Frequency Analysis of Non-Deterministic Social Choice Systems

Josh Sapira

December 10, 2023

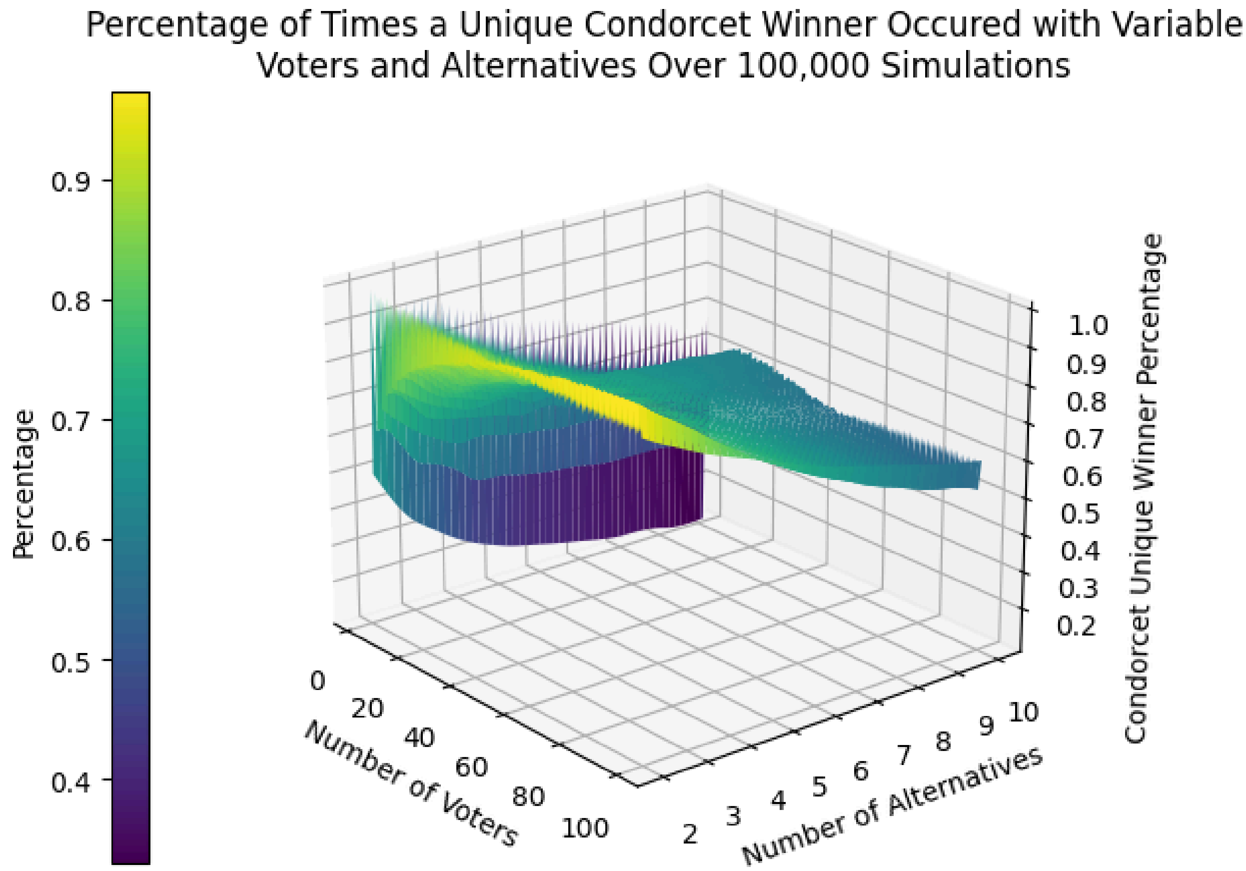
Abstract

In the past quarter-century, the United States has seen multiple massive failures of the electoral system. In this timeframe, the person elected was not the individual who received the most votes. Twice. Legal challenges have become all too common since the infamous recount in Florida during the 2000 presidential election, and with the increase in litigation has come a massive decrease in voter confidence in the results of the election. Before any sweeping changes are made to the US electoral system, there must be discussion about what this reform should aim to achieve: should we try to minimize wasted votes? Protect against tactical voting? Encourage third party participation? While this is a matter of opinion, there is one thing that any political electoral system must do, and that is select a single candidate as the winner. With this in mind, this report will analyze, compare, and discuss various social choice systems and the frequency at which they select a unique winner.

