REBUTTAL OF DEFENDANT'S EXPERT REPORTS FOR COMMON CAUSE V. LEWIS

JONATHAN C. MATTINGLY

1. Rebuttal of Common Points around the effect of political geography

A number of the experts discuss the potential effect that the North Carolina political geography has on redistricting.[1] Specifically, Dr. Hood mentions the possibility of natural packing, and Dr. Barber the urban versus rural divide. These points demonstrate the need for a methodology that accounts for this political geography; ensemble methods precisely capture it. One of the strengths of the ensemble method based on Markov Chain Monte Carlo is that it translates clearly stated redistricting criteria into a distribution on redistricting plans. The distribution on redistricting plans can distinguish between typical plans and atypical plans. This determination is fundamentally informed by the geometry of the state, its political geography, and the spatial structure of the elections used to probe the redistricting plan.

The fundamental power of the ensemble method is that it begins with a clear set of redistricting criteria as an input. It then creates a large, representative ensemble of redistricting plans which accounts for the geometry of the state and the geography of where people live and how they vote. Any collection of voting data can then be applied to this ensemble of restricting plans to obtain a collection of election results. The election results give a benchmark against which a particular redistricting may be compared under the same set of voting data. It is only the relative difference between the ensemble and the enacted plan which matters. Our ensemble of restricting plans naturally incorporates how nonpartisan redistricting criteria interact with the political geography and geometry of the state. It naturally adapts to natural packing in urban areas and other effects. It is capable of separating these natural effects from those of partisan gerrymandering. Because of this, this mode of analysis can separate bias that natural packing might induce from other effects. Additionally, none of these analyses rely on any forms of partisan symmetry or ideas of proportional representation. The ensemble method does not impose any idea of fairness nor does it select for a particular seats-to-votes curve. Rather it illuminates what the result would have typically been had only the stated redistricting criteria been utilized. It is quite possible, and often happens that the results from the ensemble method do not yield proportional representation and one party gains a natural advantage. One can then use this natural advantage as a benchmark to detect when a particular plan is biased beyond the neutral standard the ensemble establishes.

My report demonstrates that Republicans do have some natural advantage in North Carolina, but not nearly enough to explain the partisan bias of the enacted House and Senate plans. For example, in the Senate, using the AG16 election data, the ensemble shows that the Republicans will receive a majority of the seats with a minority of the statewide vote. However, the enacted plan gives the Republicans a supermajority using these same votes. [2] Similarly, in the House, using the CI12 election data with 51.8% statewide democratic votes, the ensemble shows that the Democrats win, on average, roughly 46% percent of the seats, whereas Democrats win around 42% of the seats under the enacted plan[3]. Hence there is a natural bias in favor of the Republicans, but it is significantly smaller than that of the enacted plan. These examples further demonstrate how ensemble methods are capable of separating geographic effects of natural packing from excessive partisan gerrymandering. We found similar results in our analysis of redistricting for the Wisconsin state legislature.[4]

Two of the defense's experts critique our choice of criteria in the ensemble, but take opposing views. Dr. Thornton claims that we have not considered the Reock score, incumbency, nor have achieved appropriate population deviations.[5] On the other hand, Dr. Brunell claims that our ensemble would only be valid if it *did* take partisanship into account, favoring the Republican party. He goes on to claim that we could conclude excessive gerrymandering only if the enacted plan were extreme amongst already gerrymandered plans which were designed with partisanship. We prefer to identify partisan gerrymandering by comparing a plan in questions to what would have happened if partisanship had not been explicitly considered. As discussed above, our ensemble describes what would have happened in the absence of gerrymandering, when

Date: June, 2019.

[1]For example see section V of Dr. Hood's report and Section 4 of Dr. Barber's report.

[2]See the corrected version of Figure 2 from my original report in the appendix of this report.

[3]See Figure 5 of my report; specifically 2012 Commissioner of Insurance race.

[4]See Fig 8 and 9 from "Evaluating Partisan Gerrymandering in Wisconsin" and the surrounding discussions. `https://arxiv.org/abs/1709.01596`

[5]Contrary to Dr. Thornton's claim, the initial report included everything but the Reock score, which has been included in this rebuttal; see Section 4 for details.

only the non-partisan redistricting criteria are considered.[6] Often the non-partisan redistricting criteria result in maps which favor one party or the other in the sense that one party gets more seats than predicted by proportionality and the statewide votes. The ensemble method only labels a map as an outlier when the bias far exceeds what would have naturally occurred.

2. In response to Dr. Brunell's report

We hold that one must understand what the results would have been without any partisan influence, in order to understand the effect of partisan gerrymandering. Dr. Brunell does not appear to agree with this. He instead states that it would only be fair to compare partisan gerrymandered maps against each other and to examine if a map is extreme amongst the gerrymandered plans. He goes on to claim that it would be unfair to compare a non-partisan map against a partisan map to decide if the partisan map is extreme.[7]

Dr. Brunell also makes a number of critiques on simulation methodologies which are not applicable to our method.

(1) **What distribution are we sampling?** The distribution we are sampling is completely explicit. We have described it exactly. In particular, it only includes the same widely accepted redistricting criteria as stipulated by the legislature. It is also easy to add additional factors, like considering incumbents, by adding them to the definition of the distribution. Knowing explicitly what distribution we are sampling from is one of the strengths of the Markov Chain Monte Carlo based ensemble method.

(2) **Are we sure that we have more than a few different plans in our ensemble?** We are certain that we have a large number of very different plans in our ensemble. We verified that our ensemble has 6.6×10^{86} unique plans in the House and 5.3×10^{30} unique plans in the Senate. Our ensemble are built out of an sub-ensemble on each of county-clusters. We verified that the sub-ensembles are diverse by sampling from 5 random initial conditions and ensuring the election results agree over all 5 sampling procedures. When aggregating to the statewide level, the full ensemble is exponentially diverse with a large number of plans which are very different.

(3) **Compactness.** Dr. Brunell criticizes optimizing compactness and implies that we have done so. We did not. We sampled compactness scores similar or better than those used by the Legislature. We had some better and some comparable.[8]

(4) **County-Clusters:** We have preserved county-clusters and accounted for the other redistricting criteria. Dr. Brunell does not mention this when stating how his previous criticisms do not apply to our work.

(5) **Inferring Intent:** The fact that the map-makers found such an unusual map *does* give some insight into intent. it would be highly unlikely to find such an unusually partisan map without looking for it.

Dr. Brunell's main criticism is that the enacted plan is not an extreme outlier in each and everyone of the 17 historical elections considered. This is not surprising, and we have explained this observation in some detail previously.[9].

To better understand why this is natural, consider the following example. Let us assume that a region has three varieties of people who always vote as a block and are spatially contiguous. For definiteness, lets call them red, purple, and blue people. We will assume that red always vote for the red candidate and blue for the blue candidate. Sometimes the purple vote for the red candidate and sometimes for the blue candidate. Hence, sometimes red wins two seats, and sometimes three seats, depending on how the purple people vote. Let us assume that most redistricting plans that one would naturally draw (without knowing where the red, purple, and blue people lived) would naturally produce 2 majority red districts, 2 majority blue districts, and one majority purple districts. We will call these neutral plans. Now let us consider a plan which is carefully drawn so that the purple people are never a majority but rather the purple people are split such that there are three majority blue districts and two majority red. We will call this the gerrymandered plan.

Under the gerrymandered plan the red candidates always win two of the five seats, but never more. This is typical of elections where the purple people vote with the blue people. It is typical because the majority purple district in the neutral plans would vote for the blue candidate to elect three blue candidates. On the other hand, in elections where the purple people vote with the red people, the outcome would be highly atypical as the neutral maps would have always produced three red winners but the gerrymandered plan only produces two red winners. In summary, atypical maps may lead to a typical split of elected officials under some vote counts, but not under others. It is not unusual for gerrymandered maps to sometimes produce typical results.

Furthermore, we have also observed an atypical and consistent skew in the enacted plan that favors the Republicans. For instance in Figure 5 from my report, we see that the vast majority of elections skew to the Republican party; this observation is analyzed and made formal in Figure 6.

---

[6]In one of our secondary ensembles we do protect incumbents. We do not do this in our primary ensemble. We show that the both ensembles indicate that the enacted plan is an extreme outlier.

[7]See page 3 of Dr. Brunell's report.

[8]See the Appendix Section E of my report.

[9]See the discussion on the Firewall in "Evaluating Partisan Gerrymandering in Wisconsin." `https://arxiv.org/abs/1709.01596`

When the statewide Democratic vote fraction rises above 51%, the Republicans may lose the majority as evidenced by the ensemble; in this regime the enacted plan is highly atypical and leads to more Republican seats than expected.[10] When the statewide Democratic vote fraction falls below 51%, the Republicans are nearly guaranteed to keep the majority of seats, again as evidenced by the ensemble of plans, and hence a mapmaker who is attempting a pro-Republican gerrymander need not ensure that the map acts atypically, since a typical map already delivers the majority to the Republicans.[11] In our analysis, elections which could have given a Democratic majority under the ensemble produce results more favorable to Republicans under the enacted plan. And yet, elections which are overwhelmingly Republican often produce more Democratic seats under the enacted plan than is typical under the ensemble, but never enough to dislodge the Republicans from power. This is not a flaw or contradiction in our analysis or a lack of stability as Dr. Brunell submits. Rather it shows how the enacted map responds to changes in statewide vote fraction as the Republican majority becomes threatened.[12]

The use of statewide percentages in our analysis should not be confused with proportionality in the votes to seat distribution. Rather, they are a proxy to examine how the number of expected partisan representatives changes as each party is favored. This view also reveals changes in election structure: For example, the 2012 Commissioner of Insurance and the 2008 Governor races have very similar statewide vote counts but lead to very different number of seats typically being elected. This illustrates the preceding point that proportionality is not an assumption in any way.

As explained in my report, the ranked-votes curves figures (Figures 4 and 7 in the original report) further reveal how some elections produce typical results and others do not: Where the ranked-votes curve corresponding to a certain map has a large jump when the typical ranked-votes curve does not, then any election which situates the 50% line through the jump will have atypical results for that election. If a given election does not produce an abnormal jump across the 50% line then the results will be typical, *even if the overall structure is atypical*. We have dubbed this jump "the signature of gerrymandering." It is seen in Figure 4 in the original report. More generally, if the 50% line passes through a region where the percentages in the ranked-votes curve are atypically low or high then the partisan result for the election will be atypical. Under the partisan swing hypothesis (not used in my initial report), one might imagine that the entire curve is shifted up or down by a swing in public sentiment which keeps the relative spacial structure of the votes stable. This might result in the 50% line crossing at a point where the districts are typical or atypical. Hence, some elections act differently than others. It is important to note, that we are not talking about vote-seat curves in this discussion and proportionality plays no role in the discussion.

In summary, the structure of the ranked-votes curves for a particular plan when compared to the rank-ordered marginals from the ensemble tell a lot about the structure of a plan. Using the ranked-votes curves, we can detect gerrymandered structure even when the election results in terms of seat counts are typical (because the 50% line happens to pass through a typical region of the ranked-votes curves). The atypicality would affect the seat outcomes should a shift in sentiment move the ranked-votes curves up or down so that the 50% line passes through an atypical region. Thus the number of elected officials from a given party may, at times, be typical even when the structure of the plan is atypical. For a more complete picture, we have also examined the typicality in the margins of victory. In short, it's not just about who wins, but also about "how" they win.

2.1. **More comments on stability of results.** As mentioned above, Dr. Brunell is concerned that the enacted plan is not an extreme outlier in each and every one of the 17 historical elections considered. He characterizes this as an "instability." As we explain above this stems from the misconception that a gerrymandered map will behave atypically under all different types of elections. Gerrymandered maps can behave atypically under some types of elections and typically under other types of elections. For example, a map may only become atypical when a party is in danger of losing the majority. To demonstrate this, we provide an additional perspective on Figures 2 and 5 from the initial report.

Although none of our previous analysis employed the *uniform swing hypothesis*, we now use it to examine how shifts in the statewide vote would affect the seat outcomes under different historic elections. The uniform swing hypothesis states that as statewide voter support shifts from one party to another, the vote percentages in each individual precinct shift by the same percentage. For example, the election CI12 has a statewide democratic vote fraction of 51.81%. To use the uniform swing hypothesis to create an election with a statewide percentage of 50%, we subtract 1.81% from the Democratic vote fraction in each precinct and add 1.81% to the Republican vote fraction. This will allow us to create a sequence of elections with different statewide vote fractions while preserving the spatial structure of the votes.

We pick six representative elections, one with a relatively high Democratic vote fraction (USS08), two with more balanced vote fractions (CI12 and PR12), and three with relatively high Republican vote fractions (GV12, LG16, and USS116). In particular, we include two of elections which Dr. Brunell discusses. We shift them in 1% increments to create a sequence of

[10]See Figures 2 and 5 of my report.

[11]We find that once the Democratic vote share dips below 48% to 49%, the maps are generally no longer atypical.

[12]It should be noted that the enacted Senate map shows a systematic bias to the Republican over an entire middle range of elections, at times leading to an atypical Republican supermajority.

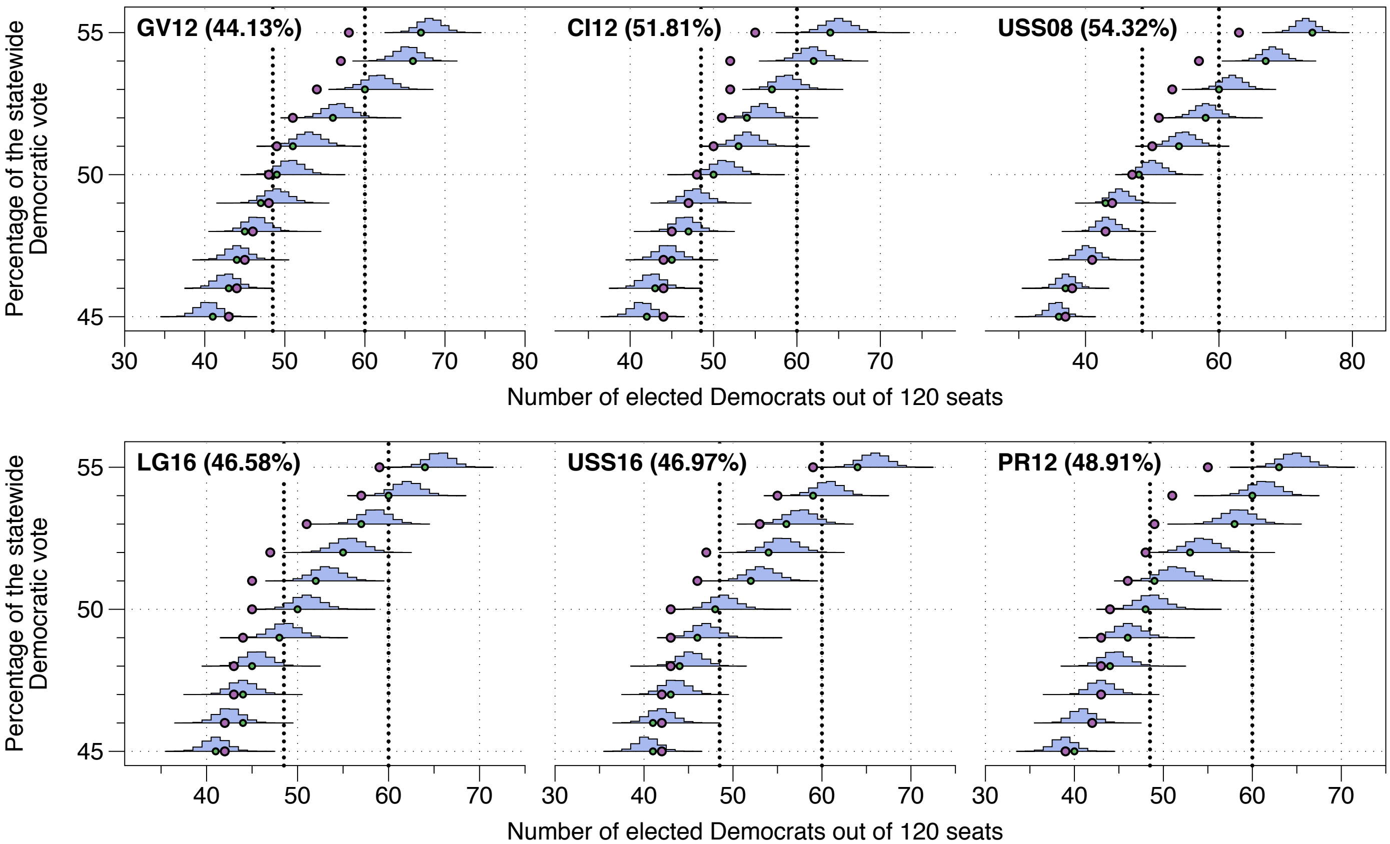


FIGURE 1. Purple dots show the enacted plan; the green dots show a plan in the ensemble. The dashed line at 60 seats shows the majority, and the dashed line at 48.5 seats shows the Republican supermajority threshold. The number of Democrats elected in the Senate which has a total of 120 seats.

elections ranging between 45% and 55% statewide Democratic fractions.[13] The results are shown in Figure 1. Purple dots represent the enacted plan and the green dots represent a selected plan from the ensemble. In each case, the enacted plan is an extreme outlier when the statewide percentage rises above 50%. This is roughly when the Republicans would typically begin to lose the supermajority under the ensemble. Under the enacted plan they never lose the supermajority at 50% of the vote. In many cases this supermajority persists well above 50% of the vote and the gap between the expected results by the ensemble and the enacted plan becomes quite large. Additionally the Republicans never lose a majority under the enacted plan, except in the USS08 election, if Democrats have 55% of the vote. This is in stark contrast to the ensemble. Under all six election structures, we see cases in which the Democrats are expected to achieve a majority of the seats under the ensemble, but receive a minority of seats under the enacted plan. We also see that the enacted plan becomes fairly typical, and even begins to elect more Democrats once the Republicans have a significantly more seats than required for a supermajority under the ensemble of plans.

These behaviors are robust and stable across the different elections considered. As already mentioned, the behavior can be understood by looking at the ranked marginal vote distribution plots. The results in Figures 2 and 5 show the same trends. There, rather then building different elections using the uniform swing hypothesis, we collect a number of historical elections with different statewide partisan vote fractions. There is more variation when using historic votes because of the differing spatial structures of the elections, but the trends are both the same and clear. Using the uniform swing hypothesis allows us to control for the spatial variation in different elections, and further demonstrates that a map may behave typically under certain types of votes, but extremely atypically under other types of votes.

3. IN RESPONSE TO DR. HOOD'S REPORT

Largely Dr. Hood is very supportive of the principles behind our approach. Dr. Hood emphases the importance of the political geography in any analysis. He also emphasizes the same nonpartisan redistricting criteria that we use with the addition of protecting incumbents, which we also include in one of our ensembles.[14] We construct our redistricting based on

[13]We verify that none of the districts vote fractions are shifted below 0% or above 100%.

[14]See Appendix Section E.25 of my report

the county-clusters which he agrees are important to preserve. Our method does not optimize the criteria, but rather we set a threshold so that each criteria is comparable or better that the enacted plan.[15] Dr. Hood claims that if one does not optimize, but sufficiently adheres to the constraints in the spirit of the enacted plan, than the maps produced would not identify the enacted plan as an outlier. This is not true. Our analysis follows his prescription and clearly identifies the enacted map as an outlier. Dr. Hood also discusses the possible importance of natural packing; the tendency of politically likeminded people to locate geographically close to each-other. Our method naturally accounts for natural packing and only identifies outliers as those maps whose partisan tilt is extreme after accounting for the effect of natural packing.

4. In response to Dr. Thornton's report

Dr. Thornton presents three major critiques of my report:

(1) She claims that the voting data used in my analysis was biased toward the Democrats.
(2) She claims the underlying criteria and assumptions used to generate the ensembles were not appropriate.
(3) She claims the enacted plan is not an extreme outlier under her own method of analysis.

The first critique is based on considering the statewide vote fractions of historic elections. Dr. Thornton suggests there is a Democratic bias in my choice of election data since the average of all of the statewide elections used is slightly Democratic. However, I used all of the major statewide elections available to me, including those considered by the legislature in the redistricting process. Furthermore, my analysis examines the enacted plan on an election-by-election basis, and hence questions of proportionality or average vote count are not relevant. One is free to consider the analysis of the enacted plan based on any given election. In fact, I demonstrate that the set of historic vote counts (favoring the Republicans) still lead to the conclusion that the enacted plan is an extreme outlier under a variety of metrics.

The critiques levied against the ensembles are either completely false or, when examined, do not change our stated conclusions. In this rebuttal, I demonstrate that we have gathered a ensemble of plans that is representative of the entire space. I also demonstrate that even accounting for *all* redistricting conditions laid out by Dr. Thornton or concerns about the ensemble, the enacted plan is still an extreme outlier.

Finally, Dr. Thornton provides analysis purporting to demonstrate the enacted plan is not an extreme outlier. Her analysis is based on an incorrect and arbitrary assumption. Her analysis does not account for the political geography of the state and instead relies on the observed average of seats won, along with assuming that a Democratic candidate has the same chance of winning in one district as they would in any other. My report shows how the standard redistricting criteria and the geometry of the state imply an natural class of distributions on redistricting plans of the state. Using the discovered distribution (rather than Dr. Thornton's imposed distribution), we determine that the enacted maps are outliers.

4.1. **In response to the chosen voting data.** Dr. Thornton claims that my results are influenced by the statewide elections chosen. [16] I have used all available historic election data from 2008 until the present from the US Senate, President, Governor, Lieutenant Governor, Commissioner of Insurance, and Attorney General races. The only reason I omitted the 2016 Commissioner of Insurance race was because it was not available in my data sources at the census block level.[17] Each election is structured differently and is a legitimate instrument to probe the features of a redistricting plan. In many of my analyses, the elections are kept separate and one is free to make conclusions on an election-by-election basis.[18] Even if a map is typical for some elections it can be an extreme outlier for others. If the map systematically subverts the intent of the voters as expressed by their votes for a particular class of elections this is undesirable and undemocratic.

Dr. Thornton goes on to claim that the (near complete) selection of relevant statewide historic elections inflates the Democratic vote share.[19] Dr. Thornton has averaged vote counts over all elections, and has declared the average vote count to be high. This statement is irrelevant to the analysis found in my report.

We have compared the enacted plan to the ensemble for each independent set of historic election data. We have not taken any aggregate average. Some of these elections have high statewide vote fractions of Republican votes; some have high statewide vote fractions of Democratic votes. My report considers all of them. This is demonstrated in Figures 2 through 6, Figure 8, each of the voting distribution figures in Section 4, pages 96-129, and elsewhere, in my report. There are a few results that are averaged over all elections, but these results are always based on the deviation from the ensemble, and are not based on the absolute number of seats won. As such they only measure the bias and not over all vote fraction.

