Evaluating Racially-Constrained Ensembles of Alternative Ohio House Districting Plans

Alan Ding

Advisor: Prof. Samuel S.-H. Wang

May 23, 2021

Abstract

A recent and particularly effective approach for identifying the presence of gerrymandering involves using Markov chain Monte Carlo simulation to generate ensembles of alternative districting plans for a state. Here, we seek to assess the racial and partisan outcomes of a set of ensembles with different constraints on the racial makeup of its districts (but are party-neutral) for Ohio House plans. We find, in particular, that race- and party-neutral ensembles generate representationally unfair maps for Black voters and Democrats and that adding racial constraints slightly depresses Democratic representation even further. Additionally, we find that all of our party-neutral ensembles will almost always generate plans with significantly negative mean-median gaps, which calls for the use of other statistics to identify intent to gerrymander when considering Ohio House plans.

1 Introduction

With the recent release of the first batch of data from the 2020 United States Census, the 2021 redistricting cycle is finally underway. States across the country will have their congressional and state legislative districts redrawn in this process. The way that districts are drawn can have substantial impacts on who in a given state is more strongly represented in Congress or in state legislative chambers through a political practice called gerrymandering. In the last couple of decades, this practice has become all the more contentious and optimized by political parties, aided both by the rapidly increasing ease of access to computing power and the deepening political divide and tension between the two major political parties in the United States.[1]

Gerrymandering can manifest itself as racial gerrymanders or partisan gerrymanders. In the former, the voting power of a racial group is diluted or strengthened by way of how districts are drawn; in the latter, the voting power of a political party can be diluted or strengthened. Especially in the case of partisan gerrymanders, the U.S. Supreme Court has not been particularly receptive to striking them down—in fact, in *Rucho v. Common Cause*,[2] the U.S. Supreme Court decided, in a slim 5-4 majority, that the question of partisan gerrymandering was not justiceable by the Supreme Court. In light of this decision, Wang et al. propose a framework that turns to state courts to litigate against district maps that are partisan gerrymanders [14]. Working with gerrymandered states individually can be furthermore advantageous when the states in question require public input on proposed maps. As it turns out, Ohio is a particularly interesting state to look at, taking these considerations into account. With the election of Judge Jennifer Brunner to the Ohio Supreme Court, the court appears to have become more receptive to taking on cases against unfairly drawn maps.[3]. Ohio

[1] https://www.theatlantic.com/politics/archive/2017/10/gerrymandering-technology-redmap-2020/543888/

[2] 139 S. Ct. 2484 (2019)

[3] Justice Brunner's campaign messaging often centers on her commitment to combating the issue of gerrymandering specifically. To take one example: https://www.judgebrunner.com/news/its-official—we-broke-gerrymandering-in-ohio—

is also a state that requires public hearings on proposed maps[4] Finally, while it may be tempting to develop ways to test for gerrymandering independent of a given state, differences in states' political and demographic geographies often render this sort of approach to be unrealistic; if not, less effective at best.

One form of public input can involve checking whether maps are fair using a particularly effective computational approach. Specifically, through a *Markov chain Monte Carlo (MCMC)* algorithm, we can obtain a neutral and representative sample of alternative maps from the space of all possible maps that are constructed in accordance with legal requirements, which we call an *ensemble*. The ensemble serves as a baseline on which statistical tests can be run to show evidence of partisan or racial gerrymandering. While MCMC simulation is not a novel computational technique, the application of this approach to redistricting is a very recent development. Notably, this technique was used in 2018 by the Metric Geometry and Gerrymandering Group (MGGG) in demonstrating that Virginia's 2011 plan for its House of Delegates was racially gerrymandered [12]. Indeed, this map was eventually struck down as a racial gerrymander in the U.S. Supreme Court in *Virginia House of Delegates vs. Bethune-Hill*.[5]

Some laws, such as the typical requirements of districts falling within some percentage of an ideal population, contiguity, or even compactness, lend themselves directly to well-defined constraints that we can program into our ensembles. For other laws, such as the Voting Rights Act (VRA), it is not straightforward or even clear how to translate the constraints that it imposes into a precisely defined condition to constrain our ensembles by. Sometimes, this is not an issue—MGGG's Virginia House of Delegates report was able to demonstrate the presence of racial gerrymandering using an ensemble generated from a Markov chain that was constrained only by contiguity as well as bounds on population deviation and compactness. Furthermore, the completely racially blind ensemble contained maps with an average of about 13 out of 33 districts (about 39% of districts) with a Black voting-age population (BVAP)

[4] https://gerrymander.princeton.edu/reforms/OH
[5] 587 U.S. ___ (2019)

proportion of at least 37%.[6] This 39% share of districts is not less than the $\approx 35\%$ proportion of Black voters living in the part of Virginia that was being examined by the report [12]. In fact, some of the neutral ensemble's maps had 15 districts with a BVAP proportion above 37%—three more than the enacted 2011 plan—which served as one of the key points in showing that the enacted plan was a racial gerrymander.

However, an ensemble that does not implement any race-based constraints might not always be able to demonstrate the racial unfairness of a hypothetical proposed map. We show that this is indeed the case in Ohio by considering the differences between ensembles of state House of Representatives districts generated by several Markov chains with different combinations of constraints. To be clear, we do not intend to show or have evidence that the Ohio House district plan enacted in 2011 is a racial gerrymander—if anything, this plan actually represents Black voters quite well. Additionally, the race-based constraints we add are quite simple and are not meant to model actual redistricting practice with respect to the Voting Rights Act.

We are also interested in seeing the implications of these constraints on partisan outcomes. We first find that a race- and party-neutral "baseline" ensemble of alternative Ohio House districting plans does not only lead to significantly less-than-representative outcomes for Black voters, but also for Democrats. We then observe that our constrained ensembles that produce more racially representative outcomes also slightly depress the Democratic seat share for Ohio House districts further relative to the baseline. Finally, we test the enacted 2011 plan and maps drawn by our ensembles (which were all party-neutral by construction) for intent and effect of partisan gerrymandering using the three tests presented in a 2016 paper by Wang [13].

[6]The choice of 37% is not arbitrary. To quote MGGG's Virginia House of Delegates report, "[n]ationally, 37% is an empirical bright line for congressional voting: 32 out of 34 current U.S. congressional districts with at least 37% BVAP had Representatives in the 115th Congress who belong to the Congressional Black Caucus, and the ratio drops off precipitously below that level." [12] In Ohio, the two districts with the lowest BVAP proportions above 37% are the 34th and the 44th (our data gives respective BVAP proportions of 39.5% and 45.6%). In 2012, the 34th district was represented by Vernon Sykes (see https://ballotpedia.org/Vernon_Sykes) and the 44th was represented by Michael Ashford (see https://ballotpedia.org/Michael_Ashford), both of whom were members of the Ohio Legislative Black Caucus. This suggests that the 37% BVAP line is likely also a reasonable threshold for ensuring that a district receives Black representation in the Ohio House.

In doing so, we surprisingly find that all of our ensembles will virtually always draw maps where *reliable-wins test* identifies them as partisan gerrymanders despite not having been programmed to consider partisan data at all.

2 Related Work

A number of reports have examined the effects of imposing different constraints on ensembles for various reasons. MGGG's Virginia House of Delegates report includes the results of a Markov chain with a constraint that accepts a map only if all of its districts have a BVAP proportion of at most 60% [12] so as to prevent the packing of Black voters into certain districts to dilute their power in others. While the report was able to obtain its result without this constraint, adding it elevated the BVAP proportion of the districts around 37% BVAP by a couple of percentage points [12].

In 2019, DeFord and Duchin studied some of the proposed redistricting rules that were under consideration to be included in a constitutional amendment in Virginia, also using MCMC methods. They find that changing the strictness of the population deviation constraints do not significantly change the partisan or racial outcomes of maps. In considering various "BVAP ensembles" where DeFord and Duchin imposed lower bounds on the number of districts with BVAP proportions above 40% (in the congressional district case) or 50% (in the state Senate case), it is shown that there is no substantial effect to partisan outcomes and that the enacted plan comports very well with the ensembles with respect to racial outcomes. Finally, the researchers note that adding constraints on preserving localities changes the partisan outcome by eliminating the most extreme of them [8].

Others have examined the effects of constraints for different purposes. For instance, a report done by Chen in 2016 looks at ensembles of maps constrained by a *minimum* amount of incumbent protection as a way to explain partisan discrepancies observed in North Carolina's congressional districting plan that favored Republicans [7]. Becker et al. examine the outcomes of a chain constrained by a much more nuanced set of criteria derived directly from the Voting Rights Act and the Constitution [6].

With this study, we contribute an examination of the effects of various racial constraints on MCMC ensembles of Ohio state House districting plans. In particular, we will take a fairly similar approach to that of DeFord and Duchin where the constraints that we evaluate do not intend to capture the legal complexities of VRA-compliance and constitutionality, but rather serve as a hypothetical model of principles that might be followed in the map drawing process. Not only is this approach appealing for its relative simplicity—there is also evidence that past state House drawing commissions have given special attention towards increasing the number of majority-Black districts.[7] It is reasonable to assume that this awareness will continue to exist in the forthcoming 2021 wave of redistricting, and so considering some of the implications of these rules will ultimately be important and valuable.

3 Data and Implementation

3.1 Demographic and election data

For the purpose of our evaluation here, we use demographic and population data from the 2010 Census at the census block level. Voting data from the 2016 presidential and U.S. Senate elections are used to evaluate the partisan effects of constraints. Unfortunately, as is the case with many other states, electoral data is not provided by official state agencies. As a result, our election numbers come from a dataset compiled and by members of the Voting Rights Data Institute (VRDI) and MGGG [1]. Additionally, these data exist only at the precinct level, so a Python package developed by MGGG called `maup` was used to disaggregate the data down to the census block level. The disaggregation was prorated by the population of the census block relative to the precinct population. While data for 2016 Ohio House and Senate elec-

[7] A point of defense raised by William Batchelder, the Republican House Speaker at the time, was that the 2011 House districting plan had doubled the number of majority-Black districts when compared to the previous enacted plan. [3]

tions also exist, we do not use those in our evaluations because of incumbency effects that are not homogeneous across the entire state.

