Comparing Empirical Performance of Automated Redistricting Algorithms Using Measures of Compactness and Partisan Fairness: A Case Study of 2021 Congressional Redistricting in Virginia

Madeleine Goertz

International Community School

AP Research

Randall S. Huberman

May 20th, 2021

Comparing Empirical Performance of Automated Redistricting Algorithms Using Measures of Compactness and Partisan Fairness: A Case Study of 2021 Congressional Redistricting in Virginia

Redistricting

- Explain what districts are
- Explain how the redistricting process generally works

Gerrymandering

- Explain what gerrymandering is.

Ways to combat gerrymandering

Automated Redistricting Algorithms

- Explain what the goals of these things are

Metrics to detect gerrymandering

- Explain what the goals of these are

Overview of Method

- Actually introduce the rest of the paper.

Literature Review

I will provide an overview of some of the current political science research into automated redistricting algorithms as well as measures of partisan fairness and of compactness as they pertain to my research.

Automated Redistricting Algorithms

The purpose of automated redistricting algorithms is to generate a set of redistricting plans that are as "impartial" as possible (Chen & Rodden, 2013). While many different algorithms have been proposed (see Altman and Mcdonald (2009), Haas et al. (2020), Lara-Caballero et al. (2019), Macmillan (2001), Weaver and Hess (1963), and Xiao (2008)), I will provide high-level overviews of two algorithms: Sequential Monte Carlo and Compact Random Seed Growth.

First, I explain how these algorithms conceptualize the redistricting problem.

Redistricting as Graph Cutting

The following redistricting algorithms all conceptualize redistricting precincts as a graph-cutting problem. For the uninitiated, a graph is a network of different interconnected points, where the points are called "nodes" and the lines connecting them are called "edges" (Fifield, Higgins, et al., 2020). Every precinct is a node, and geographically-adjacent precincts have their corresponding nodes connected by edges. Since the goal of redistricting is to assign every precinct a district, the algorithms imagine that edges between nodes are "cut" until "islands" (known as "sub graphs") are formed, which each is disconnected from the rest. The disconnected "sub graphs" then become the districts. Figure 1 provides a nice visualization of this representation with a sample set of 50 precincts(Fifield, Higgins, et al., 2020).

Sequential Monte Carlo

The second automated redistricting algorithm that I'm going to discuss is an implementation of a statistical method known as "Sequential Monte Carlo," henceforth

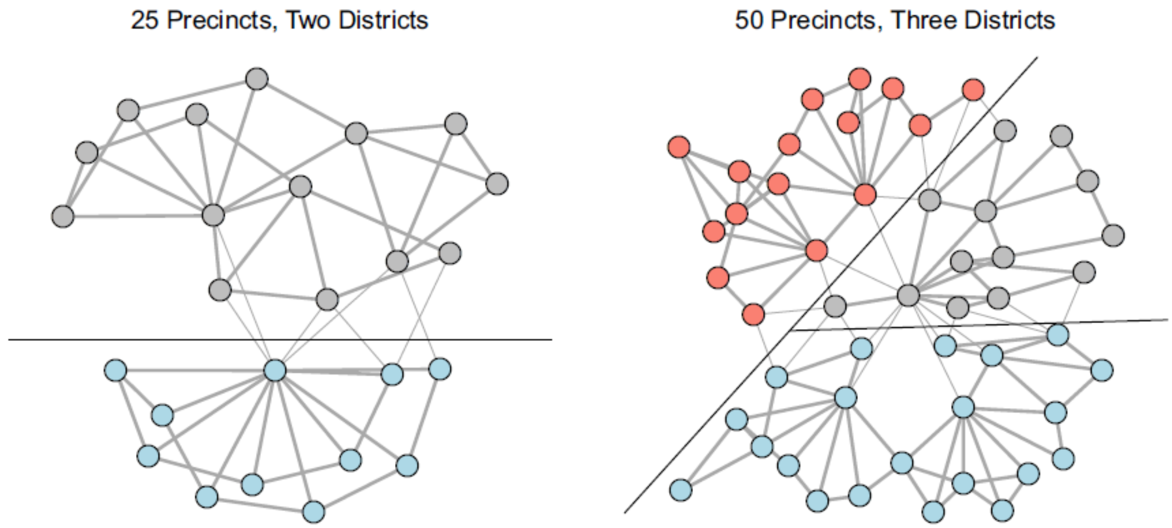


Figure 1

Representation of redistricting as graph cutting. Every node is a precinct, and nodes that share an edge are known to be adjacent precincts. The algorithms "cut away" edges between nodes until islands of districts are formed. (Fifield, Higgins, et al., 2020, p. 3)

SMC (McCartan & Imai, 2020).[1] The following overview is meant to provide a high-level understand of how the basic algorithm works. [2].

Just like MCMC, SMC conceptualizes the electoral map as a mathematical graph with precincts as nodes and edges connecting geographically-adjacent precincts. It also uses this graph-cutting concept, but SMC specifically uses something called a "spanning tree" (see Figure 2), which is a graph that is connected by the minimum number of possible edges[3].

Figure 3 visualizes the iterative splitting procedure used by SMC. First, compute

[1] I will refer to the *redistricting algorithm* that uses Sequential Monte Carlo as "SMCMC," rather than the *statistical method* Sequential Monte Carlo.

[2] Please see the section "The Proposed Algorithm" in the paper for an in-depth, mathematically-rigorous explanation. (McCartan & Imai, 2020, p. 13)

[3] Put another way, if any edge is cut from the graph, the graph will be split into two sub graphs.

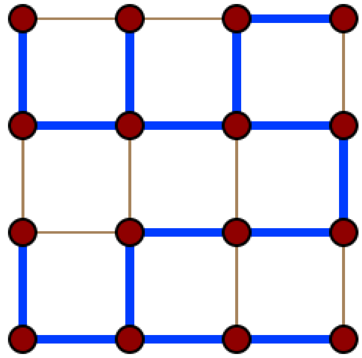


Figure 2

An example spanning tree. The spanning tree consists of all nodes and the blue edges, where the complete graph includes the grid edges as well. (Eppstein, 2007)

the deviation from the target precinct population for each subgraph generated by cutting each edge in the spanning tree. Then, randomly cut one of the edges, creating two new subgraphs[4] If the smaller subgraph meets the population and compactness requirements, then it's accepted as the first district, and the splitting procedure is repeated with the other subgraph.

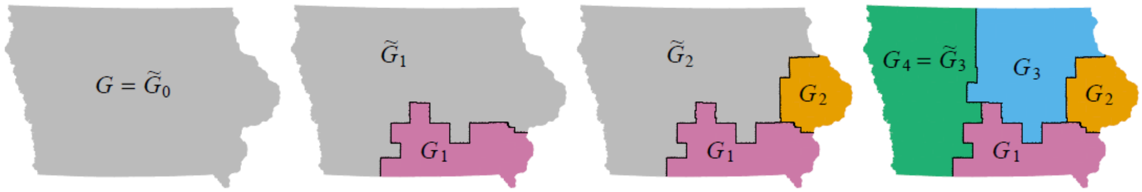
This process generates the possible redistricting plans that satisfy the requirements. For more details, please refer to McCartan and Imai (2020).

Compact Random Seed Growth

The final automated redistricting algorithm that I'm utilizing is called Compact Random Seed Growth, henceforth referred to as "CRSG" and was proposed by Chen and Rodden (2013). See the Method section for more information on how CRSG generates the starting map for MCMC. Its objective is to generate a set of districts that fall within a certain population constraint and are reasonably compact using only the geography and total population of each precinct (Chen & Rodden, 2013). The following is a high-level explanation of the algorithm.[5]

[4] Technically they're spanning trees, which are known as a spanning forest in the plural.

[5] For more details, please refer to Chen and Rodden (2013, pp. 249–50).



