forty-eight percentage points* (from 51% to 99% for Party B), these hypothetical examples assume dramatically higher levels of electoral volatility than North Carolina's electorate has actually exhibited. As noted above, the standard deviation of the two-party congressional vote share in North Carolina is just 3.1% over the last three redistricting cycles.

Furthermore—and this is a point I return to below when I discuss Trende's report—it is unremarkable that a plan's efficiency gap may vary as the electoral environment changes. Plaintiffs' three-part test is designed to take into account this possibility by invalidating a plan only when there is evidence that the plan's asymmetry is large *and durable over a range of electoral conditions*. If sensitivity testing is properly conducted (which Hood's is not), and reveals that a plan's asymmetry would disappear given plausible shifts in the electorate's preferences, then the plan would be upheld under plaintiffs' test.

It is also worth elaborating the extent to which Hood's artificial examples are patently unrealistic. The plan that supposedly represents North Carolina in the 1970s and 1980s shows the efficiency gap going from quite pro-Democratic to quite pro-Republican as Group III's affinity for Republicans increases (p. 19). But in fact, North Carolina's maps in this period exhibited large and steady pro-Democratic efficiency gaps, favoring Republicans only *once* between 1972 and 1990. Likewise, the plan that allegedly captures North Carolina from the 1990s to the present looks nothing like the state's actual maps in the 1990s and the 2000s. These maps were almost perfectly symmetric—not skewed in Republicans' favor for most preference values for Party B, as Hood's example maintains (p. 19). Hood's second hypothetical plan *does* resemble North Carolina's current congressional plan, which indeed benefits Republicans under almost all electoral conditions. This, of course, contradicts Hood's main point, because it shows that Republican control of the redistricting process—not the need to comply with the Voting Rights Act—accounts for the current plan's asymmetry.

At the end of his report, Hood makes one more comment about the efficiency gap: that it is difficult to evaluate prospectively, before an election has taken place (p. 26). That is irrelevant here, where we have observed North Carolina's current congressional plan's performance in an actual election. Hood's remark is also entirely too pessimistic, ignoring decades of advances in forecasting district-level election outcomes using previous election returns and/or demographic attributes of localities (typically, precincts). Here, for instance, plaintiffs used a regression model in their original complaint (filed before the 2016 election) that predicted the current plan's performance using 2012 precinct data. The complaint also discussed sensitivity testing that showed how the plan's performance would vary as the state's electoral environment shifted. Notably, the efficiency gap predicted by the sensitivity testing for an electoral environment like that of 2016 was accurate to *within a percentage point*. Thus not only can the efficiency gap be assessed prospectively if there has not yet been an election, it can be done so with impressive precision.

8 Trende: Efficiency Gap Methodological Choices

In my original report, I used what is known as the "full method" to calculate the efficiency gap. That is, I determined each party's wasted votes in an election, subtracted one sum from the other, and divided the resulting difference by the total number of votes cast in the election (pp. 18-19). This method incorporates variations in turnout from district to district, and has been endorsed by a federal court as "preferable because it accounts for the reality that voters do not go to the polls at equal rates across districts" (*Whitford v. Gill* (2016), p. 54). I also made certain other methodological choices in my analysis: (1) I presented efficiency gap values in percentage points rather than seats because this unit of measurement is more easily understood by most audiences (pp. 27-29). (2) To account for the reality that House delegations vary in size, I set my recommended efficiency gap thresholds so that they always represent an average partisan advantage of at least half a seat (pp. 53-54). And (3) I only considered delegations with at least seven seats because all measures of partisan asymmetry become less reliable for smaller delegations (p. 20).

In his report, defendants' final expert, Trende, does not actually disagree with any of these analytical choices. Instead, he points out that other scholars have sometimes made different choices (paras. 28-33). This is true enough, but completely irrelevant. Those other scholars' methods are not being used in this case. Mine are, and Trende apparently does not object to any of them.

