Choosing the Best Candidate: An Analysis of Voting Systems using a Monte Carlo Method

Ashton Keith

Franklin Academy High School, 42 Husky Lane, Malon
xxphignewtonxx@gmail.com

ABSTRACT: Democracy is one of the fundamental principles underpinning the United States. Recently, the plurality voting system used in our presidential elections has come under scrutiny. Problems such as the spoiler effect and the two-party system have been shown to be directly caused by plurality. Discussions over an alternative cannot determine which alternate system should be used. Every system is susceptible to unwanted paradoxes and potential for tactical voting.[1] Warren Smith used computer simulations under a Monte Carlo method with voters' feelings toward candidates based on stances on a number of issues to determine the system with maximum voter satisfaction. He found that range voting worked best in every trial.[2]

In this study, I will extend Smith's research by using more voters and basing feelings towards candidates on the distance between leanings on a political spectrum using six different voting systems. 10 million elections with 3, 4, 5, 7, 8, 10, 12, 15, 20 and 25 candidates were tested. This study determined that Copeland's Method worked best in every situation, matching a theory made by Davis, DeGroot, and Hinich,[3] followed by approval voting under manipulation and Borda under honesty. Additionally, both approval and plurality voting performed better under tactical voting while the Borda count chose the worst candidate most of the time under the same manipulation.

KEYWORDS: Statistics; Monte Carlo Study; Social Choice; Utility; Voting System; Strategic Voting; Moving Average Strategy.

Introduction.RATIONALE. Citizens believe they understand how voting works: go into a voting booth, select our favorite candidate, and the candidate that receives the most votes wins. Most will be surprised to learn that this is not the only way votes could be cast and counted and that, in fact, it's one of the more mediocre ways of choosing a candidate that satisfies the voters' wishes. Due to the controversial elections of 2000 and 2016, national attention has been focused on how exactly votes are cast. Along with the Electoral College, not addressed in this paper, the plurality voting system has come under considerable scrutiny. Some might say that the system the U.S. uses has been good enough to work for over 200 years and doesn't require a replacement. However, supposing this is true, countries that have only recently adopted democracy are at serious risk of losing it again. If voters who have never experienced a democratic election before are disappointed by the results, they may be willing to abandon democracy all together. We've seen democracies in Latin America, Africa, and the Middle East fall apart after implementing a plurality system. By finding the right voting system to use, a newly established democracy may be able to keep their voters satisfied long enough to maintain stability. Additionally, the effects of different voting systems are relevant in other fields such as computer science and machine learning, corporate decisions, and the stock market.

BACKGROUND RESEARCH. Alternatives such as range voting, Borda count, instant runoff, and the Condorcet methods have advantages over plurality, but each have their own problems. Which system selects the best candidate is hard to determine and is subject to philosophical discussion. It has been shown that no voting system is perfect: the Gibbard-Satterthwaite Theorem states that all systems discourage voting honestly in some situations.[1] Arrow's theorem states that the only ranked voting system that does not restrict how one votes, does not pick one candidate when voters unanimously prefer another, and is unaffected by adding in losing candidates, is a dictatorship.[4] Furthermore, some systems that seem reasonable are, in reality, so bad that it is better to choose a winner randomly, like bullet voting.[2]

RESEARCH GOALS. Historically, the effectiveness of voting systems has been determined through mathematical models. This study plans to run simulations on how well each system determines the best candidate and how consistently they do so. This was partially inspired by the work of Smith. However, I will base how well each voter likes a candidate on how close their political leanings are instead of stances on specific issues. This study assumes a utilitarian model of quality, or that the best voting system is that which maximizes the expected utility of society.

I compared four of the most favored voting systems in the literature: range voting,[2] Borda count,[5] Condorcet methods (here represented by Copeland's method),[6] and Hare STV,[7] along with plurality and the lottery system as references.

HYPOTHESIS. It has been shown that, under a distribution of voters symmetric across some point and a utility function that decreases with distance, Condorcet methods will have zero Bayesian regret as the number of voters goes to infinity.[3],[8]

© 2020 Terra Science and Education

DOI: 10.36838/v2i2.6

Therefore, I expect the Copeland's method to have near zero expected Bayesian regret for each situation simulated.

After Copeland's method, I predict that range voting will perform the best in deciding the most desired candidate due to the study done by Smith.[2] My research also shows that range voting has received the least criticism and those made typically deal with the potential for voters to reduce the ballot to either approval voting or plurality.