As mentioned in my report, we find historic elections in which the ensemble primarily leads to a Democratic majority but where the enacted plan leads to a Republican majority. We find historic elections in which the ensemble primarily leads to no

[15]In two county-clusters we do not satisfy the desire to be with in 5% of the ideal population. However, we show after the fact that splitting precincts to satisfy the population constraints would not change any of our qualitative conclusions.

[16]See page 5 of Dr. Thornton's report.

[17]As shown in Table 1 of Dr. Thornton's report.

[18]For example see Figure 2, 4, 5, and 7 and the tables in Figure 3 and 6 in my original report

[19]See page 8 of Dr. Thornton's report.

Republican supermajority but where the enacted plan provides a Republican supermajority. In contrast, we never see a set of vote counts that provides the Democrats a majority or supermajority in the enacted plan where it would not be expected in the ensemble. This underlines the irrelevance of this average-vote critique. For instance, even if the enacted plan produces many typical results for Republican tilted elections, the fact that it systematically under-elects Democrats in more democratically tilted elections shows that the map is an outlier. In this scenario, the map suppresses the will of the people by under electing democrats when the election favors the democrats. The fact that it might not when the election favors the Republicans, does not make this behavior excusable. As shown in Figure 5 from the report, this scenario is not hypothetical as it describes the enacted maps in the North Carolina House.[20] My analysis finds results which favor the Republicans in that the enacted plan leads to a Republican supermajority while the ensemble plans typically do not. My analysis also finds that the enacted plan does lead to atypical results which favor the Democrate in some elections, yet these have little political consequence as both the enacted plan and the ensemble of plans lead to Republican supermajorities in these cases. In all of this discussion, it is worth emphasizing that the statewide vote fraction is never used in deciding who is an outlier; it is always the behavior relative to the non-partisan ensemble.

4.2. **In response to the critiques on the ensembles of maps and its generation.** Dr. Thornton critiques the ensembles used in my analysis on a number of levels. These critiques are primarily centered around the generation of the ensemble and the criteria used to select the ensemble.

4.2.1. *Convergence of sampling.* First, Dr. Thornton claims that it is impossible to know if a sampling procedure has reached the equilibrium distribution, [21] and goes on to cite a recent article which casts doubt on the efficacy of sampling methods employed for redistricting problems. [22] The referenced article was heavily criticized in a follow-up article published by the same journal, [23] in which it was pointed out that convergence criteria for Monte Carlo Markov Chains has existed since the early 1990's and was being improperly ignored in the original Cho and Liu article. Although I agree with Dr. Thornton that there is no absolute theoretical guarantee that a chain has completely converged to a representative sample, there are rigorous convergence tests which are used with great success in practice.

In fact, convergence tests have greatly contributed to the practicality, efficacy, and abundance of using Markov Chains to assess real world problems. Such applications include computer vision, [24] medicine, [25] genetics, [26] road safety [27] and more. If Markov Chains convergence was indeed not usable because representative sampling could never be theoretically assured, then one would have to invalidate the majority of scientific conclusions reached by all the fields that have employed them. We do, however, provide rigorous guarantees that the chain will mix exponentially fast along with empirical evidence that the chain has been run long enough.

We demonstrate that the chain has been run sufficiently long by showing that the statistics of interest have little variation over five different runs starting from five different initial seeds. More precisely, in my report I adopt the convergence criteria of Gelman and Rubin.[28] Convergence results are demonstrated for each cluster[29] and discussed in some detail in Section F of the Appendix. All of the plaintiff selected clusters have ensembles whose statistics have converged to a common distribution within small margins of error in the ranked marginal distributions. Convergence results were accumulated over 5 different initial conditions: one of the initial conditions was the enacted plan, and the rest were generated by a random, local, greedy algorithm which produces four other starting points. To confirm that any remaining deviations in our sampling procedure will not affect our overall results, we examine the statewide distribution of elected Democrats under the 2012 and 2016 Presidential votes for each of the 5 initial conditions. We plot the results in Figure 2 and find nearly identical distributions for each of the 5 initial conditions over both sets of historic vote counts.

[20]We dubbed this effect the firewall effect in our analysis of the Wisconsin General Assembly in `arxiv.org/abs/1709.01596`.

[21]See page 5 of Dr. Thornton's report.

[22]See page 20 of Dr. Thornton's report and reference 27 there in: Wendy K. Tam Cho and Yan Y. Liu. *Sampling from Complicated and Unknown Distributions Monte Carlo and Markov Chain Monte Carlo Methods for Redistricting*. Physica A 506 (2018).

[23]See William T. Adler, Samuel S.-H. Wang *Response to Cho and Liu, "Sampling from complicated and unknown distributions: Monte Carlo and Markov chain Monte Carlo methods for redistricting"*. Physica A 516 (2019).

[24]e.g. Jampani, Varun, et al. "The informed sampler: A discriminative approach to bayesian inference in generative computer vision models." Computer Vision and Image Understanding 136 (2015): 32-44.

[25]e.g. Locatelli, Isabella, Paul Lichtenstein, and Anatoli I. Yashin. "The heritability of breast cancer: a Bayesian correlated frailty model applied to Swedish twins data." Twin Research and Human Genetics 7.2 (2004): 182-191.

[26]e.g. Hey, Jody, et al. "Using nuclear haplotypes with microsatellites to study gene flow between recently separated Cichlid species." Molecular Ecology 13.4 (2004): 909-919.

[27]e.g. Jang, Hakjin, Soobeom Lee, and Seong W. Kim. "Bayesian analysis for zero-inflated regression models with the power prior: Applications to road safety countermeasures." Accident Analysis & Prevention 42.2 (2010): 540-547.

[28]See reference 11 of my report; this work has been cited over 10 thousand times according to google scholar as of May, 2019.

[29]e.g. see Tables 10 through 12 of my report.

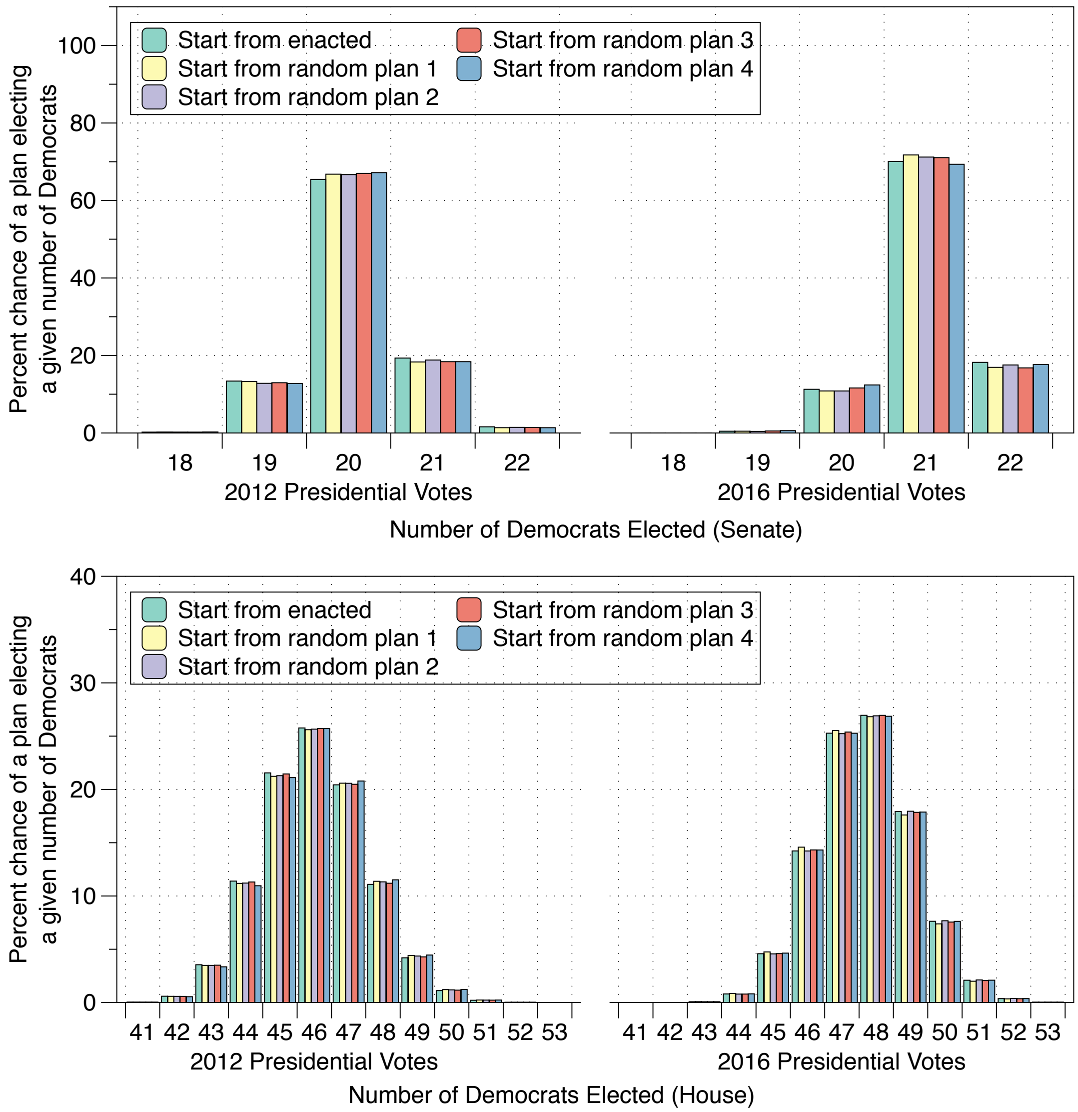


FIGURE 2. We plot the statewide distribution of elected Democrats over the 5 different initial conditions for the 2012 (left) and 2016 (right) Presidential vote counts for the North Carolina House (top) and Senate (bottom). We find nearly identical distributions for each of the 5 initial conditions across both chambers.

The county-cluster structure of the North Carolina General Assembly districts allows the districting problem to be solved independently in each of the county-clusters. This means that the problem can be broken up into a number of smaller, manageable problems. This makes the sampling highly reliable. All of the single district counties do not need to be sampled and, as we will see in the next section, a few other county-clusters are simple enough to be enumerated completely. The remaining county-clusters are then accurately sampled to high accuracy as described above. Combining the results from different county-clusters we produce 6.6×10^{86} unique plans in the House and 5.3 between 10^{78} to 10^{82} atoms. Estimates of the number of stars in the known, observable universe are around 3×10^{30}. We will see in the next section that in some cases we can calculate the results exactly and our algorithm performs extremely well.

4.2.2. *Comparison with exact results.* A few of the clusters are simple enough to enumerate all possible plans in the ensemble. We compare our ensembles generated by sampling with a Markov Chain Monte Carlo algorithm with a complete enumeration of all plans in the Duplin-Onslow, Person-Vance-Warren-Granville, and Franklin-Nash House clusters. We initially enumerate all redistricting plans which were within 10% population deviation. This produced 103,834 maps in Franklin-Nash, 4,156,485 maps in Duplin-Onslow, and 2,510 maps in Person-Vance-Warren-Granville.

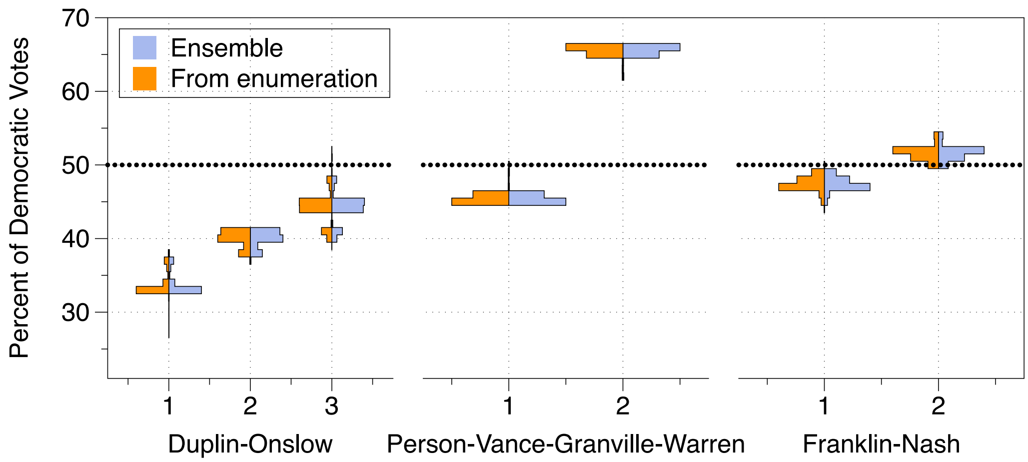


FIGURE 3. For three of the clusters, we enumerate all possible districting plans and re-weight them based on our score function. We compare this 'ground-truth' distribution to the distribution we have generated and see no discernible differences. The votes used in making this plot were taken from the 2012 Presidential race.

We then selected only the maps with less than 5% population deviation from the ideal, that satisfied the Pildes and Neimi criteria[30], in which the districts are contiguous, in which there are no double county traversals, and lastly in which precincts are kept whole. After this reduction 139 maps remained in Franklin-Nash, 5,076 maps in Duplin-Onslow , and 14 maps in Person-Vance-Warren-Granville. The resulting collection of maps are complete in that they contain all of the maps which satisfy the criteria which exist. After weighing the maps according to our score function, as is done in generating the ensemble distribution, we compare the exact distribution generated by enumeration with the sample distribution generated by our Markov Chain Monte Carlo algorithm using the 2012 Presidential votes. In Figure 3 we again compare the ranked marginal distributions ordered from least to most Democratic district. In these three clusters, there is no discernible difference between the ensemble provided by enumeration and by the ensemble we have generated. This shows that the Markov Chain Monte Carlo Methods correctly samples these distributions with high accuracy.

4.2.3. *Choice of redistricting criteria.* Next, Dr. Thornton critiques our analysis in that she says it does not fully comply with all of the stated redistricting criteria. She repeatedly asserts that because of this, it is expected that the enacted map is found to be an outlier. She claims (without justification) that if we had used what she represents as the full redistricting criteria then our results would look like the enacted map. We will see below that she is repeatedly wrong and that again and again we see that the enacted map is an outlier.

She first points out that we have two clusters in the House in which we have not complied with 5% population deviation. It appears that Dr. Thornton did not see Section E.24 of my report in which I directly address this issue and show that splitting precincts both allows one to achieve 5% population deviation and leads to qualitatively identical results as the less stringent threshold. In other words, the splitting of precincts needed to reduce the population deviation to below 5% is shown not to make any changes to the assessment that the enacted plan is an extreme outlier.

Next, Dr. Thornton points out that because my simulations do not explicitly track the Reock score, it is possible that the sampled maps go below the 0.15 Reock score threshold that Dr. Thornton says she takes from Pildes and Neimi.[31] It is correct that the initial report did not rule out this possibility. To test for the effect of complying with the Reock threshold, we have calculated the Reock score over all plans in each ensemble. We do not find a single district in any of the North Carolina Senate ensembles with a Reock score less than or equal to 0.15. In the House, we find very few districts with Reock score less than or equal to 0.15: In the Wake cluster, we find that a single district out of the 550,000 districts in the ensemble has a Reock score of less than or equal to 0.15; in the New Hanover-Brunswick cluster we find 4 of the 186,568 districts within the ensemble have a Reock score of less than or equal to 0.15; and finally, in the Mecklenburg cluster, we find 7 of the 486,588 districts in the ensemble have a Reock score of less or equal to 0.15. Removing these low Reock score would essentially not change our results at all and certainly not change our assessment that the enacted plan is an extreme outlier.

When repeating this analysis for our ensembles that explicitly preserve incumbents, we do not find a single district with a Reock score of less than or equal to 0.15. In short, accounting for a 0.15 Reock threshold will not affect our results. To

[30]We use the more stringent Polsby-Popper threshold that we used in generating our ensemble to make a fair comparison; we then verified that all of these enumerated plans had a Reock score of more than 0.15.

[31]A lower Reock score is a less compact district in some sense.

		No. elections atypically favoring Rep.		No. elections atypically favoring Dem.	
Chamber	Ensemble	Prob$< 1\%$	Prob$< 0.1\%$	Prob$< 1\%$	Prob$< 0.1\%$
Senate	Ignore municipalities	12 of 17	7 of 17	0 of 17	0 of 17
House	Ignore municipalities	13 of 17	11 of 17	0 of 17	0 of 17

TABLE 1. We count the number of elections under which the enacted plan is as or more extreme than a given percent of the ensemble of plans, favoring either the Republicans or Democrats.

contextualize the Reock scores of the ensemble with the enacted plan we examine the ranked-marginal distributions of Reock scores and compare them with the enacted plan. We present these results in the Appendix Section B. We find that for both the House and the Senate the ensemble of plans either typically has a higher Reock score or a comparable Reock score to the corresponding districts in the enacted plan. A higher Reock score signals a more compact plan.

Next, Dr. Thornton mentions that my report focuses on districting plans that preserve municipalities. She mentions that preserving precincts is more important than preserving municipalities and suggests that preserving municipalities is optional[32]. To the first point, the ensemble of plans used in my report strictly preserves precincts in all clusters with the exception of the House clusters in Wake and Mecklenburg. In these two clusters, the ensembles split the same or fewer precincts than the enacted plan. To the second point, we have generated a second set of ensembles in which municipalities are not considered and we find again that the enacted plan is an extreme outlier in favor of the Republican party[33]. Similarly to the result presented in Table 9 of my original, we count the number of considered elections under which the enacted plan is an extreme outlier in terms of the number of elected officials, and present our results in Table 1.

For both the House and the Senate, we find that the enacted plan is a 99% outlier in over two thirds of the historic vote counts under the ensemble that does not consider municipalities. In the House, the enacted plan is a 99.9% outlier in 7 of the 17 elections. In the Senate, the enacted plan is a 99.9% outlier in over half of the historic elections. Therefore, whether or not municipalities are considered as part of the redistricting process does not change our overall conclusions.

Dr. Thornton claims that I have not considered *all* the districting criteria when including the effects of incumbency. However, according to the list provided in paragraph 32 of her report, I have considered all criteria with the exception of Election Data. Though not stated in the initial report, virtually all plans in our ensemble do satisfy a 0.15 Reock threshold[34]. In the incumbent preserving ensemble, there are only three House clusters for which a complaint may be levied – in the Wake and Mecklenburg clusters we have not strictly preserved population deviation while considering incumbency protection, and in the Davie-Montgomery-Richmond-Cabarrus-Rowan-Stanly cluster we have not strictly minimized traversals. With the exception of historic election data, the incumbent preserving ensemble in the Senate complies with *all* stated redistricting criteria.

We have furthered studied the only three remaining clusters with potentially non-compliant plans. For Wake and Mecklenburg, we have confirmed that splitting precincts to achieve less than 5% population deviation does not affect the results of our primary ensemble.[35] In the original report, this analysis was not repeated for the secondary ensembles on Wake and Mecklenburg in the House that preserved incumbents. We believed that the result on the primary ensemble provides strong evidence that further splitting precincts will not significantly affect our results. In this rebuttal report, we now analyze splitting precincts in the incumbent preserving ensembles. We examine the ensemble of plans that split four or fewer precincts and achieve less than 5% population deviation and display our results in Figure 4. Similar to the results already presented in my report, we find that splitting precincts to achieve 5% population deviation does not change our conclusions.

In the Davie-Montgomery-Richmond-Cabarrus-Rowan-Stanly cluster, there are six possible county-to-county traversals. Three of them will occur in every plan: Montgomery to Richmond, Montgomery to Stanly, and Davie to Rowan. The remaining three possible traversals are Stanly to Cabarrus, Stanly to Rowan, and Cabarrus to Rowan. Each of these may occur in a districting plan. In the enacted plan only two of these remaining traversals occur; in our ensemble, 3,721 of the 7,725 plans have five total traversals, and the remaining plans have six. Using the 2016 Presidential voting data, we compare the ranked-marginal vote fraction distributions of the full ensemble to the part of the ensemble which minimizes traversals and the part of the ensemble that both minimizes traversals and accounts for incumbency. The ensemble that minimizes traversals and accounts for incumbency is comprised of 2,056 plans. We display our results in Figure 5.

We carried out additional tests in our report showing that even making the districts more compact while splitting fewer municipalities did not change our results.

[32]See paragraph 54 of Dr. Thornton's report

[33]See the appendix Section A of this report for the weights and convergence data of this new ensemble

[34]Our Poslby-Popper scores are all higher than 0.05 and we have validated that the overwhelming majority of our Reock scores are above 0.15

[35]See section E.24 of my report. Only the same number or fewer precincts are split than are split in the enacted plan.

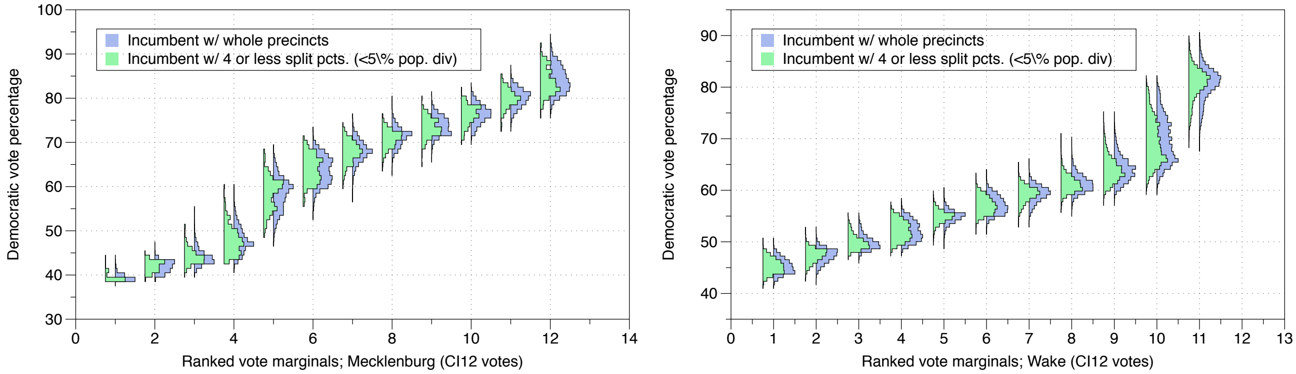


FIGURE 4. We show how splitting precincts in order to achieve 5% population deviation affects the ranked marginal vote distributions for the House clusters in Mecklenburg (left) and Wake (right). We consider the vote counts in the 2012 Commissioner of Insurance race.