It is also important not to overstate the differences between these methods. For example, there is a 0.98 correlation between the efficiency gaps that I calculate and those computed by Stephanopoulos and McGhee. Similarly, while Stephanopoulos and McGhee only consider House delegations with at least eight seats, just 49 of the 512 elections in my database involve delegations with exactly seven seats, for which this methodological choice could possibly matter.

In discussing whether the efficiency gap should be presented in percentage points or seats (para. 29), Trende also fails to notice that I used a *hybrid* approach in my report. That is, I reported the efficiency gap in percentage point terms, but I considered delegation size in setting my recommended efficiency gap thresholds. In my opinion, this approach results in the best of both worlds. It presents the efficiency gap in a format that is more intuitive for most people. But it still takes into account variations in delegation size at the critical point at which thresholds are determined for heightened scrutiny under the second prong of plaintiffs' test.

Trende makes another odd argument when he suggests that North Carolina's current congressional plan must be valid because its efficiency gap in 2016 was smaller than its predecessor's in 2012 and 2014 (para. 32). The current plan's efficiency gap was only *marginally* smaller (-19.4%) than its predecessor's (-21.4% and -21.2%), and still egregious by historical standards. That Stephanopoulos and McGhee found that the 2011 plan's asymmetry was insufficiently durable (p. 879) also says nothing at all about whether the current plan's asymmetry is likely to persist for the rest of the decade. In my report, I used several techniques to establish the resilience of the current plan's skew: sensitivity testing (pp. 54-61), a comparison between plans' first and lifetime average efficiency gaps (pp. 47-50), and a series of prognostic tests (pp. 42-47). Again, Trende does not question any of these techniques, and so presumably accepts their conclusions.

9 Trende: Efficiency Gap Ease of Calculation

Trende asserts that the efficiency gap is not easy to calculate because it requires imputations for uncontested races (para. 34) and because it is unclear what the benchmark for comparison should be in each state (paras. 35-40). To begin with, both of these claims apply to *any* measure of partisan asymmetry, not just to the efficiency gap. Whether one is using the efficiency gap, partisan bias, the mean-median difference, or any other metric, one must (1) address uncontested races and (2) set a benchmark for comparison.

With respect to the imputations, different techniques also yield virtually identical outcomes. For instance, Stephanopoulos and McGhee imputed results in uncontested races using different methods from mine (and using the simplified form of the efficiency gap too). Yet as I mentioned above, their efficiency gap estimates were almost perfectly correlated with mine. Once a sufficiently rigorous imputation approach is employed, incorporating presidential election results and incumbency information, further methodological choices make very little difference.

With respect to the benchmark for comparison, furthermore, there is an obvious one for the efficiency gap that is nearly universally employed: *zero*, the point at which both parties waste equal numbers of votes, and at which a plan treats both parties perfectly symmetrically. Would a randomly generated plan in every state have an efficiency gap of zero? As Jowei Chen and my former colleague, Jonathan Rodden, have shown, the answer may be no, though more research is necessary on this question. But whatever the answer is, it is irrelevant here. It is always appropriate to measure the efficiency gap using a zero baseline, so that the metric reveals the extent to which a plan diverges from perfect partisan symmetry.

I also note that the factor that Trende thinks might require a non-zero baseline—political geography—is fully incorporated into the *other* two prongs of plaintiffs' test. If a mapmaker intends to follow a state's geographic contours when designing a plan, then the mapmaker does *not* intend to benefit or handicap either party, and thus does not possess the requisite discriminatory purpose. Similarly, if a state's spatial patterns make it very difficult to craft a symmetric plan, then there exists a legitimate, neutral justification for a plan's asymmetry, and again the plan would be upheld. To avoid conflating the test's three prongs, it is thus important not to try to "adjust" a plan's efficiency gap based on any "natural" partisan tilt of a jurisdiction. Doing so is unnecessary given the test's other elements.