(a) Initial map (b) Iteration 1 (c) Iteration 2 (d) Iteration 3

Figure 3

Representation of splitting procedure of SMC. Every node is a precinct, and nodes that share an edge are known to be adjacent precincts. MCMC "cuts away" edges between nodes until islands of districts are formed. (McCartan & Imai, 2020, p. 14)

CRSG begins with declaring that every precinct is it's own district. A random precinct is then chosen, and then its geographically-closest[6] neighbor is merged with it, creating one fewer district. This process is repeated until you arrive at the desired number of districts. (Chen & Rodden, 2013, pp. 249–50)

After this procedure, the districts are somewhat compact due to the geographic proximity requirement, but there is no guarantee that the districts are within the required population percentage of each other.

To satisfy the population parity requirements, CRSG does the following. First, it identifies the two adjacent districts that have the greatest difference in total population. Then the precinct in the more-populous district that is furthest from the center of said district is reassigned to the less-populous district.[7] This process is repeated until all of the districts are within some desired percentage of the mean district population.Chen and Rodden (2013, pp. 249–50).

One run of CRSG will produce one set of districts, but separate runs of CRSG with the same input data may produce slightly different districts given the random choice of

---

[6] The geographically-closest precinct is the neighboring precinct with the smallest distance from its centroid to the seed precinct's centroid

[7] Provided that this reassignment doesn't break either district into parts.

districts to merge in the first pass of the algorithm.

Measures of Partisan Fairness

The following provides an overview of the measures of partisan fairness that exist within the political science literature, as well as their merits and disadvantages.

Seats-Votes Curve

Seats-votes curves are used to plot the relationship between the population vote and the power balance in a legislature (or delegation). This plot has V, the proportion of the overall votes won by the party, on the x-axis, and $S(V)$, the proportion of the seats won by the party, on the y-axis. Figure 4 illustrates several hypothetical seats-votes curves. Tufte (1973) first plotted the vote-seats relationship for elections within one state across time.

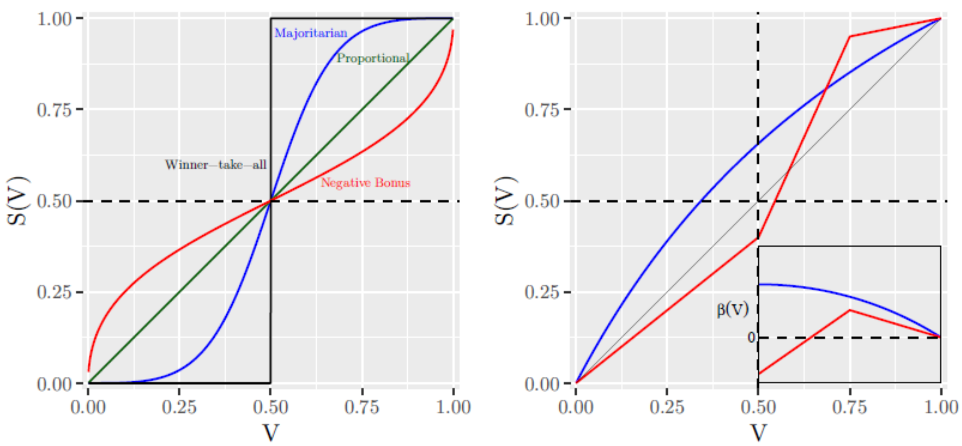


Figure 4

Types of Seats-Votes Curves. Left panel: Symmetric (fair) curves with differing levels of electoral responsiveness. Right panel: Asymmetric (biased) curves, including one consistently biased toward the Democrats (blue) and one with biases favoring different parties depending on V (red); the inset graph is for (V) for V 2 [0:5; 1] with the vertical axis scaled to be the same as the main plot, and lines color coded to the seats-votes curves. (Katz et al., 2020, p. 175)

Naturally, it's very rare to observe the necessary electoral outcomes under the same electoral system in order to determine partisan symmetry. (i.e., it's very rare for two parties to tie one year, have one win 51% of the total votes the next year, and then win 49% of the votes the following year.)

In practice, one can estimate a seats-votes curve using the principle of uniform partisan swing (Tufte, 1973).

Uniform Partisan Swing. Uniform partisan swing is the principle that when the overall vote between different elections under the same electoral system, the vote share at the district level also usually changes by the same dV. Katz et al. (2020) Empirically verified this to be true in 646 different elections.

Thus, given a list of vote proportions per district $v_1, v_2, v_3, ...$ from one election, one can adjust each vote proportion by an arbitrarily small dV until the seats share $S(V)$ is covered from 0 to 1. (Katz et al., 2020).

An example of such a curve is shown in Figure 5

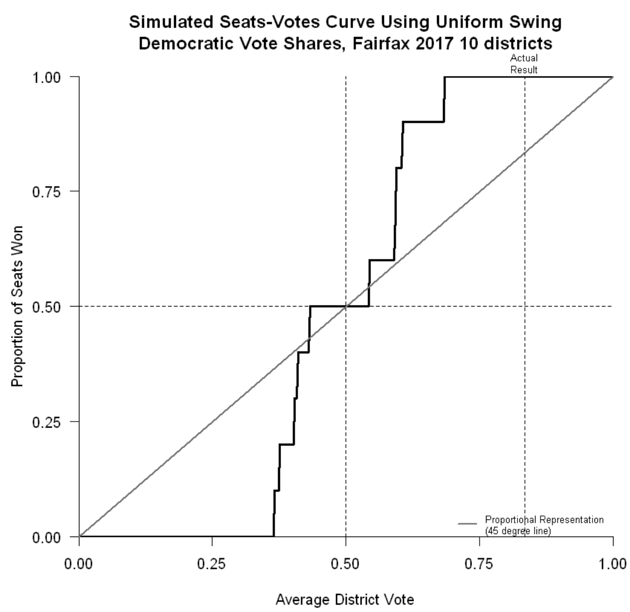


Figure 5

Sample seats-votes curve generated using uniform partisan swing. Ignore title. (Katz et al., 2020, p. 175)

The benefit of seats-votes curves is that they allow a complete view of the biases of a district map across all possible average district votes (Gelman & King, 1994). Unfortunately, generating these curves for one electoral map requires the use of statistical assumptions such as uniform partisan swing as the number of elections occurring under a since district map is usually prohibitively small (Warrington, 2018).

Partisan Symmetry

A legislative is said to have partisan symmetry if both parties can receive m proportion of the overall votes and therefore have n proportion of the seats in the legislative body.An example would be that if Republicans win 60% of the votes but control 65% of the seats, then in a symmetrical system, Democrats should also be able to control 65% of the seats by winning 60% of the votes. Katz et al. (2020)

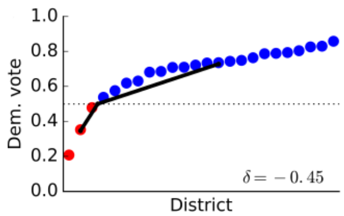
Partisan Bias

Partisan Bias is the deviation from partisan symmetry at a vote proportion of 0.5. If Democrats with 50% of the average district vote but only win 40% of the seats, there is a partisan bias of -0.1. In elections where the average vote proportion does not equal 0.5, partisan bias must be computed using the Uniform Partisan Swing assumption. (Katz et al., 2020)

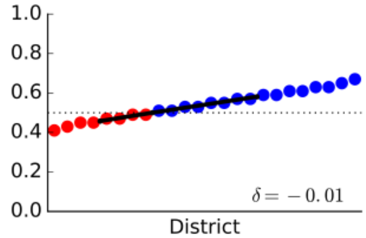
Declination

Declination is another proposed measure of partisan asymmetry that "...relies only on the fraction of seats each party wins in conjunction with the aggregate vote each party uses to win those seats" (Warrington, 2018, p. 3). Broadly, it measures the declination in the line connecting the average vote proportions of one party in districts controlled by the other party. It is essentially the normalized version of the declination measure (Katz et al., 2020). See Figure 6 for a visualization.

