Empirical Validation of Automated Redistricting Algorithms on the Virginia House of Delegates District Map

Madeleine Goertz

International Community School

AP Research

Randall S. Huberman

May 20th, 2021

Empirical Validation of Automated Redistricting Algorithms on the Virginia House of Delegates District Map

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 three algorithms: Markov Chain Monte Carlo, 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).

Markov Chain Monte Carlo

The first automated redistricting algorithm that I'm going to discuss in detail is an implementation of a statistical method known as "Markov Chain Monte Carlo," henceforth

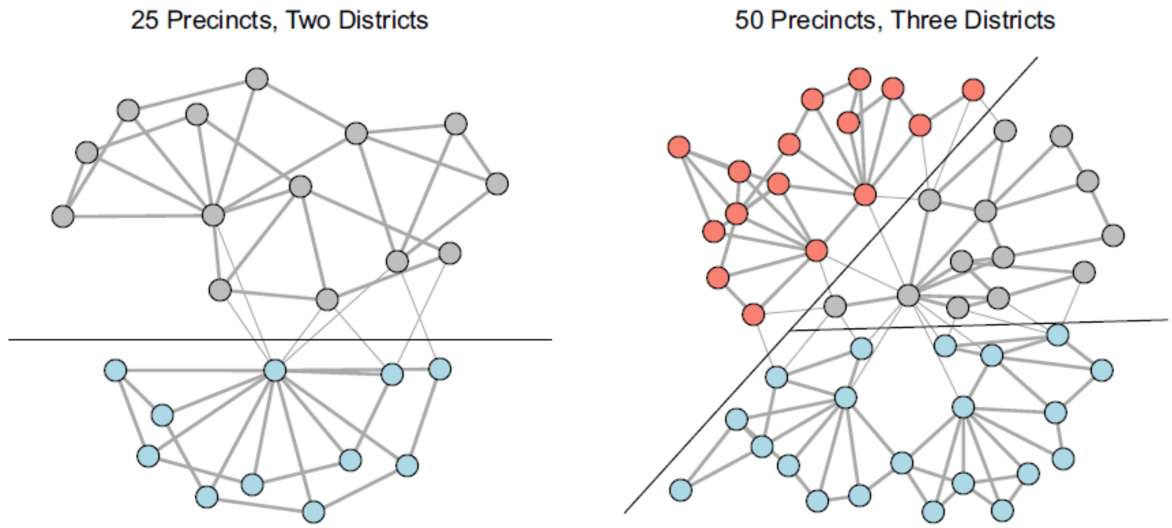


Figure 1

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)

MCMC (Fifield, Higgins, et al., 2020).[1] The following overview is meant to provide a high-level understand of how the basic algorithm works. Since the purpose of my research is to compare the *outputs*, that is the redistricting plans, generated by these algorithms, *rather than the algorithms themselves*, a rigorous understanding of the algorithms is not a prerequisite.[2].

MCMC begins with a graph representing the redistricting problems. Figure 2 visualizes the steps in this algorithm. The basic MCMC algorithm begins with a valid redistricting plan, such as the one currently in use, and randomly decides to "turn on" some edges in the graph. Then, the nodes (precincts) that are connected by these "turned on"

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

[2] Please see the section "The Proposed Methodology" in the paper for an in-depth, mathematically-rigorous explanation. (Fifield, Higgins, et al., 2020, p. 2)