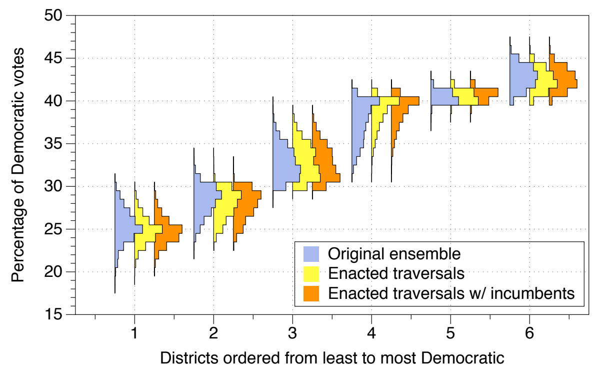


FIGURE 5. In the Davie-Montgomery-Richmond-Cabarrus-Rowan-Stanly cluster, we see consistent qualitative results when accounting for minimal traversals and incumbency. The vote counts used are from 2016 Presidential race.

In summary, Dr. Thornton wrongly states that we did not adhere to a number of criteria to which we did in fact adhere. In the case where we did not strictly satisfy the criteria, we have shown that strictly adhering to the criteria does not change our results. This directly contradicts Dr. Thornton's statements that adhering to the exact criteria used by the Legislature would produce plans with the same character as the enacted plan. Furthermore the idea that the exact criteria used by the legislature are the only legitimate point of comparison is overly rigid. By using traditional redistricting criteria in non-partisan fashion, the ensemble method produces a useful benchmark against which to compare specific plans.

Finally, Dr. Thornton suggests that the results of our ensemble may change if we made changes to our weights. The weights we chose are based on traditional redistricting criteria and are tuned to align with historical redistricting plans. It is possible, *a priori*, that changing our weights may lead to different conclusions, but we now show that not to be true for our ensembles. To test this, we reweight our ensembles to test how changing the weights would affect our qualitative results. To do this, we increase and decrease our compactness weight, population weight, and municipal weight in the score function by 20%. We then examine the number of elections in which the enacted plan is an outlier when compared to the reweighted ensembles in terms of the number of Democrats and Republicans elected. In this analysis, we freeze the clusters not selected by the plaintiffs and then produce a table similar to Table 9 from my original report (see also Tables 1 and Table 10 of this rebuttal).

Similarly to our previous results, we find a significant number of elections in which the enacted plan leads to more Republican seats than in either 99% or 99.9% of the ensemble; conversely, in nearly all cases, we find no elections in which the enacted plan leads to more Democrats than 99% of plans in the ensemble. The only exceptions are when considering an increase in the Polsby-Popper weight and in the municipality weight, in which we find that the votes of the 2012 Governor's race lead to more Democrats than over 99% of the plans in the ensemble; we note, however, that under these votes, the Republicans nonetheless have secured a supermajority in the enacted plan. We further note that the number of outlying elections we see for each of the four new ensembles remains consistent with the number of outlying elections present in the original ensemble.

		No. elections atypically favoring Rep.		No. elections atypically favoring Dem.	
Chamber	Reweighting	Prob $< 1\%$	Prob $< 0.1\%$	Prob $< 1\%$	Prob $< 0.1\%$
Senate	Original	10	8	0	0
Senate	$w_{\text{pop}} \uparrow 20\%$	10	8	0	0
Senate	$w_{\text{pop}} \downarrow 20\%$	10	8	0	0
Senate	$w_{\text{PP}} \uparrow 20\%$	11	8	0	0
Senate	$w_{\text{PP}} \downarrow 20\%$	10	8	0	0
Senate	$w_{\text{M}} \uparrow 20\%$	10	8	0	0
Senate	$w_{\text{M}} \downarrow 20\%$	10	8	0	0
House	Original	7	5	0	0
House	$w_{\text{pop}} \uparrow 20\%$	7	5	0	0
House	$w_{\text{pop}} \downarrow 20\%$	7	5	0	0
House	$w_{\text{PP}} \uparrow 20\%$	7	5	1	0
House	$w_{\text{PP}} \downarrow 20\%$	7	5	0	0
House	$w_{\text{M}} \uparrow 20\%$	6	5	1	0
House	$w_{\text{M}} \downarrow 20\%$	7	5	0	0

TABLE 2. Of the 17 considered elections, we count the number of elections in which the enacted plan is a 1% outlier and a 0.1% outlier favoring either the Republicans or the Democrats when we reweight our score function. We find two cases in which the Democrats elect more seats than 99% of plans in the ensemble, and no cases in which the Democrats elect more seats that 99.9% of plans in the ensemble. In both of these cases, the election data is taken from the 2012 Governor's race, which yields a Republican supermajority in the enacted plan. We find a significant number of elections in which the Republicans elected more seats than expected to an extreme extent. As in my original report, $w_{\text{pop}}, w_{\text{PP}}$, and w_{M} refer to the population, Polsby-Popper, and municipal weights, respectively.

4.3. **In response to Dr. Thornton's outlier analysis.** Dr. Thornton erroneously states that no outlier analysis (statistical test) was performed in my report. In fact this was the entire thrust of the report. Dr. Thornton then ignores the distribution collected and encoded in the ensemble and creates a model which is based on erroneous assumptions to perform an outlier analysis. We show that her analysis has no relation to the distribution of the ensemble, and hence no relevance. Dr. Thornton makes no attempt to respect the complicated political geography of the state, instead treating it as politically homogeneous; the ability to account for this structure is one of the strengths of our MCMC based ensemble method.

More perplexing, Dr. Thornton inverts the analysis being done and then claims it is somehow related to ours. Dr. Thornton devises a simplistic coin-flip model for which party a given district will elect. Beyond the fact that she treats all districts the same, this is not what we are doing in our analysis. We are fixing a particular vote count and then changing the districts around *en masse* to see the results of that vote count in maps that were not drawn using partisan criteria.

We hold that one must understand the partisan make up of a given election to understand how elected representatives change. The votes stay fixed in our analysis and the boundaries change. In this way we illuminate what outcome would be seen from a given set of votes in a map without any partisan input.

4.3.1. *Our statistical tests.* Dr. Thornton states that we perform no statistical test when this is in fact precisely what we do, even according to her own definition. She fixes on calculating a standard deviation for the distribution and then quantifies the number of standard deviations outside the mean the enacted plan lies. This calculation is simply a surrogate for calculating the tail or outlier probability under the assumption that the values of interest are close to being normally distributed. In the setting of a normal distribution, a plan being two or three standard deviations outside of the mean corresponds, respectively, to having 97.7% or 99.86% of plans in the ensemble be less extreme than the plan in question. Since we are ultimately interested in the chance that a random plan is as or more extreme than the enacted plan, we can simply count the number of plans from the ensemble which are as or more extreme than the enacted plan. Hence if only one plan out of a 100 is as or more extreme than the enacted plan then it is a 1% outlier. In many settings like the labor statistics setting which Dr. Thornton references to justify her analysis, an assumption that values are normally or binomially distributed might be reasonable. In this case, it is not. We preferred to directly calculate the probability of being an outlier rather than use a binomial approximation and frame outliers in terms of standard deviations.

4.3.2. *Dr. Thornton's Fictitious Distribution.* The ensemble method begins with a clear set of criteria (compactness, equal population, etc. . .) and derives the distribution of the partisan seat count and other qualities of interest for each set of votes. This distribution takes into account the shape of North Carolina, the spatially nonuniform population and political preferences across the state, the effect of cities, and the correlations between different voting blocks in different elections. From this distribution, we can calculate directly the chance of being in the tails. In other words, I use the actual distribution of simulated plans in the ensemble to calculate statistical significance.

Dr. Thornton, however, ignores the actual distribution of simulated plans I generated, and constructs her own fake distribution. It is unclear why she does not use the actual distribution she is evaluating to calculate the extremal probabilities. The very purpose of generating this ensemble was to test for the extremity of the enacted plan, all the necessary information was given in the report. Instead, Dr. Thornton assumes that the expected distribution of elected officials should be a binomial, stating that the average chance of a Democrat winning any of the 120 seats in the house is 45%. She then assumes that Democrats have the exact same 45% chance of winning every one of the 120 House districts. With this one number from our analysis, namely 45%, she claims that this binomial is relevant to critiquing our analysis. In doing so she treats all districts the same, whether they are in Wake, Durham or Davidson county. She assumes there is no geographical structure in the state and thus assumes that a Democrat running in the heart of Wake county, has the same chance of winning as a Democrat running in Davidson county. This is plainly false. We cannot assume that each Democrat, no matter their location, has a 45% chance of winning their district. Additionally, she assumes that the outcome of each district race is probabilistically independent. The use of a binomial distribution necessarily neglects any considerations of geography and political differences across the state. Furthermore it is fundamentally built on the assumption that the outcome in any district does not influence or affect the outcome in another district. This assumption of independence and the assumption of spatial uniformity lead to grossly inaccurate seat distribution and a greatly inflated standard deviation estimate. This in turn leads her to inaccurately conclude that the deviations are not extreme outliers when in fact they are extreme outliers.

To make her overestimation of the standard deviation visually apparent, we examine the binomial distribution generated by Dr. Thornton for a number of elections and compare it to the true distribution of our ensemble. We display the distributions that result from the ensemble and Dr. Thornton's assumptions in Figure 6. Dr. Thornton's distributions are very different than the ones represented by the ensemble. In particular, the width of her distributions, and hence their standard deviations, are significantly bigger than the ensemble's true distribution. This is not surprising since Dr. Thornton's distribution is built on many false assumptions and little information about the state. In particular, we see that her distribution greatly exaggerates the probability of being far from the median (or mean) of the distribution. Hence it fails to identify even the most extreme outliers. Since her distribution was constructed with little to no input about the state, the fact that it has little resemblance to the distribution calculated from the ensemble is hardly surprising.

For all of these reasons, the conclusions that Dr. Thornton draws about the enacted plan not being a statistical outlier are built on an incorrect assumption. More specifically, on page 6 Dr. Thornton reports that the Democratic seats estimated from the simulations is not statistically significantly different from the number that we calculate from the enacted map. In paragraph 68, she goes on to report that results are statistically similar when the difference of a given plan is less than two or three standard deviations from an underlying distribution. Taking a given number of standard deviations is based on the idea of a normal distribution under which two or three standard deviations would mean that 97.7% or 99.86% of plans in the ensemble are closer to the ensembles mean or median than the enacted plan.

Based on this definition, we could then calculate how many plans in the ensemble are further from the median than the enacted plan under each set of voting data. We have already done this analysis and I have presented it in the tables of Figures 3 and 6 of my report. In the North Carolina Senate, the enacted plan is more extreme than 97.7% of the plans in the ensemble in 12 of the 17 elections, and more extreme than 99.86% of the plans in the ensemble in 8 of the 17 elections[36]. In the North Carolina House, the enacted plan is more extreme than 97.7% of the plans in the ensemble in 9 of the 17 elections, and more extreme than 99.86% of plans in the ensemble in 4 of the 17 elections[37]. Furthermore we have demonstrated that both plans are extreme outliers, falling well outside of the two or three standard deviations, under a number of metrics other than only the number of elected Democrats.

The ensemble methods directly test her assumption and prove it to not only be incorrect, but also prove that the conclusions drawn from her analysis are also incorrect. In the end, Dr. Thornton's erroneous analysis is a very compelling argument for why one needs to use the ensemble methods we employ. If something as simplistic as what she proposes was in any way relevant, we would have happily used it. In fact our desire to faithfully represent the state's complex geopolitical landscape pushed us to use the ensemble method in the first place.

4.4. **Additional notes on False and Misleading Statements.** In addition to the above analysis and retorts, there are many factually incorrect statements regarding my report. I have collected a few below:

(1) In paragraph 61, in discussing the figure on page 16 of my report involving the Mecklenburg Senate Cluster, Dr. Thornton claims that "The enacted plan is either more Democratic or is within the range of possibilities given by

[36] See Corrected Figure 3 in appendix of this report which is the corrected version of Figure 3 from the original report

[37] See Figure 6 from the original report

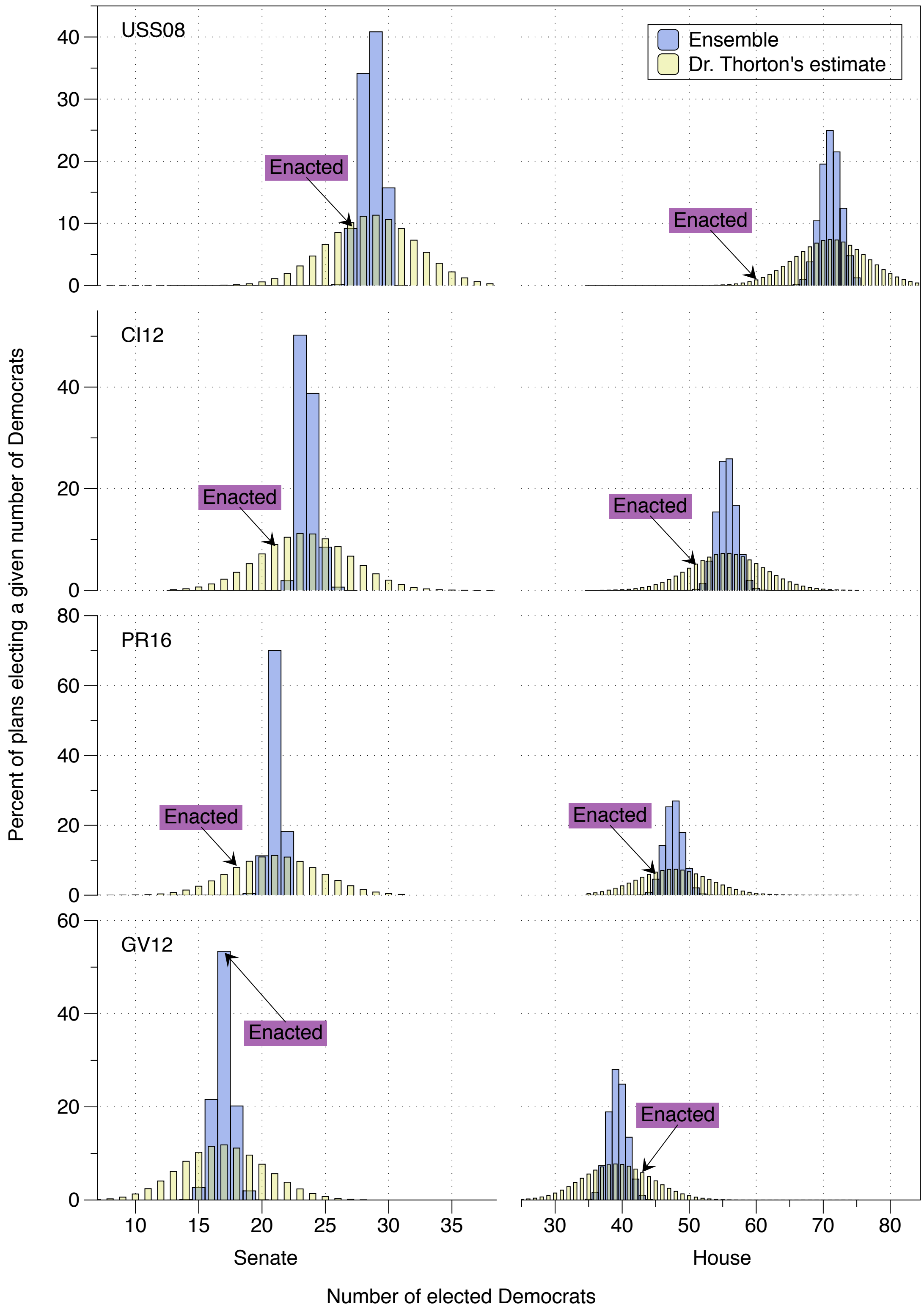


FIGURE 6. We compare Dr. Thornton's distribution with the distribution determined by the ensemble. We plot results for the North Carolina Senate (left) and House (right). We examine a range of elections which have a wide variety of statewide vote fractions; the elections considered are the 2008 United States Senate (USS08, top), 2012 Commissioner of Insurance (CI12, second from top), 2016 President (PR16, second from bottom), and the 2012 Governor (GV12, bottom). In all cases we find that Dr. Thornton's method give a variance significantly larger than that of the ensemble. In some cases, the enacted plan is an extreme outlier according to the ensemble analysis, but not according to Dr. Thornton's methodology.

the ... ensemble. This appears to be the case in nearly every instance." The figure on p. 16 in fact shows that the enacted plan is far in the tails of the marginal plots. It shows that three districts have significantly more Democrats packed in them than is almost ever seen in the ensemble, while two districts have exceptionally few democrats. This packing and cracking is repeated in many of the cluster level plots. The use of "range" by Dr. Thornton is an attempt to make the behavior seem typical when just the opposite is true. Furthermore, these plots only show marginal distributions, which is a more forgiving perspective when characterizing outliers. The paragraph in my report that precedes this figure shows that when looking at the districts in their totality, **none** of the maps in the ensemble have so few Democrats in the two most Republican districts and as many Democrats in the three most Democratic districts.

(2) In paragraph 64, "In this matter, we have an unknown pool of potential Republicans and Democrat legislators to be chosen based on an approximation of the percent of the voters who would vote Democratic and, thereby, presumably elect a Democratic legislator." Dr. Thornton seems to be advocating proportional representation and implying that we do also. We do not. We expressly only consider how the accepted, nonpartisan redistricting criteria are express through their interaction with geopolitical structure of the state.

(3) In paragraph 65, Dr. Thornton claims I have used uniform swing adjustments. I did not in my initial report, although I do use uniform swing adjustments in my rebuttal to respond to Dr. Brunell.

(4) In paragraph 67 and 68, Dr. Thornton's coin flip analysis is built on the idea of the outcome of an election in one district being unrelated to the outcome in an other district as is the case with successive coin flips. This misses many factors, including the idea that removing voters to pack one district dilutes them from another district.

(5) In paragraph 87, Dr. Thornton claims I have calculated the average number of Democratic seats across all 17 elections. While the data I have provided may be used to calculated this number, I have not done so in my report and have instead focused either on voting data from individual historical elections or on aggregate numbers that focus solely on deviations from the ensemble of plans.

5. In response to Dr. Lewis's report

Based on the legislature's redistricting criteria, we did not consider race when sampling the space of redistricting plans. The defense's expert, Dr. Lewis, has provided three analyses of selected clusters to estimate the percentage of Black citizens required in a district that will ensure a black candidate be elected. The first two analyses are based on a linear regression model which estimates voting outcomes based on demographic data; the linear model is then employed on primary elections and general elections (Tables 2 and 3 of Dr. Lewis's report, respectively). The third analysis makes several assumptions about how black voters and Democratic voters will support a minority candidate. The results of this analysis lead to a large and uncertain range of the fraction of black democratic voters needed to elect a minority candidate. For example, in Guilford county, Dr. Lewis projects that the percentage of black citizens within a district would need to be anywhere between 5% and 41% depending on the election considered.

For each of Dr. Lewis's results we first determine if the counties in the result belong to a House cluster, Senate cluster, or both. If the counties belong to a subset of a cluster, we only consider districts in a plan that strictly lie within the subset. We omit county groups that have no districts contained within them. We then ask how many districts in the enacted plan have a greater fraction of black citizen voting age population than the Dr. Lewis's reported threshold. We then count the number of plans in the ensemble that have as many (or more) districts above the threshold than the enacted plan.

We extend Tables 2 through 4 of Dr. Lewis's report for our primary ensemble in Tables 3, 4, and 5. For the majority of Dr. Lewis's results, we find that all, or nearly all, plans from our ensemble have as many districts above Lewis's reported thresholds as the enacted plan. In many of the other results from Dr. Lewis, we find that a significant fraction of the ensemble plans have as many districts above Lewis's reported thresholds as the enacted plan.

To analyze the effect of enforcing a minimum fraction of black citizen voting age population (BCVAP) on our ensembles, we consider Table 4 of Dr. Lewis's report. For the Forsyth-Yadkin and Lenoir-Pitt clusters in the House, the Plaintiffs counsel has asked me to analyze the partisan characteristics of the subset of plans in the ensemble that produce at least as many districts as the enacted plan above Dr. Lewis Black CVAP thresholds from his Table 4. We visualize these results in Figures 7 and 8, and find almost no differences in the ensembles whether we consider the thresholds proposed by Dr. Lewis or not.

		House		Senate	
Counties; Election (Year)	Black CVAP Needed (%) from Lewis	No. above BCVAP in Enacted from Lewis	% of Ensemble with the same or more above than Enacted	No. above BCVAP in Enacted from Lewis	% of Ensemble with the same or more above than Enacted
Alamance-Guilford-Randolph; Attorney General (2016)	22	-	-	2	100.0
Bladen-Greene-Harnett-Johnston-Lee-Sampson-Wayne; Attorney General (2016)	21	5	99.92	-	-
Columbus-Pender-Robeson; Attorney General (2016)	4	3	100.0	-	-
Cumberland; Attorney General (2016)	13	4	100.0	-	-
Cumberland; Commissioner of Labor (2016)	65	0	100.0	-	-
Davie-Forsyth; Attorney General (2016)	42	-	-	1	0.05
Forsyth; Attorney General (2016)	44	1	59.01	0	100.0
Forsyth; Commissioner of Labor (2016)	5	4	100.0	1	100.0
Forsyth-Yadkin; Attorney General (2016)	42	1	71.21	-	-
Franklin-Nash; Lieutenant Governor (2016)	12	2	100.0	-	-
Guilford; Attorney General (2016)	26	4	100.0	1	100.0
Guilford; Commissioner of Labor (2016)	5	6	100.0	2	100.0
Guilford; Sheriff (2014)	23	5	99.93	1	100.0
Guilford; Sheriff (2018)	30	4	100.0	1	100.0
Lenoir-Pitt; Attorney General (2016)	18	3	100.0	-	-
Nash; Sheriff (2014)	30	1	100.0	-	-
Person-Granville-Vance-Warren; Attorney General (2016)	34	1	100.0	-	-
Robeson; Sheriff (2018)	36	0	100.0	-	-

TABLE 3. We contextualize Table 2 of Dr. Lewis's report with our primary ensembles in both the North Carolina House and Senate. We examine how many districts in the enacted plan both lie within the specified counties and have a greater black CVAP than reported by Dr. Lewis for a particular election (and year). We then examine the fraction of plans in the ensemble that have the same number, or more, districts that are also above the black CVAP requirement within the specified counties.