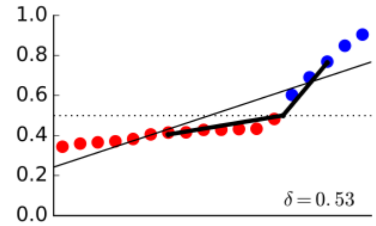
Katz et al. (2020) refutes the claim of Warrington (2018) that declination is a measure of partisan symmetry for the same reason that they refute the claim of partisan symmetry of declination, but acknowledges that it is a useful measure of the skewness of



(a) *Favoring Democrats* *($\delta < 0$)*



(b) *Fair ($\delta \approx 0$)*



(c) *Favoring Republicans* *($\delta > 0$)*

Figure 6

Sample plots illustrating the concept of declination. Plot of districts by increasing Democratic vote proportion. Points colored by party. The black lines connect the following three points: the average Democratic vote share in Republic districts, centered on the Republican districts, the same point but for the Democratic districts, and the point at average vote proportion 0.5 centered between the other two points. (Warrington, 2018, p. 6)

the distribution of district vote proportions.

Efficiency Gap

The efficiency gap is the difference between the number of wasted votes for each party, normalized to the total number of votes, where "wasted votes" are all votes for losing candidates and all votes for winning candidates over the 50%-plus-one threshold. The idea is that wasted votes are a sign of "packing" or "cracking," where a gerrymander has intentionally grouped voters together into the same district or diluted their political power across several districts. (Stephanopoulos & McGhee, 2014)

Veomett (2018) found that the efficiency gap is not a measure of partisan symmetry, as the measure becomes confused in highly non-competitive elections (e.g. party wins 80% of the vote and 100% of the seats).

Mean-Median Difference

The mean-median difference MM is defined as the difference between the mean Democratic vote share (DVS) of each district and median democratic vote share of each district (first proposed by McDonald and Best (2015)). If $MM = 0$, the districts are said to be fair.

Katz et al. (2020) finds that the mean-median difference is a reliable estimator of partisan bias when the average DVS is 0.5. In the language of seats-votes curves, this means it can detect partisan bias at the average DVS = 0.5 level, but not at any other. Essentially, they find that mean-median difference is a reliable estimate of partisan bias as the average DVS changes, but not as the seat share changes (i.e. can measure bias along x-axis but not y-axis of seats-votes curve). (Katz et al., 2020, pp. 27–9)

Measures of Compactness

A universal requirement of districts is that they be "reasonably compact." There is no legally accepted universal definition of compactness, and the scholarship is not united behind one measure either. The following is a brief overview of several different, commonly-used compactness measures that vary in approach, benefits, and weaknesses.

Polsby-Popper Score

Our first measure of district compactness is the Polsby-Popper score. Polsby and Popper (1991) introduces this measure of compactness first development in paleontology to the problem of gerrymandering. It calculates the ratio of the area of the district to the area of the circle with the same perimeter as the district. It is calculated as follows.

$$PP(d) = \frac{4\pi A(d)}{P(d)^2} \tag{1}$$

where d is the district, $A(d)$ is the area of the districts, $P(d)$ is the perimeter of the districts, and $PP(d)$ is the Polsby-Popper score (Cox, 1927; Polsby & Popper, 1991). The score will range from 0 to 1, where 0 is a lack of compactness and 1 is the most-compact district (Polsby & Popper, 1991).

Limitations and criticisms of this measure include that it is very sensitive to both geography and map resolution. Particularly near coastlines, even the most compact districts can have Polsby-Popper scores that are lower than gerrymandered districts which don't border coastlines. At finer resolutions, the same district will have a lower Polsby-Popper score as the perimeter increases. (McCartan & Imai, 2020, p. 12).

Edge-Cut Compactness

Edge-Cut Compactness takes a graph-theory perspective to district compactness. Imagine a graph where each precinct is a node, some vertex of the graph, and every adjacent precinct shares an edge, a line connecting the points. The Edge-Cut compactness score is the number of edges that must be "cut" (removed) from the original graph to form subgraphs, or districts. The theory is that compacter districts will require fewer edges to be cut. (Dube & Clark, 2016) If normalized to the number of edges and subtracted from 1, a score of 1 indicates the most compact district.

Benefits of this measure include that it is unaffected by political or natural geography and map resolution, as well as population density (McCartan & Imai, 2020, p. 11).

Method

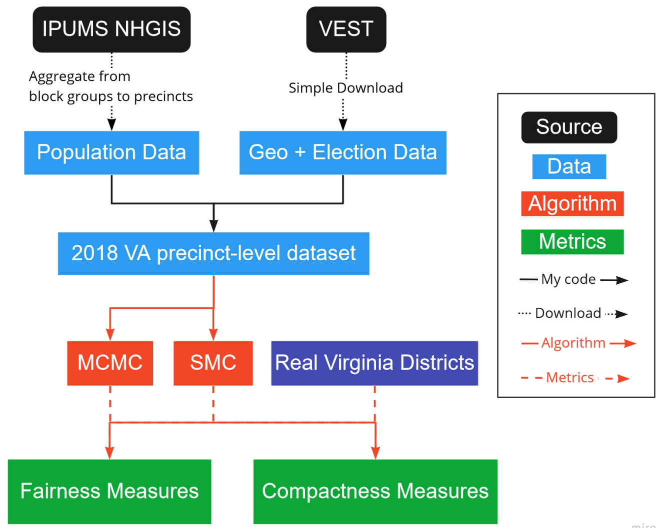


Figure 7

Graphical Overview of project

My research method simulates the 2020 redistricting of the congressional districts in Virginia using two different algorithms supplied with data from 2018. Figure 7 provides and overview of this process.

Choice of Research Method

For this study, I chose to use the experimental design method because it will allow me to isolate the hypothetical impact of the redistricting algorithm from other possible confounding variables. This method also includes the use of a control group, which allows the researcher to establish causation.

Components of Experimental Design

Experimental Units. The experimental units for this study are the complete datasets for each election year in Virginia.[8] Every row in each dataset corresponds to a precinct, the smallest geographical unit by which votes are tabulated in Virginia. For each

[8] This corresponds to the blue rectangles in Figure 7.

precinct, I also have the total population and the number of votes cast for the 2018 Democrat, Republican, and other congressional candidates. Additionally, each precinct has a polygon associated with it that represents its geographical shape.

Treatments. The treatments for this study are the two different redistricting algorithm that I'm comparing: Markov chain Monte Carlo (Fifield, Higgins, et al., 2020) and Sequential Monte Carlo (McCartan & Imai, 2020).[9] I'm using the implementations in the R programming language "redist" package (Fifield, Kenny, et al., 2020). See the Literature Review for more information . Broadly, I chose them because they are deterministic. Much of the literature focuses on creating many possible redistricting plans for a commission to choose from, but these three aim to create an "ideal" map. A redistricting plan generated by CRSG serves as the initial map for the MCMC algorithm.[10]

Response Variables. Broadly, the goal will be to evaluate how "fair" and compact each redistricting plan generated by each algorithm for each year is.[11] Refer to the Literature Review for more information on the fairness and compactness measures chosen.

Compactness Measures. I chose to compute the mean Polsby-Popper score, Fryer-Holden score, and the Edge-Cut Compactness measure for the proposed maps from each algorithm. I did not have the necessary computational resources to compute all of the measures that I describe in my Literature Review, so I elected to compute the most popular area-based measure (Polsby-Popper), the only population-based measure (Fryer-Holden), and the primary graph-theory measure (Edge Cut Compactness). While not complete, these measures provide a representative overview of the possible compactness measures.