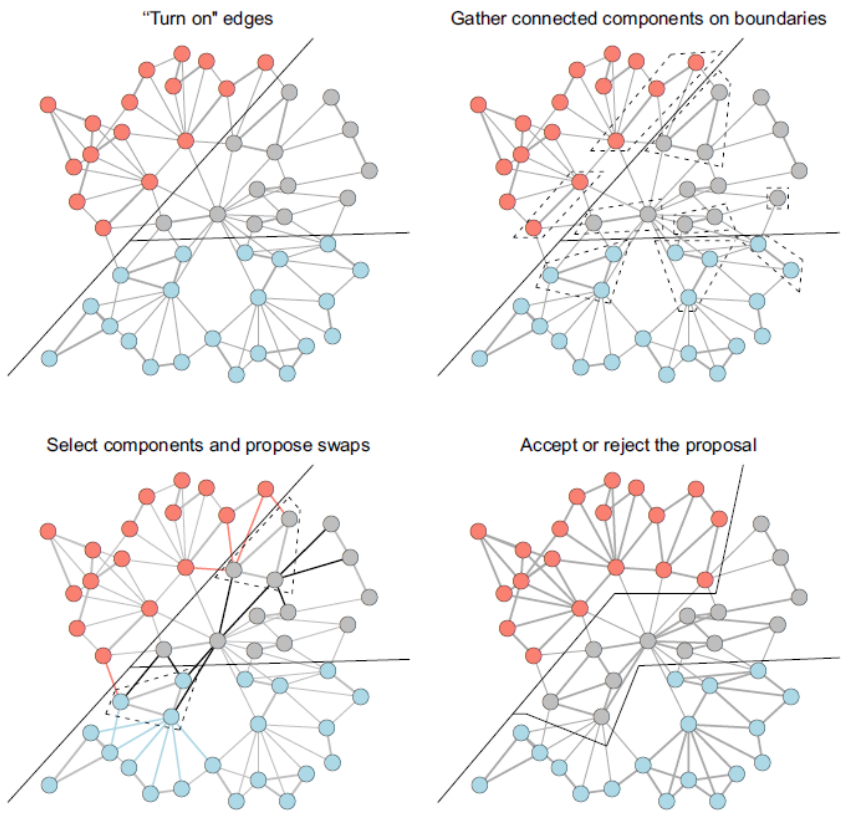


Figure 2

Representation of MCMC algorithm. (Fifield, Higgins, et al., 2020, p. 4)

edges and are located on the boundary of a district are identified. Then, these highlighted graph components are "nominated" for a swap across the district boundary based on some probability, provided that this swap would not break the district into two. Lastly, this new proposed redistricting plan is either accepted or rejected based on some "acceptance probability." This process is repeated as many times as desired. (Fifield, Higgins, et al., 2020).

MCMC in its current form allows for further constraint of this algorithm, particularly with regard to the relative population size and compactness of each district. (Fifield, Higgins, et al., 2020, p. 6) The explanation of this algorithm is beyond the scope of this paper.

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 SMC (McCartan & Imai, 2020).[3] The following overview is meant to provide a high-level understand of how the basic algorithm works. [4].

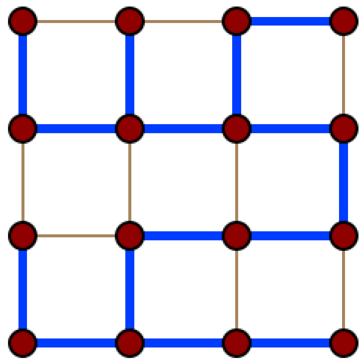


Figure 3

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)

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 3), which is a graph that is connected by the minimum number of possible edges[5].

Figure 4 visualizes the iterative splitting procedure used by SMC. First, compute the deviation from the target precinct population for each subgraph generated by cutting

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

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

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

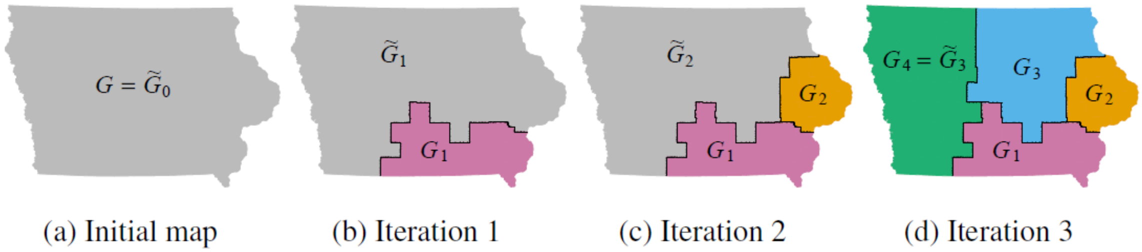


Figure 4

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)

each edge in the spanning tree. Then, randomly cut one of the edges, creating two new subgraphs[6] 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.[7]

CRSG begins with declaring that every precinct is it's own district. A random

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

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

precinct is then chosen, and then its geographically-closest[8] 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.[9] 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 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 5 illustrates several hypothetical seats-votes curves. Tufte (1973) first plotted the vote-seats relationship for elections within one state across time.