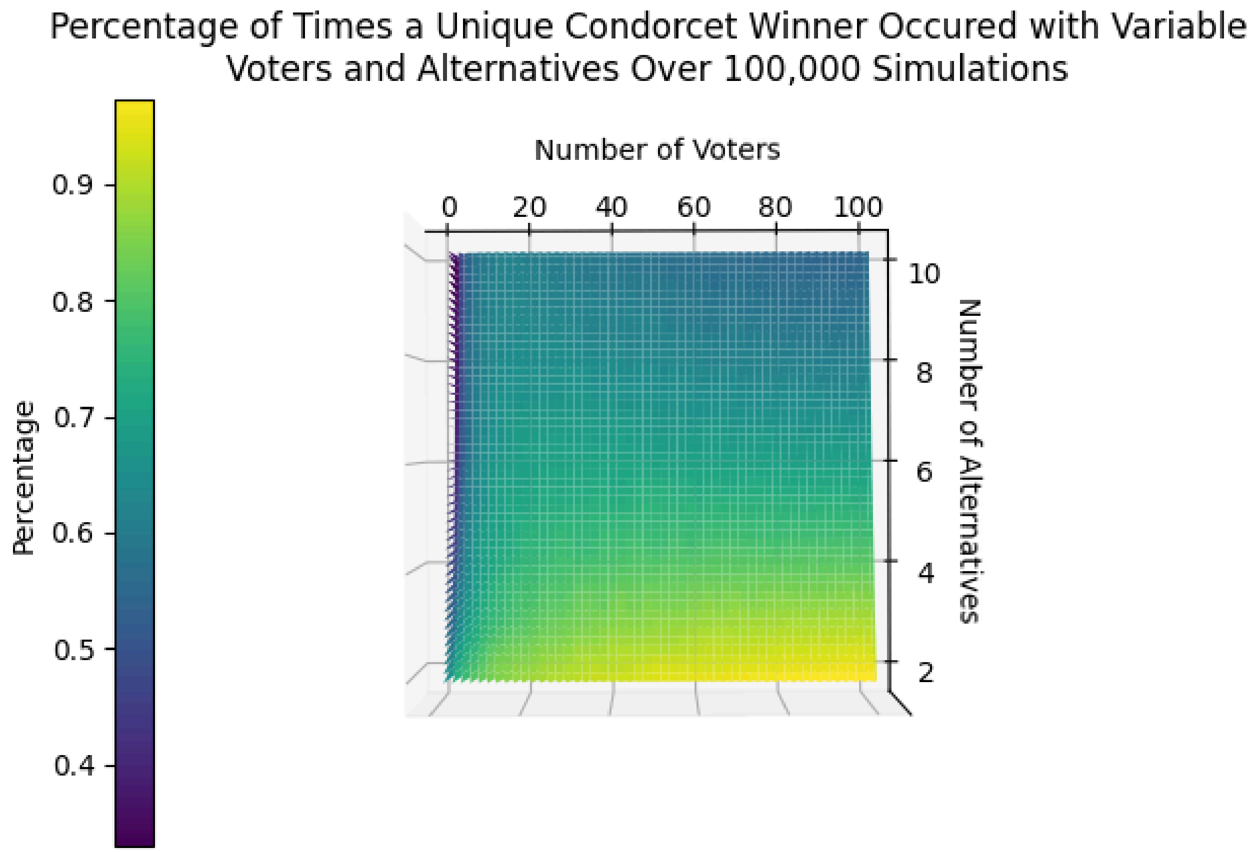
Condorcet's Method

Condorcet's Method elects the candidate(s) who receives no less than half of the votes in every head-to-head election against each other candidate, when such a candidate exists. The candidate(s) elected are deemed a “Condorcet winner”, though there is no guarantee that a candidate will be chosen. In fact, there is a well-defined case in which Condorcet's method fails to produce a Condorcet winner, and that is when the voter preferences are cyclical. For example, if, in an election between 3 candidates, the population prefers candidate A to B, candidate B to C, and candidate C to A, then no winner will be chosen.

In order to explore the frequency at which a unique winner is elected, I ran 100,000 simulated elections for every combination of population size (ranging from 2 to 100) and number of candidates (ranging from 2 to 10). The results can be seen in the 3-dimensional graph below.



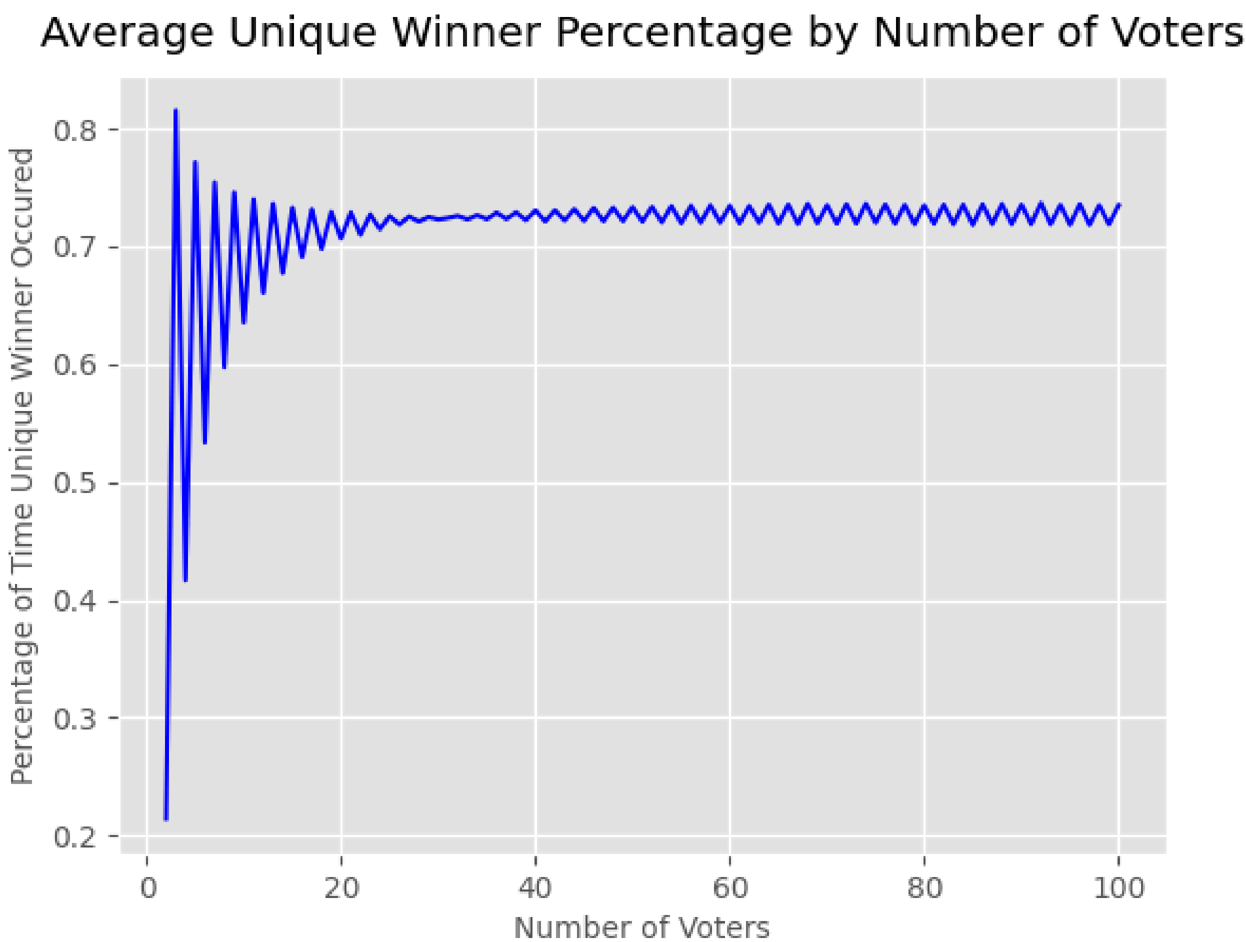
While the graph is hard to analyze without the ability to rotate the surface plot, there are some big-picture ideas we can see beginning to form. Firstly, when there are very few voters, Condorcet's Method elects a unique winner at the lowest frequencies. However, for populations of a certain minimum size, the unique winner percentage appears to be more or less stable, decreasing slightly as the number of candidates increases. These conclusions are reinforced by the following aerial view of the surface plot, which is effectively a heatmap.