Fairness Measures. I chose to compute the Partisan Bias, Declination, Efficiency Gap, Equal Population Efficiency Gap, Lopsided Wins, Mean-Median,

---

[9] This corresponds to the red rectangles in Figure 7

[10] MCMC requires a valid map to start with, and the existing map of Virginia was not valid as the population parity was no longer within 1%, as the population has changed since 2016 when the map was drawn.

[11] This corresponds to the green rectangles in Figure 7.

Responsiveness, and Tau Gap measures, all of which are described in the Literature Review, as they represent the fairness measures discussed at present in the literature (Katz et al., 2020). Computational resources were not a barrier for this step, so I was able to calculate all of the measures.

Chamber Power Balance. Since the redistricting that's occurring is hypothetical and I have precinct-level election results for each of these years, I also simulate how many seats the Democratic party would have won in the 2018 General Election under the various redistricting plans.

Control Group. The official VA Congressional district map used in the years 2016-2020[12] will serve as the control group for this experiment. I will compute the same metrics for this map as I will for my hypothetical redistricting plans.[13]

Principles of Experimental Design

The primary principles of experimental research design are randomization, replication, and local control. This is how I plan to address them.

Randomization. Every experimental unit will receive each treatment, and every experimental unit can be replicated many times without issue, so thereâĂŹs no error from a lack of randomization. Think of each treatment operating within a separate parallel universe.

Replication. Each algorithm will generate 100 sample redistricting plans. This is large enough to allow for inferences, but small enough to still be computationally feasible.

Local Control. All of the redistricting will be happening in controlled environments, so there will be no way for lurking variables to creep in and confound my results.

[12] Due to racial gerrymandering, VA had to adopt a new congressional district map in the middle of the decade ("Wittman v. Personhuballah," 2016).

[13] This corresponds to the purple rectangle in Figure 7.

Data Cleaning

To create my datasets, I cleaned and compiled three different types of data: demographic data, Geographic Information Systems (GIS) data, and election data.[14]

Demographic Data

One required piece of data in order to redistrict is demographic data at the precinct level. For my purposes, this means the total population of each precinct. In order to run the most accurate redistricting simulations, these data needed to be recent for the year being redistricted. Comprehensive population counts are only conducted by the US Census Bureau every 10 years, so I instead used the 5-year American Community Survey results at the block-group level. This is a sample survey, not a population count, but that is offset by the aggregation of sample data over a 5 year period. I downloaded this data from the IPUMS National Historic GIS project (Manson et al., 2020). Using the "maup" Python Library (Hully, n.d.), I disaggregated the data from the block-group level to the block level, prorating the demographic data based on population. This data was then aggregated up to the precinct level.

GIS Data

In order to redistrict, the algorithms need to know the shape and relative location of each precinct. In practice, this means every precinct has a "polygon" associated with it and a Coordinate Reference System that describes where these polygons fall in space. These data tables with the geometry column are known as "shapefiles." I accessed these shapefiles from the Voting and Election Science Team on their Harvard Dataverse (Voting & Team, 2019). I then merged in my precinct-level demographic data tables, so I now have shapefiles with the necessary demographic data. [15]

[14] This is an explanation of the black and blue rectangles in Figure 7 and the transitions between them.

[15] Since election administrators are free to change the precincts between elections, precinct shapefiles are unique to both a place and a time. This is why I couldn't use the 2018 precinct shapefiles with 2020 election results.

Election Data

The last necessary component needed to evaluate redistricting plans is the number of votes won by each party in each precinct in each election.[16] Conveniently, this data was already included in my shapefiles by the Voting and Election Science Team.

[16] The algorithms I'm comparing assume a 2 party system, so I only tracked Democratic and Republican votes won in each election.

Results

This section provides an overview of the results from my method.

Compactness

Figure 8 visualizes the distribution of two different compactness scores amongst plans generated by both SMC and CRSG. Subfigure (a) shows the distribution of the Polsby-Popper scores of the 100 SMC plans, subfigure (b) shows the distribution of the Edge-Cut compactness measure of the same 100 SMC plans, and subfigures (c) and (d) show the corresponding measure distributions for the plans generated by CRSG. The vertical red lines indicate the value of the measure for the existing district map, and the vertical green lines indicate the mean value of the distribution.

Partisan Fairness

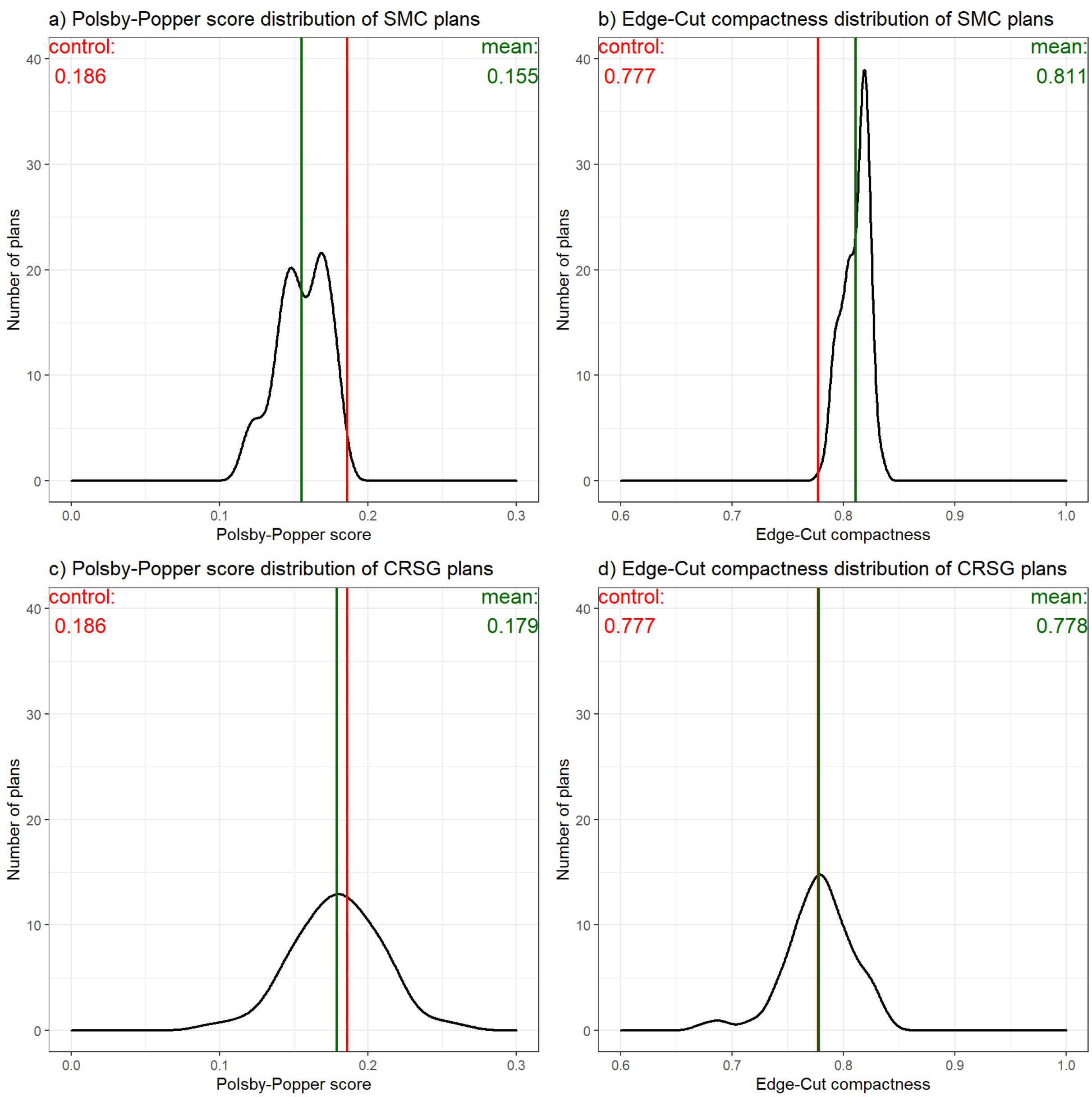
Seats-Votes Curves

Figure 9 shows the seats-votes curves (Katz et al., 2020) for the 2018 General Election under the redistricting plans generated by both algorithms and the existing map. For each plot, the x-axis plots the average of the proportion of votes won by Democrats in each district. The y-axis plots the proportion of seats won by Democrats in the delegation. Both subfigures 9a and 9b have once curve for each redistricting plan (each plot has 100 curves). The seats-votes curve for the real 2018 districts is provided for reference in subfigure 9c.