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

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

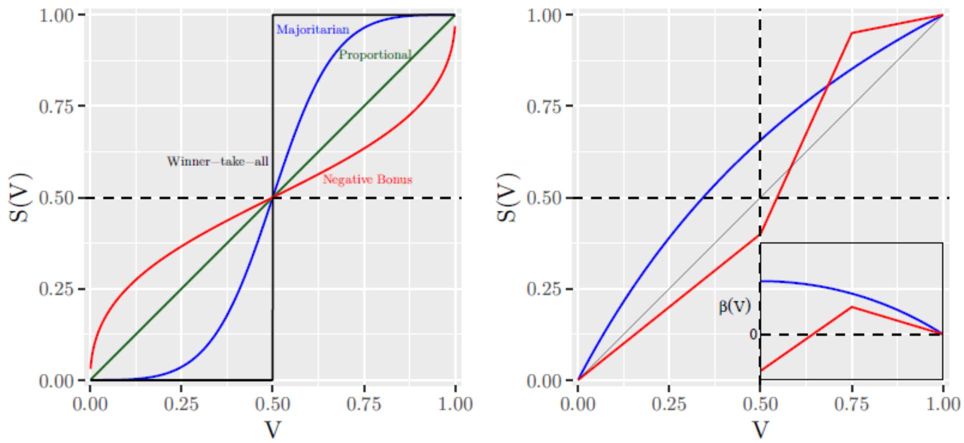


Figure 5

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 6

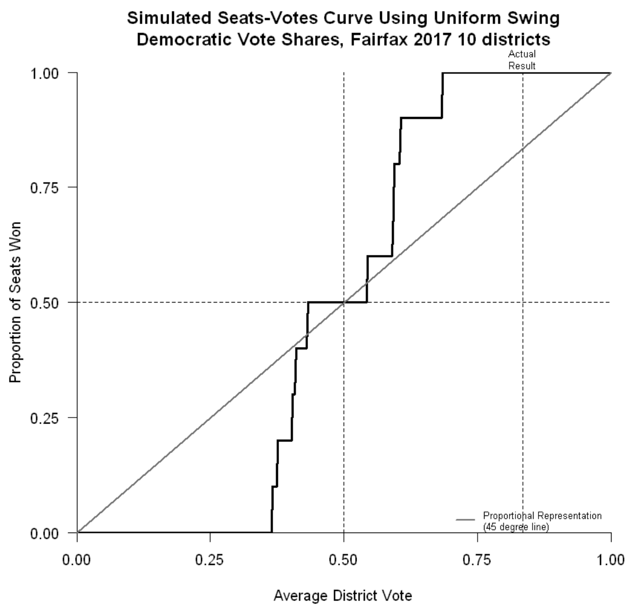


Figure 6

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)

Responsiveness

Katz et al. (2020) defines responsiveness as how much the seat proportion changes in response to a change in the average district vote. Mathematically speaking, responsiveness is directly related to the derivate (the slope) of the seats-votes curve.[10] Typically, responsiveness is measured at an average district vote of 0.5 or at the observed average district vote (Katz et al., 2020).

Lopsided Outcomes

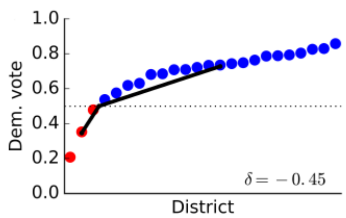
The lopsided outcomes measures, proposed by Wang (2016), is the difference in the average win size for each party. Katz et al. (2020) refutes the claim by Wang (2016) that the lopsided outcomes measure is a measure of partisan symmetry because it breaks down in noncompetitive elections. In such elections, the "packing" and "cracking" that normally indicate gerrymandering are merely a consequence of the political geography. However, the lopsided outcomes measure does serve as an indication of the skewness of the average vote distribution.

Declination

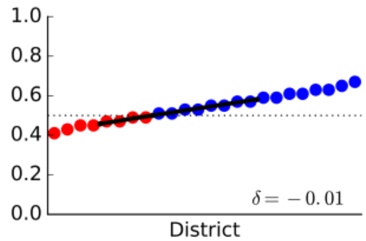
Declination is another proposed measure of partisan asymmetry that "...relies only on the fraction of seats each party wins in conjunction with the aggregate vote each party

[10] Since seats-votes curve are step functions (and are therefore not differentiable), measuring responsiveness requires applying a smoothing function to the seats-votes curve first to make it continuous (Katz et al., 2020).