3.2 Markov chain implementation

As mentioned previously, we sample from the space of permissible districting plans using Markov chain Monte Carlo (MCMC) methods, which has now become the standard sampling technique for alternative maps. In particular, we use MGGG's GerryChain package, which implements this technique in a way that facilitates generating ensembles of alternative districting plans [11]. The ensembles that we consider here are generated using a Recombination (ReCom) Markov chain. The ReCom algorithm proceeds from one state to the next by randomly selecting two adjacent districts, constructing a random spanning tree over the nodes in the combined district, and identifying an edge in the tree that separates the merged region into two new districts with roughly equal populations [9].[8] ReCom chains were empirically shown to converge quickly to a stable sampling distribution [12][9] and has since become a standard approach.

3.3 A baseline ensemble

We motivate the constraints for our "baseline" state House ensemble from the non-racial and non-partisan aspects of Ohio redistricting law. These constraints are the ones that will be kept in place for all other ensembles that we consider. While it is technically the case that all ensembles serve as baselines in some sense, in this paper, we refer to this race- and party-neutral ensemble in particular as the "baseline" which we will compare other ensembles to. From the Princeton Gerrymandering Project page on Ohio redistricting criteria:

> "In addition to the federal requirements of one person, one vote and the

[8] The state space of the Markov chain is the dual graph of the census block map. That is, nodes correspond to census blocks, and two nodes are adjacent if and only if their corresponding census blocks share a border.

> Voting Rights Act, Ohio's state constitution (Art. XI) requires that state legislative and congressional districts be compact, contiguous, and preserve whole single counties. Furthermore, for state legislative districts only, favoring an incumbent or party is prohibited, and the partisan lean of state legislative districts should be proportional to the statewide preferences of Ohio voters." [2][5]

Accordingly, the baseline constraints are the following:

- Contiguity
- Population deviation bound of 9% (above or below the average district population)
- No more than 88 county splits[9]
- A number of cut edges no more than that of the seed plan (a compactness constraint)[10]

The specific values of a 9% population deviation bound and 88 county splits were chosen because of the decision to use the 2011 enacted state House plan as the seed plan for our ensemble. This saved us the need to generate seed plans ourselves that satisfied contiguity and a given population deviation bound (functionality for doing this is implemented in GerryChain) in addition to a given county split bound (much more difficult to procedurally generate maps that satisfy all three). While the Ohio legal requirement for the amount of allowed population deviation in a district is 5% above or below the average district population, a looser bound of 9% seemed to be necessary for this initial plan to be deemed valid by GerryChain. While indeed, the political geography of Ohio is different, as mentioned before, racial and partisan

[9] If a given county intersects with k districts, we consider that county to have been split $k-1$ times. To obtain the total number of county splits we add this number of splits up for each county across all 88 Ohio counties.

[10] A cut edge is an edge incident upon two nodes that are in different districts. Intuitively, the more cut edges a map has, the longer and more winding its district boundaries are, which means that the districts are not compact.

outcomes of alternative Virginia districting maps did not seem to be affected by the strictness of the population deviation bound [8], so we have opted to settle for the 9% bound. There were a total of 88 county splits in the 2011 enacted state House plan,[11] which served as an upper bound constraint for our ensemble. Finally, the seed plan had 22,475 cut edges, which equated to about 1 in 39.3 of the 883,893 total edges in our graph. This is already reasonably compact, but since the ReCom algorithm tends to draw compact maps without being constrained to do so, the compactness constraint seemed virtually non-binding, seeing as the median number of cut edges was 19,363.5 and 99% of the ensemble's maps had fewer than 20,990 cut edges.

We run 2,500 ReCom steps to generate this baseline ensemble, which takes approximately 10 hours to complete. To decide whether we can run shorter chains and how many initial steps to discard, we examine plots of the running averages and standard deviations of the number of $\geq$50% BVAP districts and the number of Democratic districts (in the 2016 presidential election) in the chain. From the running average plots (Figures 1 and 2 on the next page), we see that these averages begin to stabilize substantially past 1,000 steps. For this reason, we assert that running 1,500 steps for the rest of our chains will give us a sufficiently accurate sample of maps. In Figures 3 and 4 (also on the next page), we look at the standard deviations for the first 1,500 steps of this baseline ensemble after removing some number of initial steps. Often the initial steps of a Markov chain are much more volatile as the chain walks towards its equilibrium distribution, so removing them tends to yield a more representative ensemble. While the number of Democratic districts did not seem to be much more volatile at first, the number of majority-Black districts took a bit more than 400 steps to stabilize. As a result, we have opted to discard the first 500 steps of the 1,500-step chains used for the rest of our state House district ensembles, giving us a collection of 1,000 maps in each ensemble.

[11] This number is low enough where minimizing the number of county splits was clearly a consideration when the enacted map was drawn. 37 counties were not split, 27 counties were split once, 6 counties each were split twice and three times, and 12 counties were split at least four times. A neutral chain without an upper bound on county splits tended to draw maps with roughly 250 total county splits.

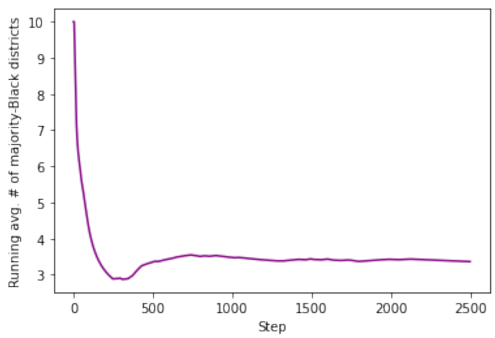

Figure 1: Running average of the number of majority-Black districts.

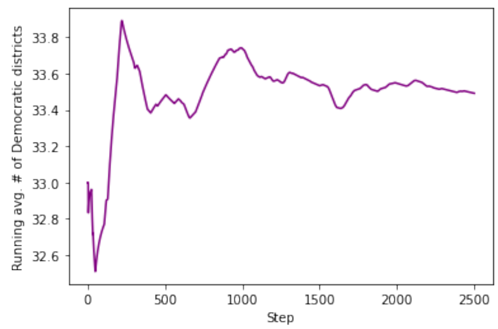

Figure 2: Running average of the number of Democratic districts.

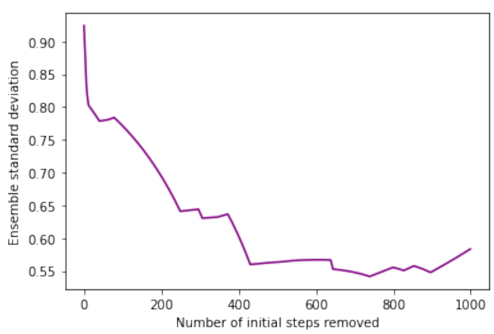

Figure 3: Ensemble standard deviation of the number of majority-Black districts based on number of removed initial steps.

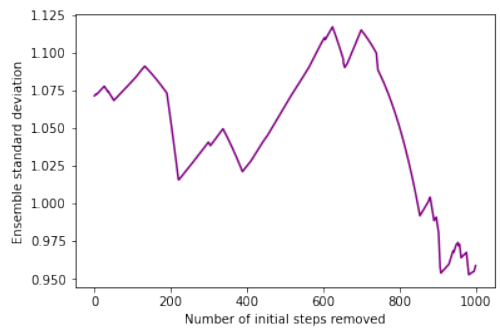

Figure 4: Ensemble standard deviation of the number of Democratic districts based on number of removed initial steps.

3.4 Additionally constrained ensembles

Other ensembles of districting plans with different combinations of constraints were generated for the purpose of comparison. All of these except the proportionality ensemble used the 2011 enacted state House plan as their seed. In particular, the following ensembles were generated and examined:

- **No packing.** In a similar vein to MGGG's Virginia House of Delegates report, we add a constraint where our Markov chain rejects any map where the district with the highest BVAP proportion has a BVAP proportion $\geq$60.5%, which is the BVAP proportion of the highest-BVAP district in the 2011 enacted Ohio state House plan.

- **No backsliding.** In light of evidence that maximizing the number of districts with a BVAP proportion of $\geq$50% has been a historical consideration of state House district drawing committees in Ohio, we add a constraint where our Markov chain rejects any map that has fewer than 10 Democratic districts with a BVAP proportion of $\geq$50%. Such a rejected map would be said to "backslide" from the 2011 enacted plan, which had exactly 10 Democratic majority-Black voter districts. This constraint reflects the idea of retrogression in voting rights law.

- **No packing or backsliding.** This ensemble requires both the no packing and no backsliding constraints from above.

- **No packing or backsliding and create opportunity districts.** This ensemble requires both the no packing and no backsliding constraints from above. In addition, this ensemble requires the creation of at least two "opportunity" districts with a BVAP proportion between 37% (inclusive) and 50% (exclusive).

- **No packing and create opportunity districts.** This ensemble requires the no packing constraint from above and at least 12 Democratic districts with a BVAP proportion $\geq$37%. We are still interested in seeing if relaxing the 50% line somewhat changes the partisan outcome of the ensemble even if it is less likely that redistricting commissions will be considering these exact constraints.

Recall that we are still requiring the contiguity, population deviation, county-split, and compactness constraints in all of these additional ensembles.

3.5 Descriptions of calculated statistics

The Black voting-age population (BVAP) proportion is the primary racial statistic that we examine. This is the ratio between the number of Black residents over 18 years of age and the total number of residents over 18 years of age in a given region. While BVAP alone does not suggest anything about VRA compliance, it is known to be considered in some form by redistricting committees[12] and has been shown to be a good proxy for the ability of a district to elect Black representatives.[13] The distributions of a few partisan statistics in our ensembles will also be examined. These are statistics that Wang identified in a 2016 article that lend themselves to a practicable demonstration of partisan gerrymandering [13]. Let m be the total number of districts in our plans. For a district i, let d_i and r_i respectively denote the total number of Democratic and the total number of Republican votes in that district, let $v_i := d_i/(d_i+r_i)$ denote the Democratic vote share in that district, and let $V := \sum d_i / \sum(d_i + r_i)$ denote the statewide Democratic vote share. Then these statistics are defined as follows:

- **Democratic seat share.** This is the proportion S_D of districts in which Democrats have won more votes than Republicans.

  $$S_D := \frac{|\{i : d_i > r_i\}|}{m}$$

  A comparison of this statistic for the enacted plan to its distribution in our ensembles exactly consitutes the *excess seats test* in Wang's article. This is a test for the presence of an effect from a partisan gerrymander. [13]

[12] See footnote 7.
[13] See footnote 6.