10 Trende: Efficiency Gap as Proportional Representation

Trende argues that the efficiency gap amounts to "proportional representation for first-past-the-post systems" (paras. 41-48) because using the *simplified* method for calculating the measure, a zero score is achieved by a two-to-one seat-to-vote relationship. But as I explained above, I employed the *full* method for computing the efficiency gap, not the simplified method, thus taking into account differences in turnout from district to district. When the full method is used, the efficiency gap is exclusively the product of district-level (not state-level) data, and it implies nothing at all about any particular seat-to-vote relationship.

It is also important to be clear about terms' definitions. Proportional representation means something specific: an electoral system in which parties' seat shares *match* their vote shares, because there is an exact *one-to-one* seat-to-vote relationship. But even using the simplified method for calculating the efficiency gap—and even assuming an efficiency gap of exactly zero—parties' seat shares *do not* match their vote shares, because there is a *two-to-one* seat-to-vote relationship. Reliance on the efficiency gap as a standard for assessing districting plans thus neither requires nor encourages proportional representation. In fact, proportional representation is *incompatible* with a consistently low efficiency gap.

Furthermore, the two-to-one seat-to-vote relationship that is implied by a zero efficiency gap (calculated using the simplified method) is not some arbitrary figure. Rather, as a federal court has pointed out, "Based on decades of observed historical data . . . with single-member, simple-plurality systems . . . we can expect that for every 1% increase in a party's vote share, its seat share will increase by roughly 2%" (*Whitford v. Gill* (2016), p. 51). In other words, the (simple-form) efficiency gap measures a plan's divergence from the seat-to-vote relationship that has actually characterized American elections for generations. This hardly amounts to a backdoor imposition or endorsement of European-style proportional representation.

11 Trende: Imputations for Uncontested Races

Trende notes that a tiny subset of my turnout imputations for uncontested races are very low (paras. 51-68). This is true. The explanation is that these races typically were *never* contested over the lives of their respective plans, and a minor glitch in my code prevented proper imputations from being produced for these cases—which number just 58 out of the nearly 8000 in my database.

Fixing these imputations makes no difference for any of my substantive conclusions, either here or in my original report. In particular, there is a correlation of *0.999* between my original efficiency gap estimates and my updated estimates based on the revised imputations. Figure 8 makes this point graphically. It is a scatter plot with the original efficiency gap estimates on the x-axis and the updated estimates on the y-axis. As is evident, almost all of the points lie exactly on the best fit line. The *only* point to change appreciably is Texas in 1972, whose revised efficiency gap is somewhat more pro-Democratic than its original efficiency gap.

Nevertheless, in an abundance of caution, all of the figures in this rebuttal report (and all of the textual discussion of these figures) use the updated rather than the original efficiency gap estimates. Again, the practical implications of this are nil.