Single-Valued Partisan Fairness Measures

Figure 8

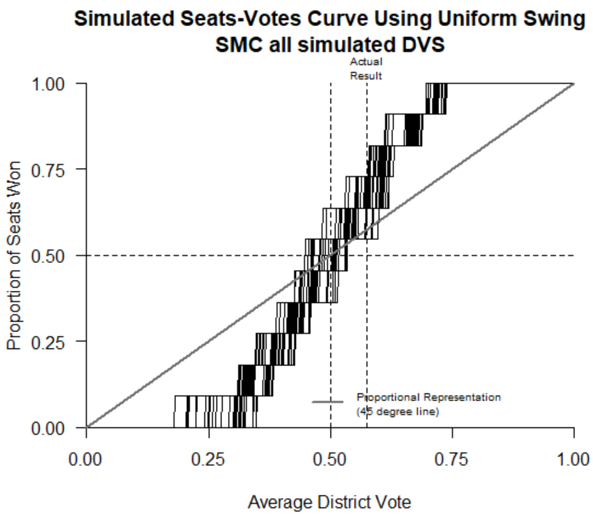
Distribution of Compactness Measures



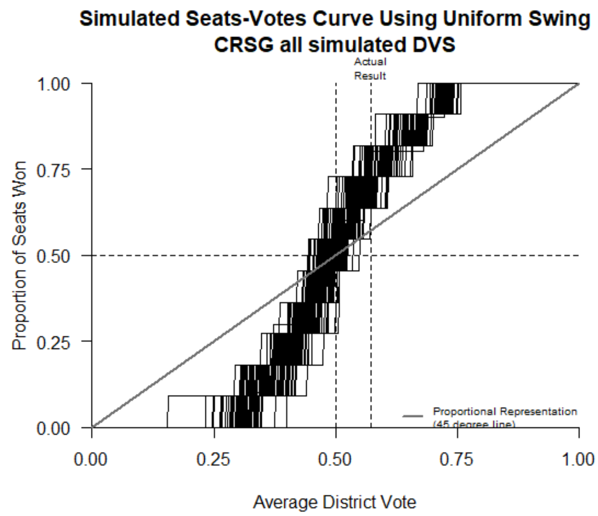
Note. The first column of plots shows the Polsby-Popper score distribution; the second column shows the Edge-Cut compactness distribution. The first row corresponds to the plans generated by SMC; the second row corresponds to CRSG. The red lines indicate the measure value for the existing districts; the green lines indicate the mean value of the distribution.

Figure 9

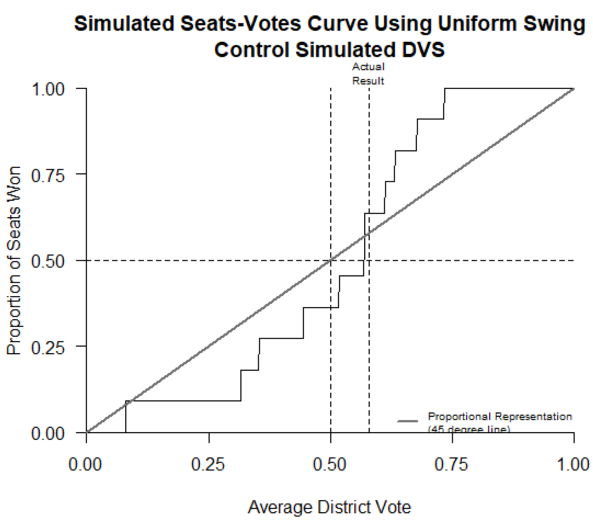
Seats-Votes Curves for SMC, CRSG, and existing plan.



(a) *SMC Seats-Votes Curve*



(b) *CRSG Seats-Votes Curve*

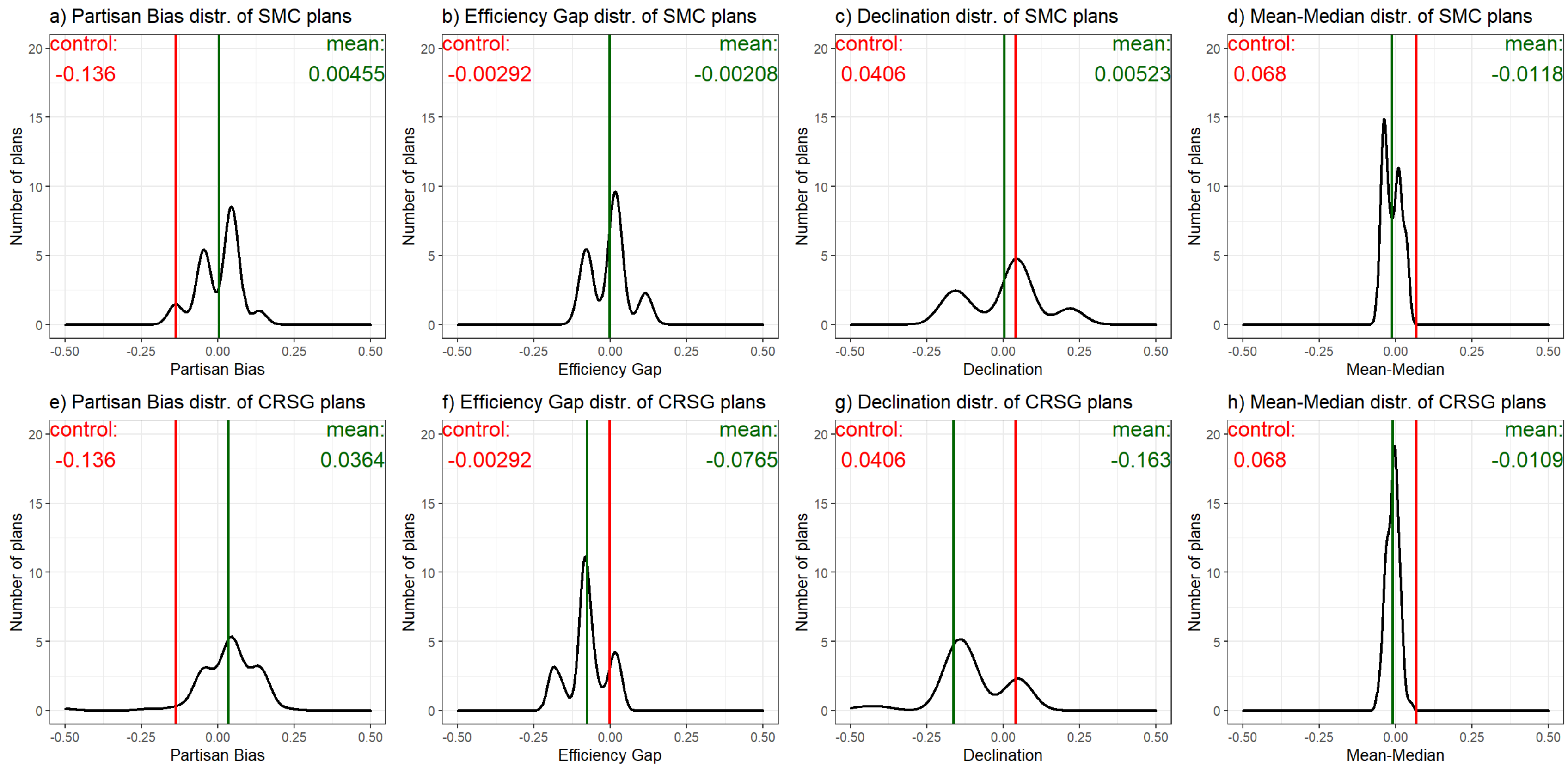


(c) *Real Seats-Votes Curve*

Note. Each plot shows the relationship between average proportion of Democratic vote share by district and the proportion of Democratic seats. Subfigure 9a illustrates this relationship for the 100 plans generated by SMC, subfigure 9b for CRSG, and subfigure 9c is for the existing plan.