		House		Senate	
Counties; Election (Year)	Black CVAP Needed (%) from Lewis	No. above BCVAP in Enacted from Lewis	% of Ensemble with the same or more above than Enacted	No. above BCVAP in Enacted from Lewis	% of Ensemble with the same or more above than Enacted
Guilford; Sheriff (2014)	43	2	8.88	1	100.0
Guilford; Sheriff (2018)	31	4	100.0	1	100.0
Nash; Sheriff (2014)	54	0	100.0	-	-
Pitt; Sheriff (2018)	28	1	100.0	-	-

TABLE 4. We contextualize Table 3 of Dr. Lewis's report with our primary ensembles in both the North Carolina House and Senate. We examine how many districts in the enacted plan both lie within the specified counties and have a greater black CVAP than reported by Dr. Lewis for a particular election (and year). We then examine the fraction of plans in the ensemble that have the same number, or more, districts that are also above the black CVAP requirement within the specified counties.

		House		Senate	
Counties	Black CVAP Needed from Lewis (%)	No. above BCVAP in Enacted	% of Ensemble with the same or more above than Enacted	No. above BCVAP in Enacted from Lewis	% of Ensemble with the same or more above than Enacted
Alamance-Guilford-Randolph	41	-	-	1	100.0
Bladen-Greene-Harnett-Johnston-Lee-Sampson-Wayne	41	1	100.0	-	-
Columbus-Pender-Robeson	26	2	16.39	-	-
Cumberland	34	2	100.0	-	-
Davie-Forsyth	41	-	-	1	1.05
Forsyth	40	1	81.58	1	6.65
Forsyth-Yadkin	41	1	76.62	-	-
Franklin-Nash	40	1	67.13	-	-
Guilford	40	4	96.67	1	100.0
Lenoir-Pitt	38	2	77.84	-	-
Nash	41	1	61.65	-	-
Person-Granville-Vance-Warren	32	1	100.0	-	-

TABLE 5. We contextualize Table 4 of Dr. Lewis's report with our primary ensembles in both the North Carolina House and Senate. We examine how many districts in the enacted plan both lie within the specified counties and have a greater black CVAP than reported by Dr. Lewis for a particular election (and year). We then examine the fraction of plans in the ensemble that have the same number, or more, districts that are also above the black CVAP requirement within the specified counties.

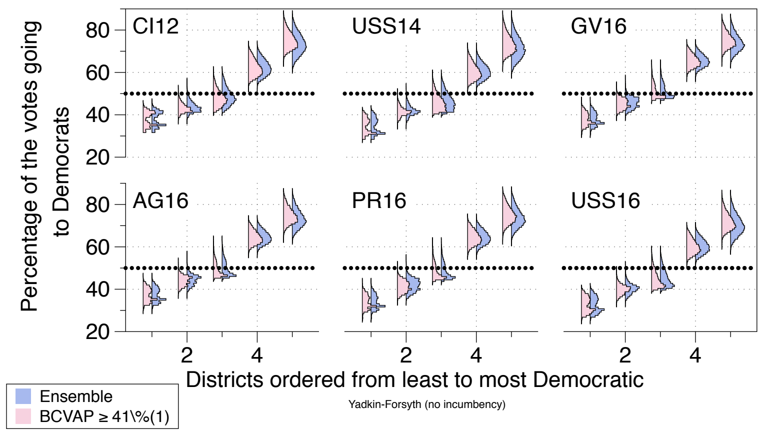


FIGURE 7. We display the original ensemble for Yadkin-Forsyth and compare it with the plans in the ensemble that have at least one district with at least 41% black citizen voting age population.

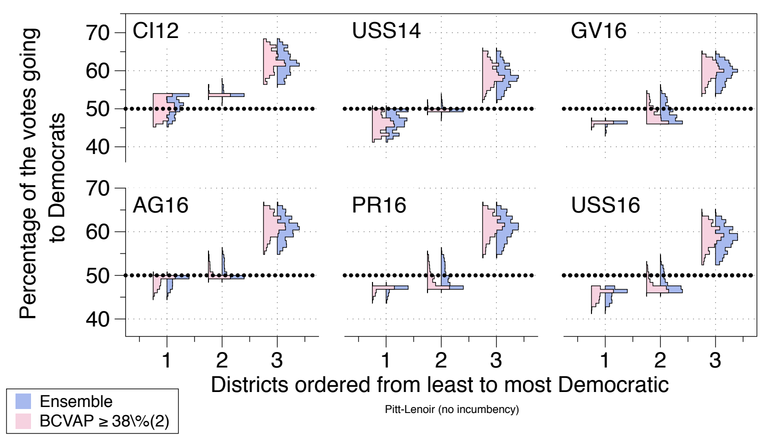


FIGURE 8. We display the original ensemble for Pitt-Lenoir and compare it with the plans in the ensemble that have at least two districts with at least 38% black citizen voting age population.

APPENDIX A. SAMPLING WITHOUT MUNICIPAL CRITERIA

In paragraph 54 of her report, Dr. Thornton suggests that municipalities do not have to be considered when redistricting. To investigate the effect that *not* considering municipalities may have on our ensembles, we have generated new ensembles on the plaintiff challenged clusters that do not consider municipal boundaries. We sample in the same way as our primary ensemble, however we set the municipal weight on the score function to be zero. The remaining weights for our score functions are presented in Tables 6 and 7 for the Senate and House, respectively. We present the convergence results for each cluster ensemble in Tables 8 and 9.

APPENDIX B. EXAMINING REOCK SCORES

We examine the statewide Reock scores and compare the Reock scores of the enacted plan with that of the ensemble. We plot the ranked marginal distributions of Reock scores for the North Carolina Senate in Figure 9 and for the North Carolina House in Figure 10. For the majority of the districts, we find that the ensemble is either comparable, or is significantly more compact than the enacted plan.

Cluster	$(w_{\text{pop}}, w_{\text{PP}})$	$(t_{\text{pop}}, t_{\text{PP}})$
Bladen-Brunswick-Pender-New Hanover	(50.0, 0.1)	(5.0, 0.179)
Buncombe-Transylvania-Henderson	(100.0, 0.4)	(5.0, 0.179)
Davie-Forsyth	(100.0, 0.15)	(5.0, 0.179)
Guilford-Alamance-Randolph	(60.0, 0.2)	(5.0, 0.179)
Lee-Sampson-Harnett-Duplin-Johnston-Nash	(100.0, 0.2)	(5.0, 0.179)
Mecklenburg	(100.0, 0.2)	(5.0, 0.179)
Wake-Franklin	(50.0, 0.3)	(5.0, 0.179)

TABLE 6. We present the weights used when re-sampling the plaintiff selected North Carolina Senate clusters without considering municipal boundaries (i.e. $w_{\text{M}} = 0$). The Table is similar to Table 6 in my original report.

Cluster	$(w_{\text{pop}}, w_{\text{PP}})$	$(t_{\text{pop}}, t_{\text{PP}})$
Alamance	(80.0, 0.1)	(5.0, 0.179)
Buncombe	(50.0, 0.2)	(5.0, 0.179)
Columbus-Pender-Robeson	(50.0, 0.1)	(5.0, 0.147)
Cumberland	(50.0, 0.2)	(5.0, 0.179)
Davie-Montgomery-Richmond-Cabarrus-Rowan-Stanly	(100.0, 0.2)	(5.0, 0.167)
Duplin-Onslow	(50.0, 0.2)	(5.0, 0.179)
Gaston-Cleveland	(50.0, 0.2)	(5.0, 0.179)
Guilford	(100.0, 0.15)	(5.0, 0.167)
Mecklenburg	(50.0, 0.2)	(10.0, 0.179)
Nash-Franklin	(50.0, 0.2)	(5.0, 0.179)
New Hanover-Brunswick	(100.0, 0.1)	(5.0, 0.179)
Person-Vance-Granville-Warren	(100.0, 0.2)	(5.0, 0.179)
Pitt-Lenoir	(100.0, 0.2)	(5.0, 0.179)
Union-Anson	(100.0, 0.2)	(5.0, 0.179)
Wake	(100.0, 0.2)	(12.0, 0.179)
Yadkin-Forsyth	(50.0, 0.1)	(5.0, 0.157)

TABLE 7. We present the weights used when re-sampling the plaintiff selected North Carolina House clusters without considering municipal boundaries (i.e. $w_{\text{M}} = 0$). The Table is similar to Table 7 in my original report.

	PR12; Max Err(%))		PR16; Max Err(%)			
Cluster	means	st. dev.	means	st. dev.	Avg. No Plans	Used Plans
Bladen-Brunswick-Pender-New Hanover	0.389	0.212	0.481	0.255	182681.0	50000.0
Buncombe-Transylvania-Henderson	0.131	0.074	0.119	0.067	3470.4	3479.0
Davie-Forsyth	0.195	0.111	0.219	0.124	19017.2	19082.0
Guilford-Alamance-Randolph	0.082	0.047	0.064	0.037	12484.6	12511.0
Lee-Sampson-Harnett-Duplin-Johnston-Nash	0.161	0.088	0.392	0.200	364.8	360.0
Mecklenburg	0.237	0.136	0.244	0.139	6412.2	6390.0
Wake-Franklin	0.141	0.082	0.145	0.084	50124.0	50361.0

TABLE 8. We present convergence data when re-sampling the plaintiff selected North Carolina Senate clusters without considering municipal boundaries (i.e. $w_{\text{M}} = 0$). The Table is similar to Table 10 in my original report.

Cluster	PR12; Max Err(%)) means	st. dev.	PR16; Max Err(%) means	st. dev.	Avg. No Plans	Used Plans
Alamance	0.263	0.150	0.220	0.125	2790.0	2826.0
Buncombe	0.145	0.083	0.211	0.121	12522.4	12431.0
Columbus-Pender-Robeson	0.148	0.085	0.174	0.099	7581.4	7598.0
Cumberland	0.230	0.131	0.269	0.153	5890.2	5869.0
Davie-Montgomery-Richmond-Cabarrus-Rowan-Stanly	0.096	0.055	0.111	0.064	13026.4	12950.0
Duplin-Onslow	0.227	0.129	0.194	0.110	5894.2	5883.0
Gaston-Cleveland	0.091	0.052	0.054	0.031	10744.4	10589.0
Guilford	0.136	0.078	0.110	0.063	9450.8	9423.0
Mecklenburg	0.228	0.131	0.253	0.144	16731.0	16652.0
Nash-Franklin	0.062	0.036	0.076	0.044	3193.0	3184.0
New Hanover-Brunswick	0.062	0.036	0.110	0.063	12142.2	12013.0
Person-Vance-Granville-Warren	0.048	0.028	0.056	0.032	26156.2	26091.0
Pitt-Lenoir	0.195	0.112	0.058	0.034	3155.8	3237.0
Union-Anson	0.410	0.231	0.358	0.202	7986.0	7915.0
Wake	0.194	0.112	0.140	0.081	50039.2	50000.0
Yadkin-Forsyth	0.288	0.164	0.325	0.185	6300.8	6326.0

TABLE 9. We present convergence data when re-sampling the plaintiff selected North Carolina House clusters without considering municipal boundaries (i.e. $w_{\mathrm{M}} = 0$). The Table is similar to Table 11 in my original report.

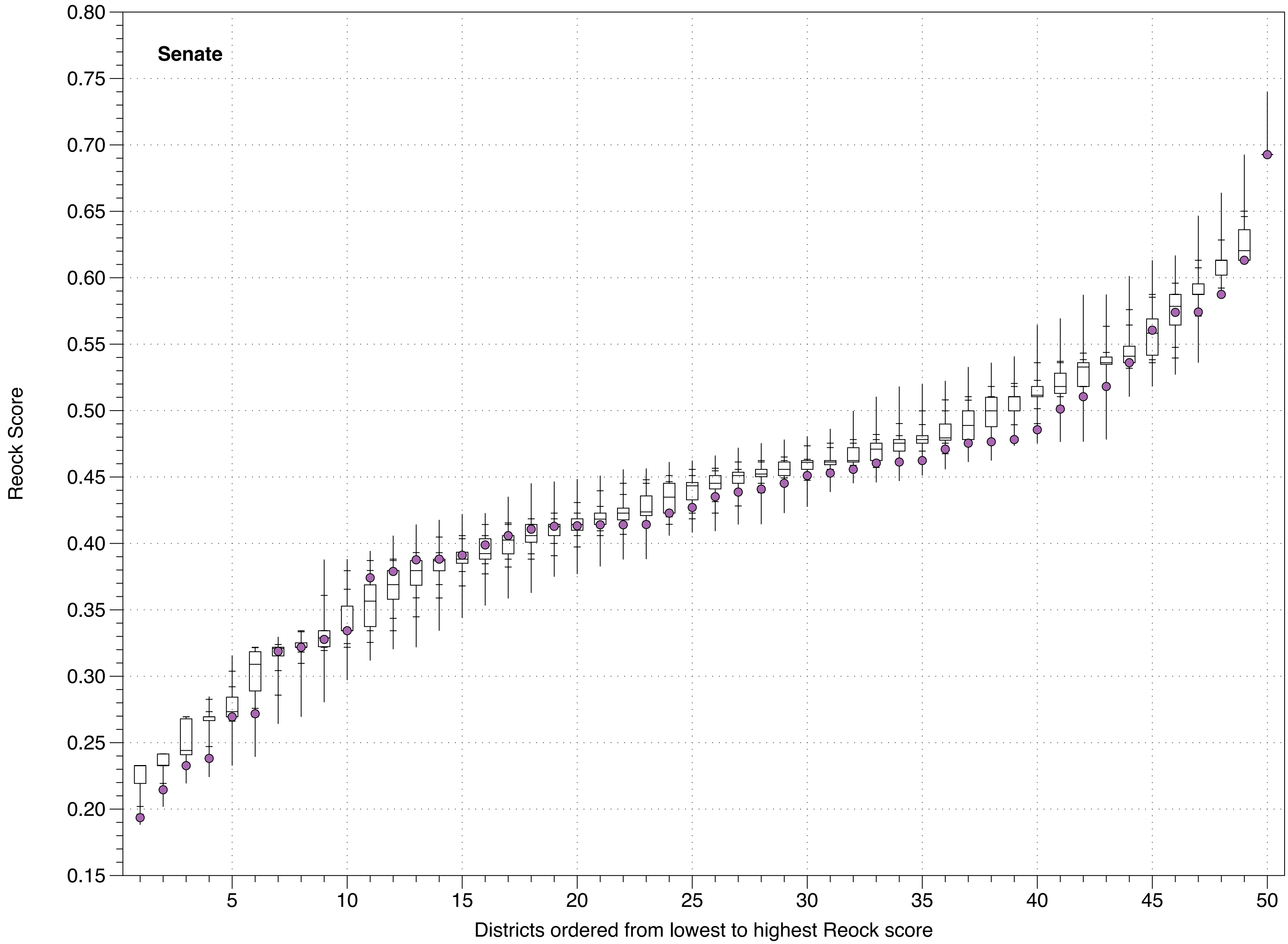


FIGURE 9. Dashes mark 2.5% and 10% outliers of the statewide ensemble. Boxes mark 50% range of the statewide ensemble. Dash in box marks median of the ensemble rank-marginal distributions. Purple circles indicate the enacted plan

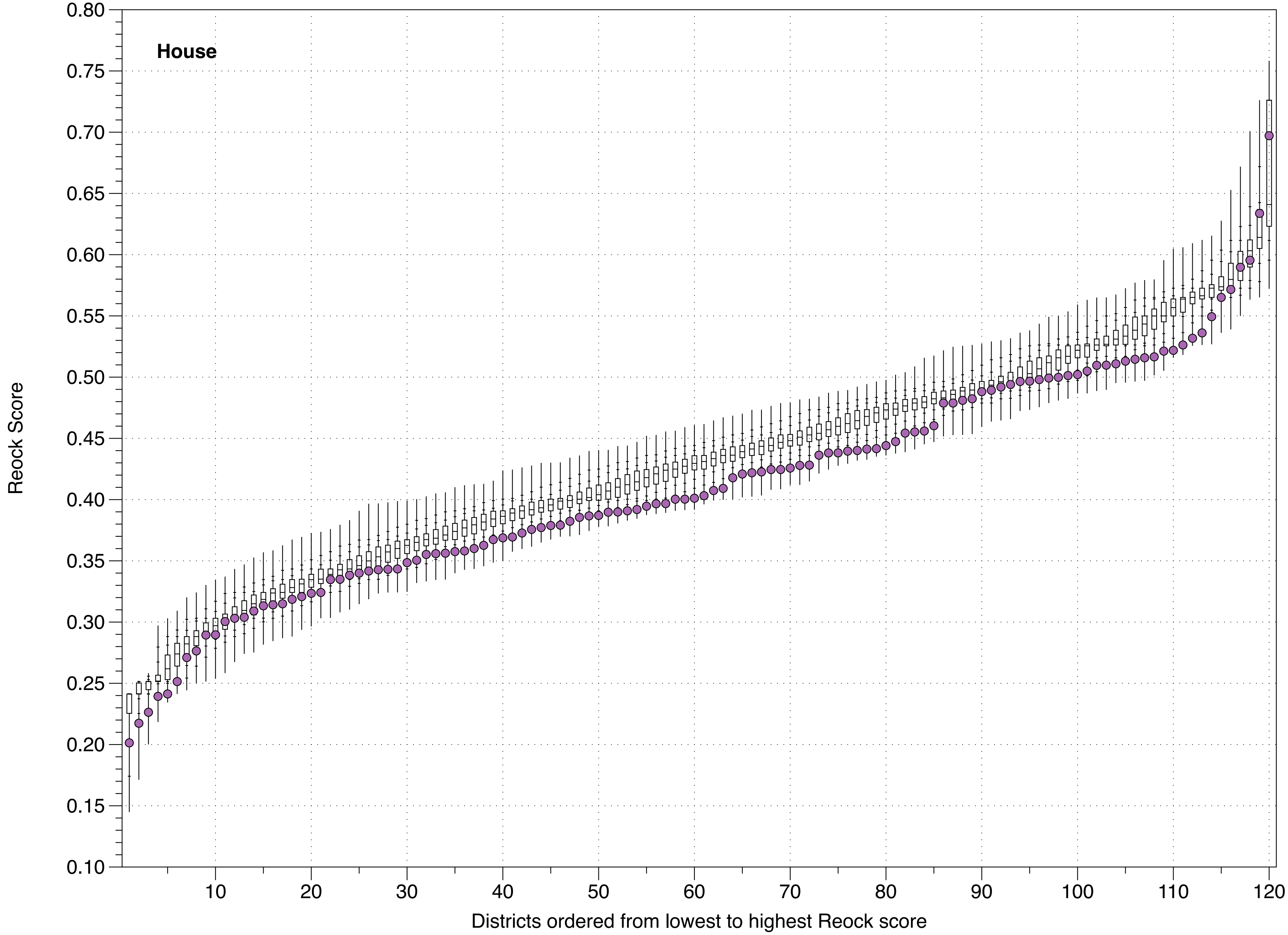


FIGURE 10. Dashes mark 2.5% and 10% outliers of the statewide ensemble. Boxes mark 50% range of the statewide ensemble. Dash in box marks median of the ensemble rank-marginal distributions. Purple circles indicate the enacted plan

APPENDIX C. ERRATA

We found a small error in one of the Tables of the original report. We correct this error in Table 10 of this report. We also correct the final paragraph on page 82 of the original report to read:

We test if these small differences affect our overall conclusions that the enacted plan is an extreme outlier in favor of the Republican Party. In the 17 considered elections, we count the number of elections in which the enacted plan leads to a number of Democratic seats that is an extreme outlier in favor of either the Republicans or the Democrats. We display the results in Table 10. In all cases we find a significant number of elections in which the enacted plan leads to more Republican seats than in either 99% or 99.9% of the ensemble; conversely, in nearly all cases, we find no elections in which the enacted plan leads to more Democrats than 99% of plans in the ensemble. The only exception is when considering unique plans in the House, in which we find that the votes of the 2012 Governor's race lead to more Democrats than over 99.4% of the plans in the ensemble; we note, however, that under these votes the Republicans have secured a supermajority in both the enacted plans and in the ensemble of plans. We further note that the number of outlying elections we see for each of the four new ensembles remains consistent with the number of outlying elections present in the original ensemble.

		No. elections atypically favoring Rep.		No. elections atypically favoring Dem.	
Chamber	Ensemble	Prob$< 1\%$	Prob$< 0.1\%$	Prob$< 1\%$	Prob$< 0.1\%$
Senate	Original	10	8 (was 9)	0	0
Senate	Unique	10	8	0	0
Senate	Incumb	11	8	0	0
Senate	Thresh	10	8	0	0
Senate	Incmb. & Thresh.	10	8	0	0
House	Original	7	4	0	0
House	Unique	7 (was 8)	5	1 (was 0)	0
House	Incumb.	7	6	0	0
House	Thresh.	7	5	0	0
House	Incmb. & Thresh.	10	6	0	0

TABLE 10. We found a small error in Table 9 of the original report. We correct the two entries above for the unique plans in the House (highlighted in red). The caption of the original table reads, "Of the 17 considered elections, we count the number of elections in which the enacted plan is a 1% outlier and a 0.1% outlier favoring either the Republicans or the Democrats. We find no election in which the Democrats elect more seats than expected to the extreme extent. We find a significant number of elections in which the Republicans elected more seats than expected to an extreme extent." To correct the caption, "Of the 17 considered elections, we count the number of elections in which the enacted plan is a 1% outlier and a 0.1% outlier favoring either the Republicans or the Democrats. Across all types of ensembles, we find only one election that leads to the enacted plan being a 99% outlier favoring the Democrats and no elections that lead to the enacted plan being a 99.9% outlier favoring the Democrats. The single election in which the Democrats are atypically favored is the 2012 Governor race in which the enacted plan leads to a Republican supermajority. In contrast, we find a significant number of elections in which the Republicans elected more seats than expected to an extreme extent."

C.1. **Corrections to the statewide Senate ensemble.** The Hoke-Cumberland cluster in the Senate was not held fixed in the way we had initially stated. In particular, when examining the statewide Senate results, we were instructed by plaintiffs lawyers to freeze the Hoke-Cumberland cluster, and we did not do this as we intended. Instead, in the section of our initial report concerning the statewide properties of the Senate when restricting to the plaintiff challenged clusters, we allowed Hoke-Cumberland to vary. We have now fixed this error. As shown below, the change to freezing Hoke-Cumberland leads to no changes in most of the results, and a small fluctuation in some of the results. Of these changes, most are so small as to be imperceptible. None of the changes that result from this error alter any of our conclusions. This error only affects the statewide analysis of the Senate (pp 4-9, Table 1, Table 3, and the Senate figures in Section G in the Appendix.). It does not affect any of the analysis of the House (pp 9-14), any statewide results that only examine the plaintiff selected clusters or the partisanship results of each cluster (pp 1445), the details of generating the ensemble (pp 45-54), any of our cluster-by-cluster results (pp 5485), nor our validations (pp 86-94).