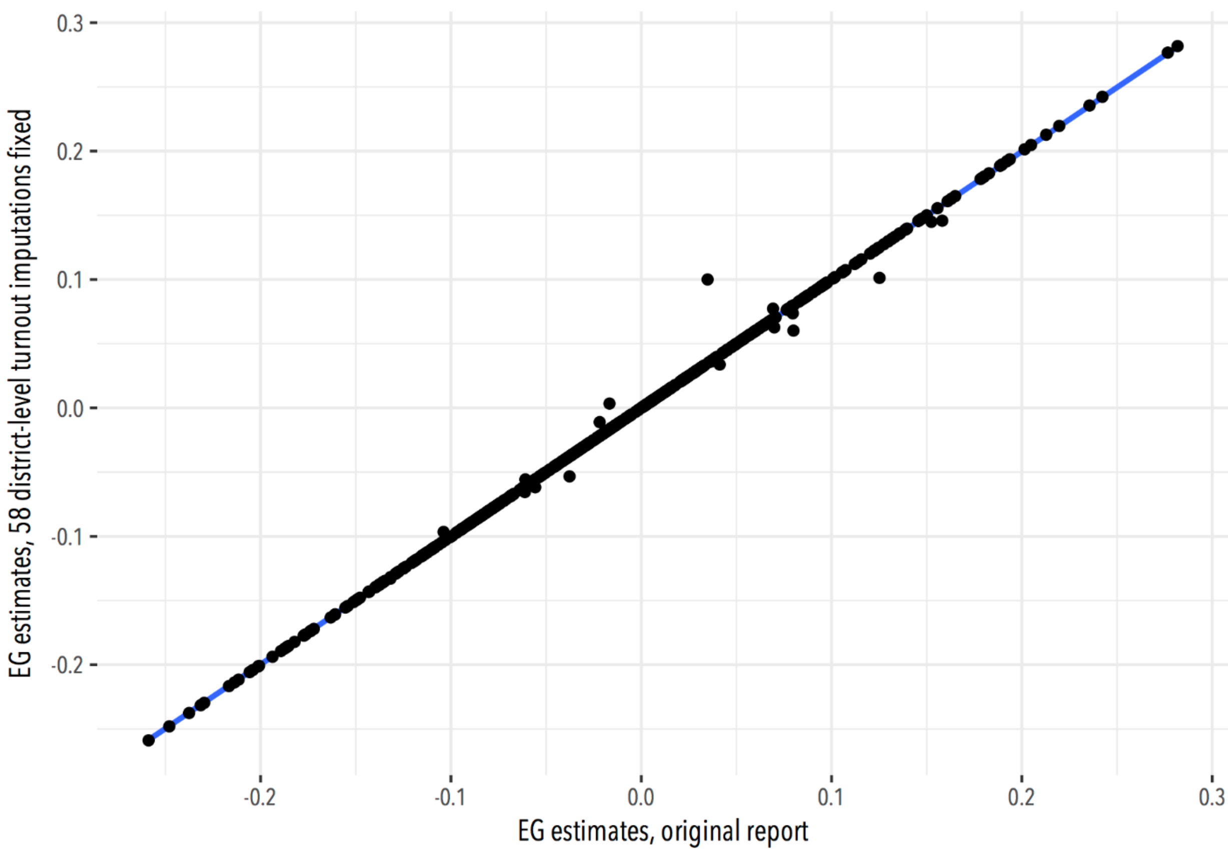

Figure 8: Updated efficiency gap estimates based on revised imputations versus original efficiency gap estimates, Congressional elections, 1972-2016.

12 Trende: Efficiency Gap Prognostic Tests

Trende states that there are errors in the prognostic tests that I conducted in my report (paras. 69-72). There are no errors. There are a few instances where the text of my report was not updated to reflect the final versions of the charts that display the results of the tests. But the charts themselves are entirely correct, and Trende does not suggest otherwise.

Trende also goes through all sixteen of these charts and paraphrases their results (paras. 74-75). As the charts present so much information that they can be difficult to process en masse, I want to focus on what I consider to be the two most probative figures.

First, the upper left pane indicates the proportion of plans that are flagged given different efficiency gap thresholds. Both over the entire 1972-2016 period (p. 44) and from 2002 to the present (p. 45), this proportion is quite low (around 25%) once my suggested 12% threshold for plans with fewer congressional seats is reached. By comparison, in the current redistricting cycle alone, more than 200 cases were filed in more than 40 states, resulting in more than 20 plans being invalidated or designed by the courts (Litigation in the 2010 Cycle). The level of potential judicial activity implied by my analysis is therefore quite low.

The second key chart (the second pane from the left in the bottom row) depicts the false positive rate. This is the proportion of plans whose *initial* efficiency gaps are above a certain threshold, but whose *average* efficiency gaps over the *remainder* of their lifetimes are relatively small (i.e., below the levels set forth in Table 2 of my original report). Both from 1972 to 2016 (p. 44) and from 2002 to 2016 (p. 45), the false positive rate plummets nearly to zero once my suggested 12% threshold for plans with fewer congressional districts is reached. In other words, there is almost no chance that courts using this threshold would ever flag a plan that would actually turn out to have a small (or opposite-signed) efficiency gap, on net, over its lifetime.