Results. For the sake of brevity while showing trends with the number of candidates, only results for 3, 15, and 25 candidates are shown. All errors represent the 99% confidence interval for the recorded averages. The highlighted rows in Table 1 are the voting systems with the least Bayesian regret among elections with more candidates and are graphed in Figure 1.

Table 1. 1000x average Bayesian regret of voters under each voting system in 3, 15, and 25 candidate elections.

Voting System	3 Candidates	15 Candidates	25 Candidates
Honest Plurality	2.11808±0.00567	25.3281±0.01941	34.8251±0.02219
MA Plurality	0.38688±0.00195	13.3740±0.01269	20.0535±0.01536
Honest Range	1.40410±0.00452	0.94987±0.00210	0.73014±0.00152
Honest Approval	2.33777±0.00726	3.21130±0.00483	2.51405±0.00349
MA Approval	0.03051±0.00027	0.23412±0.00107	0.24150±0.00097
Honest Borda	0.70203±0.00320	0.45114±0.00130	0.30524±0.00084
MA Borda	0.09226±0.00258	231.883±0.10201	289.837±0.09411
Honest Copeland	0.03050±0.00027	0.04243±0.00021	0.04338±0.00019
MA Copeland	0.09070±0.00255	1.35911±0.01584	2.19883±0.02155
Honest Hare	0.38687±0.00195	8.24258±0.00754	11.1631±0.00813
Lottery	25.9036±0.03878	53.5472±0.04567	59.5996±0.04728

In every scenario, Copeland's method resulted in the highest voter satisfaction. The range/approval and plurality voting systems performed significantly better when voters decide to vote tactically based on the polls, while the ranked voting systems all suffered under tactical voting. The plurality, Hare STV, and honest approval voting systems consistently performed worse in every scenario. The Borda count and range voting's performance improved relative to the other systems as the number of candidates grew –

as long as the voters were honest. The Borda count under manipulation suffered terribly, having an average regret of 0.28984 vs. 0.00031 under 25 candidates. The lottery system had one-fifth the regret with 25 candidates. I determined that, due to how the strategy works and that voters tend to like those higher in the polls, the very worst candidate ends up placing just above halfway in nearly every ballot, enough for them to win almost every time.

Discussion. Honest Condorcet, moving average approval, and

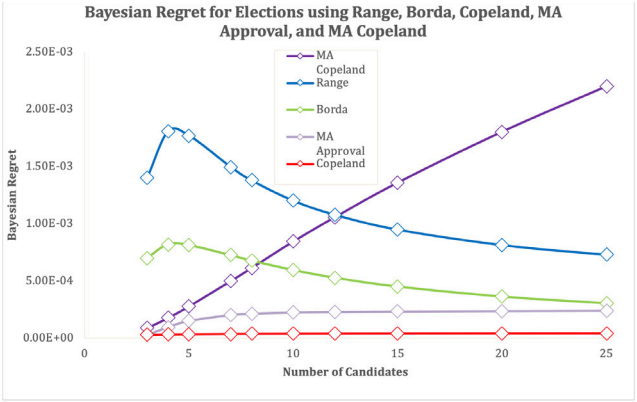

Figure 1. The average Bayesian regrets of Approval Voting with Moving Average Strategy, Honest Borda Count, Honest Range Voting, and Copeland's Method both Honest and with Moving Average Strategy.

the honest Borda methods all performed better than range voting. What benefitted each can be traced back to the model used. The Condorcet methods benefit due to the distribution of voters was perfectly symmetrical across 0; therefore, it was able to have near zero regret according to the study by Davis, DeGroot, and Hinich.[3] Honest Borda was able to benefit from the fact that it suffers from the reverse of the spoiler effect: candidates do better when a losing candidate of similar views is also running.[9] Because the views nearer the mean of the voters' will be more crowded by candidates, more preferred candidates will have an extra boost in the election. The distribution of views in Smith's study was uniform, preventing this from occurring. As for the moving average approval vote, the candidates at the top of the polls are considered first weeds out competition for the most preferred candidate who is likely to end up second or third in the ranking for honest range voting. The moving average strategy usually corrects this, with the candidate receiving a majority approval and the nature of the polls causing all other candidates to receive far less than the majority. The Borda count works in the same way, but approval voting avoids choosing the worst candidate by not being a ranked system; there is no half-way up the ballot for the worst candidate to be forced on.