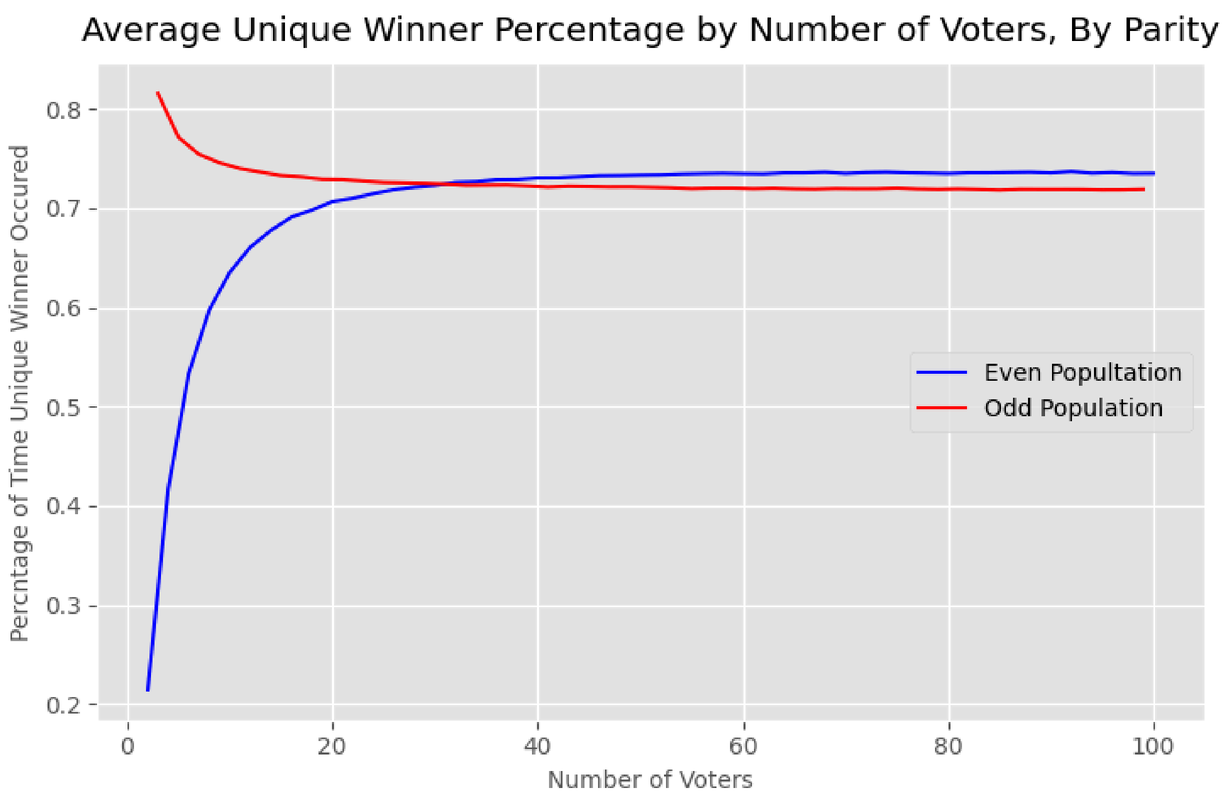
This graph confirms that the lowest unique winner percentage occurs when the number of voters is very small. However, once the population size increases beyond about 4-5 voters, we see much more encouraging results for Condorcet's Method. This heatmap also enables us to separately examine the relationship between the unique winner percentage and the number of candidates as well as the unique winner percentage and the population size.

For a fixed population size, it is evident that, as the number of candidates increases, there is a steady decrease in the unique winner frequency. This can likely be attributed to the fact that, when there are more candidates, it is more difficult for each candidate to be deemed a Condorcet winner as it requires winning more head-to-head matchups.

For any fixed number of candidates, the unique winner percentage increases as the population size increases. This can likely be attributed to the fact that cycles occur less frequently for larger population sizes, as it requires a delicate balance of votes cast to each candidate that becomes increasingly difficult to facilitate as the population increases. This pattern can be better observed in the below chart.



Clearly, there is a general increase in the unique winner percentage as the number of voters increases. However, this graph also exposes a constant oscillation of the unique winner percentage. This makes sense, though, as we would expect to see a winner far more often when the population is odd than when it is even, as head-to-head ties become impossible. In order to further explore this idea, I divided the dataset by the population's parity and graphed the results, which can be seen below.