Figure 10

Distributions of fairness measures



Note. The columns of plots illustrate the distributions of partisan bias, the efficiency gap, declination, and the mean-median difference, respectively. The first row of plots corresponds to SMC plans, the second to CRSG plans. The red lines indicate the corresponding value of the measure in the existing plan; the green lines indicate the mean value of the distribution.

Figure 10 illustrates the distributions of various fairness measures within the plans generated by SMC and CRSG. Columns of plots correspond to different measures, and rows of plots correspond to different algorithms. The red and green vertical lines indicate the control value and mean distribution value, respectively.

2018 General Election Simulation

Electoral Maps

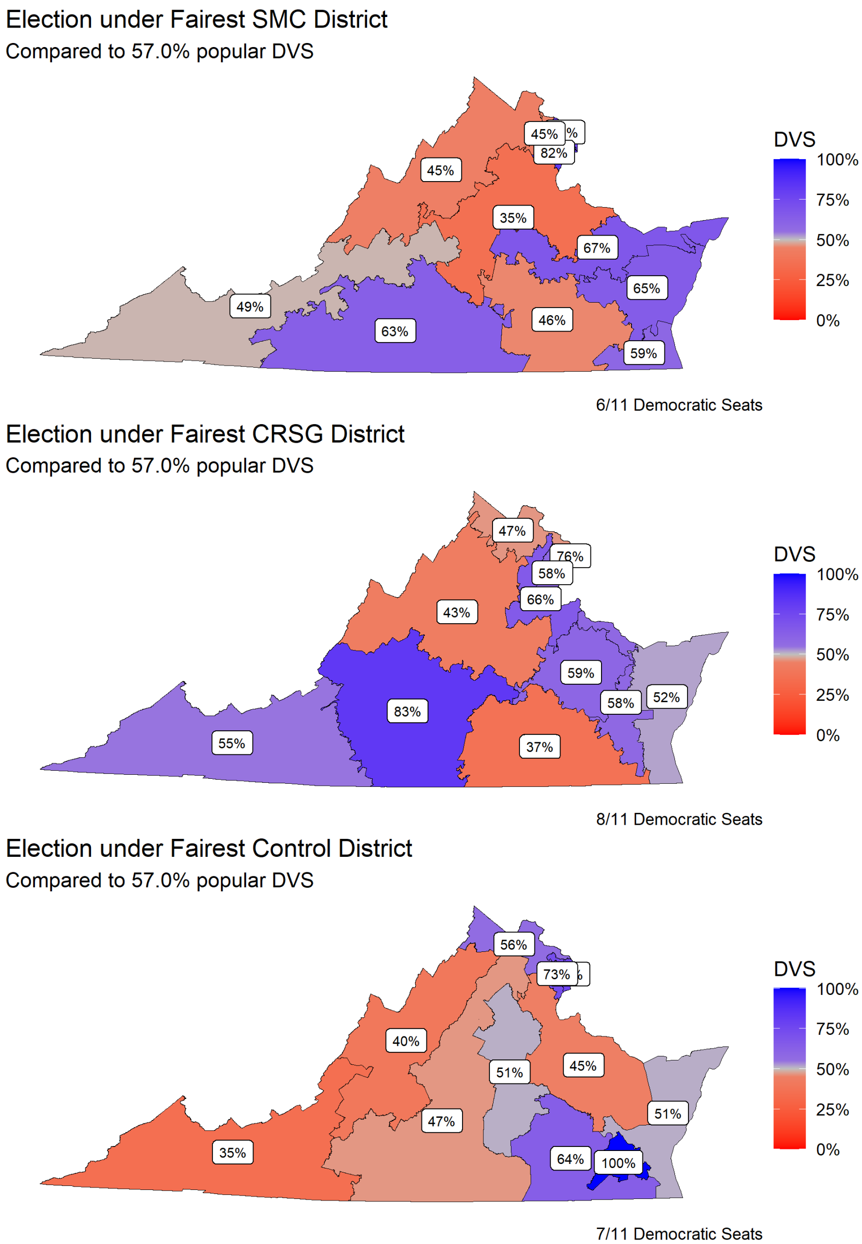
Figure 11 visualizes the outcome of a simulated 2018 General Election under various redistricting plans. Subfigure (a) was created by aggregating the precinct-level election results from 2018 to the district level using the redistricting plan generated by SMC with the least magnitude partisan bias. The average precinct-level proportion of votes won by Democrats in the district is displayed as a percentage. Values closer to 1 indicate a higher proportion of Democratic votes and are colored blue. Values closer to 0 indicate a higher proportion of Republican votes and are colored red. Purple districts are most competitive. Subfigure (b) illustrates this same simulation, only using the corresponding "fairest" redistricting plan generated by SMC. The real results from the 2018 General Election are visualized in subfigure (c) for reference.

Seats distribution

Finally, Figure 12 illustrate the distribution of seats allocated to Democrats by each algorithm, as well as the true number of seats from 2018 (indicated by the red line).

Figure 11

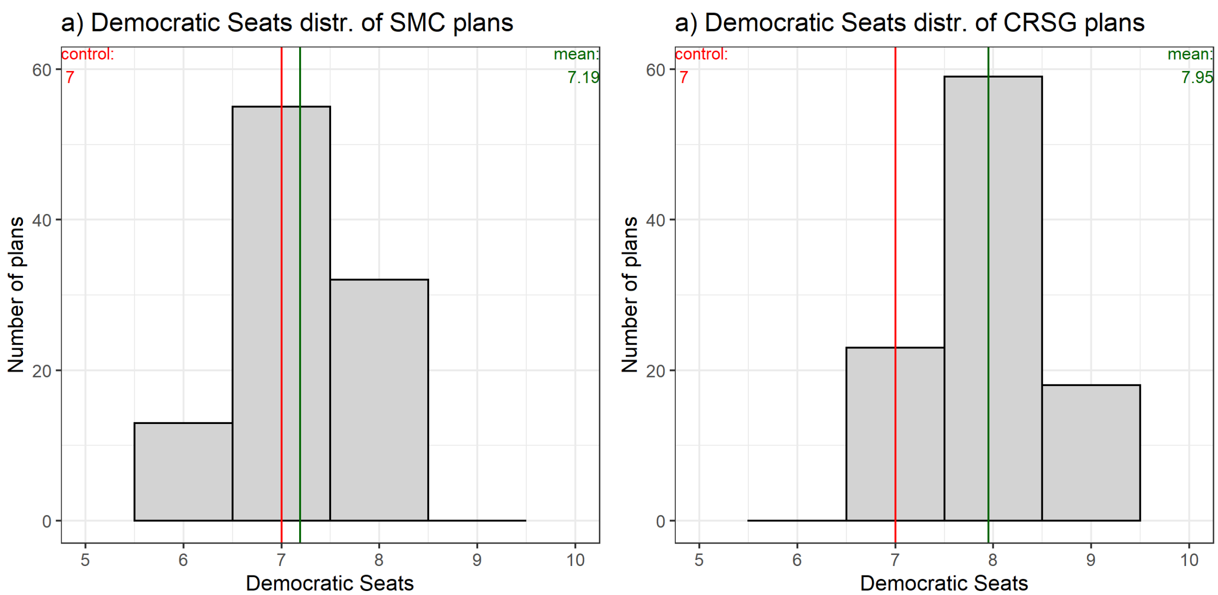
Simulated 2018 Virginia Congressional Election under SMC and CRSG



Note. Each map uses the district plan with the least-magnitude partisan bias for SMC, CRSG, and the existing plan, respectively. Navy corresponds to a greater Democratic vote proportion, and red corresponds to a greater Republican vote proportion.

