IN THE UNITED STATES DISTRICT COURT NORTHERN DISTRICT OF ALABAMA SOUTHERN DIVISION

Milligan *et al*.

Plaintiffs,

v.

Merrill *et al*.

Defendents.

REBUTTAL EXPERT REPORT

Kosuke Imai, Ph.D.

December 20, 2021

Table of Contents

I. INTRODUCTION AND SCOPE OF WORK

1. My name is Kosuke Imai, Ph.D., and I am a Professor in the Department of Government and the Department of Statistics at Harvard University. I specialize in the development of statistical methods for and their applications to social science research. I am also affiliated with Harvard's Institute for Quantitative Social Science. My qualifications and compensation are described in my initial report.

2. I understand from Plaintiffs' counsel that one of Defendants' experts offered the opinion that Mobile and Baldwin Counties are communities of interest and should not be divided across congressional districts. I also understand from Plaintiffs' counsel that there is evidence supporting the Black Belt, as defined below, as a community of interest. I express no opinions on these issues.

3. I have been asked by Plaintiff's counsel to re-run my "one-MMD (majority-minority district) simulation" from my initial report with additional weighting that encourages the algorithm to keep Mobile/Baldwin and the Black Belt together and to examine the likely effect on the range of black voting-age population (BVAP) proportion of non-MMD districts, particularly District 2. The original one-MMD simulation I conducted for my initial report generated 10,000 alternative plans, each of which was designed to have exactly one MMD with the proportion of black voting-age population (BVAP) ranging from 50% to 51%. The other six districts of each simulated plan were generated without any consideration of race. This time, however, I instructed the algorithm to generate, with a high probability, plans which keep Mobile and Baldwin Counties together and the Black Belt together. Other than this additional weight, the new one-MMD simulation procedure I employed is identical to the one used in my initial expert report. Like the original one-MMD simulation, my new 10,000 simulated plans are, on average, more compact and have no more county splits than the enacted plan.

II. SUMMARY OF OPINIONS

4. The comparison of the new one-MMD simulated plans with the enacted plan yields the following findings: the district with the second highest BVAP proportion in simulated plans achieves, on average, 6.2 percentage points higher BVAP proportion than the enacted plan. This difference is statistically significant using the conventional standard. The new one-MMD simulations generated many more plans with a greater BVAP proportion for the second highest BVAP district than my initial one-MMD simulation, which did not encourage the algorithm to avoid splitting Mobile/Baldwin Counties and the Black Belt.

5. My simulation analyses, therefore, provide evidence that race was a significant factor in drawing the enacted plan, and that, taking into account the identified communities of interest, the enacted plan is still an outlier in terms of how it cracks the Black community.

III. METHODOLOGY

6. The simulation procedure used for this report is identical to that of the one-MMD simulation from my initial report with the exception of one additional weighting I added to discourage the simulation algorithm from splitting Mobile and Baldwin Counties as well as the Black Belt. I was instructed by Plantiffs' counsel to use the following set of counties for the Black Belt: Barbour, Bullock, Butler, Choctaw, Clarke, Conecuh, Crenshaw, Dallas, Escambia, Greene, Hale, Lowndes, Macon, Marengo, Monroe, Montgomery, Perry, Pickens, Pike, Russell, Sumter, Washington, and Wilcox Counties. As standard in the literature, I used the so-called Gibbs measure to incorporate this constraint into the simulation algorithm (Autry et al. 2020; Carter et al. 2019; McCartan and Imai 2020; Kenny et al. 2021).[1] One MMD whose BVAP proportion is between 50% and 51% was generated for each simulated plan in the exactly same manner as done in the one-MMD simulation for my initial report. Finally, I used the same data set as the one analyzed in my initial report.

1. Specifically, I used the indicator variable for splitting each of these two county clusters with a penalty weight of 25.