Prior research has shown that systems similar to range, plurality, and approval voting are very susceptible to tactical voting[10],[11] but as this study shows, this is not necessarily a downside. In fact, according to a study by James Quinn, plurality, approval, and range voting were the most susceptible to tactical voting of all methods tested yet they benefitted from this manipulation. Smith's study showed manipulation performs significantly worse under every voting system,2 but this study disproved that. Borda count's quirk of almost always choosing the worst candidate appears in Quinn's study but is not further elaborated.[10] My study demonstrates the effects of manipulation depend heavily on the model voters fit into. With the issues-based utilities used in Smith's study views on candidates were heavily divided and the effect of the worst candidate always ending up in the top half of the ballots does not occur.

The conclusions made in this study and Smith's study differ in several ways. The principal difference between our studies focuses on how voters decide how much they like each candidate, or the systems' utility functions. Previous research does not consider how the views of candidates may be similar between voters and instead focuses on worst-case scenarios. This study shows that utility function, whether it involves politicians' proposed policies or the personal economic impact of each possible alternative, is very important in determining a voting system's effectiveness.

There have been several attempts to implement the Hare STV system in mayoral and gubernatorial elections in the United States, but these have often been repealed not long after being implemented. The system was repealed by voters from Burlington's mayoral elections after only two elections and is currently facing legal trouble in Maine's gubernatorial elections.[12],[13] By determining what underlies their voters' views of each candidate, a voting district can decide which voting system results in the highest voter satisfaction.

Conclusion. By using a simulation and a utilitarian approach to investigate which voting system has the highest voter satisfaction, this study reveals new information about the nature of a variety of voting systems. First, optimal strategies affect voting systems very differently depending on the cause of voters' views of each candidate. Previous work has shown that some voting systems suffer terribly from strategic voting, including bullet voting, but these findings are true regardless of what the utility function is. Borda count suffers from the same problems, electing the very worst candidate the majority of the time only when voters' opinions of each candidate are highly correlated. Second, if voters vote tactically, the performance of a voting system does not necessarily perform any worse than if voters are honest. Both plurality and range/approval voting result in greater voter satisfaction when voters act in their own best interests. Third, approval voting under manipulation and Borda under honesty are the most utilitarian voting systems with the utility function used in this study. This is contrary to Smith's study in which range voting had the best performance overall.[2] Each of the three systems were able to exploit something in the model to their advantage, Condorcet exploited the symmetric distribution, Borda used the concentration of candidates around the mean view, and Approval exploited the strategy used. Hopefully, this study will help those deciding what voting system to use.

Methods. SET UP ELECTIONS. I created three candidates who will be running in the elections, each assigned two numbers representing their political leanings. This was inspired by how The Political Compass™ assigns political persuasions, as seen in Figure [2]. These values were assigned according to a Gaussian distribution with a standard deviation of 0. 2 and mean of 0.

One thousand voters were created to vote in the elections. The same process to determine political leanings was used as for the candidates.

RUNNING THE ELECTIONS. For each voter, the utility of each candidate was calculated using the following:

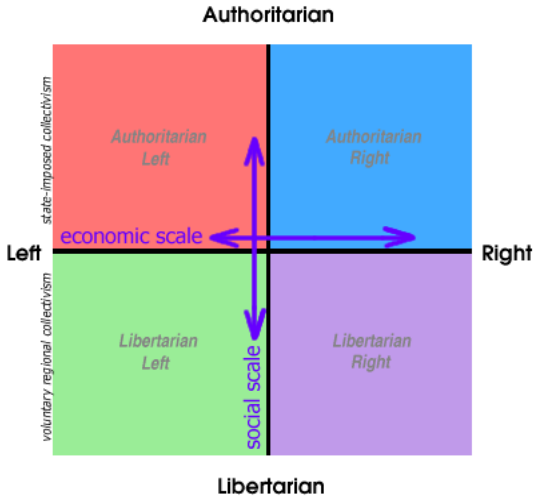

Figure 2: The 2D political spectrum that The Political Compass™ uses in assigning political leanings.[14]

$$U_C = -||\vec{v} - \vec{C}||^2,$$

where U_C is the utility of the candidate for that voter and v and c are the vectors representing the views of the voter and candidate, respectively. A higher utility represents a greater desire that a candidate wins.