- **Mean-median gap.** This is the difference Δ_M between the median and mean Democratic vote shares across all districts.

$$\Delta_M := \text{median}\{v_i\}_{i=1}^m - \text{mean}\{v_i\}_{i=1}^m$$

  Note that a positive number for this statistic suggests a partisan gerrymander in favor of Democrats. While it is not necessary to realize the distribution of this statistic with an ensemble for the purpose of deeming a map to be a partisan gerrymander, this is still useful for analyzing the partisan effects of the racial constraints that we impose on our ensembles. Without considering any ensembles, this is the statistic used for Wang's *reliable-wins test*, a test for the intent to gerrymander [13].

- **Mean win gap.** This is the difference Δ_W between the mean Republican vote share across all districts that Republicans have won and the mean Democratic vote share across all districts that Democrats have won.

$$\Delta_W := \text{mean}\{r_i : d_i < r_i\} - \text{mean}\{d_i : d_i > r_i\}$$

  Again, a positive number for this statistic suggests a partisan gerrymander in favor of Democrats, and a grouped t-test suffices (that is, generating an ensemble of maps is not necessary) for deeming a map to be a partisan gerrymander using this statistic. Without considering any ensembles, this is the statistic used for Wang's *lopsided outcomes test*, another test for the intent to gerrymander [13].

4 Results

4.1 BVAP proportions in the ensembles

Figures 5 and 6 are plots of the ensembles' BVAP vectors. The color of the boxplots correspond to the particular ensemble that they are for, and the boxplots themselves are plotting the distribution (over all maps in an ensemble) of the BVAP proportions of the x^{th}-ranked district by lowest BVAP. In each boxplot, the box captures the range of values between the 25^{th} and 75^{th} percentiles, and the whiskers span the 1^{st} and 99^{th} percentiles. The dotted horizontal line in the plot is at 37%, which, as explained in Footnote 6, is around the minimum BVAP for a district to comfortably elect Black representatives. Only the districts with the 40 highest BVAP proportions are plotted.

As emphasized in DeFord and Duchin's work with imposing constraints on ensembles of Virginia maps, the neutral baselines for redistricting metrics can vary drastically from state to state based on its laws and its underlying political geography [8]. In our baseline ensemble, we find that this holds, especially when considering the BVAP distributions of the maps' districts. Recall that in MGGG's Virginia report (which considered a part of southeastern Virginia whose BVAP was around 35%), the number of districts in a neutral ensemble with their BVAP above 37% was roughly proportional to the region's BVAP. Here, we find that this is not at all the case. Ohio's Black population in 2010 was 13.1%[14], but the baseline ensemble's maps contained an average of 7.48 Democratic districts out of 99 (7.56%) with BVAP proportions above 37%. Ohio's political geography causes plans generated without factoring in race (but follow non-racial redistricting laws) to lead to maps that would likely elect substantially fewer Black legislators to the Ohio House relative to the Black proportion of Ohio's population.

Somewhat surprisingly, the no-packing ensemble did not improve upon its maps' average number of Democratic districts with BVAP above 37%, with an average of 7.26 districts (7.33%). Looking at the purple boxplots in Figure 5, we see that while the no-packing constraint successfully depresses the BVAP proportions of the top two or three districts relative to the baseline, the remaining $\geq$37% BVAP districts do not have elevated BVAP proportions relative to the baseline. Instead, we observe that the elevation in BVAP occurs between the 80^{th} and 89^{th} sorted districts which had BVAP proportions much lower than 37% in both the baseline and no-packing ensembles. Finally, as one would expect, the ensemble that also includes the 12-opportunity-district constraint (that is, maps must have at least 12 Democratic districts with $\geq$37%

[14]https://www.census.gov/quickfacts/OH

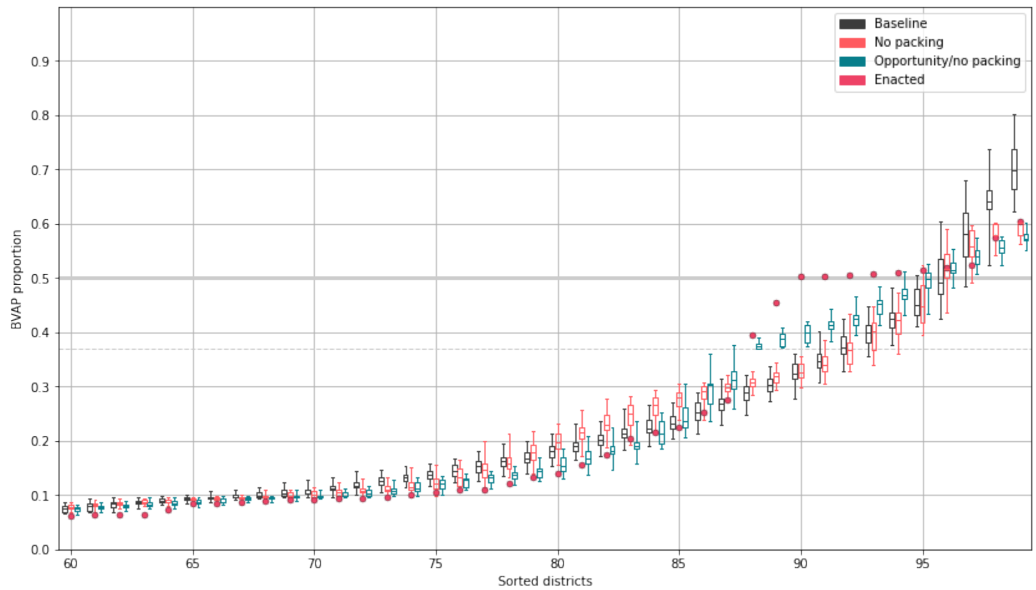

Figure 5: Plot of BVAP vectors for the baseline, no-packing, and opportunity/no-packing ensembles.

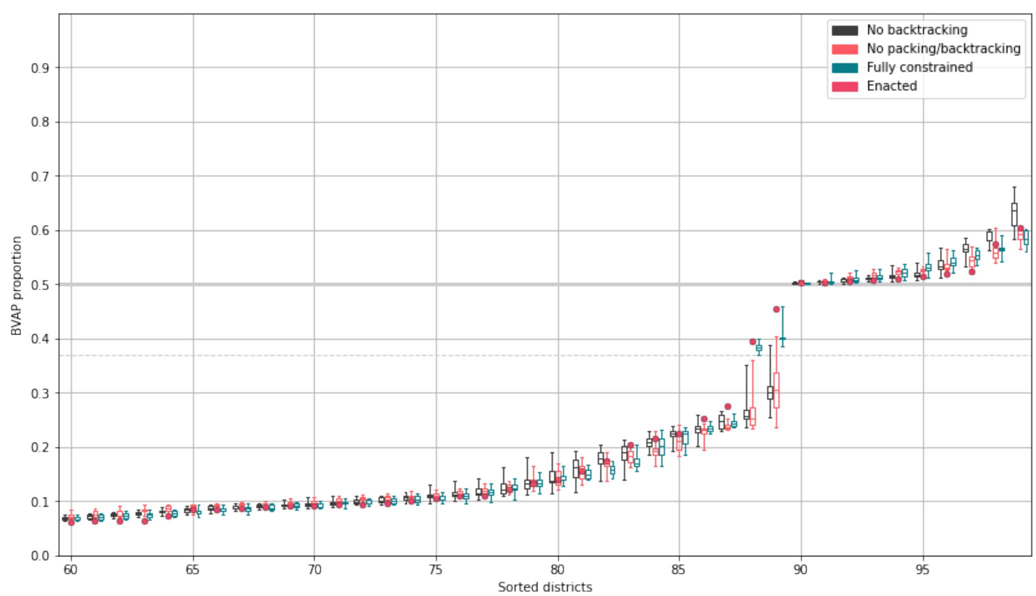

Figure 6: Plot of BVAP vectors for the no-backsliding, no-packing/no-backsliding, and fully constrained ensembles.

BVAP) performs much better with respect to representation, with an average of 12.03 districts (12.15%).

Unsurprisingly as well, the three ensembles with a no-backsliding constraint yield much more racially representative maps. More interestingly, however, each of them tend to do the "bare minimum". All 1,000 maps in all three ensembles contained exactly 10 majority-Black Democratic districts. The no-backsliding ensemble created an average of 0.183 Democratic districts with BVAP proportions between 37% and 50%, the no-packing and no-backsliding ensemble created an average of 0.086 such districts, and the fully constrained ensemble created the bare minimum of two such districts in every one of its maps. The addition of the no-packing constraint to the no-backsliding ensemble again made virtually no impact on the representativeness of its maps. While the no-packing constraint depresses BVAP proportions in its top three districts relative to the no-backsliding ensemble, it is unclear which districts are systematically elevated as a result. We finally remark that the enacted plan follows the fully constrained ensemble remarkably closely in terms of Black representation, suggesting that maximizing the number of majority-Black districts and then creating as many opportunity districts as possible was a likely awareness that the state House redistricting committee in 2011 had.

4.2 Partisan outcomes of the ensembles

Figures 7 and 8 are plots of the ensembles' Democratic vote share vectors for the 2016 presidential election and Figures 9 and 10 are plots for the 2016 U.S. Senate election. The boxplots are set up in exactly the same way as in Figures 5 and 6 where the boxes span the 25$^{\text{th}}$ and 75$^{\text{th}}$ percentiles, and the whiskers span the 1$^{\text{st}}$ and 99$^{\text{th}}$ percentiles. Only the districts with the 60 highest Democratic vote shares are plotted. We use this section to obtain an idea of the ensembles' overall tendencies for allocating the Democratic vote across its maps' districts and to check the effectiveness of each additionally constrained ensemble for explaining vote share "anomalies" from the baseline ensemble.

