Estimating the Probability of a Voting Cycle

Chang Geun Song

(Last update, October 9, 2021)

Abstract

Theoretically, Condorcet paradox exists and occurs, but much less in practice than it predicted. This paper targets estimating the probability of the paradox, being closer to what the data tells. Survey data is the best alternative to actual election data under the ranked voting system. We use German Politbarometer, which offers two benefits for empirical analysis of voting systems; containing scoring data and the enormous observation size. We suggest different statistics and models based on cardinality. Specifically, considering a 'median' of collected evaluations as a significant factor in predicting the winner of one-to-one comparison, approximating the probability of a cycle from the probability of two sets of three events occur. The model predicts a significantly lower (but positive) voting cycle frequency than popular models, a probabilistic model with IC and IAC assumptions. Our approach involves 1) assigning three candidates presumed positions of first, second and third 2) noting the gaps between pairs of candidates in apparent estimated merit, and then 3) computing the probability that the three pairwise comparisons will have a combination of outcomes that results in a cycle.

1. Introduction

The decision-making process of a group is not like that of the individual. While considering that people have different views and preferences, one needs to pick the best choice in a given situation. A natural and acceptable way to aggregate the opinions of the members of the group is voting. The most common electoral system, plurality rule, is used worldwide and continuously. Easy to understand and to do are advantages of plurality rule, but it also has flaws. Thus, researchers have suggested many alternatives to better aggregate individuals' preferences and reach better outcomes for the groups that use them. Almost all of them are turned out to be demonstrably better than the plurality rule in many respects.

Nevertheless, no voting rule is "ideal." In particular, no voting rule is free from Condorcet paradox, the case that the aggregated preference is cyclic or transitivity is not valid: Suppose there is a group and three available options, A, B, and C. The group prefers A to B, B to C, and C to A. If the group decides to use pairwise comparisons sequentially to find the best option for the group, the winner depends on the order of the comparisons, and it allows the possibility for the agenda setter to manipulate the result. The more severe problem is, for any option wins under ANY voting rule, there will always be another option with a majority over the chosen if a voting cycle exists. Thus, voting provides necessarily imperfect estimates of the socially best option, and the impact of Condorcet paradox has been discussed as a serious potential peril to the coherence of democratic decision-making. More terrifying is that some theories and probabilistic models predicted the frequent occurrence of the paradox.

Researchers also tried to find empirical evidence. However, it was hard to detect Condorcet paradox in practice as often as previous studies insisted. Rather, the voting cycle is very rarely happening. Nevertheless, one might disagree with the gap between the theories and the real world because the number of empirical researches is limited. According to **Van Deemen (2014) [16]**, only 47 empirical studies related to it, involving 265 elections in total, from 1955 to 2010. Moreover, almost all of these studies dealt with a few elections.

The main reason for a limited number of studied elections is that available data are fairly scarce. There have been countless elections in human history, and, for many of them, the data exists. However, most elections are held under the plurality rule, which can only reveal how many (presumably, first-choice) votes each candidate had. This information is insufficient for constructing each voter's specified preference on candidates and so aggregated preference. Moreover, this prevents determining the paradoxes or opportunities for strategic voting.

In the present paper, we use a part of German Politbarometer, survey data. The data include respondents' scoring data of politicians. Using the data, we construct individuals' reported preference orderings and check the frequency of voting cycles in synthetic elections. One might be concerned that people who participate in a poll or survey may have significantly different preferences than people who do not. So that an analysis based on the survey data will not reflect the whole population, and it is not a good way to approximate the preference orderings of voters. However, it is impossible to know voters' rankings unless voters report their rankings of candidates in elections, so survey data is generally the best available actual data except for the actual election data.

Additionally, the lack of data from actual elections makes it interesting to use survey data. A huge benefit of using Politbarometer data is its vast size of observations: We analyze 1,022 synthetic elections. Another advantage of using poll or survey data is that respondents have no incentive to report false preferences intentionally, and so less distortion in aggregated preference. Since it is merely answering a survey and nothing changes with regards to one's welfare. However, that also causes some meaningless or careless responses, and we screen them out.