Using the utilities of each candidate for each voter, ballots were filled out according to the following voting strategies: honest range, honest plurality, honest ranked [Borda count, Copeland's method, and Hare STV], and honest approval. The winner was determined under each of the following voting systems with their corresponding ballots: plurality voting, range/approval voting, Borda count, Copeland's method, Hare STV, and lottery. Using the results of the plurality, range, Borda count, and Copeland's method elections, ballots for each voter were filled out under their moving average strategy. The winner was determined according to the corresponding voting system.

DOI: 10.36838/v2i2.6

I calculate the Bayesian Regret, or the sum of the winner's utility to all of the voters minus the sum of the most preferred candidates, of each of the elections run. This can be found using the following equation:

$$D = ||\vec{W} - \vec{u}||^2 - ||\vec{B} - \vec{u}||^2,$$

where D is the Bayesian Regret of the election and W, B, and U are the vectors representing the winning candidate's views, most preferred candidate's views, and the mean of the voters' views, respectively. The most preferred candidate has the views closest to the mean of the voters.

REPEAT. For the elections with three candidates, a total of 10 million elections were run and their regrets analyzed for each of the voting systems and strategies. The most utilitarian voting system under the model used ends up producing the lowest average regret overall.

More elections were run using 4, 5, 7, 8, 10, 12, 15, 20, and 25 candidates using the same method used for 3 candidates. By doing this, I was able to determine whether a voting system's performance stays consistent in large and small elections.

The source code used in this study is available on GitHub™ at the following link:

https://github.com/AshtonKeith/Voting-System-Simulations

VOTING SYSTEMS.

A. Plurality voting is the system most people are familiar with. Each voter casts a vote by marking a single candidate from a list; the candidate with the most votes wins the election. This system is known to suffer from the spoiler effect, an inevitable trend towards only two candidate races and failing to meet the independence of irrelevant alternatives which means that removing a losing candidate from the ballots might change who wins the election.

B. In range voting each voter gives a rating for every candidate on a ballot. The candidate with the highest average rating wins the election. This system is known to fail the later-no-harm criterion, which means that voters will be encouraged to give some of their non-favorite candidates the lowest possible ratings, as well as the Condorcet criterion where a candidate may win even though a majority of voters prefer a different candidate. Approval voting is equivalent to range voting except every rating is restricted to 1 or 0. The paradoxes of range voting also apply to approval voting.

C. In Borda count each voter ranks the candidates in order of most to least favorite. Each candidate gets a vote for every other candidate they beat on each ballot; ex: if there are five candidates, the top preference gets four votes, second preference gets three, etc. The candidate with the most votes wins. This system does not satisfy the Condorcet criterion nor the independence of irrelevant alternatives. Voters are sometimes encouraged not to rank their top preference first.

D. In Copeland's method each voter ranks their candidates in order of preference. After all of the ballots are cast, each candidate is compared pairwise with every other. A candidate receives a point for every other candidate they rank higher than in a majority of ballots. The candidate with the most points wins. This voting system is the only one listed here that satisfies the Condorcet criterion and was chosen for this purpose. It is not independent of irrelevant alternatives, fails later-no-harm, and sometimes encourages voters to not rank their favorite candidate first. In fact, sometimes voters are better off not voting at all than to vote honestly.

E. In Hare STV, each voter ranks their candidates in order of preference. Votes are assigned to the top preference of each ballot and the candidate with the least votes is eliminated. The votes are redistributed to the top preference still running on each ballot. This continues until there is one candidate left standing. This voting system fails to satisfy the Condorcet criterion, independence of irrelevant alternatives, and the consistency criterion. Voters are sometimes discouraged from ranking their favorite candidate first or even from voting at all, and a voter may actually cause their favorite candidate to lose by ranking them higher on their ballot. Despite all of this, this voting system is the most popular system for head-of-state elections, behind plurality.

F. In lottery, a voter is chosen at random and their top preference wins the election. This system is not taken seriously and is only included for theoretical reasons. Although this voting system fails several criteria, it has one very important property: voters are always encouraged to vote honestly. Therefore, we can predict exactly how every voter is going to fill out their ballots if we know their views of each candidate.