In the 2016 presidential election, 51.7% of voters in Ohio voted for Republican

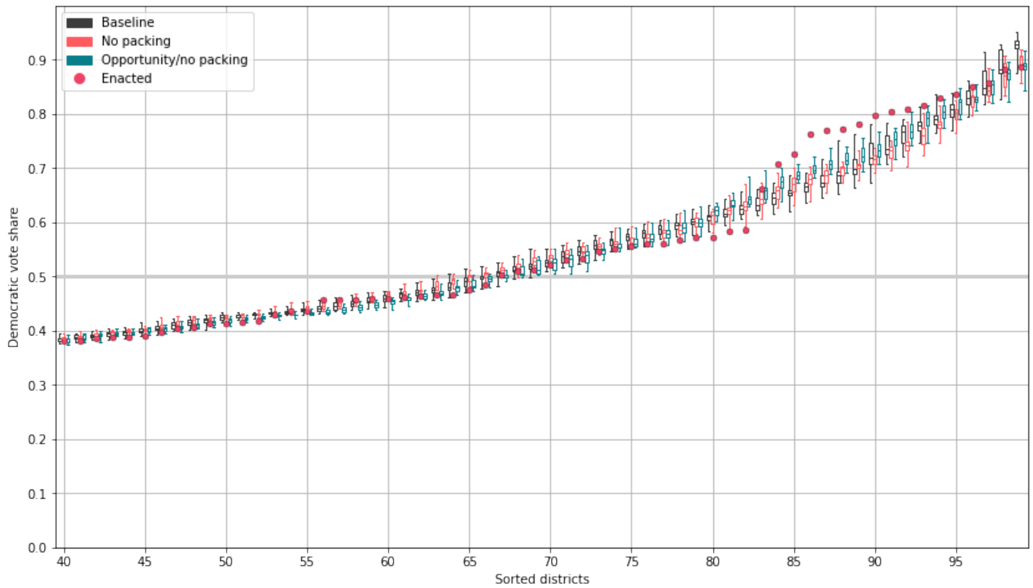

Figure 7: Plot of Democratic vote share vectors using 2016 presidential election data for the baseline, no-packing, and opportunity/no-packing ensembles.

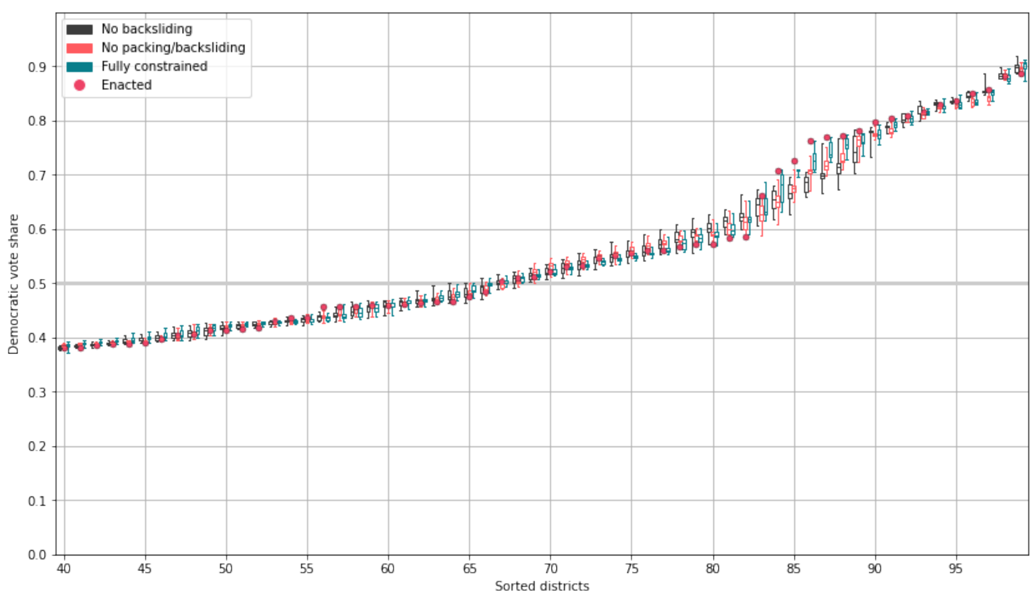

Figure 8: Plot of Democratic vote share vectors using 2016 presidential election data for the no-backsliding, no-packing/no-backsliding, and fully constrained ensembles.

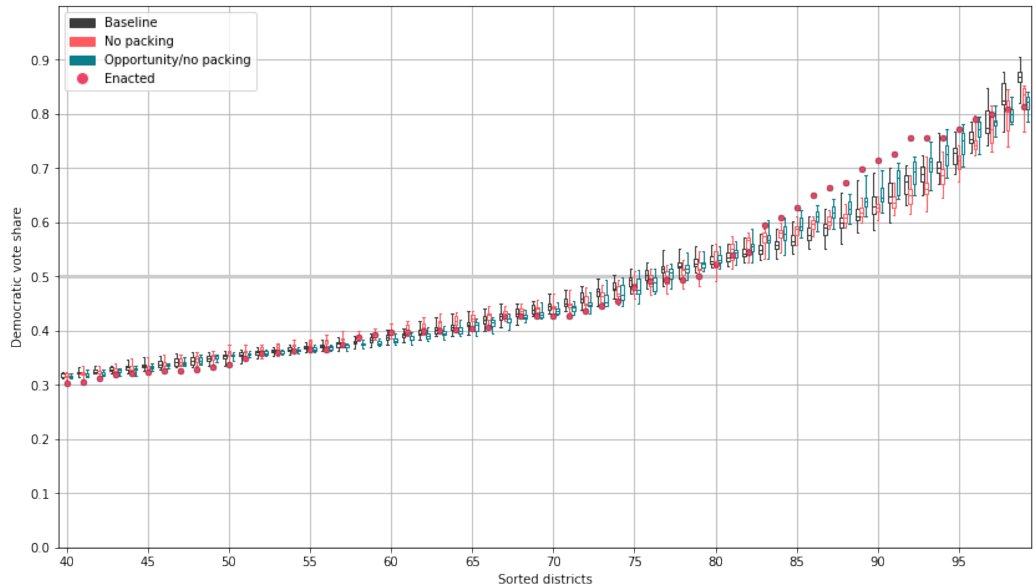

Figure 9: Plot of Democratic vote share vectors using 2016 U.S data. Senate election for the baseline, no-packing, and opportunity/no-packing ensembles.

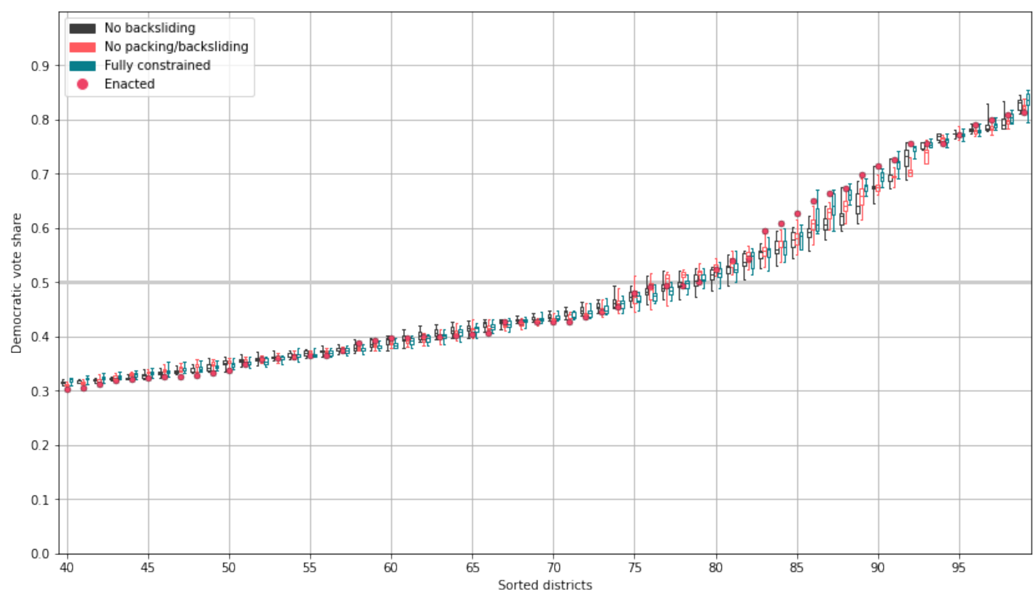

Figure 10: Plot of Democratic vote share vectors using 2016 U.S. Senate election data for the no-backsliding, no-packing/no-backsliding, and fully constrained ensembles.

Donald Trump over the 43.6% who voted for Democrat Hillary Clinton.[15] In the enacted plan, we notice an unusually steep increase in the Democratic vote shares between the 82nd and 86th sorted districts for the 2016 presidential election. This is paired with a depression in the Democratic vote shares between the 75th and 82nd districts (relative to the baseline ensemble). From the enacted plan plot alone, it is unclear what causes the increase. In fact, the constraints imposed on the three ensembles in Figure 7 also fail to explain increase, as the ensembles' plots themselves exhibit a much more steady increase around these districts. Additionally, in looking at Figure 8, the no-backsliding ensemble also fails to capture the anomaly almost entirely. The no-packing and no-backsliding ensemble exhibits only very slightly elevated BVAP proportions around the 86th sorted district. The fully constrained ensemble appears to explain the anomaly most closely, therefore suggesting that it is likely the product of trying to create sufficiently many majority-Black districts *and* subsequently trying to create opportunity districts, which interacts with racial voting patterns.

The 2016 U.S. Senate election in Ohio was between Democrat Ted Strickland and Republican Rob Portman. Portman was the incumbent candidate and won the election with 58.0% of the votes (Strickland received 37.2%),[16] a much larger margin than Trump's win over Clinton in the state for the presidential election. In the U.S. Senate election, relative to the baseline ensemble, we observed elevated Democratic vote shares between the 83rd and 96th sorted districts and depressed Democratic vote shares between the 70th and 79th sorted districts (which happen to have vote shares just below 50%). Again, the no-backsliding, no-packing/no-backsliding, and fully constrained ensembles (Figure 10) tend to explain the anomaly better than the baseline, no-packing, and opportunity/no-packing ensembles (Figure 9), and the fully constrained ensemble comes the closest.

[15] https://ballotpedia.org/Presidential_election_in_Ohio,_2016
[16] https://ballotpedia.org/United_States_Senate_election_in_Ohio,_2016

4.3 Distributions of partisan statistics by ensemble

The purpose of this section is three-fold. We briefly examine the partisan outcomes of the 2011 enacted House plan using the three tests as laid out by Wang in his 2016 article—the excess seats test, the lopsided outcomes test, and the reliable-wins test [13]. We then consider differences in the partisan outcomes between ensembles. Finally, we place our ensembles in the context of Wang's tests and examine the extent to which maps in our ensembles, which do not intend to draw partisan gerrymanders by construction, can nevertheless be considered partisan gerrymanders by these tests.