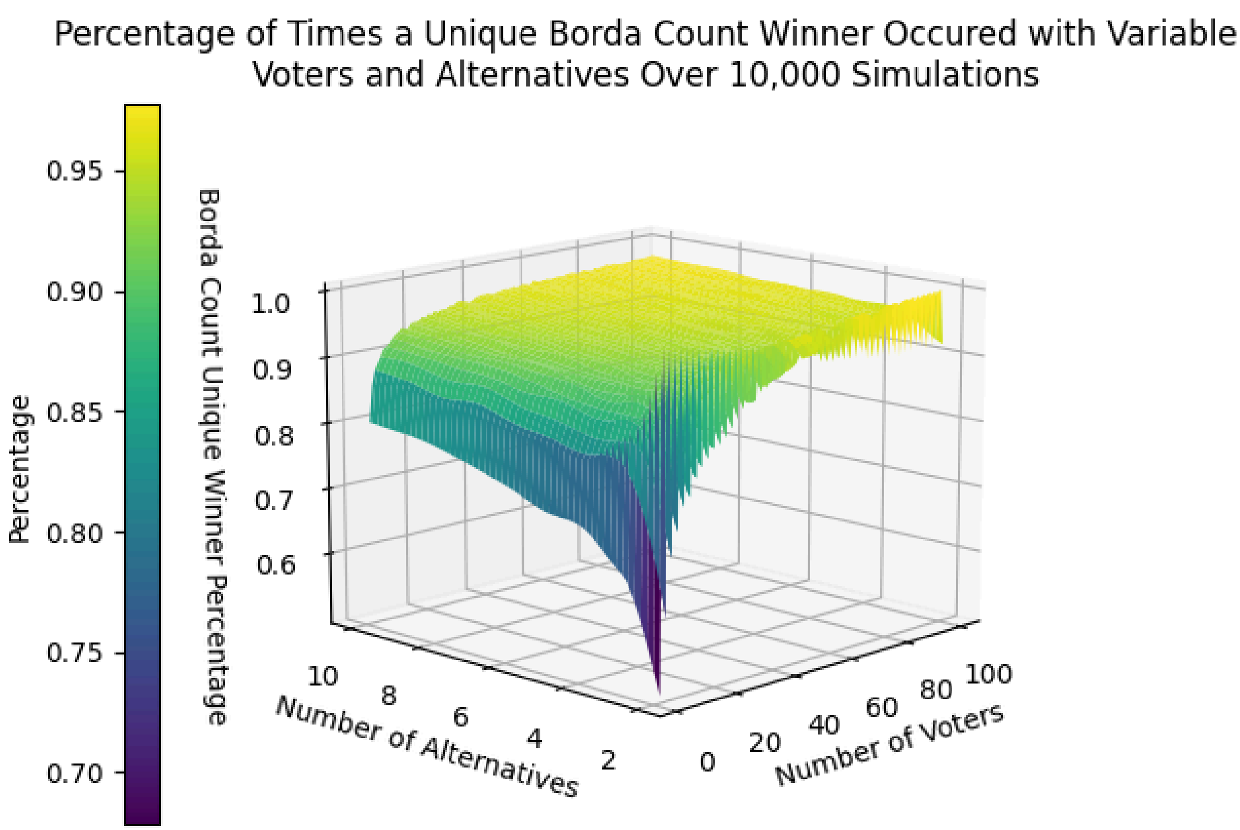


It is near-insane to imagine a single voter can have such a great impact on the success of Condorcet's method, but this graph demonstrates a definitive difference between even voting populations and odd ones. What I find most interesting, though, is that, as the number of voters grows, the even populations produce a unique winner at an increasing frequency while the odd populations produce a unique winner at a decreasing frequency. Eventually (with a population size of about 30 voters), the even populations begin producing unique winners at a higher rate than the odd populations. This is likely because, for larger population sizes, the odd populations are producing more than one winner at higher frequencies, decreasing the rate at which it elects a unique winner.

Borda Count

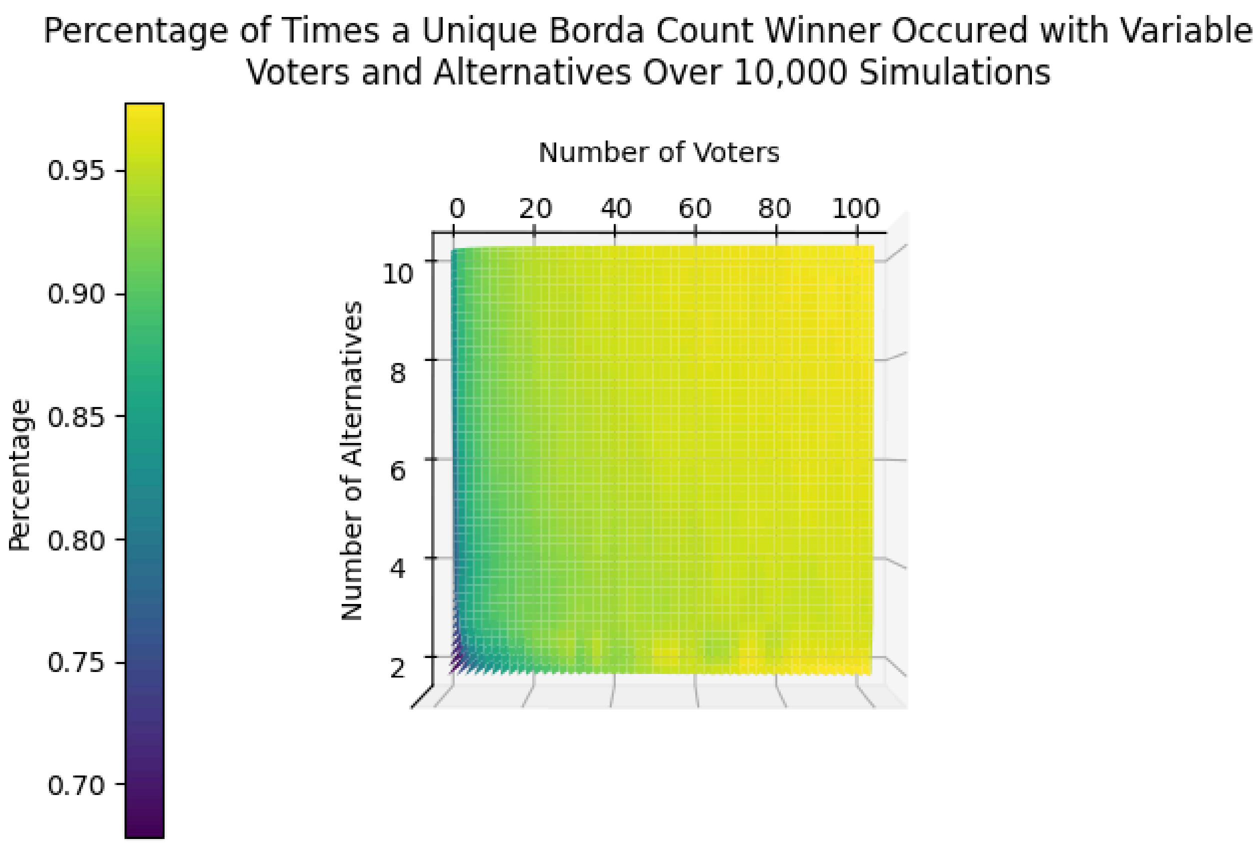
The Borda Count is a ranked-voting system that assigns points to candidates based on their ranking on each ballot. The winner is declared to be the individual with the most points. However, it is possible for two or more candidates to receive the same number of points while receiving more points than every other candidate. In this case, multiple individuals are output as the winner, in which case the Borda Count fails to identify a unique winner.

Like I did for Condorcet's Method, in order to explore the frequency at which a unique winner is elected, I ran 10,000 simulated elections for every combination of population size (ranging from 2 to 100) and number of candidates (ranging from 2 to 10). The results can be seen in the 3-dimensional graph below.