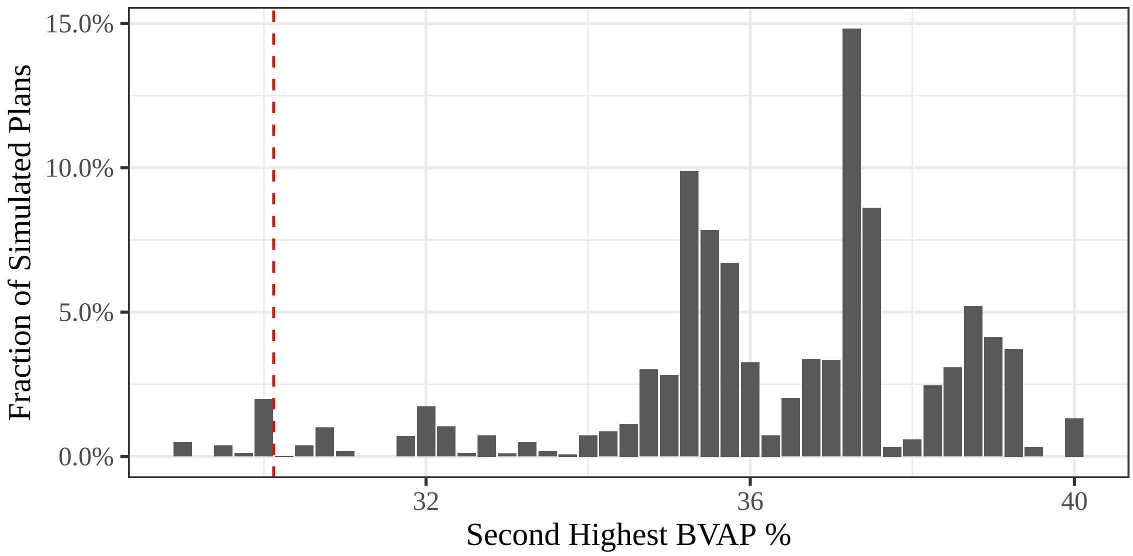

Figure 1: The second highest Black voting age population (BVAP) proportion (after the simulated majority-minority district) in each simulated plan. The vast majority of simulated plans have greater BVAP than the enacted (red).

IV. EVALUATION OF THE ENACTED PLAN

7. Using the redistricting simulation methodology described above, I evaluate evidence regarding whether race was a primary factor in drawing the enacted plan. In Appendix A, I demonstrate that the simulated plans are on average at least as compact as the enacted plan based on the standard compactness measures. Appendix B further shows that all of the simulated plans have fewer than or equal to the number of county splits the enacted plan does. In addition, like the original one-MMD analysis conducted for my initial report, all simulated plans have at most one incumbent located in any given district.

8. I can easily generate additional plans by running the algorithm longer, but for the purpose of my analysis, 10,000 simulated plans for each set will yield statistically precise conclusions. In other words, generating more than 10,000 plans, while possible, will not materially affect the conclusions of my analysis.

9. Figure 1 shows the distribution of BVAP proportion for the district that has the second highest BVAP proportion under each simulated plan. Note that under more than 99% of the simulated plans, District 2 has the second highest BVAP proportion. When compared to the en-

acted plan (represented by the red dashed line), under the simulated plans, this district has a much higher BVAP proportion with a maximum value of 39.9%. Although all of non-MMD districts were generated without using any information about race, the simulation plan has, on average, the second highest district-level BVAP proportion at 36.3%, which is 6.2 percentage points higher than the corresponding BVAP proportion under the enacted plan (30.1%). Only 3% of the simulated plans have the second highest district-level BVAP proportion to be less than the one for the enacted plan. In other words, this difference between the simulated plans and the enacted plan is statistically significant.

10. When compared to the original one-MMD simulation reported in my initial report, this new one-MMD simulation generated many more plans with a greater BVAP proportion for the second highest district-level BVAP proportion. This implies that keeping Mobile and Baldwin Counties together and the Black Belt together is likely to significantly increase the second highest district-level BVAP proportion.

Pursuant to 28 U.S.C. § 1746, I hereby declare under penalty of perjury that the forgoing is true and correct:

Executed, this day, December 20, 2021, in Cambridge, Massachusetts.

Kosuke Imai, Ph.D.

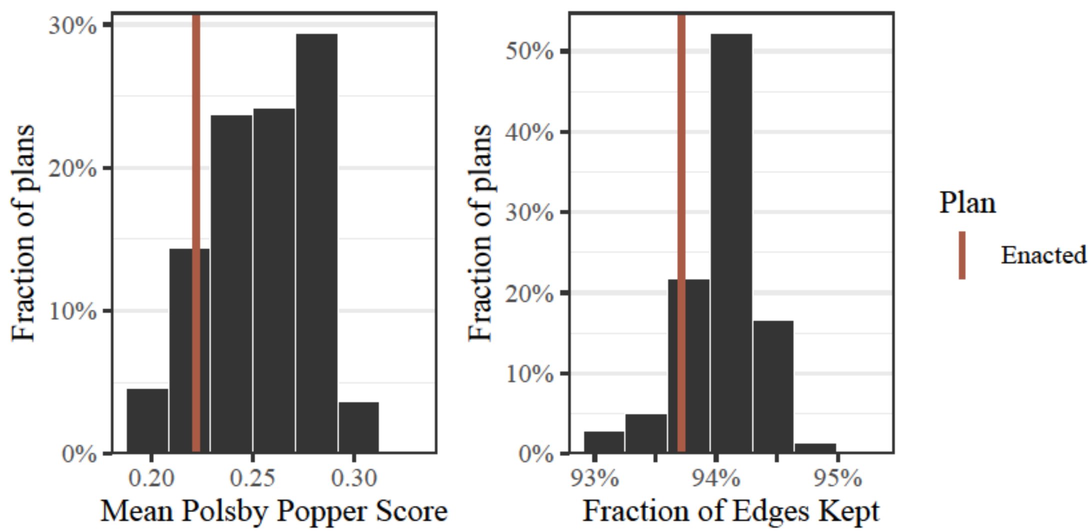

Figure 2: The compactness of the one-MMD simulated plans according to two measures – Polsby-Popper compactness (left) and fraction of edges kept (right). In general, simulated plans are as compact or more compact than the enacted plan.