4.3.1 Democratic seat share and the excess seats test

33 of Ohio's 99 House districts had more voters who voted for Democrat Hillary Clinton than Republican Donald Trump in the 2016 presidential election. We find that this is not statistically significantly different from the mean number of these districts in any of the ensembles. Between the ensembles themselves, all but the no-packing ensemble saw decreases in their Democratic seat shares (by only one seat or so, at most). The no-packing constraint, however, also had no real effect on the number of majority-Black or opportunity districts. The magnitude and direction of the effect of imposing the more binding racial constraints appears to be roughly consistent with DeFord and Duchin's findings in Virginia [8]. Additionally, the fully constrained and the opportunity/no-packing ensembles saw less of a reduction in their average Democratic seat share, with both averages almost exactly equal to the 33 Democratic districts of the enacted plan.

For the 2016 U.S. Senate election, 21 of Ohio's 99 House districts went for Democrat Ted Strickland over Republican Rob Portman. Here, we find that this is statistically significantly less than the mean number of districts in the baseline and the opportunity/no-packing ensembles (respectively, a 2.38-sigma and a 2.08-sigma difference), and not significantly different in any of the other ensembles. In fact, the fully constrained ensemble, which has comported most closely with the enacted plan in every statistic we have calculated thus far, once again lines up extremely well with the

Democratic Ohio House Seats Using 2016 Presidential Election Data

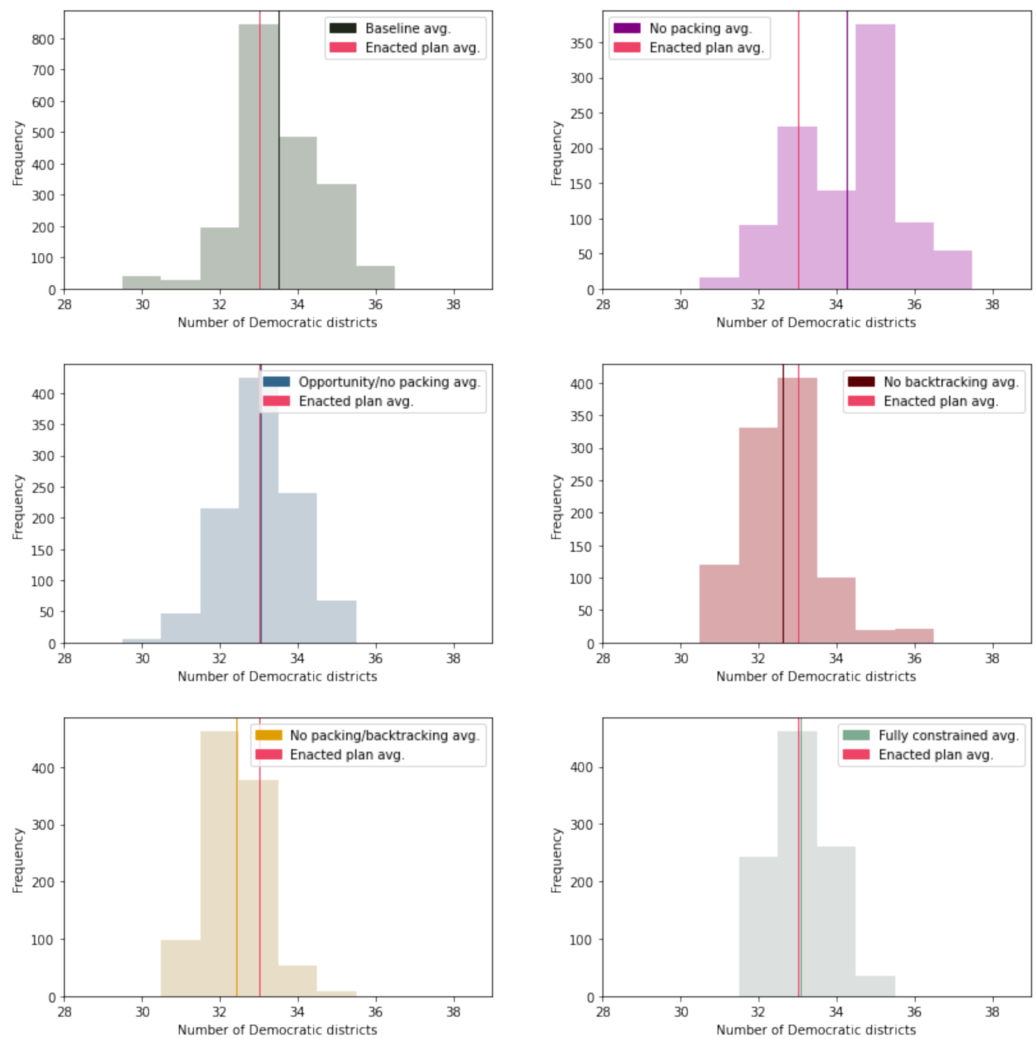

Figure 11: Histograms for the distribution of the number of seats going for Democrat Hillary Clinton in the 2016 presidential election in each ensemble.

Democratic Ohio House Seats using 2016 U.S. Senate Election Data

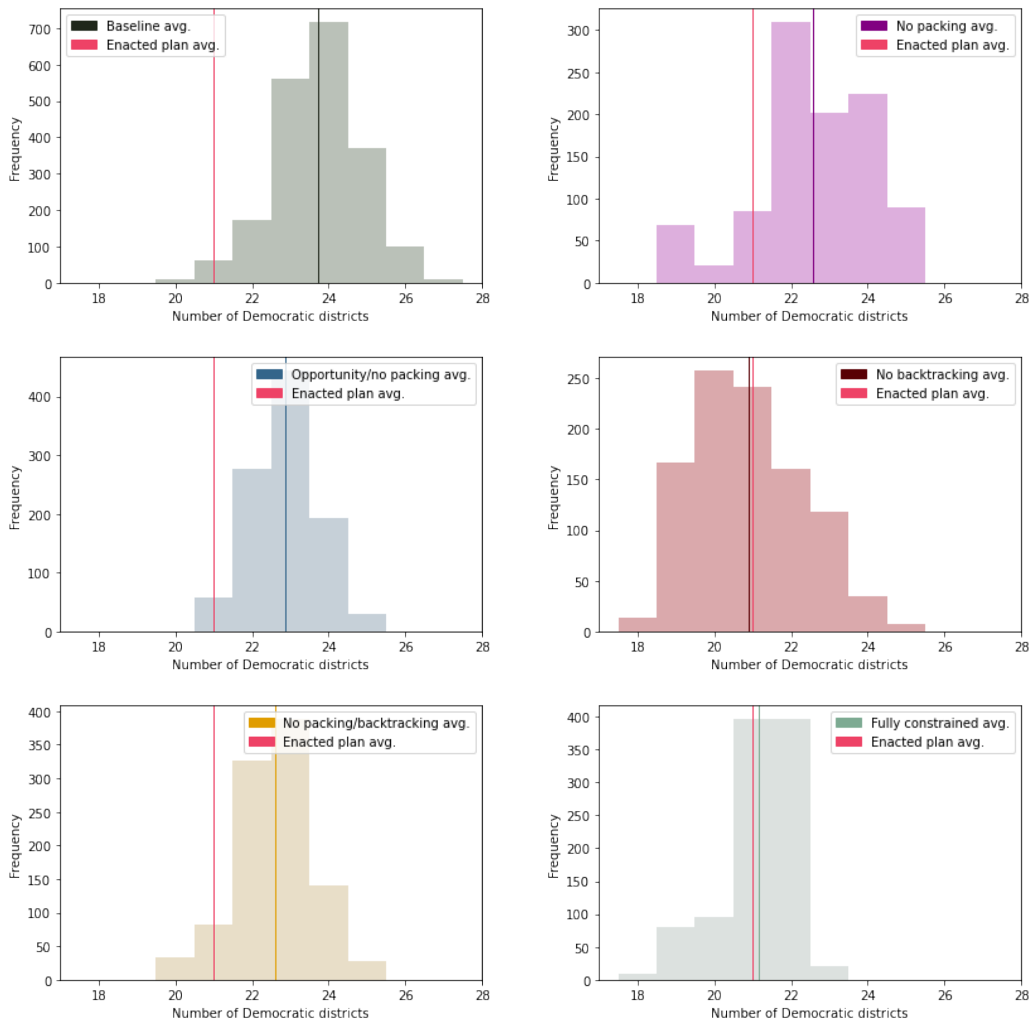

Figure 12: Histograms for the distribution of the number of seats going for Democrat Ted Strickland in the Ohio 2016 U.S. Senate election in each ensemble.

enacted plan, which differs from the ensemble's mean by just 0.154 districts. Thus, whether intentional or not, a partisan gerrymander (relative to a racially neutral ensemble) appears to be the effect of the racial constraints when we consider voting data from the 2016 Senate election. The differences between the ensembles are far more drastic with the Senate elections, with the no-backsliding and fully constrained plans yielding almost 3 fewer Democratic districts than the baseline. Finally, it is notable that the opportunity/no-packing ensemble exhibits the highest Democratic representation aside from the baseline, which is consistent with its behavior with the presidential election data.

Note that even the race- and party-neutral baseline Markov chain draws maps that give fewer districts than what would be expected from a *double-bonus norm*. A double-bonus norm dictates that a party that gets 50% of the votes gets 50% of the seats, and for every percentage point above (or below) 50%, the party gains (or loses) 2% of the seats. This norm (rather than a proportional norm where the seat share and the vote share are the same) has been empirically observed to hold by Goedert in 2014 over the past 40 years of data for elections with an outcome near 50-50 [10].[17] Once again, let us consider 2016 presidential election, where 51.7% of voters in Ohio voted for Republican Donald Trump over the 43.6% who voted for Democrat Hillary Clinton.[18] Out of the voters who voted either Republican or Democratic, Clinton received 43.6% / (43.6% + 51.7%) = 45.75% of the vote. Then, under a double-bonus norm, Clinton would have won in 41.5% of the 99 House districts (i.e. in about 41 of them), which is 5 more districts than the map drawn by our baseline ensemble that yields the strongest Democratic representation, about 7.5 more than the ensemble mean, and 8 more than the enacted plan. We can similarly show that in the Senate election, Strickland would have won in about 27 districts via a double-bonus norm,[19]

[17] The political geography of most states does not allow for a proportional seats-votes norm to arise out of neutral redistricting.

[18] See footnote 15.