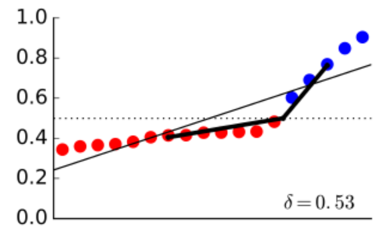
uses to win those seats" (Warrington, 2018, p. 3). Broadly, it measures the declination in the line connecting the average vote proportions of one party in districts controlled by the other party. It is essentially the normalized version of the declination measure (Katz et al., 2020). See Figure 7 for a visualization.



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



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



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

Figure 7

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)

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 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).

Equal Population Efficiency Gap

After proving that the efficiency gap is not a measure of partisan symmetry, Katz et al. (2020) suggests a *equal population efficiency gap*, which is the efficiency gap calculated under an assumption of equal population in each district, rather than total votes cast irrespective of district. This measure produces a linear seats-votes curve with a slope of 2, which does represent partisan symmetry, but it only captures one specific case of symmetry.

Figure 8 illustrates how the equal population efficiency gap could deem a map to be fair, but the complete seats-votes curve reveals the strong partisan bias towards Republicans.

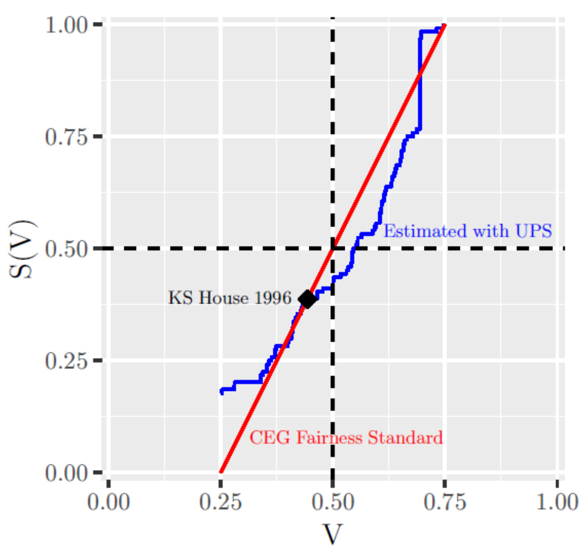


Figure 8

Seats-votes curve of the equal population efficiency gap and the Kansas House Election results of 1996, estimated using Uniform Partisan Swing.

Tau Gap

The tau gap, a variation of the efficiency gap proposed by Warrington (2018), takes a different view on wasted votes. Instead of defining every vote by the winning party in a winning district over the "50%-plus-one" hurdle to be wasted, the order the votes, and declare votes 0-50% not wasted, 50-τ% partially wasted, and τ-100% wasted, where τ is an chosen percentage between 50% and 100%. This is supposedly more in-line with the idea that it would be very risky for a party to aim for 50%-plus-one votes, and that "buffer votes" more closely resemble reality (Warrington, 2018).

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).

Schwartzberg Score

The Schwartzberg score is very similar to the Polsby-Popper Score, but it calculates the ratio of of the perimeter of the district to the perimeter of a circle with the same area as the district (Schwartzberg, 1966). It is calculated as follows.

$$S(d) = \frac{P(d)}{2\pi\sqrt{\frac{A(d)}{\pi}}} \tag{2}$$

where $S(d)$ is the Schwartzberg score. $S(d)$ ranges from 0 to 1, with 0 again being the least compact and 1 being the most compact district.

While the Schwartzberg measure is very sensitive to almost all forms of noncompactness (eg. spikes, cross-shapes, etc.) (Polsby & Popper, 1991, p. 301), it is hampered by the same issues as the Polsby-Popper score, namely that it incorrectly characterizes coastal districts as gerrymandered and is sensitive to resolution.

Length-Width Score

The Length-Width score requires a minimum-bounding rectangle to be drawn around the district. The length-width score is the ratio of the shorter side of the rectangle to the longer side (i.e. a district encapsulated by a square would have a score of 1) (Harris, 1964). Just like the previous two measures, the Length-Width score ranges from 0 to 1, with a score of 1 indicating the most compact district.