In combination, these two figures establish that my recommended thresholds are quite conservative. Relatively few plans would be flagged under these thresholds—far fewer than are targeted under other redistricting causes of action. And virtually all of the plans that would be flagged would be "hits," in the sense that they would, in fact, go on to significantly benefit the same party that enjoyed an advantage in the first election after the lines were drawn.

13 Trende: Efficiency Gap Anecdotes

Trende devotes the bulk of his report to a lengthy series of anecdotes about the efficiency gap (paras. 76-140). All of these anecdotes have exactly the same structure. Either a plan exhibits large efficiency gaps when we might have expected it to exhibit small ones (because it was designed by a court, a commission, or divided government). Or a plan exhibits small efficiency gaps (or even efficiency gaps favoring the opposing party) when we might have expected it to exhibit large ones (because it was designed by a single party). Or a plan's efficiency gap fluctuates significantly from election to election.

First, it is important to stress that these are unrepresentative anecdotes, not a comprehensive analysis of the data. My systematic analysis of a large number of elections leads me to conclusions that contradict Trende's selected examples. For instance, I showed that unified Democratic control of redistricting leads to a pro-Democratic shift in the efficiency gap, while unified Republican control leads to a pro-Republican shift (p. 34). These results are especially strong in more recent years, when redistricting technology has improved. Similarly, I showed in several ways that plans with large initial efficiency gaps tend to remain skewed in the same direction over their lifetimes. These plans generate very small rates of false positives (pp. 44-45); they exhibit high correlations between their initial and remainder-of-plan average efficiency gaps (pp. 48-49); and their partisan tilts persist even in the face of large changes in the statewide electoral environment (pp. 57). Trende never disputes any of these analyses (or their conclusions), choosing instead to highlight a handful of cases from the several hundred in my database.

Second, while Trende discusses the efficiency gap, none of his critiques actually have anything to do with the efficiency gap *per se*. Rather, they are simply points that apply to *all* measures of partisan asymmetry, and indeed all metrics that are based on observed election results. Parties sometimes lose seats they expected to win; they sometimes win seats their opponents hoped they would lose; and there is a reasonable amount of volatility from one election to the next. These factors affect the efficiency gap, true, but they also affect *every* measure that is derived from parties' votes and seats. There is simply nothing that is specific to the efficiency gap in Trende's analysis.

Third, Trende's examples conflate the efficiency gap with the rest of plaintiffs' three-part test, which requires (1) discriminatory intent, (2) a large and durable discriminatory effect, and (3) a lack of a neutral justification before liability is found. When the entire test proposed by plaintiffs is taken into account, not one of Trende's examples is problematic. For instance, consider Trende's many examples involving a party that intended to gerrymander but whose efforts led to small or even unfavorable efficiency gaps. Under plaintiffs' test, such plans would—correctly—be left undisturbed for want of discriminatory *effect*.

Likewise, consider Trende's anecdotes of large efficiency gaps arising even in the absence of a deliberate attempt to benefit or handicap a party. Here too there is no difficulty under plaintiffs' test. Plans of this kind would be sustained because discriminatory *intent* was absent.

Last, we have Trende's examples of efficiency gaps fluctuating from one year to another. And again, plaintiffs' test produces the right result. These plans would be validated because their discriminatory effect, even if deliberate and occasionally large, is not *durable* enough.

It is ironic, then, that Trende criticizes plaintiffs for being "reductionist" and for "suggest[ing] that gerrymandering can be summarized by a single statistic" (para. 89). Plaintiffs advocate a standard that encompasses several different concepts, and in which the efficiency gap is just one applicable measure for just one part of one prong (the severity of a plan's partisan asymmetry). If anything, Trende's anecdotes actually demonstrate the functionality of plaintiffs' multi-pronged test, in which the actual values of the efficiency gap are but one component.