[19] Note that Goedert actually fits a probit model over the seats-vote relationship since the natural seats-votes curve tends to follow more of a sigmoidal curve, becoming less steep as one moves away from 50-50 [10]. As a result, since Strickland's vote share was sufficiently low, 27 districts is fewer than the number of districts Strickland would win under the more accurate probit norm.

which we obtain in only 8 of the 2,000 maps in our baseline ensemble (and is 6 more districts than what we obtain from the enacted plan).

4.3.2 The mean-median gap and the reliable-wins test

We only calculate and plot the mean-median gap for the presidential election because the discrepancy between the Democratic and Republican vote shares in the U.S. Senate election is large enough to require a different test for the presence of partisan gerrymandering [13]. Recall that positive values of the mean-median gap correspond to bias in favor of Democrats (and negative values correspond to bias in favor of Republicans). Relative to all ensembles except the no-backsliding and the opportunity/no-packing ones, the enacted plan has a statistically significantly more negative mean-median gap. Notably, we observe significance for both the baseline ensemble (2.82-sigma difference) *and the fully constrained ensemble* (2.30-sigma difference) which thus far has otherwise tracked very closely with the enacted plan. Among the different ensembles, aside from the no-packing ensemble, the racial constraints seem to lead to a mean-median gap that is 0.5% to 0.7% more negative (in absolute terms). That is, the racially constrained ensembles' maps (again, minus the no-packing ensemble) all exhibit fairly similar means within about 0.2 percentage points of each other. They instead differ in their variance, with more strictly constrained ensembles have a lower mean-median gap variance. The differences in variance better explain why the enacted plan's mean-median gap was only significantly different from the average mean-median obtained in two of these four ensembles.

To run Wang's reliable-wins test directly, we first find that the standard deviation of the Democratic vote share across all 99 House districts in the 2016 presidential election is 18.2%. The standard error of our mean-median gap estimator is then $18.2\% \times 0.756/\sqrt{99} \approx 1.38\%$ [13]. Thus, a districting plan with a mean-median gap outside of $\pm 1.96 \times 1.38\% \approx \pm 2.71\%$ is far enough from 0 for the reliable-wins test to identify the plan as a partisan gerrymander at a 5% significance level. This not only means the enacted plan's mean-median gap of −4.60% has was identified by the test as a Republican gerrymander, but also, that virtually every map in every ensemble

Mean-Median Gaps in Ensembles of Alternative Ohio House Plans

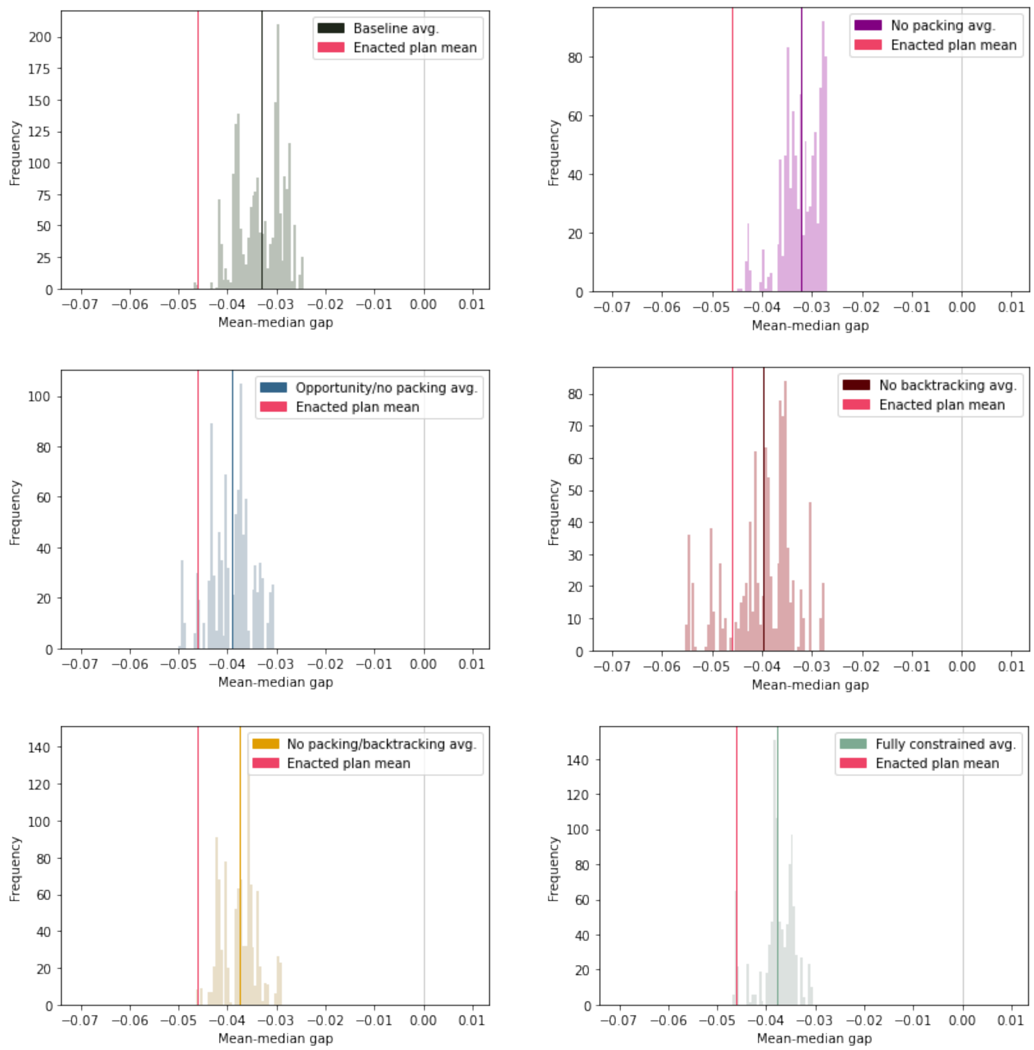

Figure 13: Histograms for the distribution of the mean-median gap statistic in each ensemble.

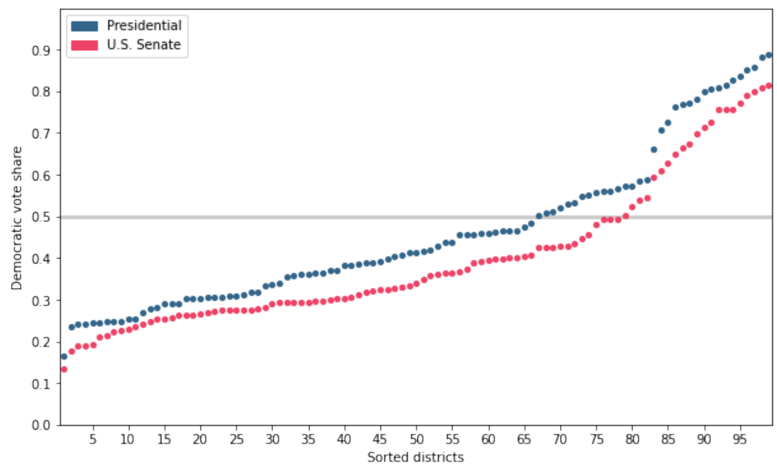

Figure 14: Plot of Democratic vote shares in the presidential election compared to the U.S. Senate election by sorted district, demonstrating the incumbency effect.

would also be identified as Republican gerrymanders. In other words, the political geography of Ohio exists in such a way where almost all party-blind alternative maps would have median districts with a lower Democratic vote share than average, which gives Republicans more districts where they can secure “narrow-but-reliable” wins [13].

4.3.3 The mean win gap and the lopsided outcomes test

Once again, we only consider the presidential election. The incumbency effect for the U.S. Senate election seems to affect each district roughly equally, lowering its Democratic vote share by about 7%. This reduces the average Democratic win and increases the average Republican win, giving us a mean win gap whose positive bias can be almost entirely explained by incumbency instead of more nuanced and less uniform differences in the Democratic vote share by region. The enacted plan’s mean win gap is not significantly different from the means of any ensemble except for the no-packing ensemble, which exhibited a 2.22-sigma difference. As with the mean-median gap and the presidential Democratic seat share, racially constrained ensembles other than the no-packing ensemble exhibited more negative mean win gaps;

Mean Win Gaps in Ensembles of Alternative Ohio House Plans

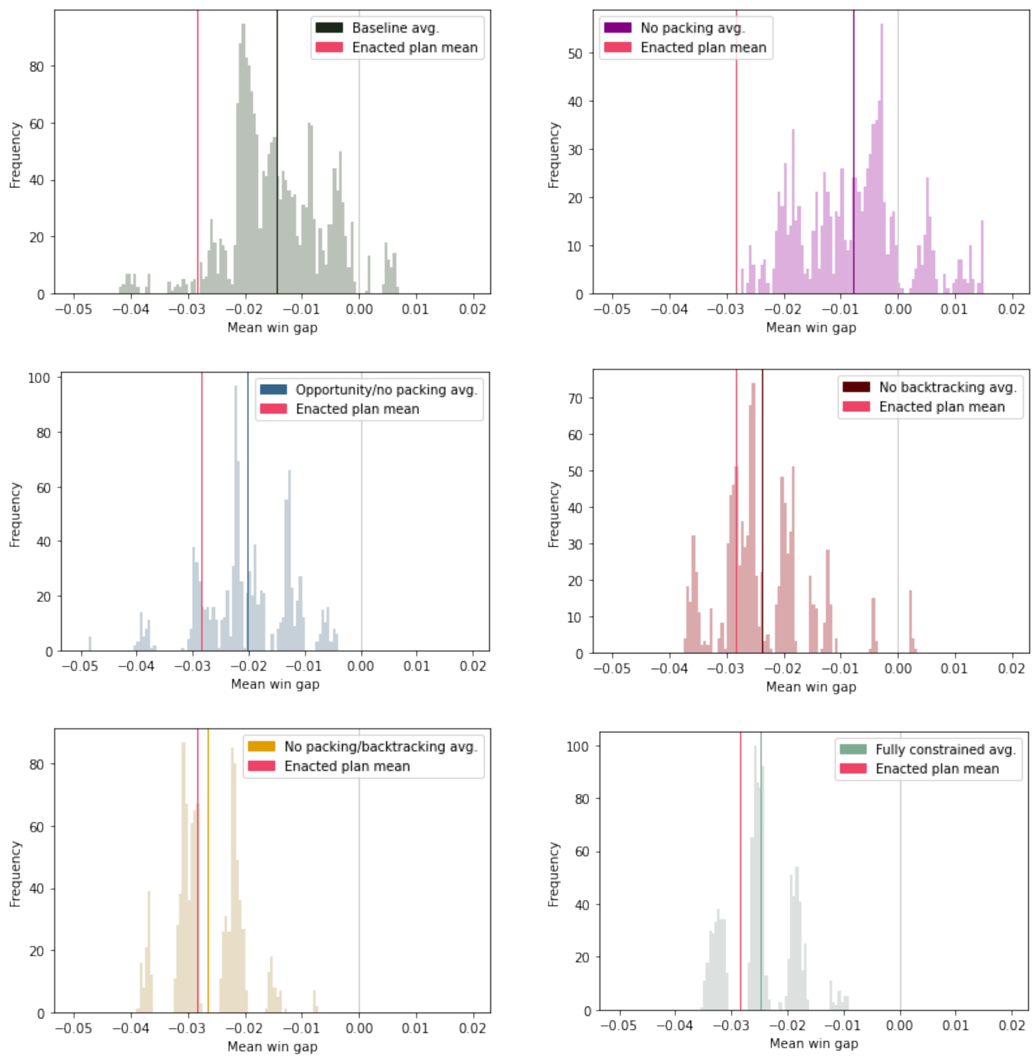

Figure 15: Histograms for the distribution of the mean win gap statistic in each ensemble.

here, by about 0.7% to 1.2% (again, in absolute terms). Once more, out of these four ensembles, the opportunity/no-packing ensemble' maps appear to have the smallest average mean win gap.