Criticisms of this measure include that it is not sensitive to perimeter changes, meaning that a square district and a square district of the same size with many indentations would have the same compactness score (Polsby & Popper, 1991).

Convex Hull Score

The Convex Hull score calculates the ratio between the area of the district and the area of the convex hull of the district. A convex hull is the smallest possible polygon with entirely acute interior angles that contains the desired points. It again ranges from 0 to 1, with 1 being the most compact district. (Maceachren, 1985)

Criticism of the Convex Hull score primarily address that coastal districts may have lower scores even though they may be as compact as possible given their geographic constraints.

Reock Score

The Reock Score is the ratio between the area of the district and the area of the smallest possible circle that encloses it, known as the minimum bounding circle. It is very similar to the Length-Width score and the Convex-Hull score, and it also ranges from 0 to 1, with 1 indicating the most compact district, a circle. (Reock, 1961)

Criticism of the Reock Score include that it does not account for internal indentations associated with "salamander-like" districts, just like the Length-Width score (Maceachren, 1985).

Boyce-Clark Index

The Boyce-Clark Index scores the compactness of districts using the length of the radii from the geometric centroid, or "center of mass", of the district. It is computed as

$$1 - \sum_{1}^{16} \{ \frac{|\frac{r_i}{\sum_i r_i} * 100 - 6.25|\}}{200} \tag{3}$$

,

where r_i are the 16 individual radii. A perfect circle will have a Boyce-Clark index of 1. (Boyce & Clark, 1964; Fifield, Kenny, et al., 2020)

Criticisms of this measure include that it would score districts that must naturally be concave, possibly due to coastal geography, lower than is fair. Since the resolution is 16 radii, it is also sensitive to the orientation of the district (Boyce & Clark, 1964).

Fryer-Holden Score

The Fryer-Holden score takes a different approach to quantifying compactness, as it takes the population of each precinct into account. This score is "based on the distance between voters within the same political district in a state relative to the minimum such distance achievable by any districting plan in that state" (Fryer & Holden, 2007, p. 2). A score of 1 indicates a compact district.

This measure is robust to some of the weaknesses of the previous measures, namely that it isn't affected by coastal geography, map resolution, or population density. It can also be compared across states. (Fryer & Holden, 2007)

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

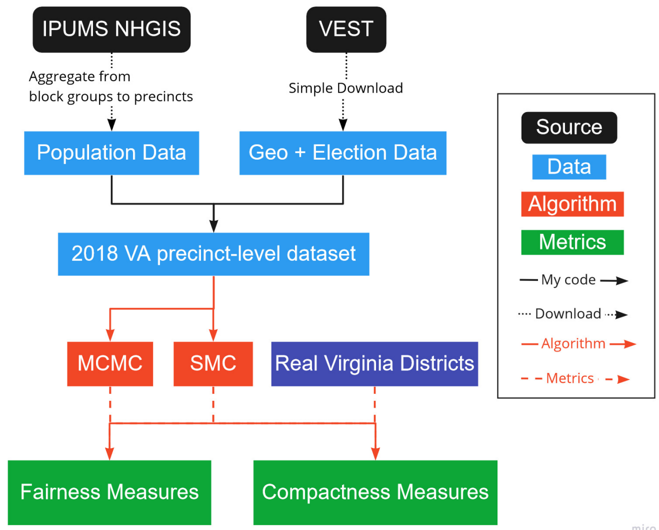


Figure 9

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 9 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.[11] Every row in each dataset corresponds to a precinct, the smallest geographical unit by which votes are tabulated in Virginia. For each

[11] This corresponds to the blue rectangles in Figure 9.

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).[12] 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.[13]

Response Variables. Broadly, the goal will be to evaluate how "fair" and compact each redistricting plan generated by each algorithm for each year is.[14] 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,

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

[13] 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.

[14] This corresponds to the green rectangles in Figure 9.

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[15] 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.[16]

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.

[15] Due to racial gerrymandering, VA had to adopt a new congressional district map in the middle of the decade (Breyer, 2016).

[16] This corresponds to the purple rectangle in Figure 9.

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.[17]

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. [18]

---

[17] This is an explanation of the black and blue rectangles in Figure 9 and the transitions between them.

[18] 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.[19] Conveniently, this data was already included in my shapefiles by the Voting and Election Science Team.