14 Trende: Average Efficiency Gap

Trende argues that the average efficiency gap exhibited by a plan over its lifetime is not a meaningful statistic, and that I have "switch[ed] [my] inquiry" by examining the average efficiency gap in this litigation (para. 143). This criticism is misguided on several fronts. First, there is no more simple statistical summary of a variable than its mean. A plan's average efficiency gap reveals which party that plan benefited and by how much, on net, over the entire time the plan was in effect. Second, I analyzed the average efficiency gap extensively in the Wisconsin case to which Trende refers. Indeed, the court favorably cited my finding that "[g]iven historical trends . . . Wisconsin's plan would have an average pro-Republican efficiency gap of 9.5% for the entire decennial period" (*Whitford v. Gill* (2016), p. 51). And third, in my report in this case, I did not even examine "the sign of the 'average' efficiency gap" (para. 143). Rather, I examined the direction and magnitude of plans' *remainder-of-plan* average efficiency gaps—that is, their average efficiency gaps *not including* their initial efficiency gaps. This more rigorous approach properly declines to incorporate a plan's first score into any distillation of its subsequent performance.

Trende also displays a table listing plans that had efficiency gaps of the same sign in each election in which they were in effect (paras. 145-46). While this is not my preferred method for evaluating the durability of a plan's asymmetry, the table nevertheless further illustrates the workability of plaintiffs' test. There exist numerous maps (38 percent of the ones in the table, according to Trende) that were designed by a single party, that had large initial efficiency gaps, and that continued to exhibit efficiency gaps in the same direction in each subsequent year. These are exactly the maps that would be unlawful under plaintiffs' test. Conversely, there also exist numerous maps where all of these elements were *not* aligned—for instance, because a map had a small initial efficiency gap, or was not designed by a single party. These maps properly would not be struck down by plaintiffs' test.

15 Trende: Sensitivity Testing

In my report, I emphasize the need to establish that a plan's asymmetry will persist over its lifetime before deeming that plan suspect. One of the ways this can be shown is through sensitivity testing, and Trende, to his credit, carries out sensitivity testing for several states using their 2012, 2014, and 2016 election results (paras. 147-66). Unfortunately, he makes methodological mistakes in conducting his sensitivity testing, and then draws the wrong conclusions from it.

Trende's analytical errors are twofold. First, he uses *national* (rather than *state*) election data to decide by how many points to shift each state's vote share in each direction (para. 149). This is problematic when a party performs especially well or poorly in a given state, and also when a given state's own electoral history looks different from that of the country as a whole. For example, 2012 was quite a good Democratic year in Illinois (unsurprising, given that the Democratic candidate for president hailed from the state), meaning that it is unreasonable to shift the Democratic vote share by up to *another* eight points in a Democratic direction, thus simulating a bigger Democratic wave than has ever occurred over the last half-century. Similarly, southern states like Alabama, Georgia, and Texas have experienced much more electoral volatility than the nation generally over the last fifty years. It is therefore inappropriate to model their swings as if they were the same as those of the rest of the country.

Second, Trende makes no effort to consider the *likelihood* of different electoral outcomes. That is, he treats epic waves like those of 1974 and 1994 as being as plausible as any other result. But this is obviously wrong; waves that large occur only very rarely, and most elections are much closer to the historical mean. In my sensitivity testing, in contrast, I was careful to note the electoral swings that North Carolina has actually witnessed from 1972 to 2016, in order to help assess the likelihood that any particular electoral environment will arise over the rest of the decade (p. 59).

Interpreting his data, Trende again stresses the fact that the efficiency gap can change from election to election. This is obvious, however, and does not require sensitivity testing to be demonstrated. The point of sensitivity testing is to help us determine if a plan that has exhibited a large initial efficiency gap will continue to benefit the same party given plausible shifts in the state's electoral conditions. Arizona's map, for instance, had a double-digit pro-Democratic efficiency gap in 2012. But according to Trende's sensitivity testing, this districting plan would advantage Republicans if their vote share increased by just three percentage points (para. 150). This is strong evidence that the map's asymmetry is not durable enough to be actionable.

Conversely, plans including Florida's, Michigan's, Ohio's, Pennsylvania's, and Virginia's exhibited very large efficiency gaps in 2012 that would *not* dissipate given realistic changes in the statewide electoral environment. The sensitivity testing thus confirms the durability of these gerrymanders, and indicates that the discriminatory effect prong of plaintiffs' test is satisfied.