V. APPENDIX

A. Compactness of the Simulated Plans

1. I now show that the simulated plans are more compliant than the enacted plan. I use the average Polsby–Popper (Polsby and Popper 1991) and edge-removal (DeFord, Duchin, and Solomon 2021; McCartan and Imai 2020) scores, two commonly-used quantitative measures of district compactness. Figure 2 also shows that according to the Polsby–Popper and edge-removal scores, the new one-MMD simulated plans are, on average, more compact than the enacted plan.

B. County Splits of the Simulated Plans

2. Figure 3 presents the number of counties split within the MMD (left plot) and the total number of counties split (right plot). The figure shows that when compared to the enacted plan, all of the new one-MMD simulated plans have fewer or equal number of county splits within the MMD and across all districts.

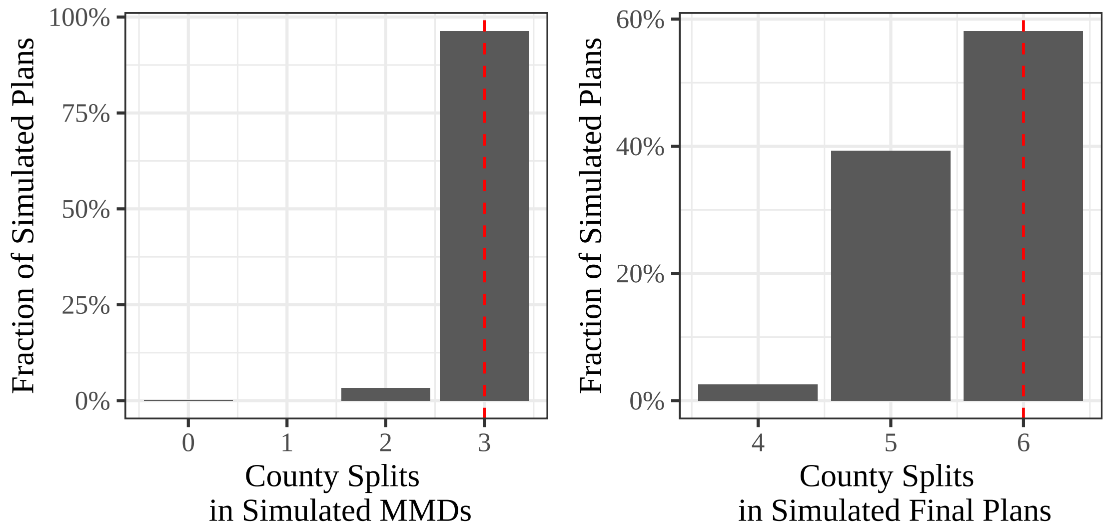

Figure 3: The number of county splits in each simulated majority-minority district (left) and in the complete simulated plans (right). All simulated plans used in the analysis have the same number or fewer splits than the enacted plan (red).

C. References

Autry, Eric, Daniel Carter, Gregory Herschlag, Zach Hunter, and Jonathan Mattingly. 2020. "Multi-scale merge-split Markov chain Monte Carlo for Redistricting." *arXiv preprint arXiv:2008.08054.*

Carter, Daniel, Gregory Herschlag, Zach Hunter, and Jonathan Mattingly. 2019. "A Merge-Split Proposal for Reversible Monte Carlo Markov Chain Sampling of Redistricting Plans." *arXiv preprint arXiv:1911.01503.*

DeFord, Daryl, Moon Duchin, and Justin Solomon. 2021. "Recombination: A Family of Markov Chains for Redistricting." Https://hdsr.mitpress.mit.edu/pub/1ds8ptxu, *Harvard Data Science Review* (March 31, 2021). https://doi.org/10.1162/99608f92.eb30390f. https://hdsr.mitpress.mit.edu/pub/1ds8ptxu.

Kenny, Christopher T., Shiro Kuriwaki, Cory McCartan, Evan Rosenman, Tyler Simko, and Kosuke Imai. 2021. "The Use of Differential Privacy for Census Data and its Impact on Redistricting: The Case of the 2020 U.S. Census." *Science Advances* 7, no. 41 (October): 1–17.

McCartan, Cory, and Kosuke Imai. 2020. "Sequential Monte Carlo for sampling balanced and compact redistricting plans." *arXiv preprint arXiv:2008.06131.*

Polsby, Daniel D, and Robert D Popper. 1991. "The third criterion: Compactness as a procedural safeguard against partisan gerrymandering." *Yale Law & Policy Review* 9 (2): 301–353.