[19] 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.

Overall Measures

	alg	control	mcmc	smc
type	measure			
	ecc	0.777428	0.774128	0.816295
compact	fh	2.15416e+21	1.66564e+21	1.84849e+21
	pp	0.185835	0.172610	0.165511
	dseats	7.00000	9.00000	7.28000
	bias	-0.136364	-0.136364	0.0436364
	dec	0.0406080	-0.485947	-0.0304248
	effgap	-0.00292268	-0.179412	-0.0170277
fair	effgapeqpop	-0.0220841	0.172966	0.0204446
	lopwin	0.0886207	0.000318324	0.0664807
	meanmed	0.0680291	0.0533195	-0.0159417
	resp	0.00000	9.09091	0.363636
	taugap	-1.14986	-0.0848905	-0.368612

Table 1

Mean Fairness and Compactness Measures for MCMC, SMC, and Control

Table 1 shows the mean fairness and compactness measures for the redistricting plans generated by the two algorithms and the existing congressional districts.

For the Edge-Cut Compactness and Fryer-Holden score, the table contains the mean of the respective measures for the 100 plans generate by the algorithms. For the Polsby-Popper score, the table contains the mean of the district-level mean Polsby-Popper score for all 100 plans. The compactness measures were normalized so that values of 1 indicate more compact districts and values of 0 represent the least compact districts.

The fairness measures in the table are all the mean of said fairness measure for each redistricting plan generated by each algorithm. "dseats" is the average number of seats that Democrats would have held had the 2018 General Election taken place with said redistricting plan. All other remaining fairness measures have been normalized, where values of 0 indicate fairness and deviation in either direction indicates a partisan imbalance.

Compactness

Partisan Fairness

Figure 11 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 11a and 11b 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 11c.

2018 General Election Simulation

Figure 12 visualizes the outcome of a simulated 2018 General Election under various redistricting plans. Subfigure 12a was created by aggregating the precinct-level election results from 2018 to the district level using the redistricting plan generated by MCMC with the least 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 12b 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 12c for reference.