Our goal is to report the empirical result and then suggest a distinct way to explain Condorcet paradox and the departure of the result from the predictions of the existing models. The modeling process is based on actual data, not like the probabilistic model with cultural assumptions. What makes our work different from the others the most is that we directly connect the cardinal information to the probability of winning in the pairwise election and establish a model based on that to estimate the frequency of the voting cycle. Specifically, we introduce original statistics for discrete data, fraction-based median, which contains information of the standard median and the distributions of observations. Then, we check the predictability of the gap in the fraction-based median of scores of two candidates for winning in pairwise comparisons. Finally, we consider a cycle as a consecutive three head-to-head election and calculate the probability by multiplications of three estimated probabilities.

Contents

The remaining parts of this paper are as follows: in section 2, a literature review, we discuss some well-known probabilistic models and describe the departure of predictions of these models for the occurrence of Condorcet paradox in empirical work. Section 3 describes the data that are used, the methods of analysis, and the results. Finally, section 4 summarizes works and proposes future research.

2. Literature Review

Condorcet paradox was initially discussed in the late 18th century by Condorcet. Since then, there have been fewer results from empirical work than from theoretical approaches. Researchers have also analyzed elections generated from a hypothetical, probabilistic model because of the scarcity of large data sets of the relevant sort.

Firstly, we will outline widely using probabilistic models and then introduce empirical works on Condorcet paradox.

Probabilistic model

Individuals' preference orderings determine the aggregate preference ordering through the application of a voting rule, and the choice of the voting rule affects who will win.

When we have n alternatives, we have $n!$ possible strict orderings. Furthermore, the frequency distribution over the strict orderings would be expressed by a vector with $n!$ components, such that the sum of components is 1.

Impartial Culture (IC) model

Guilbaud (1952) [8] proposed the impartial culture (IC) model, which, for every voter, assigns equal probability to every possible preference ordering. For example, suppose there are six (strict) rankings ABC, ACB, BAC, BCA, CBA, and CAB for three options A, B and C. Hence, the IC model assumes that every individual would have each preference ranking with 1/6 probability.

Impartial Anonymous Culture (IAC) model

The impartial anonymous culture (IAC) model was developed by **Kuga and Nagatani (1974) [9]** and **Gehrlein and Fishburn (1976)** **[3]**.

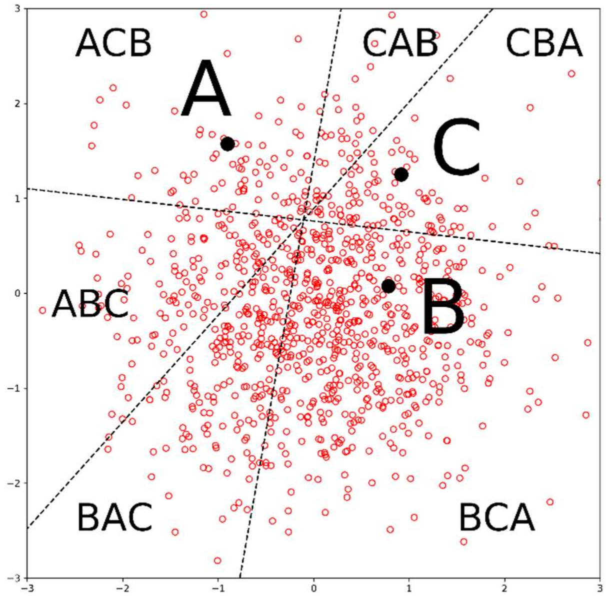
For the case of two alternatives A and B with three voters, there are 2 possible preference orderings, AB and BA, and $2^3 = 8$ possible preference profiles. Then, we can categorize the profiles into four "classes" by their composition without considering the order of voters.

AEC1: AB x3 - (AB, AB, AB)
AEC2: AB x2, BA x1 – (AB, AB, BA), (AB, BA, AB), (BA, AB, AB)
AEC3: AB x1, BA x2 – (AB, BA, BA), (BA, AB, BA), (BA, BA, AB)
AEC4: BA x3 - (BA, BA, BA)

The class is called *anonymous equivalence class* (AEC). The IAC model assumes that each class of preference profiles has an equal probability of realization, 1/4 for the given example.