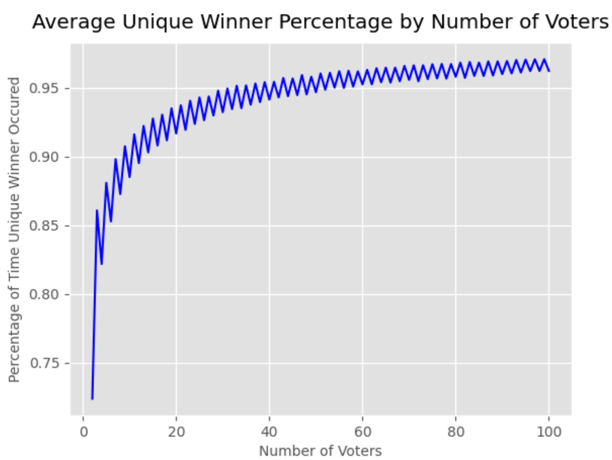
We see a very low trough occur at the minimum input values for number of candidates and number of voters. However, the unique winner percentage seems to climb quite quickly with an increase in both the number of candidates and the number of voters. What is encouraging, though, is that there appears to be a large range of inputs that produce a high unique winner percentage. This can be further explored by the top-down view of the surface plot below.

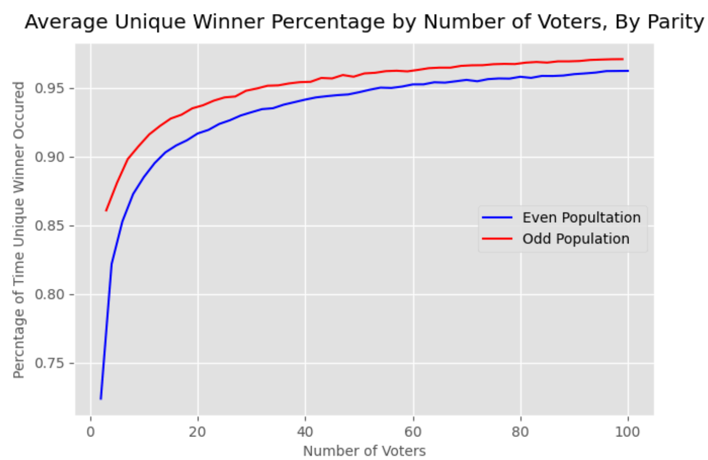
PPa



Past about 20 voters, the unique winner percentage almost never falls below 95%. For small populations, though, we see very low unique winner percentages, peaking at about 85%. When we fix the population size, there seems to be a slight increase in unique winner percentage as the number of candidates increases. Interestingly, the rate at which the winner percentage increases is smaller for larger populations, though this may be attributable to the fact that the unique winner percentage for 2 candidates is higher for larger populations. This is likely because with more candidates, it is harder for ties to occur – and because it is impossible for the Borda Count to produce no winners, ties are the only situation in which a unique winner is not chosen.

If we fix the number of candidates instead, we see an increase in the unique winner percentage as the population grows. Similarly to how increasing the number of candidates lessened the likelihood of ties occurring, increasing population size means less ties occur, which manifests in a higher unique winner percentage. This trend can be further examined in the charts below, which show the aggregate unique win percentage and the frequencies grouped by population parity.



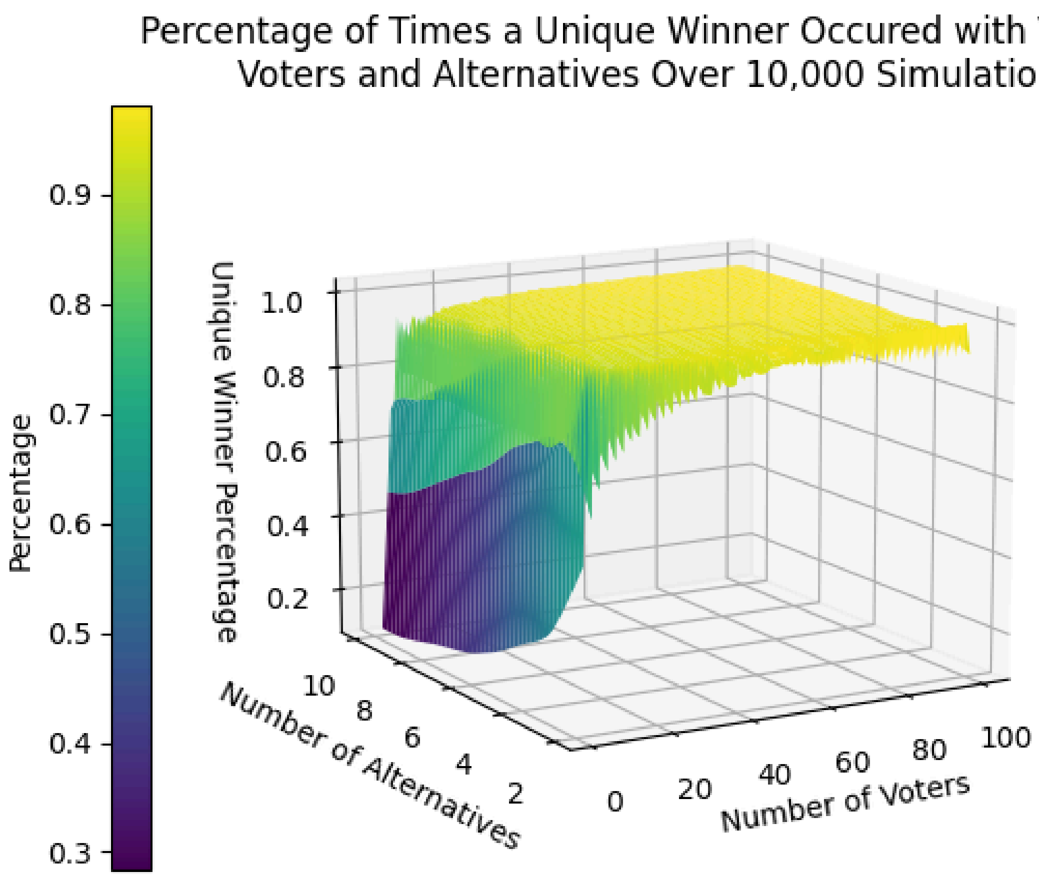


This chart confirms what we saw in the heatmap: an increase in the number of voters corresponds to an increase in the unique winner percentage. The oscillation in the aggregate graph is to be expected, as we established differences in even and odd populations and their likelihood of producing ties. Interestingly, though, the graphs of the unique winner percentage for even and odd populations follow very similar shapes, just translated slightly.