To run Wang's lopsided outcomes test directly, we perform a paired t-test on the set of Democratic vote shares in the districts that Democrats win against the set of Republican vote shares in the districts that Republicans win. We get an insignificant difference ($p \approx 0.27$). Out of all 6 ensembles, only five maps in the opportunity/no-packing ensemble ended up having significantly different average wins between Democrats and Republicans 5% significance level (all with $p \approx 0.04$).

4.4 Achieving better racial and partisan outcomes

We finally present a contiguous *and* reasonably compact map that achieves better racial representation (more opportunity districts) and partisan representation (much closer adherence to a double-bonus norm). The existence of such a map shows that, in practice, fairer racial outcomes do not need to lead to less fair partisan outcomes when manually drawing maps. This map was obtained from making a few adjustments to an existing map on Dave's Redistricting App (DRA) with districts that achieve proportional representation.[20] Partisan outcomes for this map conform nearly perfectly to a double-bonus norm. Using DRA's default composite 2016-2020 election results, the 45.92% two-party Democratic vote share yields an expected 41.57 Democratic seats (41.99%; a double-bonus norm yields 41.84% of seats). Additionally, the map also has 12 Black voter opportunity districts, which is the same number of such districts in the 2011 enacted plan. Our map, as well as a rather comprehensive set of metrics for the map, can also be found on Dave's Redistricting App (see footnote for URL).[21]

[20] https://davesredistricting.org/maps#viewmap::27fc7c3a-c9cd-4144-babb-498888a25968
[21] https://davesredistricting.org/join/fd4b5a33-950b-41e4-a629-3b9f147a6db7

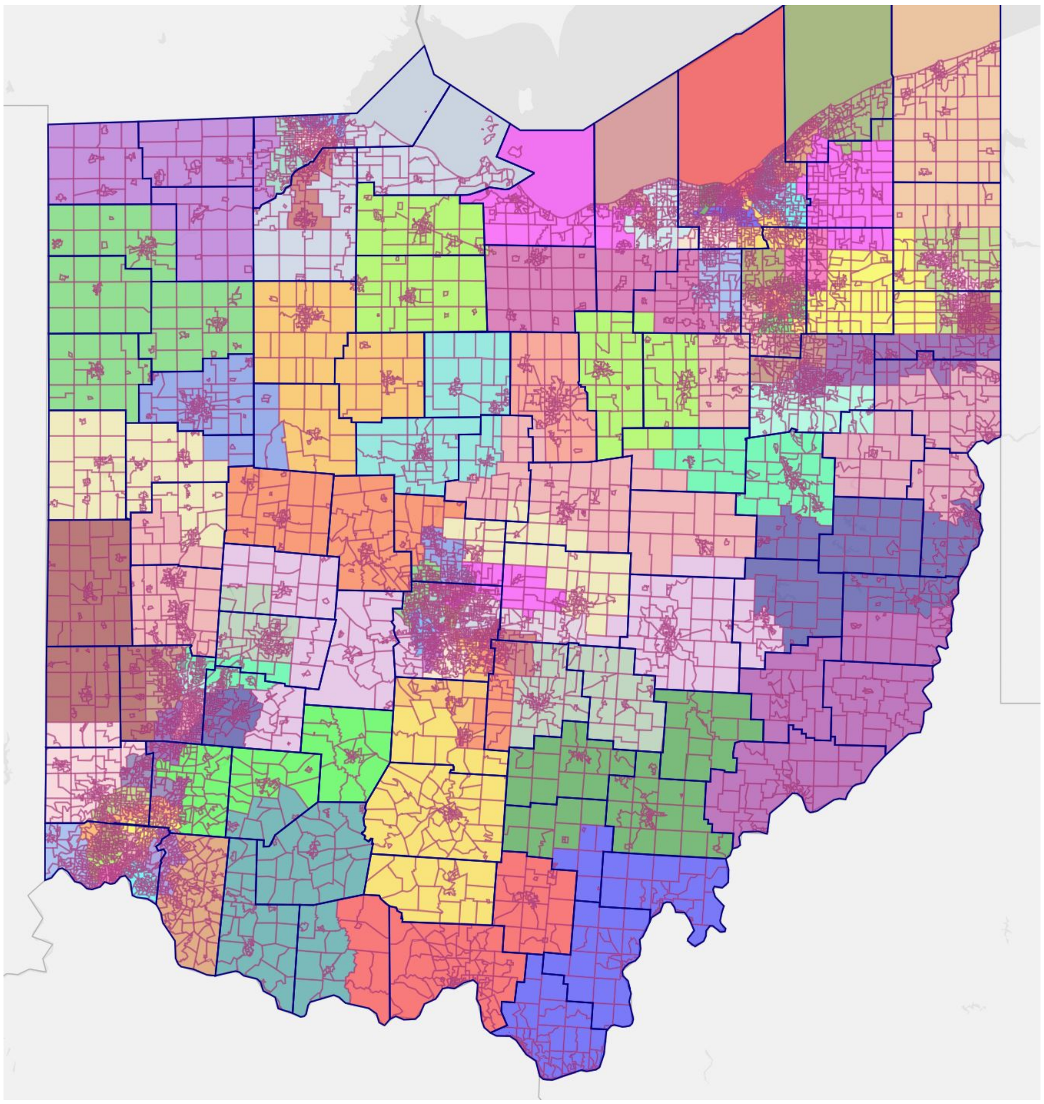

Figure 16: A plan for Ohio state House districts with improved racial *and* partisan representation over maps in our baseline ensemble.

5 Conclusions

This study has primarily served as an evaluation of racial and partisan outcomes of several ReCom Markov chain-generated ensembles of alternative plans for Ohio House of Representatives districts. The ensembles we examine differ only in the racial constraints imposed on its maps. We first find that completely race-neutral ensembles generate maps with a disproportionately low number of districts with a Black voting-age population proportion high enough to reliably elect Black representatives (37%), a pure consequence of Ohio's political geography. Additionally, we find that the racially constrained ensembles did not tend to contain maps that were significantly more racially representative than what was required by the constraints.

The different combinations of racial constraints also enabled us to test hypotheses about considerations that the redistricting committee had when drawing the 2011 enacted House district plan. We find that the means of our fully constrained ensemble track the enacted plan closely in every statistic we examine except for the mean-median gap, suggesting that creating many majority-Black districts *and* a few opportunity districts as well was a priority.

We also generally find that ensembles with constraints that resulted in better racial representation contained maps that were slightly more biased against Democrats relative to the baseline ensemble, noting that we find slightly lower numbers of Democratic districts and more negative values of the partisan statistics we consider. Out of these racially constrained ensembles, our ensemble that constrains for opportunity districts instead of majority-Black districts consistently exhibits slightly lower levels of bias against Democrats. However, we find that the natural Democratic disadvantage is much greater. Not only does the baseline ensemble already exhibit an incredibly significant representative deviation from a double-bonus norm—nearly all of its maps have a mean-median gap low enough to be considered outside the zone of chance by Wang's reliable-wins test [13]. The test therefore may not be sufficient to show intent to gerrymander for drawing Ohio House plans, given the behavior of neutral ensembles there. Of course, this result does not invalidate this test in general, but rather, emphasizes the need to consider multiple indicators for gerrymandering in practice. We finally note that the 2011 enacted plan is not identified as a partisan gerrymander using the excess seats and mean wins test, but is deemed a Republican gerrymander using the reliable-wins test. Despite the fact that almost all of our maps across all 6 of our ensembles are also identified as Republican gerrymanders using the reliable-wins test, the −4.60% mean-median gap that the plan exhibits is also statistically significantly lower than the baseline and fully constrained ensemble means (respectively, −3.32% and −3.76%), as well as the means of two of the other ensembles.

The work done here gives us some insight into the implications of Ohio's political geography on redistricting for state House districts. These results may prove to be useful for participants in the public input process during Ohio's redistricting cycle in 2021. However, much more can still be learned. Immediate future work that expands on these findings here could apply the same analyses to other states with drastically different political geographies and landscapes, or to Ohio state Senate and congressional district plans, or to different elections in different years, or with the recently released 2020 Census data. A particularly interesting tidbit of Ohio redistricting law for the 2021 cycle is the requirement that "the partisan lean of state legislative districts should be proportional to the statewide preferences of Ohio voters" [2][5]. Since proportional representation does not tend to occur our of party-neutral redistricting,[22] we are certain that examining an ensemble of maps with representationally proportional districts would reveal interesting and relevant conclusions for the upcoming redistricting cycle. Different constraints that more directly model constitutionality and VRA-compliance could also be explored and applied to redistricting in Ohio. A more direct examination of Ohio's political geography could be fruitful as well. The takeaways from such work might strengthen, add nuance to, or simply refute the results of this study. Regardless, any work towards trying to eliminate gerrymandering and other issues with representation in the United States is work that brings us closer to "a more perfect Union." [4]

[22]See footnotes 17 and 19.