Normal spatial model

The normal spatial model was developed by **Good and Tideman (1976) [7].** This model assumes an attribute space, and an individual's preference ordering is determined by the distances from the options to his or her most desired option, with closer options ranked higher. The Normal spatial model assumes that the voters' most desired options are drawn from a (multivariate) normal distribution. The hyperplanes that bisect line segments connecting all pairs of candidates divide the space into sectors corresponding to different strict preference orderings. The integration of the voters' probability density over all such sectors specifies the probabilities. For three candidates, one needs at least two dimensions to represent all six possible, strict rankings.



<figure 1, normal spatial model, 3 candidates, 2-dimensional>

The probabilities implied by the models

The normal spatial model implies that as the number of voters approaches infinity, the probability of a voting cycle approaches 0, since a candidate closer to the mode of the voters' locations will always have more votes than a candidate who is farther from the mode. **Gehrlein (2002, 2006) [4][5]** provides calculations for the probabilities of the existence of the Condorcet winner under different assumptions on culture. In the three-candidate case, it is equivalent to the probability of non-existence of the Condorcet paradox. When the number of voters goes large enough, the IC model and the IAC model yield probabilities of Condorcet paradox decreases and converges to 8.77% and 6.25%, respectively. On the other hand, the estimation of the frequency of voting cycle by normal spatial model, vary depends on the data, is much lower than that of IC and IAC model and it converges to 0 as the number of voter goes infinity.

Empirical works

We will briefly summarize the framework and results of empirical works that provide the frequency of Condorcet paradox. We purposely omit relevant ones published before 2010 (cf. **Gehrlein and Lepelley (2011: 13)** [6]and **Van Deemen (2014, Sect. 5) [16]**).

Voting cycles in simulations based on real data

Popov et al. (2014) [13] used five-candidate American Psychological Association (APA) presidential election data to compare and evaluate various voting systems, including the Plurality rule, Instant Runoff Voting, or Vorda Rule. The data include voters' full or partial ranking of candidates. With 12 elections, they generated three thousand bootstrapped ballot profiles for each. The frequency of Condorcet paradox was infrequent, lower than 0.3% of total samples.

Darmann el al. (2019) [1] used online survey data collected two weeks before the 2015 parliamentary elections in the Austrian federal state of Styria. Questions asked respondents to assign points ranges in [-20,20]. The authors reported 13 voting cycles out of 1000 bootstrapped profiles but 0 voting cycles of three candidates who win the most votes under plurality rule.

Another notable empirical study is **Tideman and Plassmann (2012) [15].** They used actual election data from Electoral Reform Society, and survey data from the American National Election Studies. They showed that the spatial model fits the data better than other probabilistic models, including IC and IAC. Authors estimated the share of preference orderings by spatial model and simulated 1,000,000 elections for each different number of voters. For 1,000 voters, the voting cycle existed 0.53% of simulated elections, and it converged to 0% as the number of voters increases.

Voting cycles in real data

Mattei (2011) [11] measured the Condorcet efficiency (the probability of selecting a Condorcet winner) of several voting rules and calculates how often a voting paradox happens in three alternative and four-alternative elections that are constructed from a vast data set consisting of individual ratings of 17,770 movies. The ratings came from 480,189 distinct users of Netflix, who were asked to assign stars from 1-5 (with half stars permitted) for how well they liked the movies they watched[1,2]. The author made three subsets by randomly drawing 2,000 movies from the 17,770 movies. He then generated all possible three-alternative elections and four-alternative elections in each subset. A derived election was included in the analysis only if it had more than 350 voters. The share of three-alternative elections that exhibited Condorcet paradox in the three independent subsets were 0.041%, 0.044%, and 0.056% in each subset. In **Mattei et al. (2012) [12]**, they extended the work. They made ten subsets with the same data set with non-overwrapping 1,777 movies for each and checked the voting cycle for all possible triples and quadruples in each subset. The frequency of the cycles in the top three candidates in four alternative elections was lower than 0.11% of the total.