Since the statewide Senate results depend on behavior in all the clusters, we have repeated those calculations. Having reanalyzed each result that is affected by freezing Hoke-Cumberland, we find only insignificant or minimal changes. Other than some small numerical fluctuations, these changes have not affected any substantive conclusion, or even the substance of any sentence, in our report. We have also replotted the figures which were affected by freezing Hoke-Cumberland. As stated, most changes are imperceptible to the eye. Those that did change did not affect any of our conclusions.

To demonstrate the sparsity of the overall effects of freezing Hoke-Cumberland, we reproduce the entirety of Section 3.1 from my original report that has been, in any way, affected and highlight changes to the text in blue. Although this reproduction involves multiple pages in this Errata, the actual changes to text are few in number, and the substance of each sentence is unchanged. The included figures have also changed, but the conclusions drawn by them have not.

Following the corrections to Section 3.1 of my original report, we reproduce the affected Appendix Tables 1, 3 from my report; we include the unchanged entries for ease of comparison and to demonstrate the limits of the effect. In Table 1, none of the quartiles positions change by more than one and, as seen by the updated Figure 3 and Section 3.1, do not affect any of our conclusions. Table 3 presents the corrected numerical data that is presented in the corrected Figure 2 which provides the same overall structure as the original figure. Finally, we present the new statewide Senate box plots from Section G. The differences are inconsequential as can be noted by how the results of there analysis does not change in the updated Section 3.1.[38]

In short, although we have included a number of changes in the interest of correctness, the effect of this errata is insignificant to any of the conclusions we have drawn.

C.1.1. *Within the section "3.1 North Carolina Senate".* ... we test if the enacted plan is atypical of the ensemble by determining the net number of elections for which a plan skews in favor of the Democratic or Republican Party.

To calculate this number, we select a plan in the ensemble and count the number of elections (out of the 17 historical elections considered) in which that plan elects more Democrats than the median number of elected Democrats in the ensemble; we then subtract the number of elections in which that particular plan elects fewer Democrats than the median number of elected Democrats in the ensemble. We then cycle through all of the elections under consideration. For example, in the enacted plan, 15 of the 17 elections lead to fewer elected Democrats than the median number of elected Democrats over the ensemble, and 0 of the 17 elections lead to more elected Democrats than the median number of elected Democrats: this would provide a value of $1 - 15 = -14$ (see the 'seat shift' column in the table of Figure 3). We repeat this procedure for each plan in the ensemble. By contextualizing the net median skew within the ensemble of plans, we measure partisan bias in a way that is adapted to the geographic structure in the votes across the state. The baseline for what is typical is set by the emergent ensemble of plans. To display this baseline, we plot the probability of seeing a certain net number of elections favoring each party and contextualize the enacted plan in this distribution (see the upper right figure in Figure 3).

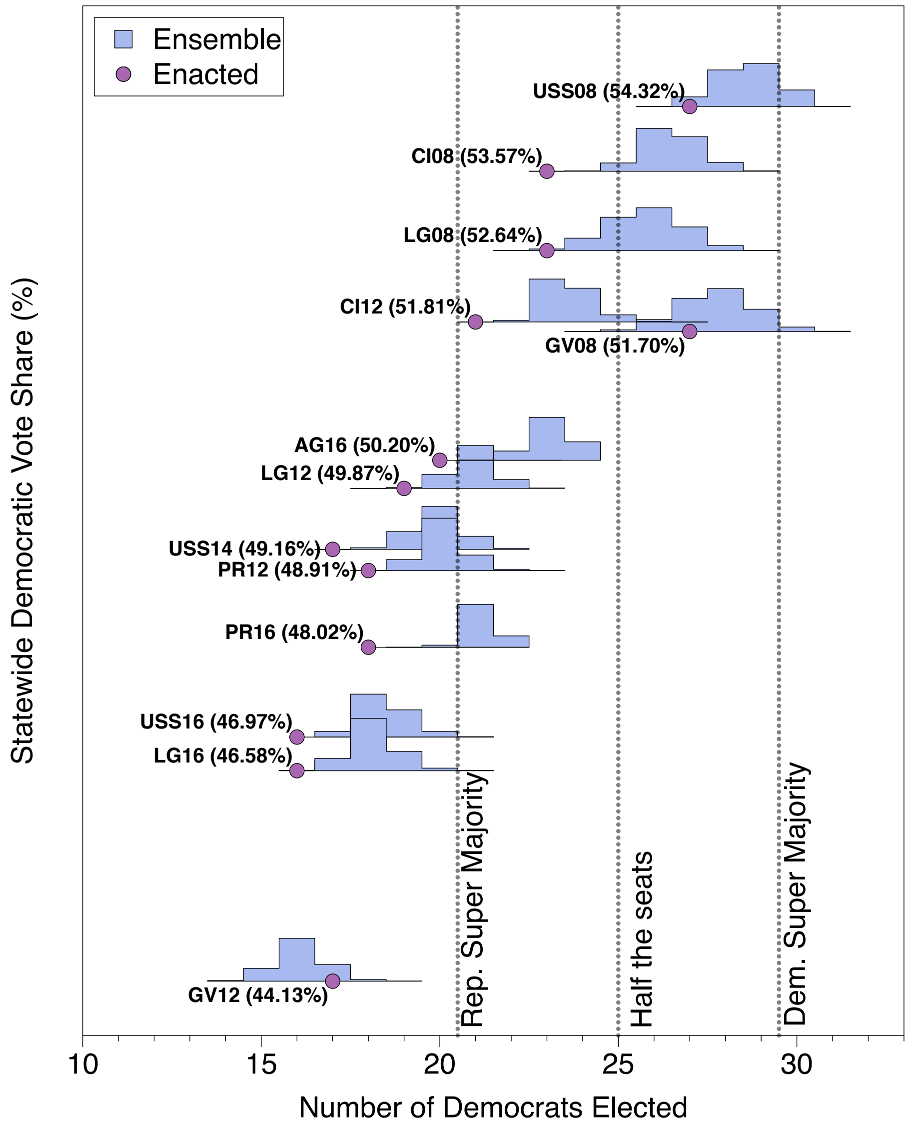
We find that less than 0.005% of plans in the ensemble have a net median skew of -14 or less, meaning that it is highly atypical to find a neutrally selected plan from the ensemble which favors the Republican Party as consistently as does the enacted plan.

Although the enacted plan favors the Republicans in an atypically large number of elections, it may be the case that degree of skew within any particular election is typical of the ensemble. To test this we ask how often plans in the ensemble are as far, or farther, away from the median number of elected Democrats in the ensemble of plans. Under many historic vote counts, the number of Democrats elected in the enacted plan is abnormally low with respect to the ensemble. In 12 of the 17 examined elections, less than 2% of the plans in the ensemble give a skew from the median that is the same or greater than is seen in the enacted plan. In 8 of the 17 examined elections (including 4 of the 5 elections in 2016), less than 0.04% of the plans in the ensemble give a skew from the median that is the same or greater than is seen in the enacted plan (see the '%' column in the table in Figure 3). In short, we find that the number of seats elected by the enacted plan is an extreme outlier with respect to the ensemble in the majority of considered elections.

The enacted plan, in contrast, deviates from the median number of elected Democrats, on average, by just under two seats in favor of the Republicans (see the 'seat shift' column and last row in the table of Figure 3). We find that this difference is greater than all plans in the ensemble (see the bottom right figure of Figure 3). The ensemble shows a typical average deviation of plus or minus half of one seat, and can range as high as plus or minus one seat, but we never see any plan that favors the Republican Party to the same extent with which the enacted plan does according to this metric. We see that in all five 2016 elections, the shift is 2 or 3 seats in favor of the Republican Party.

To understand the above results, it is not enough to understand *who* won the election, but we must also understand the *margin* of the partisan victory. We compare the typical margins of victory in the ensemble with the enacted plan by taking a similar approach to that taken previously (see [1,2,3]): for a given election and a given plan, we rank the districts from least to most Democratic. By examining this ranking over a number of plans in the ensemble, we examine the typical Democratic/Republican vote fractions in the most Democratic district, the second most Democratic district, etc... and similarly in

[38] See Corrected Figure 4 below and accompanying text; note that the results have not changed(as can be quickly gleaned from noting the lack of blue text).



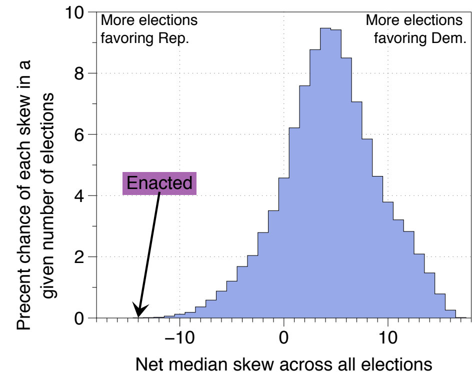
CORRECTED REPORT FIGURE 2. Each blue distribution represents the range of Democratic seats won in the ensemble of plans, given a set of historic votes; the height is the relative probability of observing the result. We only include a selection of the historic vote counts for clarity. Other vote counts either lead to qualitatively similar conclusions (e.g. USS10), or are well above the displayed range of statewide vote fractions (AG08). We label each distribution with an abbreviation for the votes used along with the Democratic statewide vote percentage. Abbreviations contain the year of the last two characters, and AG for Attorney General, USS for United States Senate, CI for Commissioner of Insurance, GV for Governor, LG for Lieutenant Governor, and PR for United States President.

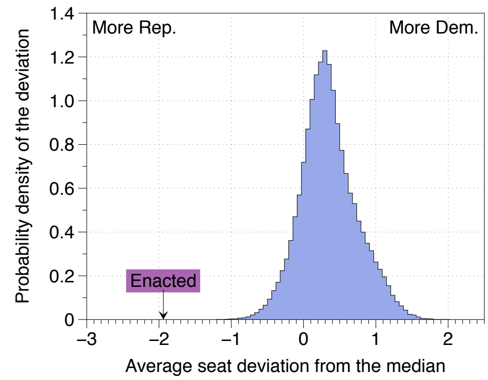
the most Republican district, the second most Republican district, and so on. We display the resulting ranked-vote marginal distributions for the middle 20 districts in two of the elections in Figure 4; we display the full ranked marginal distributions across all other elections in the appendix (see Section G).

The ranges of each ranked district are represented by box-plots in Figure 4. In these box-plots 50% of all plans have a corresponding ranked district that lies within the central box; the median is indicated by the line through the center of the box; the tick marks denote location of the 1% and 99% outliers; the extent of the lines outside of the boxes represent the full range of results observed in the ensemble (i.e. the maximum and minimums). We compare the ranked-votes curve of the enacted plan with the ranked-votes marginal distributions. There are 50 seats. Any dot (or box) that lies above the 50% line on the vertical axis will elect (or typically elect) a Democrat; any dot (or box) that lies below the 50% line will elect (or typically elect) a Republican.

The enacted plan has significantly fewer democratic votes in the middle districts than is typical of the ensemble. In the 2008 Commissioner of Insurance race, we see that the Democrats would typically win about 26 seats, as the (24th ranked marginal distribution straddles the 50% line), but that the enacted plan leads to 23 elected Democrats, with corresponding

Election	Median Dems. in Ensemble	Dems. elected in enacted plan	Seat shift	% of plans that are as far or farther from the median
USS10	15	15	0	-
GV12	16	17	D+1	24.6%
LG16	18	16	R+2	0.20%
USS16	18	16	R+2	0.013%
PR12	20	18	R+2	0.21%
LG12	21	19	R+2	1.28%
USS14	20	17	R+3	0.01%
PR16	21	18	R+3	0.0023%
PR08	22	19	R+3	0%
GV16	21	19	R+2	0%
CI12	23	21	R+2	0.01%
AG16	23	20	R+3	4.4e-3%
CI08	26	23	R+3	0.00%
LG08	26	23	R+3	1.49%
GV08	28	27	R+1	39.8%
USS08	29	27	R+2	9.32%
AG08	40	39	R+1	40.1%
Average	-	-	R+1.94	-



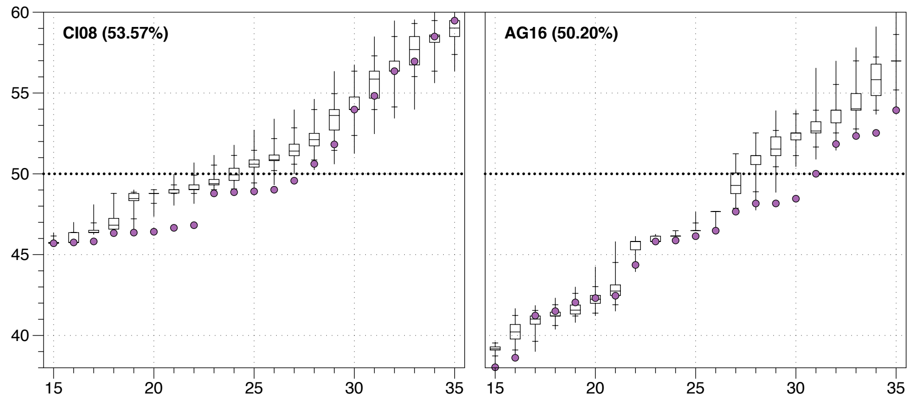


CORRECTED REPORT FIGURE 3 In the table (left), we list the median number of elected Democrats over the plans in the Senate ensemble for each considered election. We then list the number of Democrats that would have been elected by the enacted plan for each election. We calculate the difference between the median number of Democrats and the Democrats that would have been elected by the enacted plan, and then display the chance that a random plan from the ensemble would be as far or farther than the enacted plan is away from the median. We then plot the net median skew (out of 17 elections) that give a favorable result to the Democrats versus those that give a favorable result to the Republicans; a skewed result is one in which a party wins more than the median number of seats (top right; the number of skewed elections for the enacted plan is 15 for the Republicans and 1 for the Democrats over the 17 considered elections). Next, we show the amount the plans in the ensemble deviate from the median number of seats in the ensemble, averaged over the set of historic vote counts. In both cases, we find that the enacted plan is an extreme outlier when compared to the ensemble of plans.

ranked districts having far fewer Democratic votes than is typical. When the statewide vote share shifts toward the Republicans, as in the 2016 Attorney General election, the Democrats tend to win fewer seats, both in the ensemble and in the enacted plan, however the depression in the middle districts remains fairly consistent, as does the disparity between the expected number of seats going to the Democrats and the number of Democratic seats gained in the enacted plan.

To test the observation that the middle districts have an abnormally low Democratic vote fraction, we consider the 20th through 30th least Democratic districts and consider the average vote share in this range. We then compare average vote share in the ensemble with the average vote share in the enacted plan. We choose the ranges at 20 and 30 because this is where power shifts to give a supermajority for either party. In 14 of the 17 elections, we see that fewer than 0.0005% of the plans in the ensemble have fewer average Democrats in these middle districts than the enacted plan; in the AG08 and LG08 elections, we find that fewer than 0.1% of the plans in the ensemble have fewer Democrats in the middle plans than the enacted plan. There is only one anomalous election —GV08 —in which roughly 80% of all plans in the ensemble have more Democrats in the middle districts than in the enacted plan (see Section G).

The above differences between the enacted plan and the ensemble can affect majority representation. In the 2008 Commissioner of Insurance election the Democrats are expected to gain a majority of the seats in over 90% of the plans in the ensemble; in the 2008 Lieutenant Governor race the Democrats are expected to gain a majority of the seats in over 59% of the plans; yet in both sets of election data, the Republicans win a majority of the seats under the enacted plan. In contrast, we find no elections under which the Republicans would have been expected to receive a majority under the ensemble, but would not receive a majority in the enacted plan.



CORRECTED REPORT FIGURE 4 Over two elections, we plot the typical range of the 15th least Democratic district to the 35th least democratic district. The ranges are represented by box-plots: 50% of all plans in the ensemble have a corresponding ranked district that lies within the box; the median is given by the line within the box; the ticks mark the 1% and 99% quartiles; the extent of the lines outside of the boxes represent the range of results observed in the ensemble. We compare the ranked-votes curve of the enacted plan (purple dots) with the ranked-votes marginal distributions (box plot). There are 50 seats; any dot (or box) that lies above the 50% line on the vertical axis will elect (or typically elect) a Democrat; any dot (or box) that lies below the 50% line will elect (or typically elect) a Republican.

The differences in the ensemble and the enacted plan also can affect the supermajority. Under one of the elections (LG12) the ensemble of plans yields a Republican supermajority in less than 30% of the plans, yet the enacted plan provides a Republican supermajority with 19 and 17 Democratic seats, respectively. In PR08, the Republicans do not gain a supermajority in over 98.9% of plans in the ensemble, whereas they do gain a supermajority under the enacted plan. In two other elections (AG16, GV16, ~~PR16~~) the Republicans do not gain a supermajority in over 99.2% of the plans in the ensemble, whereas they gain a supermajority in each of these elections under the enacted plan.

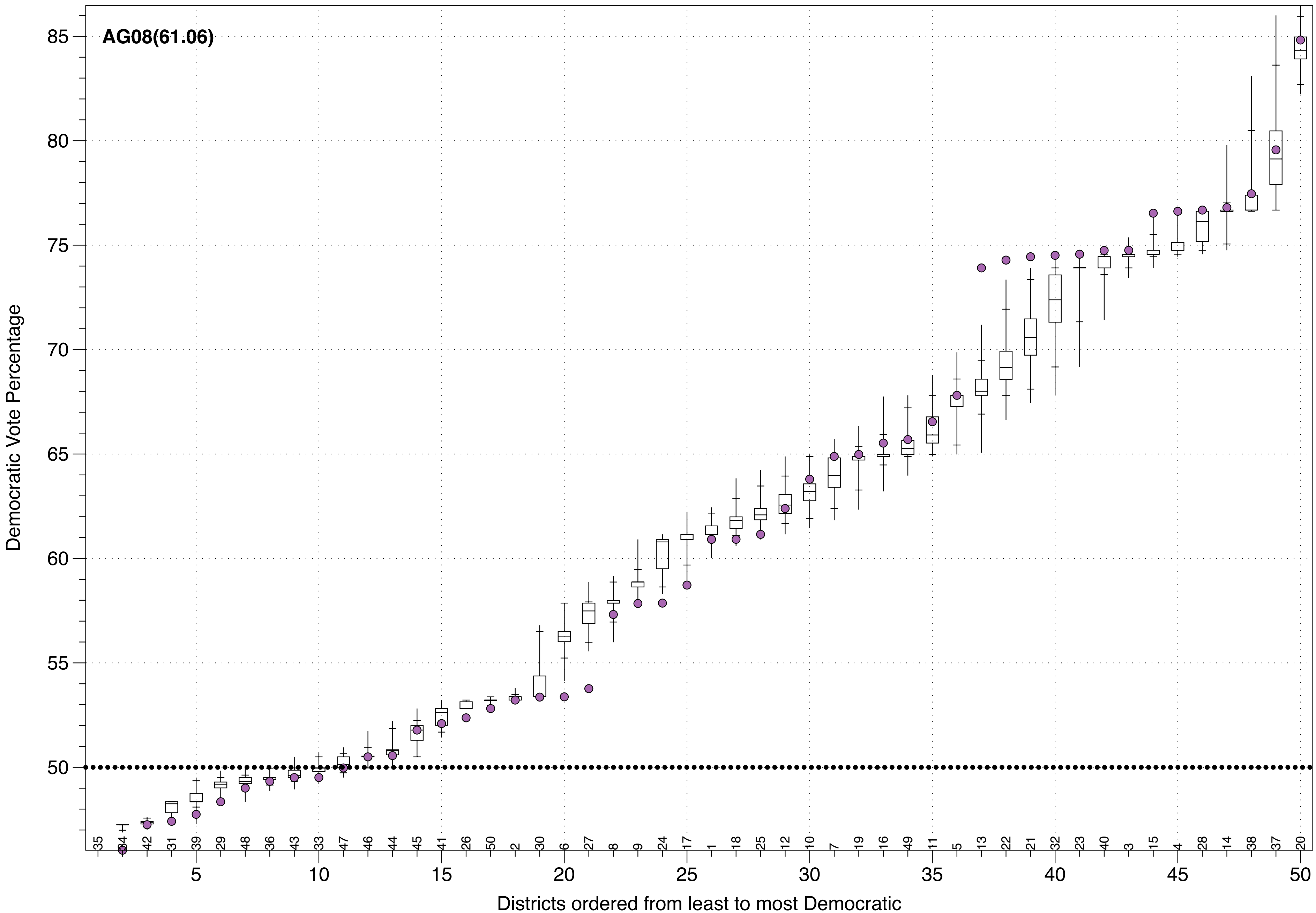
C.1.2. *Other corrected Figures and Tables.*