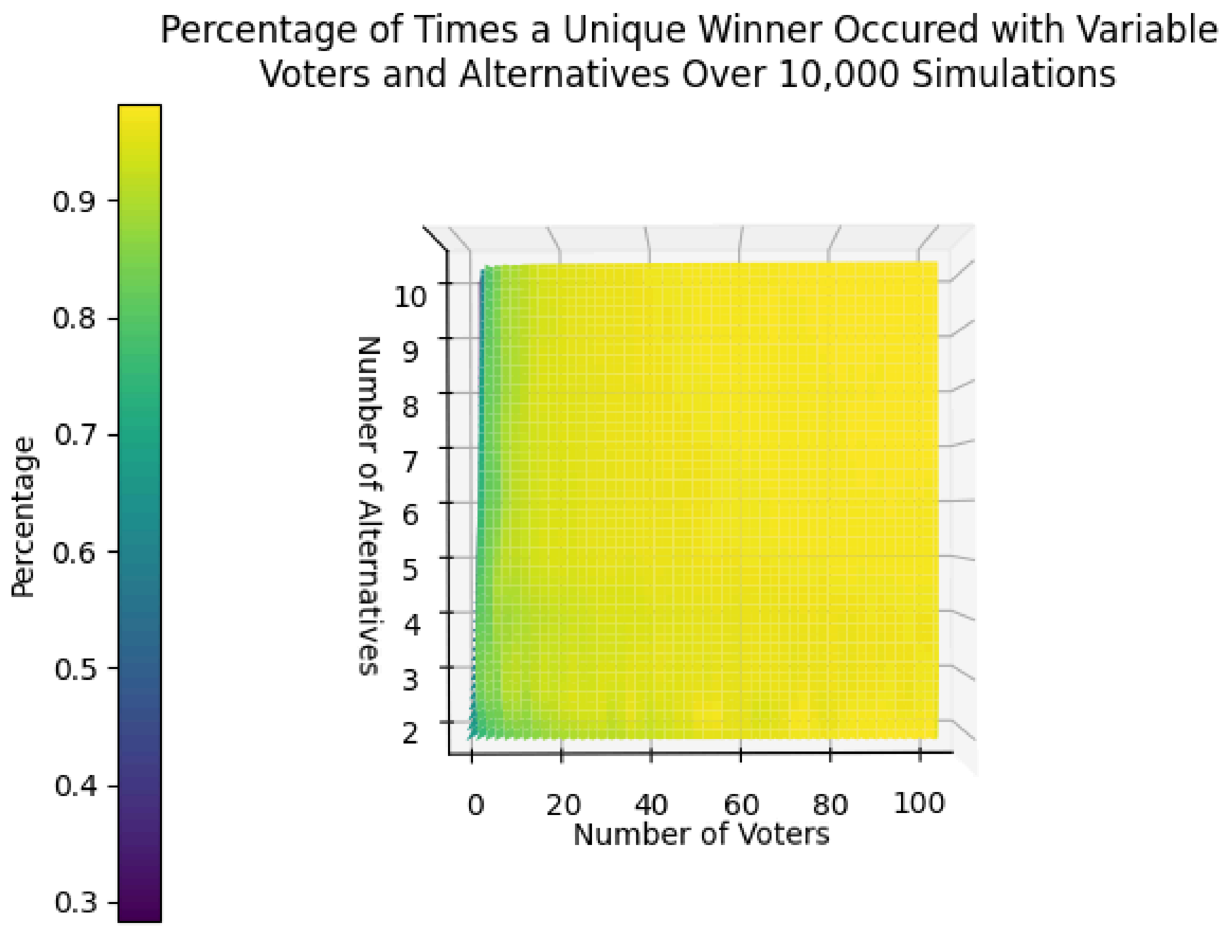
Instant Runoff

In an instant runoff election, voters rank all candidates and the candidate who receives the fewest first place votes is removed from every voter's ballots. The results are then retallied with the new field of candidates, and this process is repeated until a winner is declared. It is impossible for no winner to be chosen, but there are scenarios in which the entire field receives the same number of first place votes at any stage of the process, in which case multiple candidates are declared winners.

The relationship between the unique winner frequency, the number of candidates, and the voting population can be observed in the graph below.

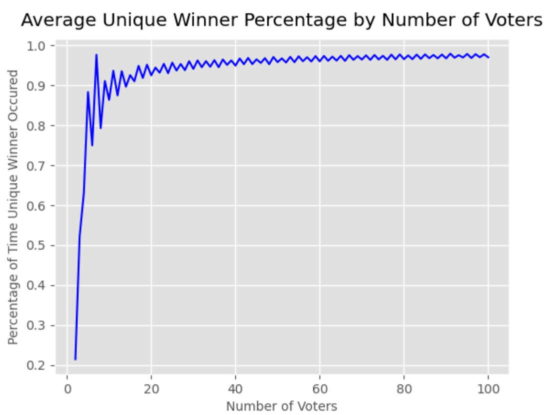


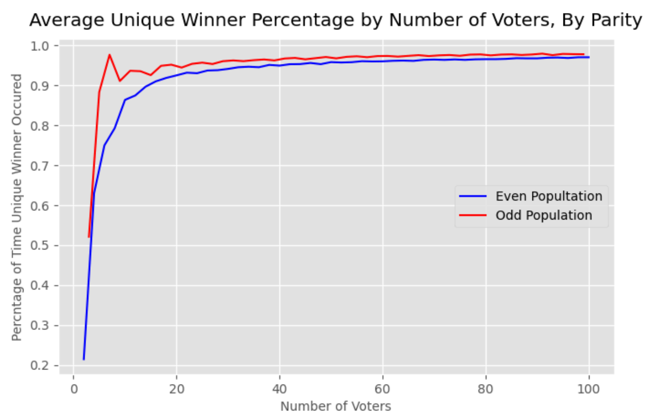
We see very lower unique winner frequencies for small population sizes, but those values quickly rise and level out as the population size increases. It appears, in this case, that the population size seems to be more influential than the number of candidates in regards to producing the highest unique winner percentage. This can be confirmed in the aerial-view of the 3-dimensional surface plot



For small, fixed populations, there appears to be a small increase in unique winner percentage associated with an increase in the number of candidates. For other fixed population sizes, though, there is little-to-no increase in unique winner percentage as the size of the candidate field grows. It makes sense that increasing the number of candidates would have little impact on the unique winner percentage as the system works by iteratively removing candidates, turning large candidate fields into small ones before determining a winner.