References

[1] Ohio Election Shapefiles. [Online]. Available: https://github.com/mggg-states/OH-shapefiles

[2] Princeton Gerrymanding Project - Ohio Page. [Online]. Available: https://gerrymander.princeton.edu/reforms/OH

[3] Redistricting in Ohio after the 2010 census. [Online]. Available: https://ballotpedia.org/Redistricting_in_Ohio_after_the_2010_census

[4] The Constitution of the United States: A Transcription. [Online]. Available: https://www.archives.gov/founding-docs/constitution-transcript

[5] The Ohio Legislature - Ohio Constitution - Section XI.03. [Online]. Available: https://www.legislature.ohio.gov/laws/ohio-constitution/section?const=11.03

[6] A. Becker, M. Duchin, D. Gold, and S. Hirsch, "Computational Redistricting and the Voting Rights Act," 2020.

[7] J. Chen. (2016) Expert Report of Jowei Chen, Ph. D. [Online]. Available: https://s10294.pcdn.co/wp-content/uploads/2016/05/Expert-Report-of-Jowei-Chen.pdf

[8] D. DeFord and M. Duchin, "Redistricting Reform in Virginia: Districting Criteria in Context," 2019.

[9] D. DeFord, M. Duchin, and J. Solomon, "Recombination: A family of Markov chains for redistricting," 2018.

[10] N. Goedert, "Gerrymandering or geography? How Democrats won the popular vote but lost the Congress in 2012," vol. 1, no. 1, p. 2053168014528683, publisher: SAGE Publications Ltd. [Online]. Available: https://doi.org/10.1177/2053168014528683

[11] Metric Geometry and Gerrymandering Group, "Gerrychain." [Online]. Available: https://github.com/mggg/GerryChain

[12] ——, "Comparison of Districting Plans for the Virginia House of Delegates," 2018.

[13] S. S.-H. Wang, "Three Tests for Practical Evaluation of Partisan Gerrymandering," 2016.

[14] S. S.-H. Wang, R. F. Ober, and B. Williams, "Laboratories of Democracy Reform: State Constitutions and Partisan Gerrymandering," 2019.

A Code for executing Markov chain simulations

```
##### IMPORTS #####
from functools import partial
import json
from datetime import datetime
import pickle
import numpy as np
import pandas as pd
import scipy.stats as st
import geopandas as gpd
import matplotlib.pyplot as plt

import gerrychain
import maup

from gerrychain import (
    Election,
    Graph,
    MarkovChain,
    Partition,
    accept,
    constraints,
    updaters,
    tree,
)

from gerrychain.metrics import efficiency_gap, mean_median
from gerrychain.proposals import recom
from gerrychain.tree import recursive_tree_part
from gerrychain.updaters import cut_edges
##################

election_names = ["PRES16", "SEN16", "USH16", "SSEN16", "STH16"]
election_columns = [
    ["PRES16D", "PRES16R"],
    ["SEN16D", "SEN16R"],
    ["USH16D", "USH16R"],
    ["SSEN16D", "SSEN16R"],
    ["STH16D", "STH16R"],
]

vap_columns = ['totVAP', 'WVAP', 'BVAP', 'AVAP', 'NatVAP', 'HVAP']
pop_col = 'tot'
```

```python
county_col = 'COUNTY'
assignment = 'HDIST'

shapefile_path = './data/oh_blocks.shp'
ensemble_output_path = './ensembles/sth_1500step_nopacking_opportunity.pkl'

# construct GerryChain Graph object, set null values to 0
graph = Graph.from_file(shapefile_path)
gdf = gpd.read_file(shapefile_path)

for x in election_columns:
    for i in range(len(graph.nodes)):
        if np.isnan(graph.nodes[i][x[0]]):
            graph.nodes[i][x[0]] = 0
        if np.isnan(graph.nodes[i][x[1]]):
            graph.nodes[i][x[1]] = 0

# calculate ideal population for districts
num_districts = gdf[assignment].nunique()
ideal_population = sum(gdf["tot"].fillna(0.0)) / num_districts

# construct initial Partition
elections = [
    Election(
        election_names[i],
        {"Democratic" : election_columns[i][0], "Republican" : election_columns[i][1]},
    )
    for i in range(len(election_columns))
]

my_updaters = {
    "population": updaters.Tally(pop_col, alias="population"),
    "cut_edges": cut_edges,
    "num_county_splits": updaters.county_splits(
        "num_county_splits",
        county_field_name=county_col
    ),
}

vap_updaters = {vap_column : updaters.Tally(vap_column) for vap_column in vap_columns}
my_updaters.update(vap_updaters)

election_updaters = {election.name : election for election in elections}
my_updaters.update(election_updaters)

initial_partition = Partition(graph, assignment=assignment, updaters=my_updaters)
```

```python
# construct MarkovChain and define necessary constraints
proposal = partial(
    recom, pop_col=pop_col, pop_target=ideal_population, epsilon=0.02, node_repeats=2
)

# upper-bound the number of cut edges by the number of cut edges in the initial partition
compactness_bound = constraints.SelfConfiguringUpperBound(
    lambda p: len(p["cut_edges"])
)

# upper-bound the number of county splits by the number of splits in the initial partition
def num_co_splits(partition):
    return sum(
        [len(partition['num_county_splits'][key].contains) - 1
         for key in partition['num_county_splits']]
    )

co_split_bound = constraints.SelfConfiguringUpperBound(num_co_splits)

# lower-bound the number of 37%-50% BVAP districts
def num_37_50_bvap_districts(partition):
    bvap = np.array(list(partition['BVAP'].values()))
    pct_bvap = bvap / np.array(list(partition['totVAP'].values()))
    df = pd.DataFrame(
        {
            'BVAP': pct_bvap,
            'PRES16': list(partition['PRES16'].percents('Democratic'))
        }
    )
    cond = (0.37 <= df['BVAP']) & (df['BVAP'] < 0.5) & (df['PRES16'] >= 0.5)
    return cond.sum()

bvap_37_50_bound = constraints.SelfConfiguringLowerBound(num_37_50_bvap_districts)

# lower-bound the number of Democratic majority-Black districts
def num_50plus_bvap_districts(partition):
    bvap = np.array(list(partition['BVAP'].values()))
    pct_bvap = bvap / np.array(list(partition['totVAP'].values()))
    df = pd.DataFrame(
        {
            'BVAP': pct_bvap,
            'PRES16': list(partition['PRES16'].percents('Democratic'))
        }
    )
    cond = (df['BVAP'] >= 0.5) & (df['PRES16'] >= 0.5)
```

```python
    return cond.sum()

50plus_bvap_bound = constraints.SelfConfiguringLowerBound(num_50plus_bvap_districts)

# lower-bound the number of opportunity districts
def num_opp_districts(partition):
    bvap = np.array(list(partition['BVAP'].values()))
    pct_bvap = bvap / np.array(list(partition['totVAP'].values()))
    df = pd.DataFrame(
        {
            'BVAP': pct_bvap,
            'PRES16': list(partition['PRES16'].percents('Democratic'))
        }
    )
    cond = (df['BVAP'] >= 0.37) & (df['PRES16'] >= 0.5)
    return cond.sum()

opportunity_bound = constraints.SelfConfiguringLowerBound(num_opp_districts)

# upper-bound the % BVAP of all districts by the max % BVAP in the initial plan (60.5%)
def max_bvap(partition):
    bvap = np.array(list(partition['BVAP'].values()))
    pct_bvap = bvap / np.array(list(partition['totVAP'].values()))
    return np.max(pct_bvap)

max_bvap_bound = constraints.SelfConfiguringUpperBound(max_bvap)

# comment out constraints that are not used
chain = MarkovChain(
    proposal=proposal,
    constraints=[
        constraints.within_percent_of_ideal_population(initial_partition, 0.09),
        compactness_bound,
        co_split_bound,
        bvap_37_50_bound,
        50_plus_bvap_bound,
        opportunity_bound,
        max_bvap_bound,
        constraints.no_more_discontiguous,
    ],
    accept=accept.always_accept,
    initial_state=initial_partition,
    total_steps=1500,
)

# execute chain
```

```
generated_chain = []
for partition in chain:
    d = {key: partition[key] for key in my_updaters.keys()}
    generated_chain.append(d)

# process and pickle chain
def election_results_dict(partition, election_name):
    return {
        'wins': partition[election_name].wins("Democratic"),
        'percents': partition[election_name].percents("Democratic"),
        'mean_median': partition[election_name].mean_median(),
        'mean_thirdian': partition[election_name].mean_thirdian(),
        'efficiency_gap': partition[election_name].efficiency_gap(),
        'partisan_bias': partition[election_name].partisan_bias(),
        'partisan_gini': partition[election_name].partisan_gini(),
    }

def process_partition(partition):
    return {
            'num_county_splits': num_co_splits(partition),
            'totVAP': partition['totVAP'],
            'WVAP': partition['WVAP'],
            'BVAP': partition['BVAP'],
            'AVAP': partition['AVAP'],
            'NatVAP': partition['NatVAP'],
            'HVAP': partition['HVAP'],
            'PRES16': election_results_dict(partition, 'PRES16'),
            'SEN16': election_results_dict(partition, 'SEN16'),
            'USH16': election_results_dict(partition, 'USH16'),
            'SSEN16': election_results_dict(partition, 'SSEN16'),
            'STH16': election_results_dict(partition, 'STH16'),
        }

processed_chain = []
for i in range(len(generated_chain)):
    processed_chain.append(
        {
            'num_county_splits': num_co_splits(generated_chain[i]),
            'num_cut_edges': len(generated_chain[i]['cut_edges']),
            'totVAP': generated_chain[i]['totVAP'],
            'WVAP': generated_chain[i]['WVAP'],
            'BVAP': generated_chain[i]['BVAP'],
            'AVAP': generated_chain[i]['AVAP'],
            'NatVAP': generated_chain[i]['NatVAP'],
            'HVAP': generated_chain[i]['HVAP'],
            'PRES16': election_results_dict(generated_chain[i], 'PRES16'),
```

```
            'SEN16': election_results_dict(generated_chain[i], 'SEN16'),
            'USH16': election_results_dict(generated_chain[i], 'USH16'),
            'SSEN16': election_results_dict(generated_chain[i], 'SSEN16'),
            'STH16': election_results_dict(generated_chain[i], 'STH16'),
        }
    )

with open(ensemble_output_path, 'wb') as f:
    pickle.dump(processed_chain, f)
```