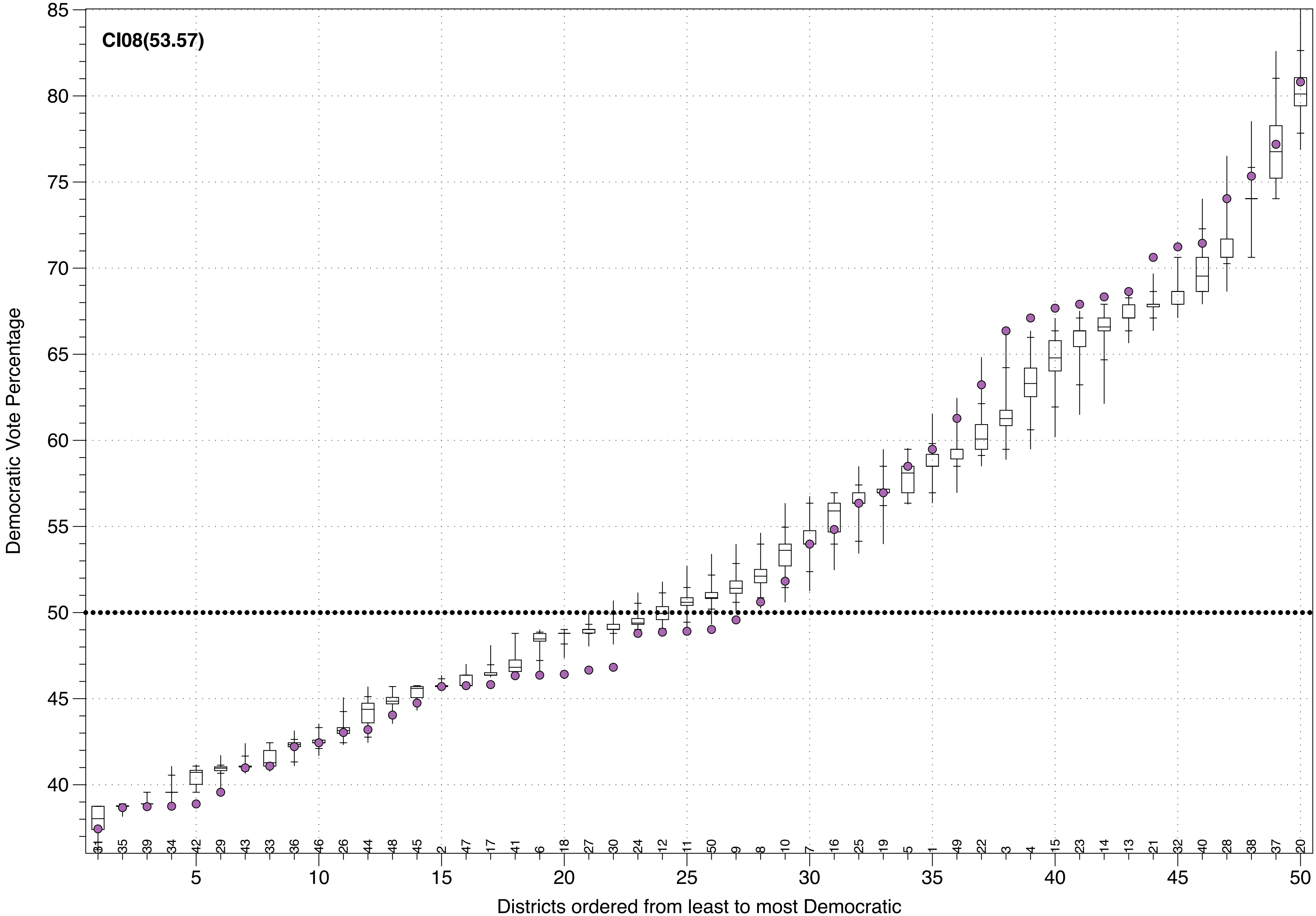
			Less than $P\%$ of plans have fewer than S seats			Less than $P\%$ of plans have more than S seats		
Elec.	Enacted	Ensemble Median	S ($P = 1\%$)	S ($P = 5\%$)	S ($P = 25\%$)	S ($P = 25\%$)	S ($P = 5\%$)	S ($P = 1\%$)
USS10	15	15	14	15	15	16	16	17
GV12	17	16	15	16	16	16	17	18
LG16	16	18	17	18	18	19	19	20
USS16	16	18	18	18	18	19	20	20
PR12	18	20	19	20	20	20	21	22
LG12	19	21	20	21	21	21	22	22
USS14	17	20	19	20	20	20	21	22
PR16	18	21	21	21	21	21	22	22
PR08	19	22	21	22	22	23	23	24
GV16	19	21	21	21	21	22	22	22
CI12	21	23	23	23	23	24	25	25
AG16	20	23	22	23	23	24	24	24
CI08	23	26	26	26	26	27	28	28
LG08	23	26	24	26	25	26	27	28
GV08	27	28	26	28	27	28	29	30
USS08	27	29	28	29	28	29	30	30
AG08	39	40	39	40	39	40	41	41

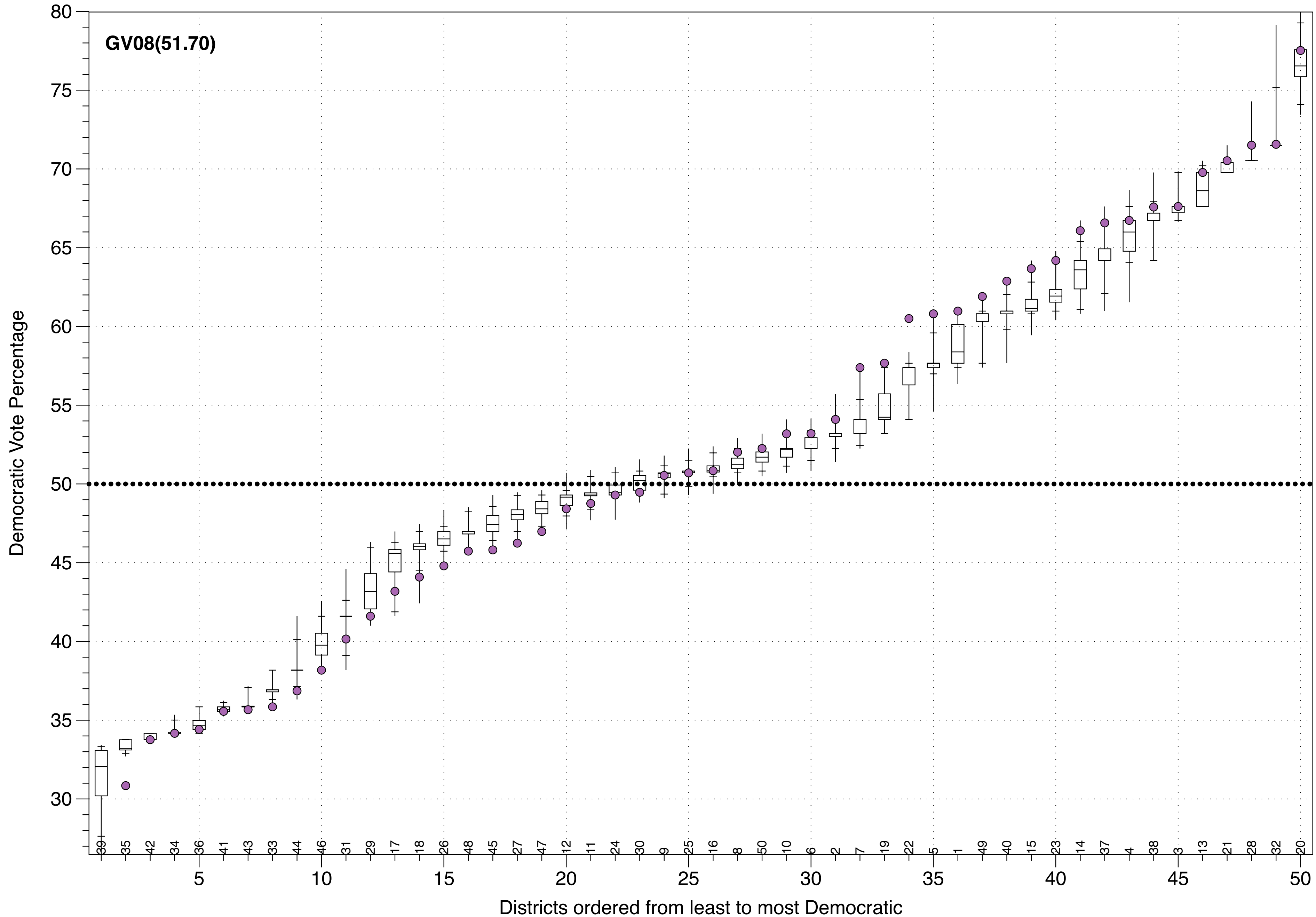
CORRECTED REPORT TABLE 1. We summarize the range of expected partisan results in the enacted plan and the ensemble for the North Carolina Senate. The results are from our primary ensemble that respects municipalities, population, compactness and whole county/traversal criteria.

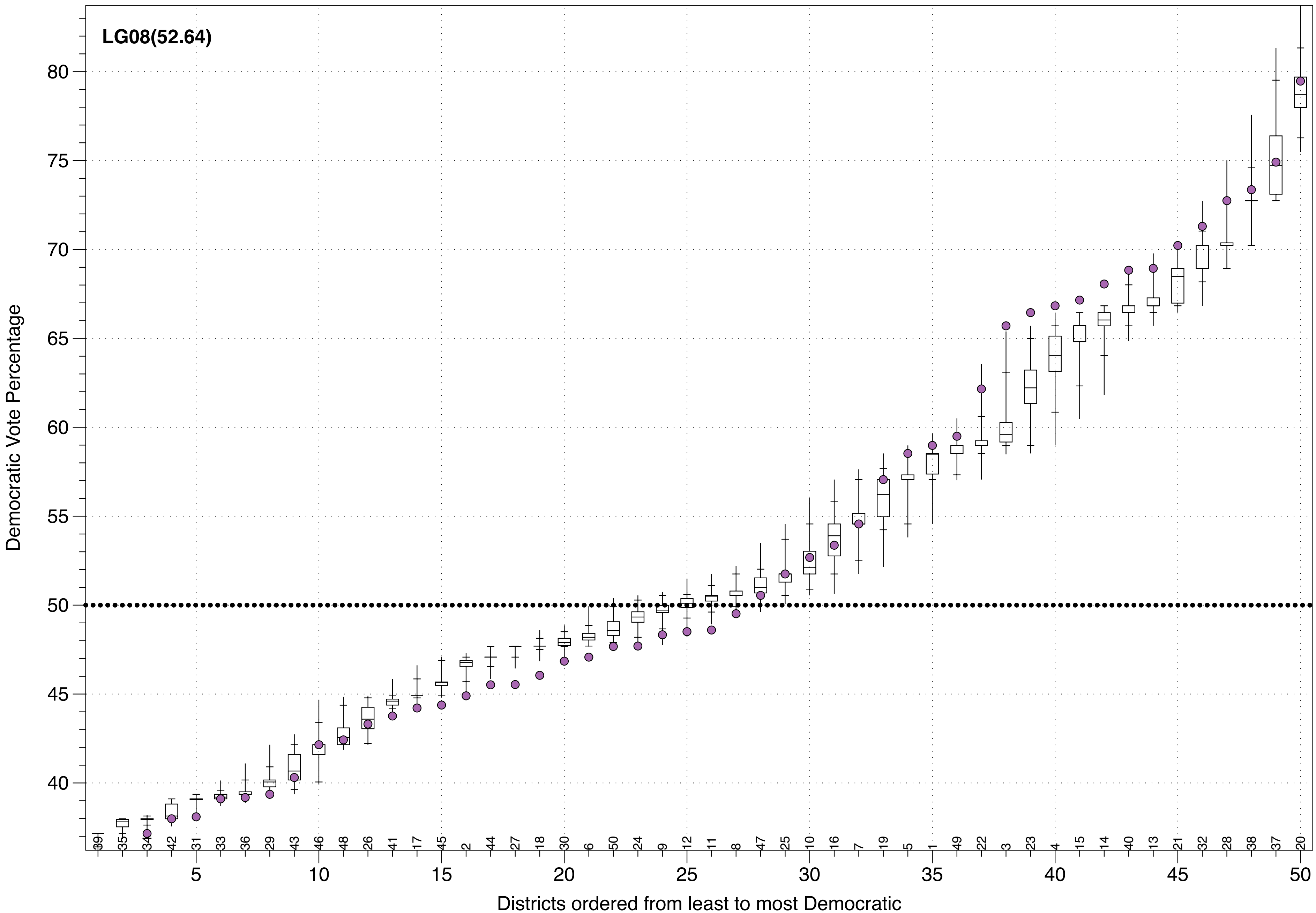
Elec.	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42
USS10	1.85	18.6	51.4	26.6	1.30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
GV12	1e-4	0.91	17.2	57.1	22.3	2.20	0.06	1e-3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LG16	0	0	0	0.20	14.1	60.2	22.3	2.84	0.12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
USS16	0	0	0	1e-2	6.77	52.5	33.1	7.02	0.48	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PR12	0	0	0	0	1e-3	0.21	13.3	65.4	19.3	1.59	0.02	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LG12	0	0	0	0	0	1e-2	1.28	20.7	64.0	13.4	0.50	1e-2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PR16	0	0	0	0	0	1e-3	0.05	3.83	75.7	20.3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
USS14	0	0	0	0	0.01	1.98	22.8	56.1	17.5	1.40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
GV16	0	0	0	0	0	0	0	0.75	59.3	39.8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PR08	0	0	0	0	0	0	0	1.02	19.4	49.7	26.0	3.75	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AG16	0	0	0	0	0	0	0	1e-3	0.15	12.9	60.6	26.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CI12	0	0	0	0	0	0	0	0	0.01	1.89	50.2	38.7	8.48	0.62	1e-2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CI08	0	0	0	0	0	0	0	0	0	0	1e-2	0.43	8.67	45.3	36.0	8.94	0.54	0	0	0	0	0	0	0	0	0	0	0	0	0
LG08	0	0	0	0	0	0	0	0	0	0.03	1.45	10.3	28.2	35.5	20.2	4.07	0.04	0	0	0	0	0	0	0	0	0	0	0	0	0
GV08	0	0	0	0	0	0	0	0	0	0	1e-3	0.06	1.44	9.81	28.5	37.0	19.0	3.75	0.24	1e-3	1e-7	1e-12	0	0	0	0	0	0	0	0
USS08	0	0	0	0	0	0	0	0	0	0	0	0	0	0.15	9.17	34.1	40.8	15.7	1e-2	0	0	0	0	0	0	0	0	0	0	0
AG08	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3.11	37.0	44.2	14.8	0.69

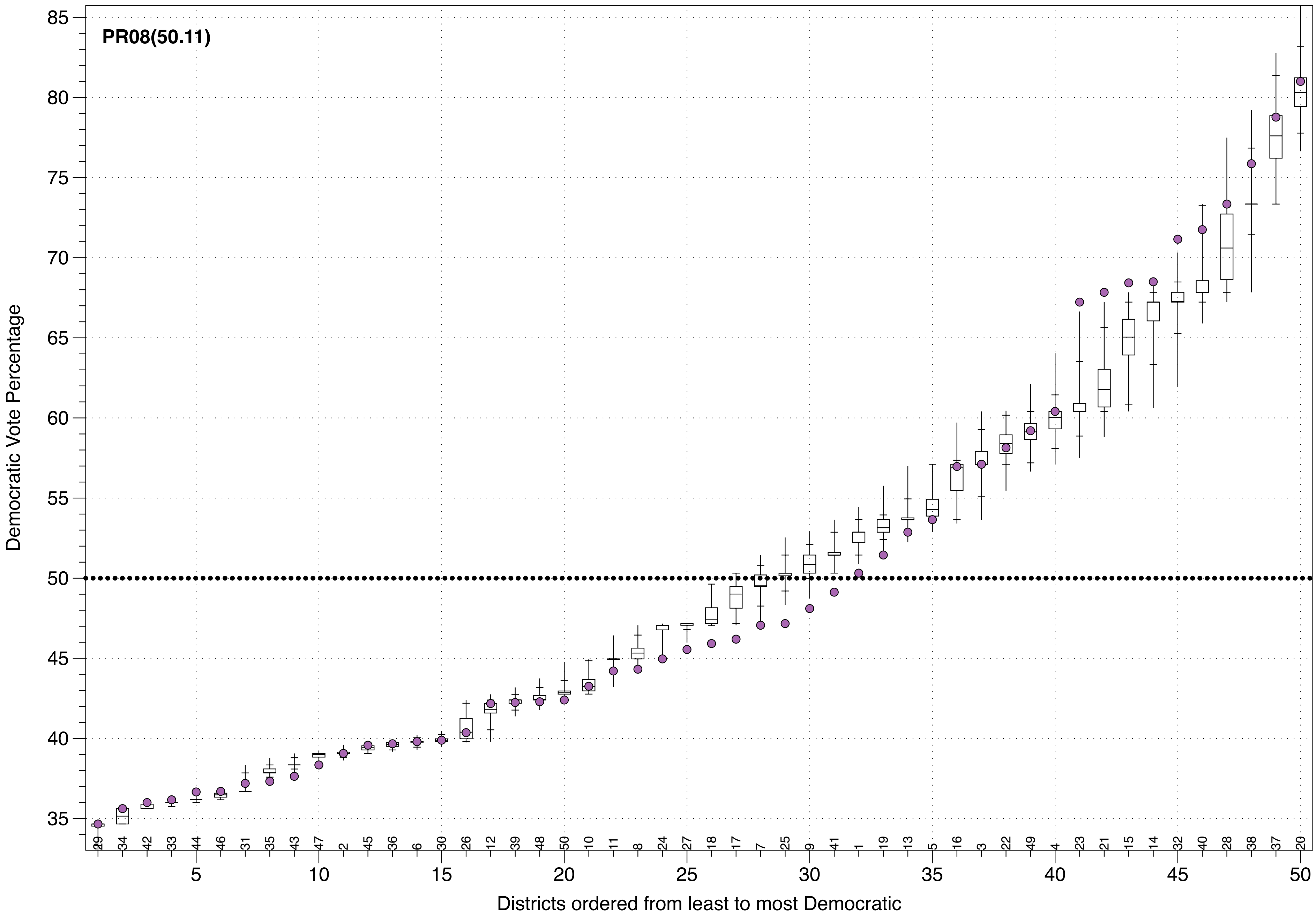
CORRECTED REPORT FIGURE 3; corrected rows are highlighted in the election column. We display the percent chance of electing a given number of Democrats for a given election in the North Carolina Senate. The column headers show the number of Democrats elected; the row headers show the election considered. We use our primary ensemble that respects municipalities, population, compactness and whole county/traversal criteria. The result observed in the enacted plan is highlighted in purple.

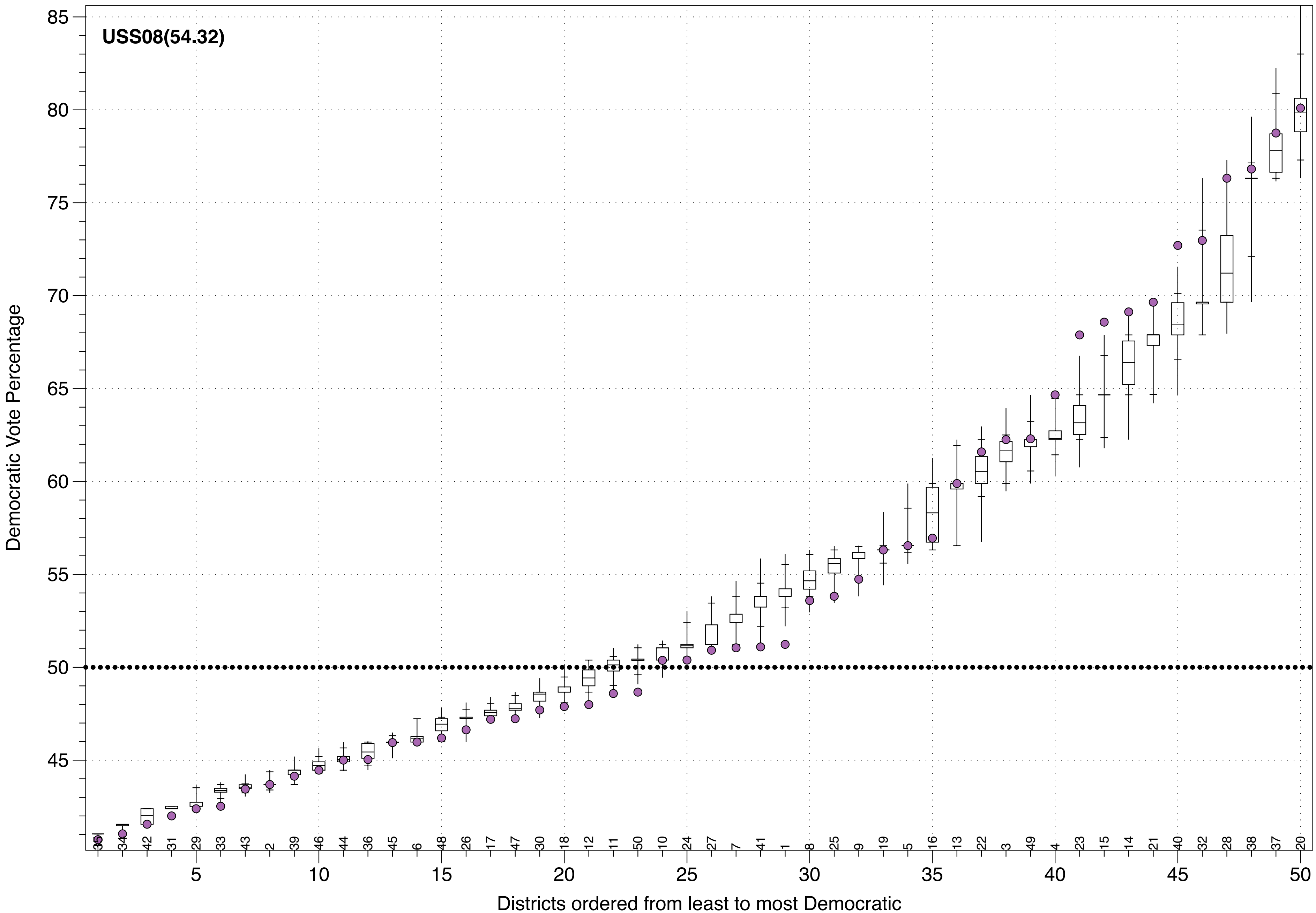

AG08(61.06)
Democratic Vote Percentage
85
80
75
70
65
60
55
50
35 34 42 31 39 29 48 36 43 33 47 46 44 45 41 26 50 2 30 6 27 8 9 24 17 1 18 25 12 10 7 19 16 49 11 5 13 22 21 32 23 40 3 15 4 28 14 38 37 20
5
10
15
20
25
30
35
40
45
50
Districts ordered from least to most Democratic