Figure 12

Distributions of Democratic seats under SMC and CRSG



Note. a) shows the distribution of the number of seats allocated to Democrats by the 100 SMC plans; b) shows the corresponding distribution for CRSG plans. The red and green lines indicate the control and mean distribution value of the number of Democratic seats, respectively.

Discussion

The goal of my research is to compare two different automated redistricting algorithms in an empirical context and to evaluate the results using a variety of compactness and partisan fairness standards that have been established in the literature.

I will begin by analyzing the compactness metrics, then the partisan fairness metrics, and finally the simulated election results. I will end by discussing the limitations of my findings.

Compactness

Beginning with compactness measures, I'm analyzing the results presented in Figure 8.

Polsby-Popper score

The Polsby-Popper score compares the area of the district to the area of a circle with the same perimeter as the district (Polsby & Popper, 1991). It ranges from 0 to 1, where higher values indicate a compacter district.

The SMC plans averaged a Polsby-Popper score of 0.155, which is slightly less compact than the real district (0.186). The CRSG plans also averaged a score just slightly less than the control plan (0.179). Through the lens of this measure, CRSG produced plans as compact as the control, and SMC's plans were slightly less compact.

Edge-Cut compactness

The second compactness metric measured was the edge-cut compactness score, a measure grounded in graph theory. It measure the proportion of edges that had to be cut from the initial precinct graph to form the district, and has been normalized to the total number of districts[17] (Dube & Clark, 2016). This measure isn't affected by geographic features such as coastlines or mountain ranges, which makes it a more reliable measure of compactness than Polsby-Popper to some (see McCartan and Imai (2020)).

[17] For an explanation of the graph of precincts, please see the Redistricting as Graph Cutting section.

Through the lends of edge-cut compactness, SMC plans were more compact than the control ($0.811 > 0.777$), and CRSG plans were just as compact as CRSG plans ($0.778 \approx 0.777$).

Conclusion

Both compactness measures did not reach the same conclusions. Polsby-Popper found the SMC plans to be less compact on average than the control, while edge-cut compactness found SMC to be more compact. This is likely due to the fact that Polsby-Popper scores are very sensitive to resolution and geography (McCartan & Imai, 2020). Virginia is a states with high-perimeter borders, defined by the Blue Ridge Mountains and the Chesapeake Bay. Additionally, Accomack County includes the Virginia Barrier Islands, which are separately from the mainland of the state ({United States Geological Survey}, 2021). All this is to say that the very uneven border of Virginia due to geography likely interfere with the ability of the Polsby-Popper score to accurately quantify compactness.

With that in mind, if we only consider edge-cut compactness, SMC was able to generate compacter plans the CRSG or the control. However, both distributions were reasonably compact, and there are no worrying differences between compactness between the three proposals.

Partisan Fairness

Moving on to the partisan fairness results, I'll first analyze the seats-votes curves, and then the single-value measures outlined in the Literature Review.

Seats-Votes Curves

Figure 9 includes the seats-votes curves that I'm going to analyze. When viewed as a whole, the 100 curves for SMC in a) form a Majoritarian seats-votes curve (see Figure 4). Partisan symmetry is the symmetry about the point $(0.5, 0.5)$ on this graph (Katz et al., 2020), and the SMC plan curves display partisan symmetry. The same is true for the 100 curves from CRSG in b). This indicates that both algorithms generated plans that, when taken together, do not advantage one party over the other.

Compared to the curve for the existing district map in 2018 (shown in c)), the two algorithms generated more-symmetrical plans. The control plan is still Majoritarian, but it favors Republicans from DVS $[0, 0.53]$ and Democrats from $[0.53, 1]$. To illustrate this, we can see that if Democrats won 50% of the average district vote, they'd win about 38% of the seats, not 50%.

In summary, both algorithms generated more symmetrical plans on average when compared to the existing plan.

Single-Valued Measures

Next I will analyze the partisan bias, efficiency gap, declination, and mean-median difference score distributions for the two algorithms. These results are visualized in Figure 10.

Partisan Bias. Partisan bias is the distance from partisan symmetry at an average DVS of 0.5. On a seats-votes curve (see Figure 9), this is the difference between 0.5 and the y-coordinate of the curve at $x = 0.5$. It ranges from -0.5 to 0.5, with values of 0 being least biased. (Katz et al., 2020)

As shown in a), the SMC plans on average were very fair ($0.0045 \approx 0$). e) shows this to be true for the CRSG plans as well, though they do slightly favor Democrats ($0.0364 > 0$). Both algorithms generated less biases plans on average than the control, which significantly favored Republicans ($-0.136 < 0$).

These results align with the the conclusions drawn from the seats-votes curve. Interestingly, the control plan is biased towards Republicans, but this is not evident in the election map, as the actual DVS is greater than 0.5. Nevertheless, both algorithms generated less-biased plans on average than the control plan.

Efficiency Gap. The next measure is the efficiency gap, which measure the difference in "wasted votes" between majority-Democrat and majority-Republican districts, normalized to the total vote count. It ranges from -1 to 1. (Stephanopoulos & McGhee, 2014)

All three sets of plans; SMC, CRSG, and the control; have an efficiency gap within the acceptable margin (± 0.07).

The efficiency gap doesn't find evidence of significant differences in the number of wasted votes between Democratic and Republican districts between the algorithms and the control. This aligns with the literature, as most criticisms of the efficiency gap point out that is mischaracterizes "wasted votes" in highly non-competitive elections (see Veomett (2018) and Katz et al. (2020)). However, Virginia is a very competitive state, so this does not appear to apply in this situation.

Declination. The next measure is declination, which quantifies the concavity of the the line on the plot of districts by ascending DVS; see Figure 6 for a visualization. It ranges from -1 to 1, with values of 0 indicating fairness. (Warrington, 2018)

The SMC plans are on average fair ($0.00523 \approx 0$), as is the control plan ($0.0406 \approx 0$). However, the declination of the CRSG is biased towards Democrats ($-0.163 < 0$). Seeing as the seats-votes curves for SMC and CRSG were found to be symmetrical, these results agree with the findings of Katz et al. (2020) that declination is not a measure of symmetry, but rather of the skewness of the district vote proportion distribution.

This district-level vote proportion distribution appears to be skewed towards Democrats on average in the CRSG plans, and centered in the average SMC plans and the control plan.

Mean-Median Difference. The next measure is the mean-median difference, which is the difference between the mean district DVS and the median district DVS. It ranges in value from -1 to 1, with values of 0 indicating a non-skewed distribution (McDonald & Best, 2015). Katz et al. (2020) finds that at an averaged DVS of 0.5, the mean-median difference measures partisan bias.

The result that the control plan slightly favors Republicans ($0.068 > 0$) while the average SMC plans and CRSG plans are basically fair ($-0.0118 \approx 0, -0.0109 \approx 0$), aligns with the results from the partisan bias measure, which supports the finding of Katz et al. (2020).

Conclusion

When viewed together, the partisan fairness measures illustrate the following conclusions. SMC generated fair and symmetrical districts from the perspective of all four measures. CRSG also generated on average fair districts, with only a skew in the district-level Democratic vote proportion. Both algorithms produced on average fairer districts than the control plan, which was found to be biased towards Republicans by all fairness measures.