Voting honestly in range voting and approval voting is not well defined. Voters can put down the exact utility each candidate has for them in range voting, but voters often only know the utility of one candidate compared to another. For this to work, all voters need to agree on what to compare the candidates against, which is not feasible. For this study, I defined honest range voting as marking the utilities of each candidate given by the simulation, transformed linearly so that a voter's favorite candidate receives a 1 and their least favorite candidate a 0. In approval voting I defined the threshold for marking 1 as the average utility of all of the candidates running. If a voter thinks a candidate is better than average, a one is put on the ballot for them.

I will also be using the moving average strategy for plurality, approval, Borda, and Copeland. Each voter considers each candidate in order of their likelihood of winning, starting with the top two. The voter gives the most votes available to the candidate they prefer more and the least to the one they prefer less. Each successive candidate is compared with all of the ones considered before them. If said candidate is liked more than the considered candidates on average, then the maximum number of votes available is given, and the least number of votes otherwise. This strategy was proven to be the optimal strategy for voters under plurality, range/approval, and Borda by Smith.[2]

Acknowledgements. I thank my Biology teacher and Science National Honors Society advisor, Ms. Rogers, for repeatedly reviewing the research design and final paper and pushing for my project to get presented to a wider community.

I also thank Mary Eileen Wood and the Terra Northeast Fair Judges for their interest and feedback, for choosing my project as a Ying Scholars Semifinalist, and for suggesting that I submit to this journal when my application for the Ying Scholarship failed to go through.

References.

(1) Gibbard, A. Manipulation of Voting Schemes: A General Result. Econometrica 1973, 41 (4), 587.

(2) Smith, W. D. Range voting, 2000. RangeVoting.org - Center for Range Voting. https://www.rangevoting.org/WarrenSmithPages/homepage/rangevote.pdf (accessed Mar 30, 2019).

(3) Davis, O. A.; DeGroot, M. H.; Hinich, M. J. Social Preference Orderings and Majority Rule. Econometrica 1972, 40 (1), 147-157.

(4) Arrow, K. J. A Difficulty in the Concept of Social Welfare. Journal of Political Economy 1950, 58 (4), 328–346.

(5) Saari, D. G. Geometry of Voting. Handbook of Social Choice and Welfare 2011, 897–945.

(6) Voting Research - Voting Theory. https://www.princeton.edu/~cuff/voting/theory.html (accessed Mar 30, 2019).

(7) Bartholdi, J.; Orlin, J. Single Transferable Vote Resists Strategic Voting. Social Choice and Welfare 1991, 8 (4).

(8) Smith, W. D. Best Voting Systems in D-Dimensional Politics Models. RangeVoting.org - Center for Range Voting. https://www.rangevoting.org/BestVot2.html (accessed Nov 08, 2019).

(9) Olson, B. Election Methods In Space! https://bolson.org/voting/sim_one_seat/www/ (accessed Dec 15, 2019).

(10) Quinn, J. Voter Satisfaction (VSE) FAQ | James Quinn. https://electionscience.github.io/vse-sim/VSE/ (accessed Nov 10, 2019).

(11) Balinski M.; Laraki R. Election by Majority Judgment: Experimental Evidence. In: In Situ and Laboratory Experiments on Electoral Law Reform; Dolez B., Grofman B., Laurent A., Eds.; Studies in Public Choice 25; Springer: New York, 2011; pp 13-54. https://www.fc.up.pt/dmat/engmat/2012/seminario/artigos2012/bruno_neto/ElectionByMajorityJudgment(ExperimentalEvidence)Final.pdf (accessed Dec 14, 2019).

(12).McCrea, L. Burlington Voters Repeal Instant Runoff Voting. https://vprarchive.vpr.net/vpr-news/burlington-voters-repeal-instant-runoff-voting/ (accessed Mar 30, 2019).

(13)Seelye, K. Q. Ranked-Choice Voting System Violates Maine's Constitution, Court Says. https://www.nytimes.com/2017/05/23/us/maine-ranked-choice-elections-voting.html?_r=0 (accessed Mar 30).

(14)About the Political Compass. https://www.politicalcompass.org/analysis2 (accessed Mar 30, 2019).

Author. Ashton Keith graduated from Franklin Academy High School in June 2019 and is currently attending SUNY Binghamton as a math and physics major. He has always been interested in math, statistics, and programming, and is grateful to have found a project that incorporates all three.

DOI: 10.36838/v2i2.6