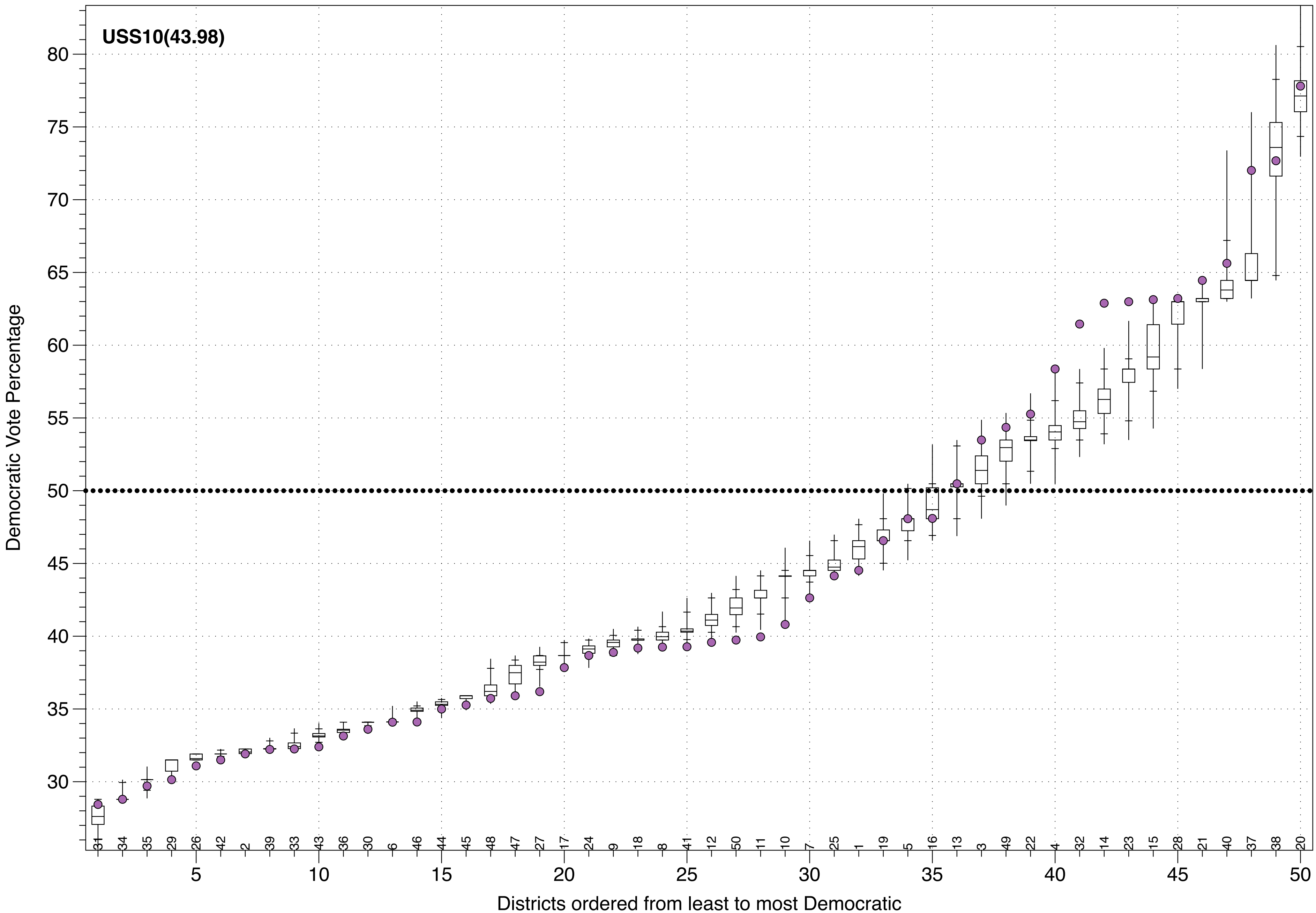

Cl08(53.57)
Democratic Vote Percentage
Districts ordered from least to most Democratic

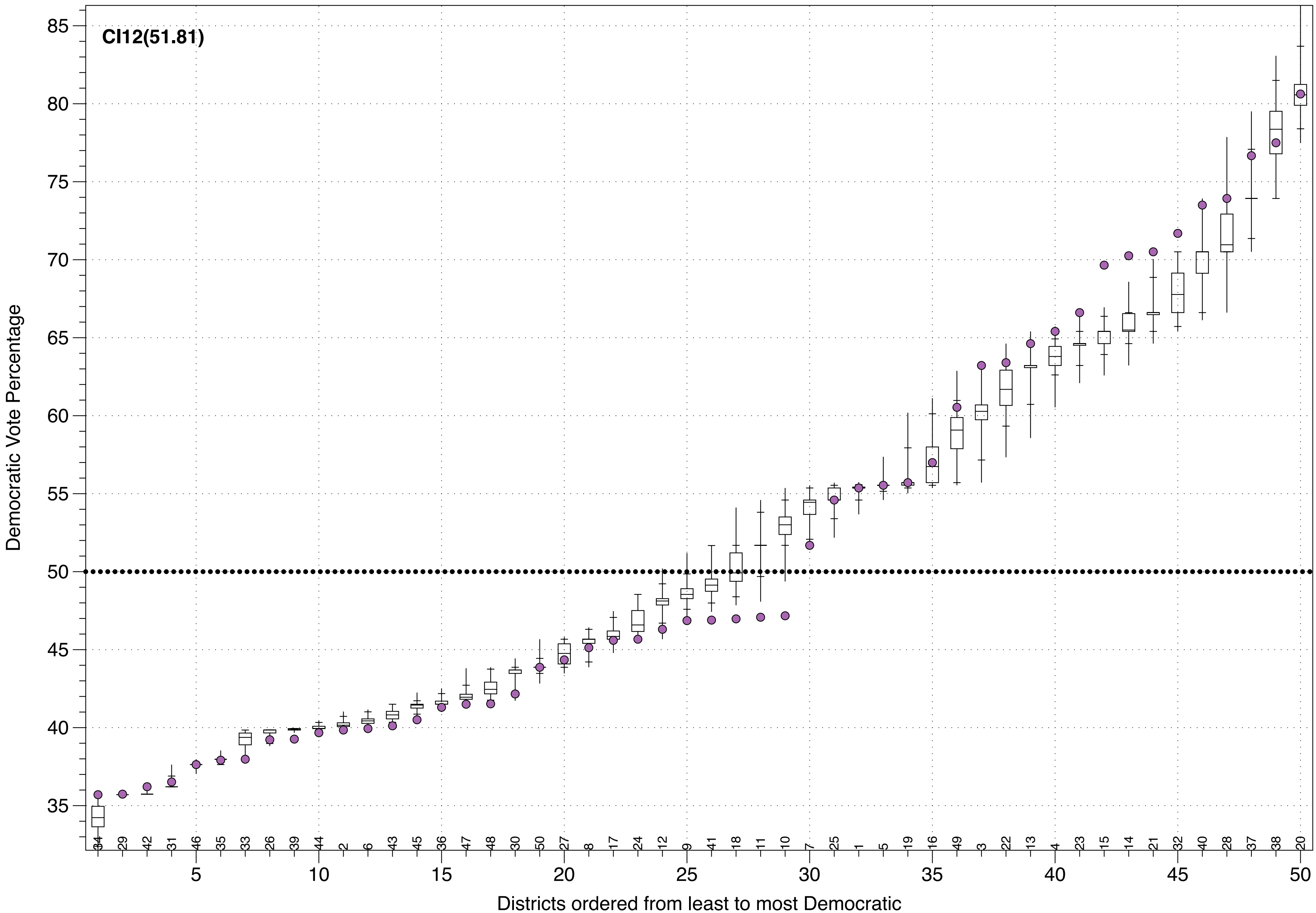

GV08(51.70)
Democratic Vote Percentage
80
75
70
65
60
55
50
45
40
35
30
39 35 42 34 36 41 43 33 44 46 31 29 17 18 26 48 45 27 47 12 11 24 30 9 25 16 8 50 10 6 2 7 19 22 5 1 49 40 15 23 14 37 4 38 3 13 21 28 32 20
5
10
15
20
25
30
35
40
45
50
Districts ordered from least to most Democratic

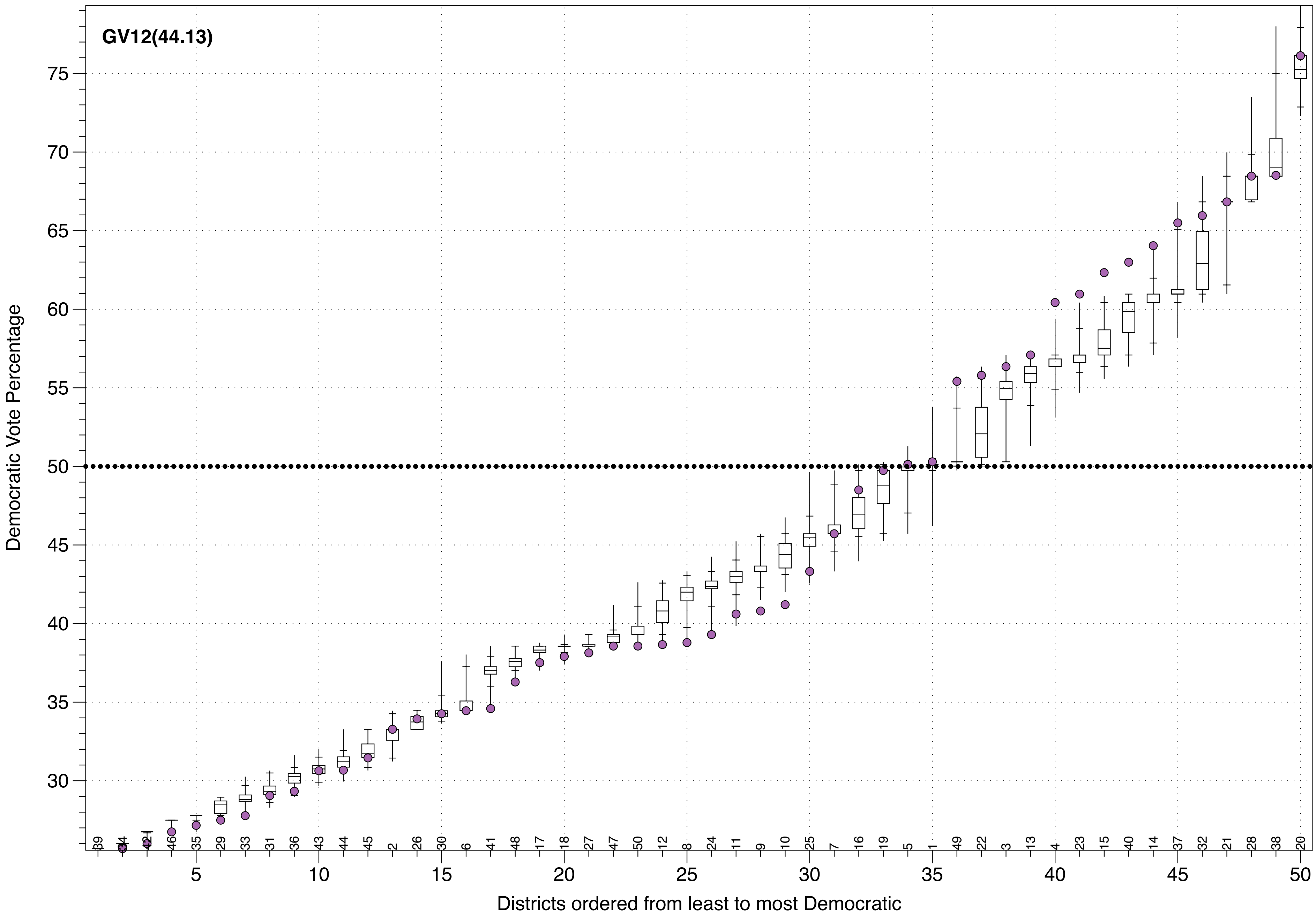

LG08(52.64)
Democratic Vote Percentage
40
45
50
55
60
65
70
75
80
39 35 34 42 31 33 36 29 43 46 48 26 41 17 45 2 44 27 18 30 6 50 24 9 12 11 8 47 25 10 16 7 19 5 1 49 22 3 23 4 15 14 40 13 21 32 28 38 37 20
5
10
15
20
25
30
35
40
45
50
Districts ordered from least to most Democratic

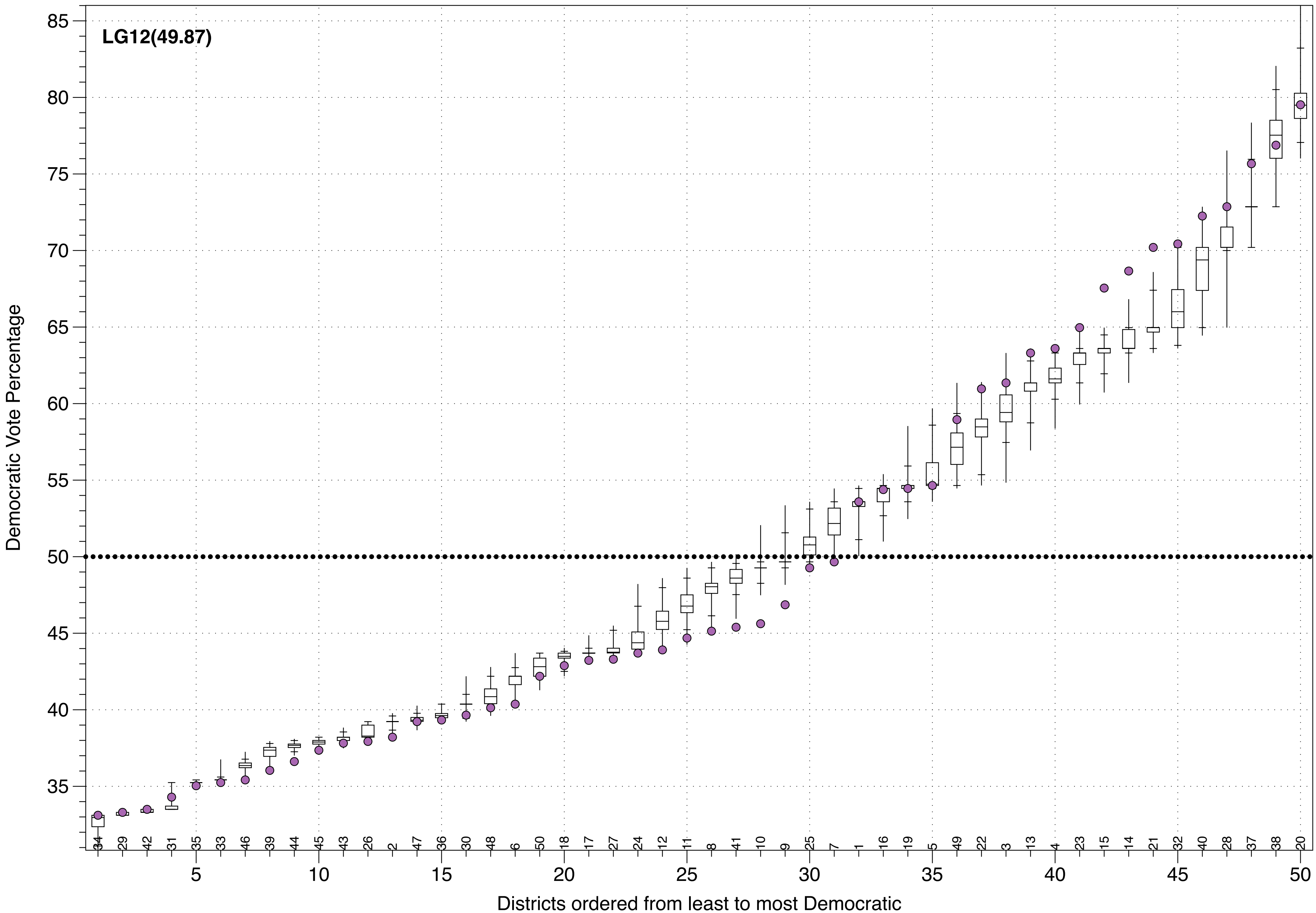

PR08(50.11)
Democratic Vote Percentage
85
80
75
70
65
60
55
50
45
40
35
29 34 42 33 44 46 31 35 43 47 2 45 36 6 30 26 12 39 48 50 10 11 8 24 27 18 17 7 25 9 41 1 19 13 5 16 3 22 49 4 23 21 15 14 32 40 28 38 37 20
5 10 15 20 25 30 35 40 45 50
Districts ordered from least to most Democratic

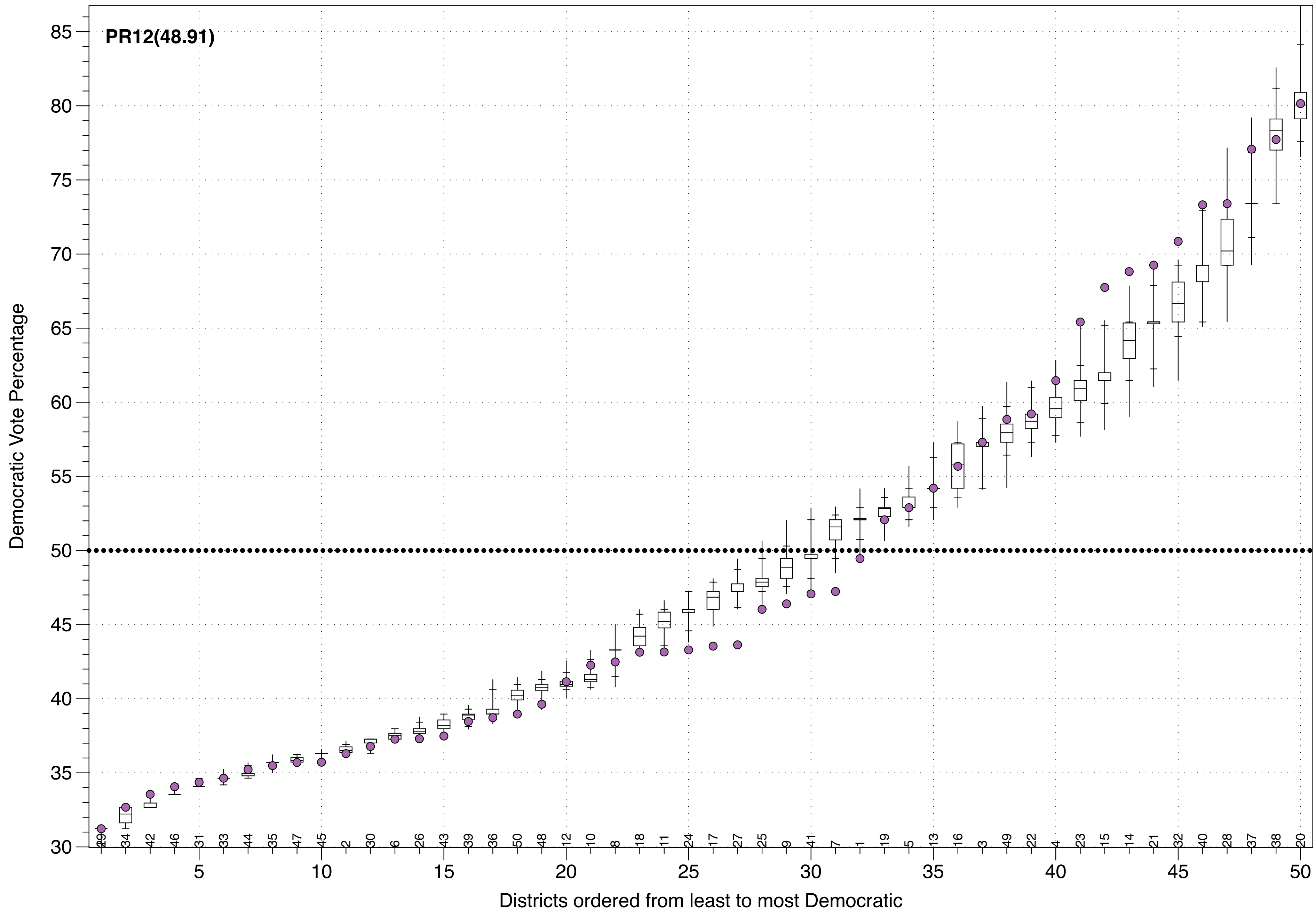

USS08(54.32)
Democratic Vote Percentage
45 50 55 60 65 70 75 80 85
Districts ordered from least to most Democratic
5 10 15 20 25 30 35 40 45 50
34 42 31 29 33 43 2 39 46 44 36 45 6 48 26 17 47 30 18 12 11 50 10 24 27 7 41 1 8 25 9 19 5 16 13 22 3 49 4 23 15 14 21 40 32 28 38 37 20

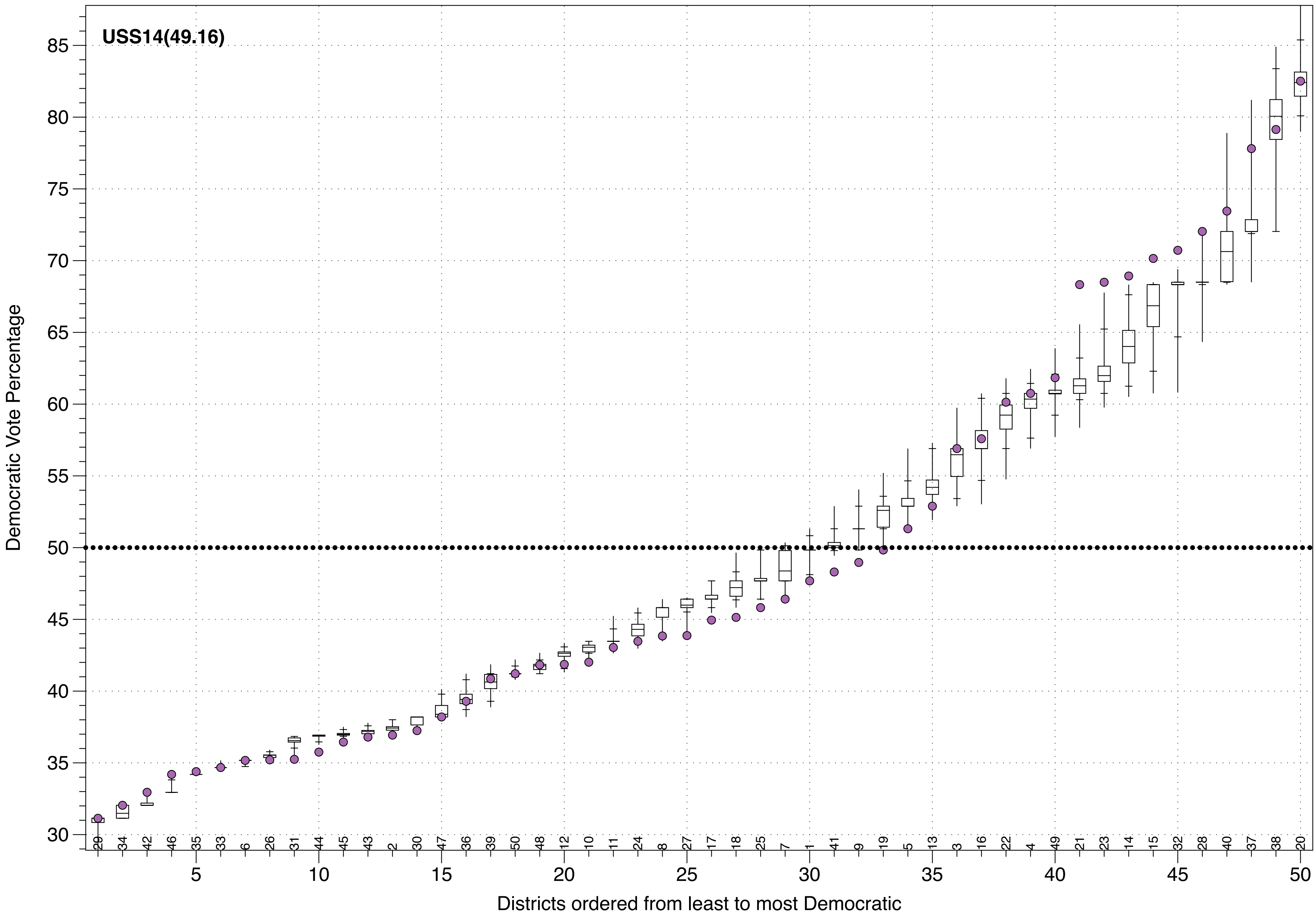

USS10(43.98)
Democratic Vote Percentage
30
35
40
45
50
55
60
65
70
75
80
31 34 35 29 26 42 2 39 33 43 36 30 6 46 44 45 48 47 27 17 24 9 18 8 41 12 50 11 10 7 25 1 19 5 16 13 3 49 22 4 32 14 23 15 28 21 40 37 38 20
5
10
15
20
25
30
35
40
45
50
Districts ordered from least to most Democratic