For a fixed number of candidates, there seems to be a small initial increase associated with the increase in population size, but this increase quickly diminishes as the unique winner percentage remains more or less stable. This is visible in the graphs below, which show the average unique winner percentage for a given population size with a fixed number of candidates.



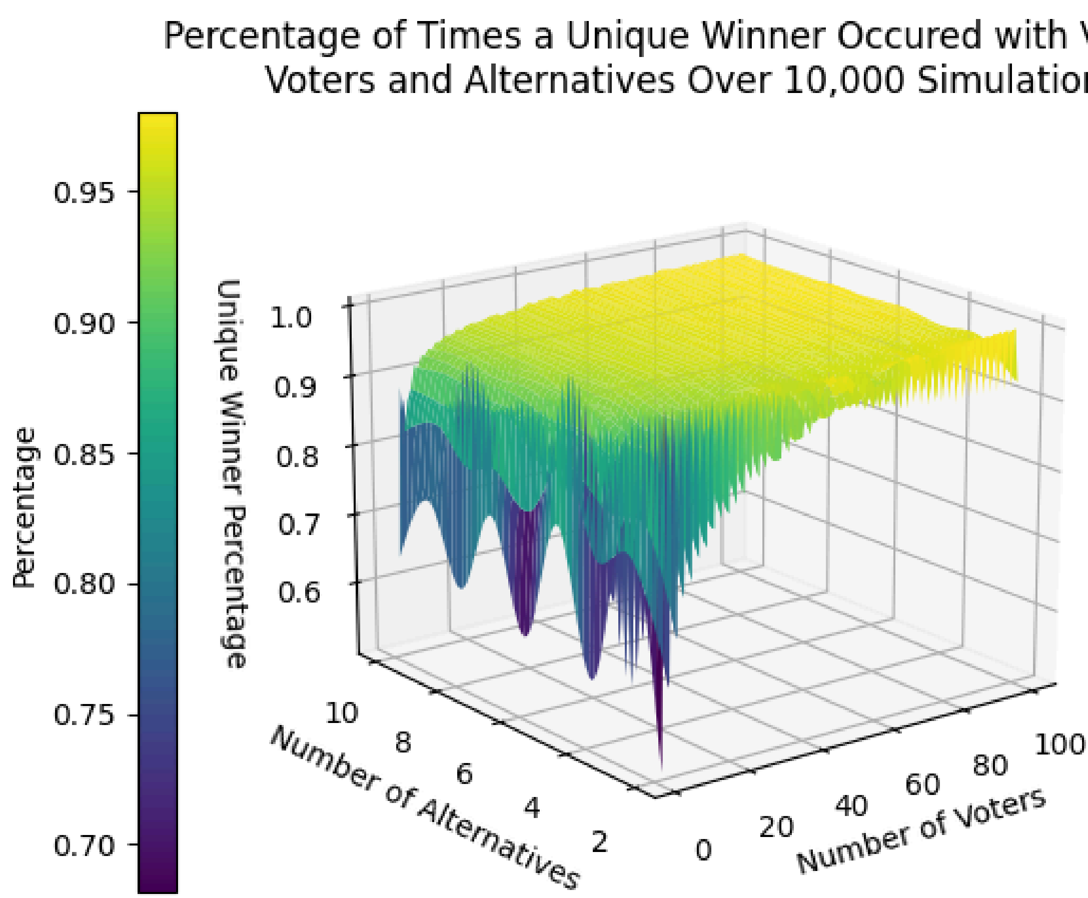


The instant runoff still exhibits some difference between even and odd populations, but the degree to which the two populations produce different results is much smaller than the previous cases. Interestingly, when the population size reaches about 20 voters, the unique winner percentage more or less levels off. This makes sense intuitively, because for small populations, it is more likely for ties to occur, especially when the candidate field is cut down to just a handful of candidates.