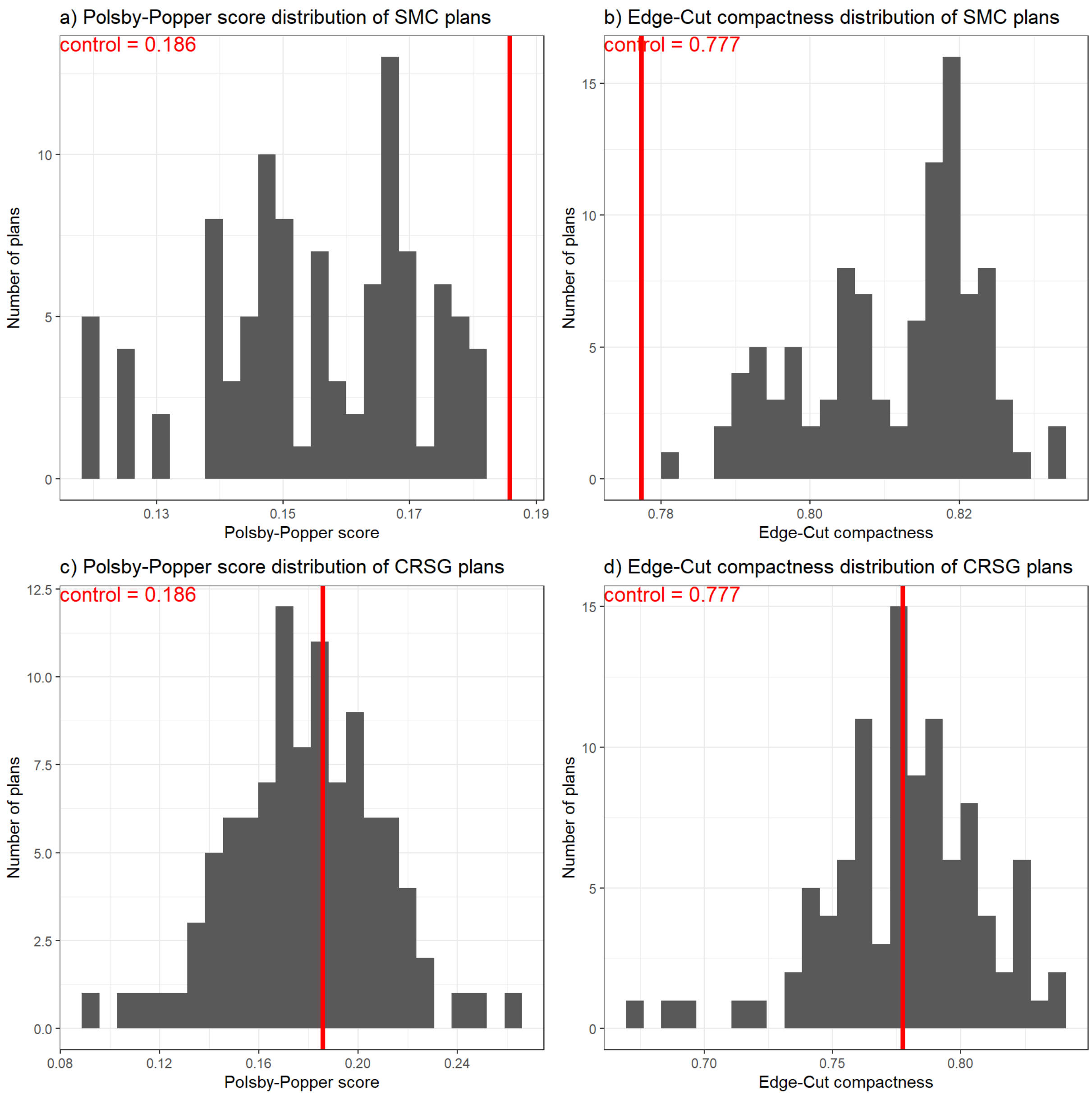
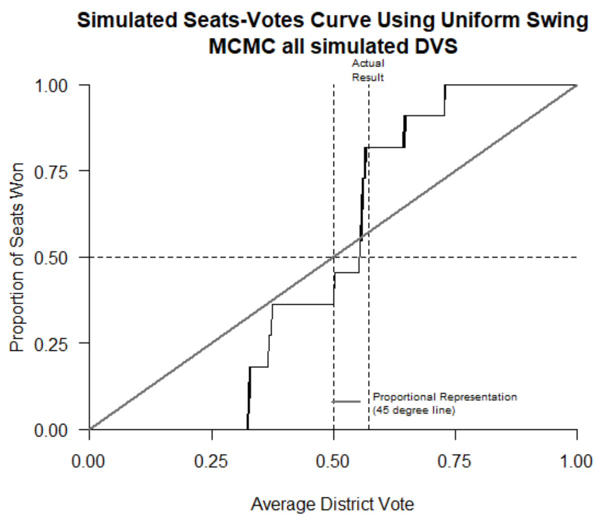
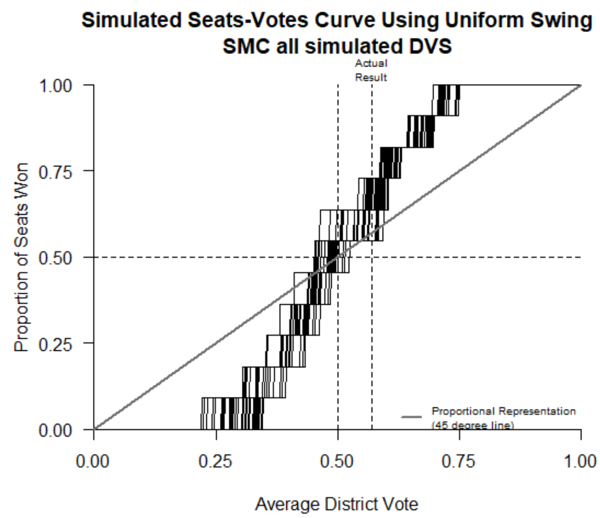


