An Exploration of a Potential Solution to Gerrymandering

Lisa W. Kay[1], Shane P. Redmond[1]
[1]Eastern Kentucky University, 521 Lancaster Avenue, Richmond, KY 40475

Abstract
Elections in the United States, such as for the U.S. House of Representatives, typically divide a state into *n* voting districts with equal populations. It has been shown that this method does not tend to produce proportional representation, where the percent of the state's districts won by a certain political party matches the total statewide percent of the vote won by that party. For example, in Wisconsin's 2016 congressional elections, Republicans won 62.5% of districts but had only 45.8% of the statewide vote in that election. This phenomenon has frustrated some citizens' sense of fairness and has led to charges of gerrymandering by one party or another. A hypothesized solution to this issue is the suggestion that increasing the number of voting districts makes it more likely that elections will be proportional. This project investigates this hypothesis through simulated elections to track the percentage of elections that are proportional as the number of districts varies.

Key Words: gerrymandering, elections

1. Introduction

Elections in the United States, such as for the U.S. House of Representatives, typically divide a state into *n* voting districts with roughly equal populations. It has been shown that this method does not tend to produce proportional representation, where the percentage of the state's districts won by a certain political party matches the total statewide percentage of the vote won by that party. For example, in Wisconsin's 2016 congressional elections, Republicans won 62.5% of districts but had only 45.8% of the statewide vote in that election. This phenomenon has frustrated some citizens' sense of fairness and has led to charges of gerrymandering by one party or another. A hypothesized solution to this issue is the suggestion that increasing the number of voting districts makes it more likely that elections will be proportional. This project investigates this hypothesis through simulated elections to track the percentage of elections that are proportional as the number of districts varies.

The authors wanted to determine how often the percentage of seats won is "about the same" as the percentage of the vote across the state. Thus, a bound was set (for example, say the election is proportional when the two percentages differ by five percentage points or less), it was determined how many simulated elections fell within that bound, and that was converted into a percentage. Let $P(n, m)$ be the percentage of elections in a state with n districts that fall within m percentage points of being proportional. Another potential measure of fairness is the efficiency gap (Stephanopoulos & McGee, 2014). The aim was to run simulations that would calculate $P(n, m)$ and the efficiency gap as n increases but m

stays fixed. The authors hypothesized that $P(n, m)$ should increase as n increases, but they were interested in seeing by how much and what the gains look like for large values of n, and they wanted to explore what happens when m varies. They considered n as large as 200, although this could be too large in practice; according to the United States Census Bureau (2011), California has the most districts with 53 while a few states have only one district. Also, not many elections have outcomes in which one party earns 95% of the vote, so the authors wanted to run some simulations for which the election outcomes were restricted to more realistic intervals (e.g., instead of choosing a random number between 0 and 100 for each district's percentage, choose a random number between 30 and 70).

Samuel S.-H. Wang (2016) utilized simulations to randomly select "combinations of districts from around the United States that add up to the same statewide vote total for each party" (p. 1289) in order to examine the effects of gerrymandering. Mira Bernstein (2017) conducted simulations using a uniform distribution on the interval (0, 1), a truncated normal distribution with mean 0.55 and standard deviation 0.2, and a normal distribution with mean 0.5 and standard deviation 0.05. Bernstein also utilized 2016 Florida election data to conduct simulations. The present study includes similar simulations but focuses on the effect of varying the number of districts.

2. Methods

One million simulations were conducted with different combinations of the following: number of districts (5, 10, 20, 30, 40, 50, 75, 100, 200); margin (3, 5); distribution used to generate district percentages—uniform distribution on the interval (0, 100), uniform distribution on the interval (30, 70), truncated normal distribution (mean 50, standard deviation 17, minimum 0, and maximum 100); and data from recent elections (2006, 2008, 2010, 2012, 2014, 2016). For each case, R was used to generate the following: $P(n, m)$, the percentage of simulated elections in a state with n districts that fall within m percentage points of being proportional, and the mean of the absolute values of the efficiency gaps of the simulated elections in a state.

3. Results

R was used to simulate data from two different uniform distributions (runif), a truncated normal distribution (rtruncnorm), and a data set composed of actual district election percentages. The mean (50) and standard deviation (17) of the truncated normal distribution were estimates based upon the data gathered from recent elections. Election data for Republican candidates were gathered from Wikipedia (n.d.). Elections were omitted under the following circumstances: a candidate ran unopposed, there were multiple Republican candidates, there was not both a Republican and a Democrat candidate, and more than 5% voted for third-party candidates. A summary of results is given in Table 1.