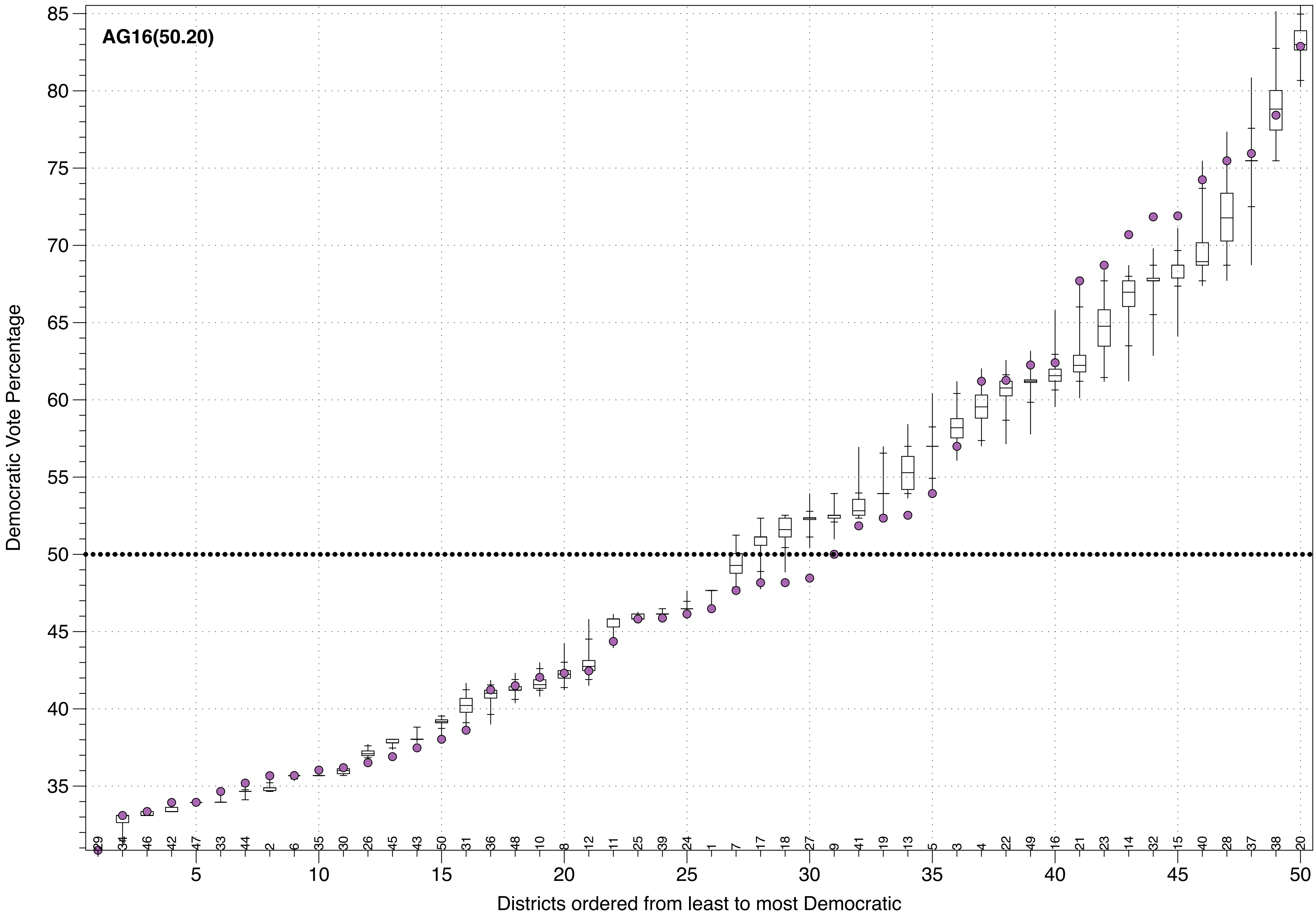

CI12(51.81)
Democratic Vote Percentage
Districts ordered from least to most Democratic
35
40
45
50
55
60
65
70
75
80
85
5
10
15
20
25
30
35
40
45
50
34 29 42 31 46 35 33 26 39 44 2 6 43 45 36 47 48 30 50 27 8 17 24 12 9 41 18 11 10 7 25 1 5 19 16 49 3 22 13 4 23 15 14 21 32 40 28 37 38 20

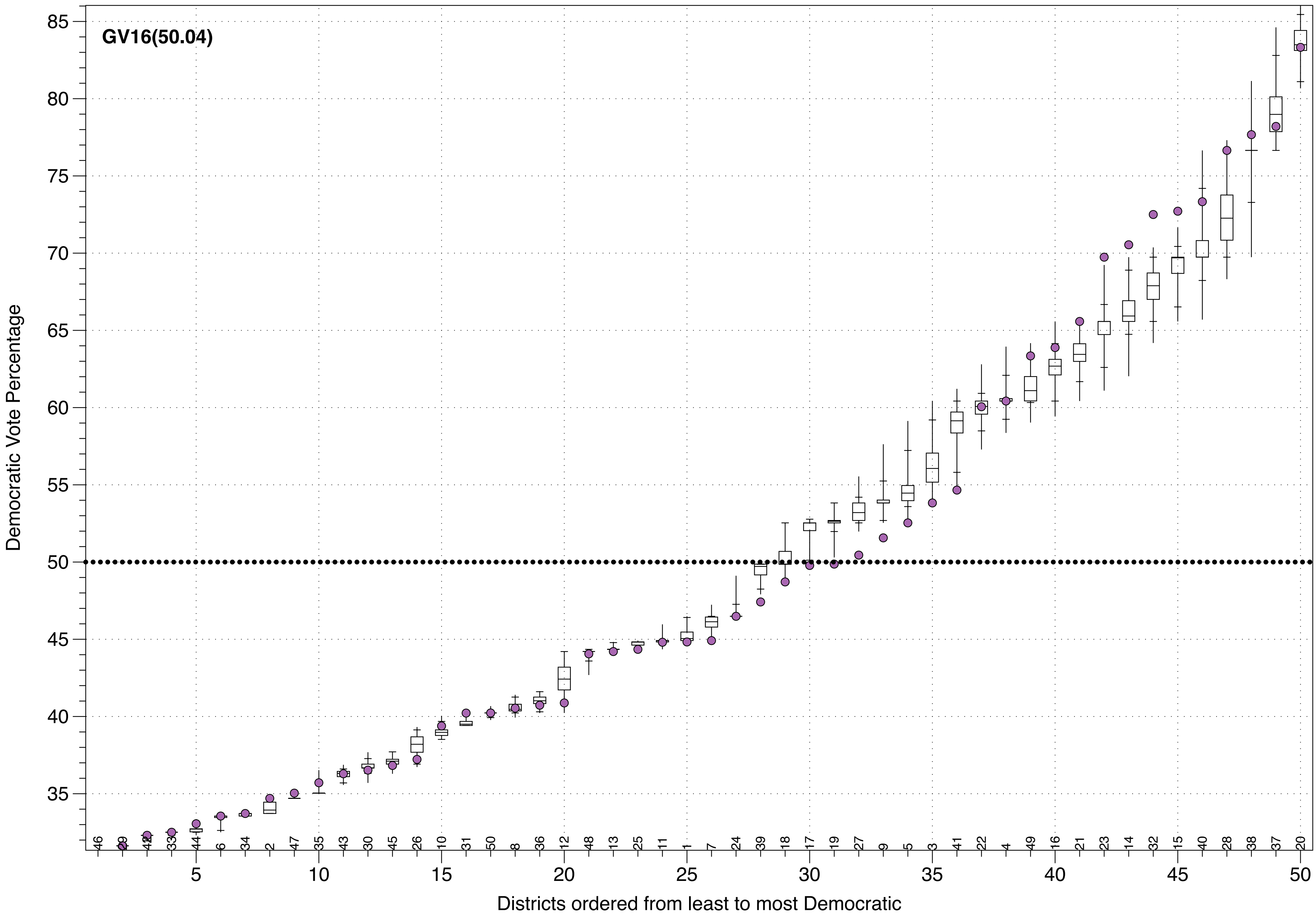

GV12(44.13)
Democratic Vote Percentage
30
35
40
45
50
55
60
65
70
75
39 34 42 46 35 29 33 31 36 43 44 45 2 26 30 6 41 48 17 18 27 47 50 12 8 24 11 9 10 25 7 16 19 5 1 49 22 3 13 4 23 15 40 14 37 32 21 28 38 20
5
10
15
20
25
30
35
40
45
50
Districts ordered from least to most Democratic

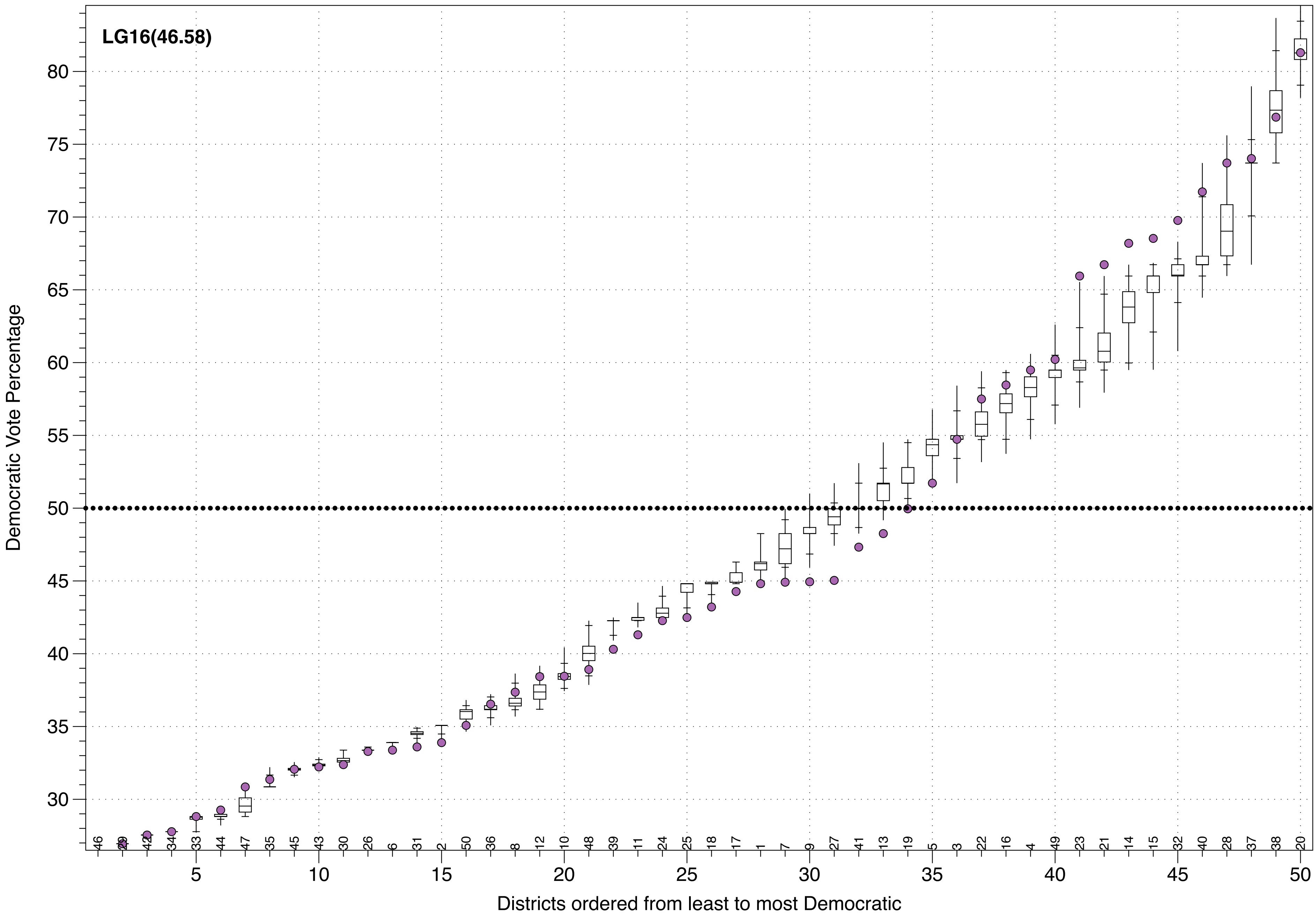

LG12(49.87)
Democratic Vote Percentage
Districts ordered from least to most Democratic

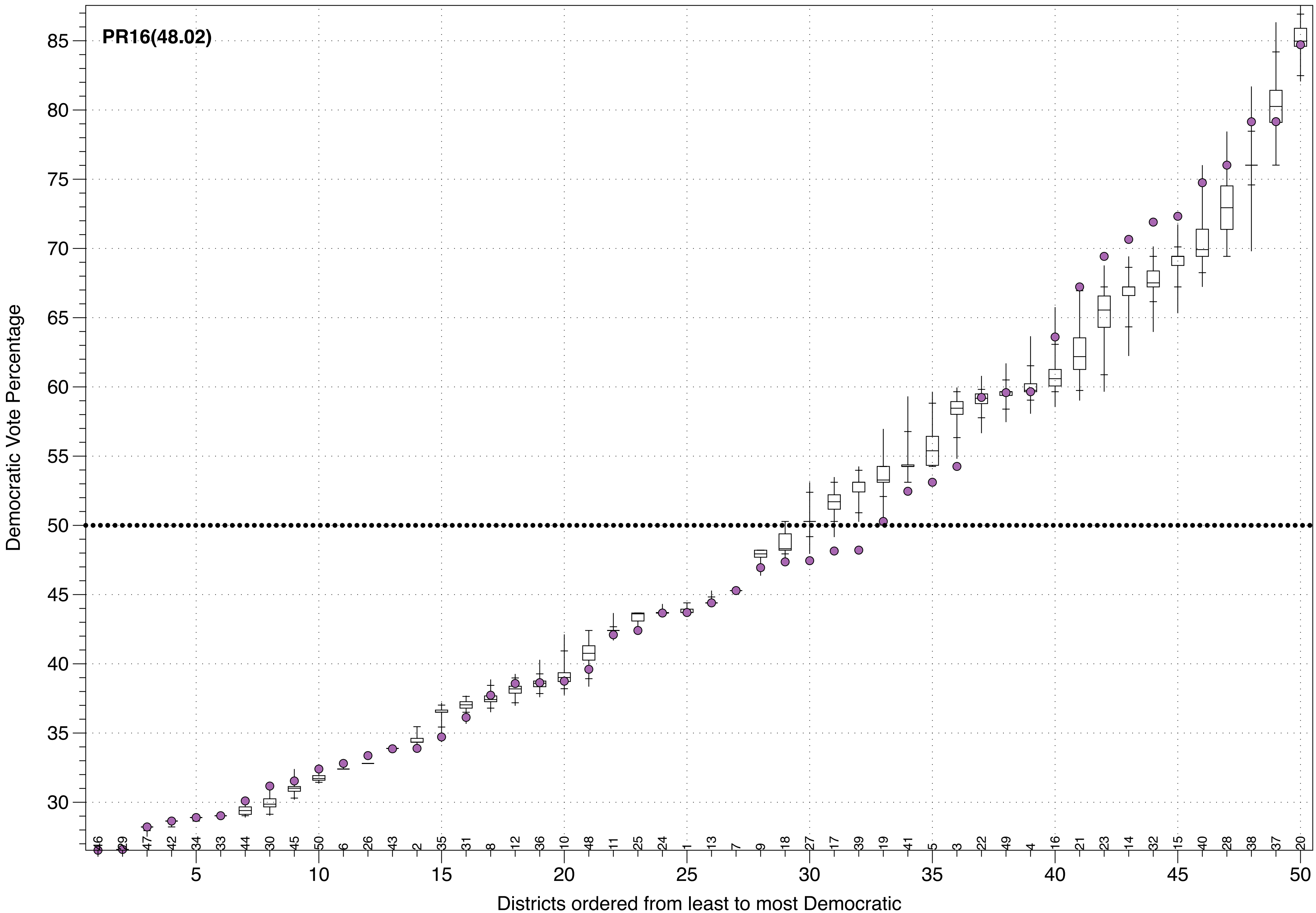

PR12(48.91)
Democratic Vote Percentage
85
80
75
70
65
60
55
50
45
40
35
30
29 34 42 46 31 33 44 35 47 45 2 30 6 26 43 39 36 50 48 12 10 8 18 11 24 17 27 25 9 41 7 1 19 5 13 16 3 49 22 4 23 15 14 21 32 40 28 37 38 20
5
10
15
20
25
30
35
40
45
50
Districts ordered from least to most Democratic

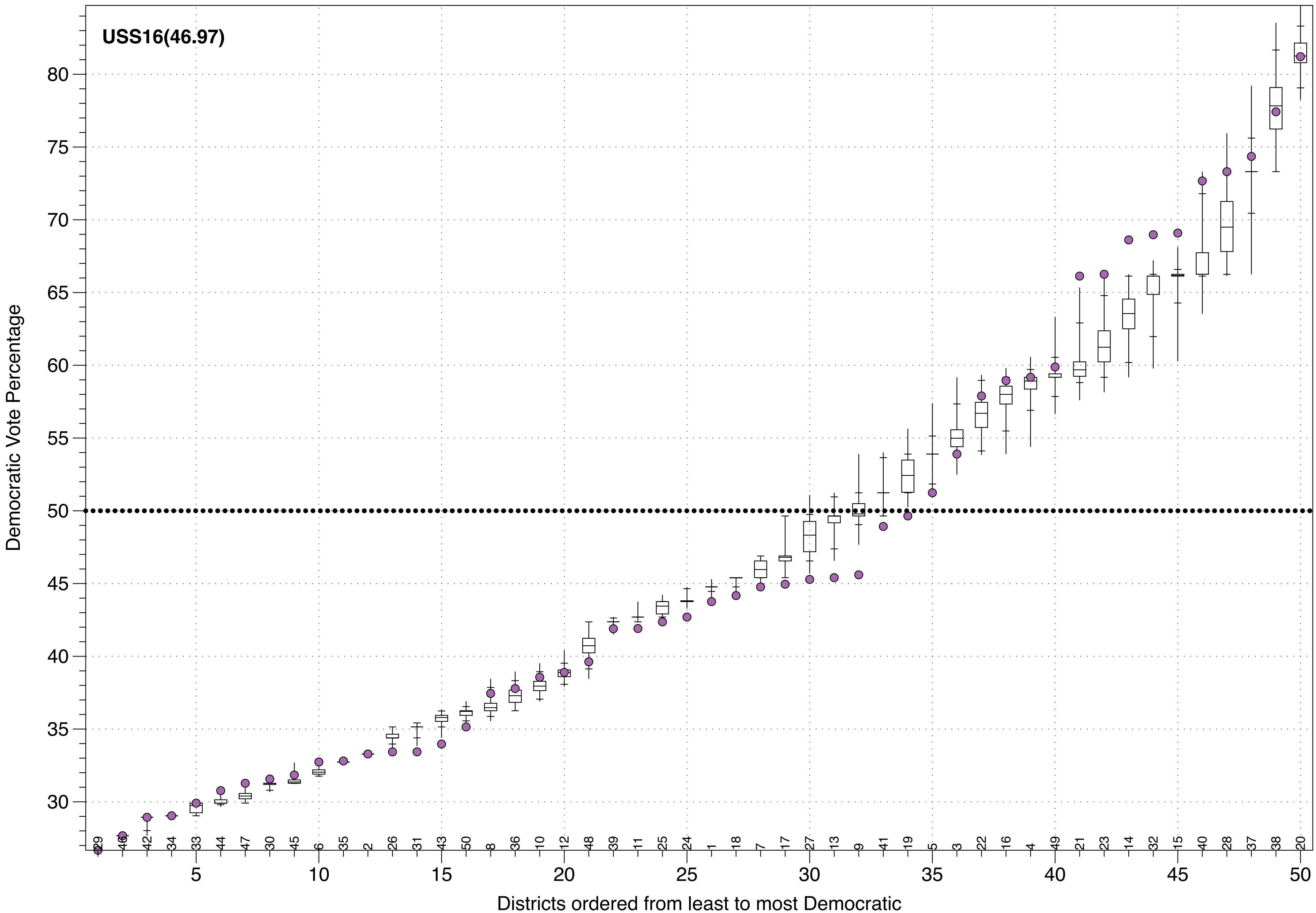

USS14(49.16)
Democratic Vote Percentage
30
35
40
45
50
55
60
65
70
75
80
85
29 34 42 46 35 33 6 26 31 44 45 43 2 30 47 36 39 50 48 12 10 11 24 8 27 17 18 25 7 1 41 9 19 5 13 3 16 22 4 49 21 23 14 15 32 28 40 37 38 20
5
10
15
20
25
30
35
40
45
50
Districts ordered from least to most Democratic


AG16(50.20)
Democratic Vote Percentage
85
80
75
70
65
60
55
50
45
40
35
29 34 46 42 47 33 44 2 6 35 30 26 45 43 50 31 36 48 10 8 12 11 25 39 24 1 7 17 18 27 9 41 19 13 5 3 4 22 49 16 21 23 14 32 15 40 28 37 38 20
5
10
15
20
25
30
35
40
45
50
Districts ordered from least to most Democratic


GV16(50.04)
Democratic Vote Percentage
Districts ordered from least to most Democratic


LG16(46.58)
Democratic Vote Percentage
30
35
40
45
50
55
60
65
70
75
80
46 29 42 34 33 44 47 35 45 43 30 26 6 31 2 50 36 8 12 10 48 39 11 24 25 18 17 1 7 9 27 41 13 19 5 3 22 16 4 49 23 21 14 15 32 40 28 37 38 20
5
10
15
20
25
30
35
40
45
50
Districts ordered from least to most Democratic


PR16(48.02)
Democratic Vote Percentage
Districts ordered from least to most Democratic



USS16(46.97)
Democratic Vote Percentage
30
35
40
45
50
55
60
65
70
75
80
29 46 42 34 33 44 47 30 45 6 35 2 26 31 43 50 8 36 10 12 48 39 11 25 24 1 18 7 17 27 13 9 41 19 5 3 22 16 4 49 21 23 14 32 15 40 28 37 38 20
5
10
15
20
25
30
35
40
45
50
Districts ordered from least to most Democratic

I certify that the foregoing is true and correct to the best of my knowledge.

Sincerely,

sig.pdf

Jonathan C. Mattingly
June 7th, 2019.
Durham, NC