Figure 10

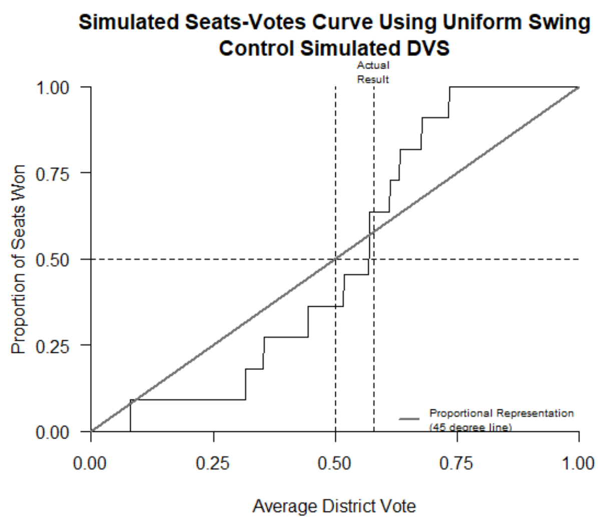
Distributions of compactness measures.



(a) *MCMC Seats-Votes Curve*



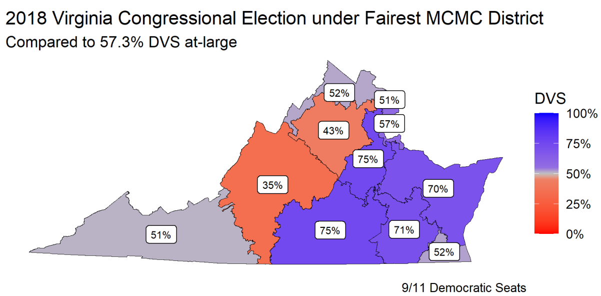
(b) *SMC Seats-Votes Curve*



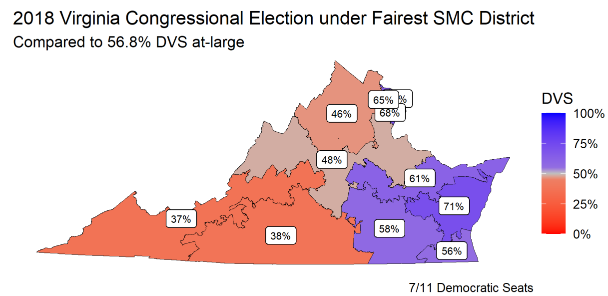
(c) *Real Seats-Votes Curve*

Figure 11

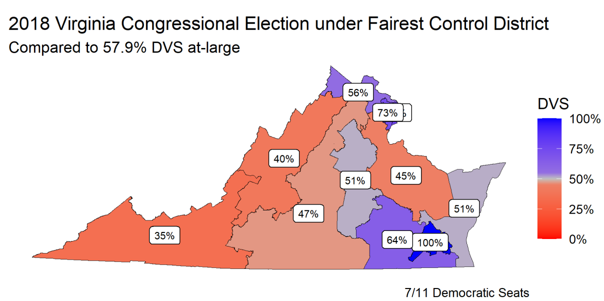
Seats Votes Curves



(a) *2018 Election under Fairest MCMC Map*



(b) *2018 Election under Fairest SMC Map*



(c) *2018 Election under Existing Map*

Figure 12

Simulated Elections under different Redistricting Plans write!

Write!

References

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

Boyce, R. R., & Clark, W. A. V. (1964). The Concept of Shape in Geography. *Geographical Review*, *54*(4), 561. https://doi.org/10.2307/212982

Breyer, S. (2016). Wittman v. Personhuballah.

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

Fryer, J., & Holden, R. (2007). Measuring the Compactness of Political Districting Plans. *The Journal of Law and Economics*, *54*. https://doi.org/10.1086/661511

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

Harris, C. C. (1964). A Scientific Method of Districting. *Behavioral Science*, *9*(3), 219–225. https://doi.org/10.1002/bs.3830090303

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

Maceachren, A. M. (1985). Compactness of Geographic Shape: Comparison and Evaluation of Measures. *Geografiska Annaler: Series B, Human Geography*, *67*(1), 53–67. https://doi.org/10.1080/04353684.1985.11879515 _eprint: https://doi.org/10.1080/04353684.1985.11879515

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

Reock, E. C. (1961). A Note: Measuring Compactness as a Requirement of Legislative Apportionment. *Midwest Journal of Political Science*, *5*(1), 70–74. https://doi.org/10.2307/2109043

Schwartzberg, J. (1966). Reapportionment, Gerrymanders, and the Notion of Compactness. *Minnesota Law Review*, *50*, 443–452. https://scholarship.law.umn.edu/mlr/1701

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

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

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

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

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