Table 1. Election Simulation Results

Results of 1,000,000 Simulations *n* = number of districts, *m* = margin									
		Uniform(0, 100)		Uniform(30, 70)		Truncated Normal		Election Data	
n	*m*	*P*(*n*, *m*) %	Mean Absolute Efficiency Gap	*P*(*n*, *m*) %	Mean Absolute Efficiency Gap	*P*(*n*, *m*) %	Mean Absolute Efficiency Gap	*P*(*n*, *m*) %	Mean Absolute Efficiency Gap
5	3	17.89	0.1041	1.58	0.1162	9.52	0.1104	10.33	0.1086
10	3	25.39	0.0732	22.20	0.0817	19.77	0.0776	19.13	0.0831
20	3	35.50	0.0517	24.29	0.0575	27.28	0.0547	25.34	0.0677
30	3	42.96	0.0421	31.55	0.0469	33.36	0.0446	29.26	0.0622
40	3	48.65	0.0365	35.94	0.0406	38.19	0.0386	31.81	0.0595
50	3	53.59	0.0327	39.86	0.0363	42.28	0.0345	33.52	0.0581
75	3	63.13	0.0266	47.81	0.0297	50.59	0.0282	35.82	0.0564
100	3	70.04	0.0230	54.07	0.0257	57.04	0.0244	36.27	0.0559
200	3	85.81	0.0163	70.56	0.0182	73.58	0.0173	33.77	0.0555
5	5	29.35	0.1040	7.90	0.1163	18.25	0.1105	19.08	0.1087
10	5	41.05	0.0732	26.56	0.0816	31.28	0.0777	31.20	0.0830
20	5	55.91	0.0516	42.84	0.0576	44.07	0.0546	41.18	0.0676
30	5	65.58	0.0421	49.62	0.0469	52.63	0.0447	47.03	0.0622
40	5	72.67	0.0365	56.39	0.0406	59.41	0.0386	50.95	0.0595
50	5	77.89	0.0326	61.73	0.0363	64.73	0.0345	53.67	0.0581
75	5	86.58	0.0266	71.56	0.0296	74.66	0.0282	57.85	0.0565
100	5	91.70	0.0231	78.34	0.0257	81.28	0.0244	60.35	0.0559
200	5	98.58	0.0163	92.00	0.0181	93.74	0.0173	65.72	0.0556

4. Conclusions

The simulation results indicate that $P(n, m)$, the percentage of simulated elections in a state with n districts that fall within m percentage points of being proportional, tends to increase as the number of districts increases. One exception is the case in which the simulations are based on election data, the margin is three percentage points, and the number of districts is quite high (200). $P(n, m)$ is higher for a margin of five percentage points than for a margin of three percentage points, as expected. $P(n, m)$ is fairly high (> 90%) for 200 districts and a margin of five percentage points for the two uniform distributions and the truncated normal distribution; it is close to 1 for the uniform distribution on the interval (0, 100). It is close to 66% for the election data with a margin of five percentage points and 200 districts.

The mean of the absolute value of the efficiency gap decreases as the number of districts increases. The mean of the absolute value of the efficiency gap falls below 0.02 for the two uniform distributions and the truncated normal distribution with 200 districts, but not for the election data.

Initial results suggest that increasing the number of districts might result in greater fairness, as measured by $P(n, m)$ and the efficiency gap. (The desirability of having more politicians is another matter.) It is worth noting that the election data set may not be particularly well modeled by the uniform or truncated normal distributions.

5. Suggestions for Future Studies

One possibility for future studies would be to run simulations with other models besides the uniform and the truncated normal distributions. Considering other reasonable values for the margin might also be worth exploring. There are other existing potential measures of fairness that could be examined, and creating an original measure is another possibility.

References

Bernstein, M. (2017, November). *Measures of partisan fairness*. Paper presented at Geometry of Redistricting conference, Durham, NC. Retrieved from <https://sites.duke.edu/gerrymandering/files/2017/11/MB-duke-slides.pdf>

Stephanopoulos, N. O. & McGhee, E. M. (2015). Partisan gerrymandering and the efficiency gap. *University of Chicago Law Review*, *82*: 831-900.

United States Census Bureau. (2011, November). Congressional apportionment. Retrieved from <https://www.census.gov/prod/cen2010/briefs/c2010br-08.pdf>

United States House of Representatives election, 2006. (n.d.) In *Wikipedia*. Retrieved July 22, 2018, from <https://en.wikipedia.org/wiki/United_States_House_of_Representatives_elections,_2006>

United States House of Representatives election, 2008. (n.d.) In *Wikipedia*. Retrieved July 22, 2018, from <https://en.wikipedia.org/wiki/United_States_House_of_Representatives_elections,_2008>

United States House of Representatives election, 2010. (n.d.) In *Wikipedia*. Retrieved July 15, 2018, from <https://en.wikipedia.org/wiki/United_States_House_of_Representatives_elections,_2010>

United States House of Representatives election, 2012. (n.d.) In *Wikipedia*. Retrieved July 15, 2018, from <https://en.wikipedia.org/wiki/United_States_House_of_Representatives_elections,_2012>

United States House of Representatives election, 2014. (n.d.) In *Wikipedia*. Retrieved July 15, 2018, from <https://en.wikipedia.org/wiki/United_States_House_of_Representatives_elections,_2014>

United States House of Representatives election, 2016. (n.d.) In *Wikipedia*. Retrieved July 15, 2018, from

<https://en.wikipedia.org/wiki/United_States_House_of_Representatives_elections,_2016>

Wang, S. S.-H. (2016). Three tests for practical evaluation of partisan gerrymandering. *Stanford Law Review*, *68*, 1263−1321.