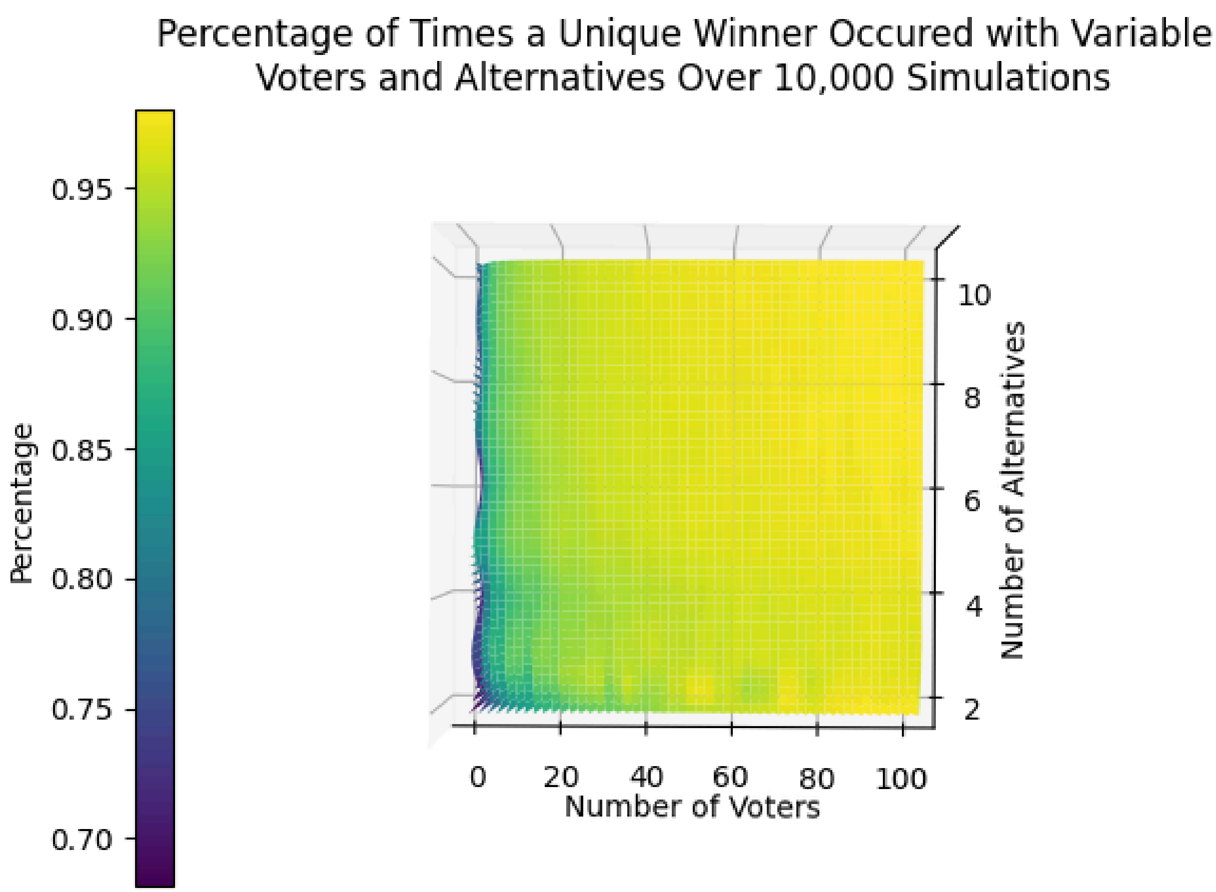
Coombs Method

Coombs' Method is nearly identical to an instant runoff, only instead of candidate(s) with the fewest first place votes being eliminated, candidates who receive the most last place votes are removed from the ballots.

The results of the simulated elections using Coombs' Method for every combination of population size (ranging from 2 to 100) and number of candidates (ranging from 2 to 10) can be seen in the graph below.

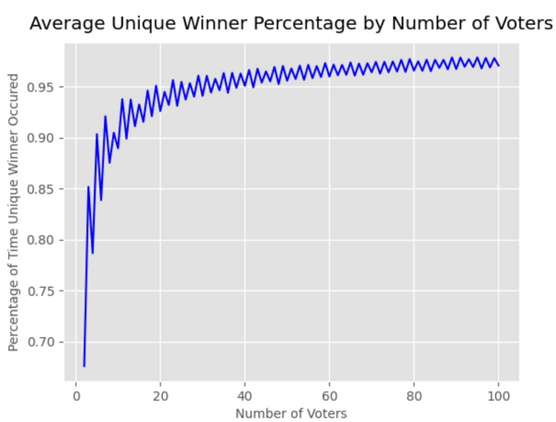


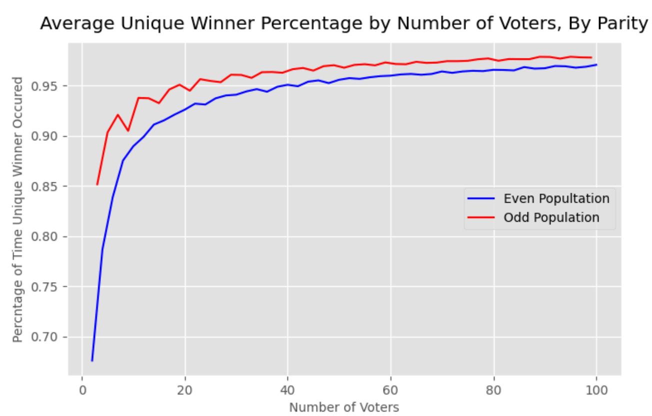
The surface plot, like the graph produced by the instant runoff, shows lower unique winner frequencies for small population sizes, though those values quickly rise and level out as the population size increases. We do, however, see a sinusoidal pattern when fixing the number of voters and varying the number of candidates. It appears that having an odd number of candidates produces a unique winner at a higher frequency than when there are an even number of candidates. This can be confirmed by getting a top-down view of the surface plot, which is shown below.



For fixed population sizes, we see a small increase in unique winner percentage as the number of alternatives increases. This makes sense as more candidates will likely produce fewer ties, but candidates are removed throughout the process so the impact of having more candidates is greatly diminished.

For a fixed number of candidates, there appears to be a general increase in the unique winner frequency associated with an increase in the number of voters. The rate at which the frequency increases, though, decreases as population size increases. This should be attributable to the fact that an increase of one voter represents a larger percentage increase for a small population than it does a big one, so the impact is more visible when the population size is smaller. This phenomenon can be seen in the graphs below, which chart the unique winner percentage across all possible numbers of candidates against the size of the population.





This graph confirms what we saw from the surface plot: as the number of voters grows, the unique winner percentage increases as well, though the rate at which it increases slows with the increase in population. Again, this decreasing rate of change is likely attributable to the fact that increasing a small population by 1 voter represents a larger percentage increase than adding 1 voter to a large population.

Comparison

Having examined multiple social choice systems, it is time to compare the results. The following table tabulates the unique winner frequency across all 10 million simulations as well as the unique winner percentage grouped by population parity.

	Total Population	Even Populations	Odd Populations
Condorcet	71.42%	70.21%	72.65%
Borda	94.16%	93.27%	95.08%
Instant Runoff	93.82%	92.17%	95.49%
Coombs’	94.92%	93.83%	96.04%

As we would expect, the unique winner percentage is higher for odd populations than it is for even populations, with the values differing by 2-3 percentage points. In practice, though, there is no accurate way of knowing whether a voting population is even or odd, and the impact of the parity of the population is significantly smaller for larger populations. So, on the scale of the United States electorate, the even and odd distinction is less relevant than the aggregate results, which would suggest Coomb’s method is the most likely to produce one clear winner.

Conclusion

This report is not, by any means, a fleshed out document meant to be taken as scripture. My only goal was to introduce an idea of which systems may be considered “practical”. At no point do I even begin to consider other properties that may be desirable in an electoral system, like promoting third-party participation or encouraging people to vote. Deciding which social choice system is “best” is a large-scale debate that may never be resolved. The pursuit of an ideal social choice system remains a long, windy road while the conclusions drawn in this report represent just a few small steps down the path.