Also worth noting here is that Trende's analysis entirely corroborates my own findings about the persistence of the asymmetry of North Carolina's current congressional plan. Trende shows that the plan's efficiency gap would become even more pro-Republican if the statewide vote swung by up to six points in a Democratic direction. Democrats would have to improve their 2016 showing by *nine percentage points*—a truly enormous shift seen only once in the last half-century—to encounter anything other than a double-digit efficiency gap against them (para. 165). This is the most relevant chart in Trende's discussion—a vivid demonstration of the Republican advantage in the current North Carolina plan and entirely consistent with plaintiffs' claims.

16 Trende: Efficiency Gap and Competitiveness

Using an artificial example of thirteen districts decided by one vote each, Trende observes that small vote shifts from year to year can produce large changes in the efficiency gap (paras. 169-172). Again, I note that any plan in which the efficiency gap could swing so significantly from election to election would not be one in which the durability component of the discriminatory effect prong was satisfied. Moreover, if a state *aimed* for competitiveness when it designed its plan (as Washington in the 1990s apparently did), then the state would *not* have sought partisan advantage, and so the discriminatory intent prong would not be met.

Trende then points out, using another pair of toy examples, that uncompetitive plans can have low efficiency gaps (at least over narrow vote share ranges) (para. 176). This only proves that the efficiency gap is not a measure of competitiveness—which is a "feature" of the metric, not a "bug." The Supreme Court has expressed interest in the concept of partisan asymmetry. But it has *not* suggested that it wishes to police *bipartisan* gerrymanders that protect incumbents from *both* parties. It would therefore be improper if the efficiency gap incorporated competitiveness in addition to partisan skew.

17 Trende: Efficiency Gap and Party Control

Finally, Trende argues that single-party control is not a significant driver of the efficiency gap (paras. 178-81). But in making this argument, Trende does not challenge or address my regression analysis finding that single-party control is indeed a significant driver, especially in more recent years (p. 34). Instead, he simply notes that there are divergences between the efficiency gap's trends over time and the temporal patterns of party control over redistricting (paras. 179-81). That there are divergences is undeniable—and unsurprising, since no one has asserted that party control is the *only* explanation for the efficiency gap, such that the efficiency gap follows party control in lockstep. Other factors, such as national electoral tides (and variations in the rates with which individual states follow national trends), state-specific variations in political geography, candidate quality, and so on, surely play a role as well. But equally obviously, the existence of divergences between the two time series does not mean that single-party control is not a key determinant of the efficiency gap. Discriminatory intent can be an important cause of a plan's discriminatory effect without being its sole source.

Simon Jackman

Simon Jackman

April 17, 2017

References

Gelman, Andrew and King, Gary. 1994. "A Unified Method of Evaluating Electoral Systems and Redistricting Plans" 38(2) *American Journal of Political Science* 514-554.

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

King, Gary and Browning, Robert X. 1987. "Democratic Representation and Partisan Bias in Congressional Elections" 81(4) *American Political Science Review* 1251-1273.

Levitt, Justin, "Litigation in the 2010 Cycle" *All About Redistricting*, available at http://redistricting.lls.edu/cases.php.

McDonald, Michael D. and Best, Robin E. 2015. "Unfair Partisan Gerrymanders in Politics and Law: A Diagnostic Applied to Six Cases" 14(4) *Election Law Journal* 312-330.

McGhee, Eric. 2014. "Measuring Partisan Bias in Single-Member District Electoral Systems." *Legislative Studies Quarterly* 39(1): 55-85.

Rodden, Jonathan. 2010. "The Geographic Distribution of Political Preferences 13 *Annual Review of Political Science* 321-340.

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

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

Cases

LULAC v. Perry, 548 U.S. 349 (2006).

Whitford v. Gill, ___ F. Supp. 3d ___, 2016 WL 6837229 (W.D. Wis. Nov. 21, 2016)