Election Simulation

Using these plans, I also simulated the 2018 Virginia Congressional election using the plans generated by SMC and CRSG with the lowest-magnitude partisan bias. The election maps are shown in Figure 11, and the distribution of the number of seats allocated to Democrats by each algorithm is shown in Figure 12.

The proportion of the popular vote that went to the Democratic party in 2018 was 0.57. Under proportional representation, this means the Democrats should control 6.27 seats. SMC's least-biased plan allocated Democrats 7 seats on average, which is close to proportional representation. CRSG allocated on average 8 seats to the Democrats, which is biased in their favor. The Control plan also allocated 7 seats to Democrats.

Evaluating redistricting plans solely based on the number of seats awarded doesn't show the full picture. For instance, the control map was found to be biased towards Republicans (calculated by generating a seats-votes curve using the uniform partisan swing assumption). However, since the average district vote was 0.57, this discrepancy isn't observed. The seats-votes curve and single-valued measures of partisan fairness are more-reliable views of the underlying imbalances than a simulated election.

Limitations

The limitations of the compactness measures largely are due to their sensitivity to geography. Additionally, the seats-votes curves are very sensitive to the idiosyncrasies of the particular election since the number of seats is so small. Each algorithm generated 100 districts, so the possibility of randomly generated biased districts could be reduced by having a larger sample size. Finally, every district vote proportion was calculated using only the votes for the Democratic and Republican parties, ignoring third-party candidates, though this is unlikely to have had a significant effect as the United States has first-past-the-post elections.

Conclusions

There were not significant differences in the average compactness of plans generated by the algorithms and the control plan. The compactness measures provided additional evidence that the Polsby-Popper score cannot accurately measure compactness when faced with districts that have lengthy perimeters due to their geography.

SMC was found to generate fair, symmetrical districts. CRSG districts were mostly fair, though they showed a slight Democratic bias in some cases. The control districts were found to be biased towards Republicans by all measures of fairness.

Write!

References

Altman, M., & Mcdonald, M. P. (2009). Bard: Better Automated Redistricting. *Journal of Statistical Software*.

Chen, J., & Rodden, J. (2013). Unintentional Gerrymandering: Political Geography and Electoral Bias in Legislatures. *Quarterly Journal of Political Science*, *8*(3), 239–269. https://doi.org/10.1561/100.00012033

Cox, E. P. (1927). A Method of Assigning Numerical and Percentage Values to the Degree of Roundness of Sand Grains. *1*(3), 179–183.

Dube, M., & Clark, J. (2016). Beyond the Circle: Measuring District Compactness Using Graph Theory.

Eppstein, D. (2007). English: A 4x4 Grid Graph and One of Its Spanning treesFrançais : Une Grille 4x4 et Un de Ses Arbres Couvrants. Retrieved March 1, 2021, from https://commons.wikimedia.org/wiki/File:4x4_grid_spanning_tree.svg

Fifield, B., Kenny, C. T., McCartan, C., Tarr, A., Higgins, M., Kawahara, J., & Imai, K. (2020). Redist: Simulation Methods for Legislative Redistricting. Retrieved January 29, 2021, from https://CRAN.R-project.org/package=redist

Fifield, B., Higgins, M., Imai, K., & Tarr, A. (2020). Automated Redistricting Simulation Using Markov Chain Monte Carlo. *Journal of Computational and Graphical Statistics*, *0*(0), 1–14. https://doi.org/10.1080/10618600.2020.1739532

Gelman, A., & King, G. (1994). Enhancing Democracy Through Legislative Redistricting. *American Political Science Review*, *88*(3), 541–559. https://gking.harvard.edu/files/abs/red-abs.shtml

Haas, C., Hachadoorian, L., Kimbrough, S. O., Miller, P., & Murphy, F. (2020). Seed-Fill-Shift-Repair: A Redistricting Heuristic for Civic Deliberation. *PLoS ONE*, *15*(9), 1–34. https://doi.org/10.1371/journal.pone.0237935

Hully, M. (n.d.). Maup: The Geospatial Toolkit for Redistricting Data. Retrieved February 18, 2021, from https://github.com/mggg/maup

Katz, J. N., King, G., & Rosenblatt, E. (2020). Theoretical Foundations and Empirical Evaluations of Partisan Fairness in District-Based Democracies. *American Political Science Review*, *114*(1), 164–178. https://doi.org/10.1017/S000305541900056X

Lara-Caballero, A., de-los-Cobos-Silva, S. G., Mora-Gutiérrez, R. A., Rincón-García, E. A., Gutiérrez-Andrade, M. Á., & Lara-Velázquez, P. (2019). Multiobjective Genetic Algorithms for Reinforcing Equal Population in Congressional Districts. *Mathematical Problems in Engineering*, 1–14. https://doi.org/10.1155/2019/2825854

Macmillan, W. (2001). Redistricting in a GIS Environment: An Optimisation Algorithm Using Switching-Points. *Journal of Geographical Systems*, *3*(2), 167. https://doi.org/10.1007/PL00011473

Manson, S., Schroeder, J., Van Riper, D., Kugler, T., & Ruggles, S. (2020). National Historical Geographic Information System: Version 15.0. https://doi.org/10.18128/D050.V15.0

McCartan, C., & Imai, K. (2020). Sequential Monte Carlo for Sampling Balanced and Compact Redistricting Plans. *arXiv:2008.06131 [cs, math, stat]*. Retrieved January 18, 2021, from http://arxiv.org/abs/2008.06131

McDonald, M. D., & Best, R. E. (2015). Unfair Partisan Gerrymanders in Politics and Law: A Diagnostic Applied to Six Cases. *Election Law Journal: Rules, Politics, and Policy*, *14*(4), 312–330. https://doi.org/10.1089/elj.2015.0358

Polsby, D. D., & Popper, R. D. (1991). The Third Criterion: Compactness as a Procedural Safeguard against Partisan Gerrymandering. *Yale Law & Policy Review*, *9*(2), 301–353. Retrieved March 12, 2021, from http://www.jstor.org/stable/40239359

Stephanopoulos, N., & McGhee, E. (2014). *Partisan Gerrymandering and the Efficiency Gap* (SSRN Scholarly Paper No. ID 2457468). Social Science Research Network. Rochester, NY. Retrieved March 26, 2021, from https://papers.ssrn.com/abstract=2457468

Tufte, E. R. (1973). The Relationship between Seats and Votes in Two-Party Systems. *American Political Science Review*, *67*(2), 540–554. https://doi.org/10.2307/1958782

{United States Geological Survey}. (2021). The National Map Services. Retrieved March 31, 2021, from https://apps.nationalmap.gov/services/

Veomett, E. (2018). Efficiency Gap, Voter Turnout, and the Efficiency Principle. *Election Law Journal: Rules, Politics, and Policy*, *17*(4), 249–263. https://doi.org/10.1089/elj.2018.0488

Voting, & Team, E. S. (2019). 2018 Precinct-Level Election Results. https://doi.org/10.7910/DVN/UBKYRU

Warrington, G. S. (2018). Quantifying Gerrymandering Using the Vote Distribution. *Election Law Journal: Rules, Politics, and Policy*, *17*(1), 39–57. https://doi.org/10.1089/elj.2017.0447

Weaver, J. B., & Hess, S. W. (1963). A Procedure for Nonpartisan Districting: Development of Computer Techniques. *The Yale Law Journal*, *73*(2), 288–308. https://doi.org/10.2307/794769

Wittman v. Personhuballah. (2016).

Xiao, N. (2008). A Unified Conceptual Framework for Geographical Optimization Using Evolutionary Algorithms. *Annals of the Association of American Geographers*, *98*(4), 795–817. https://doi.org/10.1080/00045600802232458