Kurrild-Klitgaard (2014) [10] provided possible evidence of voting cycle in practice based on the result of a Monmouth University Poll in March and April 2015, which asked head-to-head comparison of four potential candidates (Jeb Bush, Chris Christie, Ted Cruz, and Scott Walker) for 2016 Republican party presidential primaries. The result showed a top voting cycle of

---

1 One can find detailed information at https://www.netflixprize.com/.

2 The rating by stars has been abolished and replaced by ratings of “like” or “dislike.”

three candidates. Feizi et al. (2020) [2] also analyzed several polls before the actual election. The polls began some weeks before the 2017 Iranian presidential election. The polls included the pairwise comparison of candidates. The authors concluded that there is no voting cycle of three leading candidates in every poll.

3. Analysis

Description of the data

We use German Politbarometer data, which are freely available to the public at the Politbarometer website. These data come from surveys of German voters, conducted approximately monthly, beginning in 1977 and ending in December 2019. The surveys contain many questions about demographic characteristics, region of residence, and weekly working hours.

We focused on questions related to a *skalometer* (in English, consists of “scale” and “meter”) of politicians. Each participant is asked to independently assign to each politician one of the 11 integers from -5 to 5; higher integers mean the participant thinks higher of the politician. Thus, we have the cardinality as well, which is uncommon, and this provides the preference order on the politicians and degree of likeness or dislikeness.

We treat the data as separate surveys by regions where participants are living (the former West or East Germany) and by months. We also divide the data when the set of politicians for the questions varies even in the same region and the same month. After cleaning the data, we have 1,022 usable surveys.

In each survey, at least four and at most 21 politicians are evaluated (10.62 on average), and the number of participants in a survey ranges from 179 to 3,345 (930.05 on average). The politicians in the questionnaire consist of influential politicians at the time of the survey. There is no regular rule for forming the pool; some politicians continue to the following survey, and others do not. Some excluded politicians are re-included in later surveys' pools. In some surveys, former politicians or famous politicians from countries other than Germany are also on the list[3].

For convenience, we confine our analysis to triples of candidates, the minimum number of alternatives needed for Condorcet paradox. We construct triples from each survey for all possible combinations of three politicians in the survey. To be specific, we can make four triples from four politicians, ten triples from five politicians, and $n(n-1)(n-2)/6$ triples from n politicians.

For every triple, we checked for the existence of a cycle by paired comparisons. We exclude participants who did not provide a score for both politicians. As a result, there are 206 majority cycles (including semi-cycle), out of 181,579 triples, a rate of 0.113%, or about 1 out of 881, showing a significant departure from the predictions of statistical models.

Framework for analysis

1. Decompose a voting cycle into three probabilistic events

To check whether there is a voting cycle, we need to know the majorities in three paired comparisons. Define A**p**B as the strictly prefer A to B, n[A**p**B] as the number of the voter who prefer A to B., and d_{AB} as the difference between n[A**p**B] and n[B**p**A].

---

[3] Ronald Reagan, Margaret Thatcher, George W. Bush, Michail Gorbachev, Tony Blair Jacques Chirac, and Vladimir Putin are on the list on some surveys.

For arbitrarily labeled alternatives **A**, **B**, and **C**, as table 1 shows, there are eight possible combinations of the sign of d_{ij} for $i \neq j \in \{\mathbf{A}, \mathbf{B}, \mathbf{C}\}$, and they correspond to the six possible strict preference orderings and two cyclical orderings of the three options.

	d_{AB}	d_{BC}	d_{AC}	preference ordering
sign of difference in majority	+	+	+	ABC
	+	+	-	ABCA (cycle)
	+	-	+	ACB
	+	-	-	CAB
	-	+	+	BAC
	-	+	-	BCA
	-	-	+	CBAC (cycle)
	-	-	-	CBA

<table 1, eight possible profiles of signs of d_{ij} and implied preference orderings>

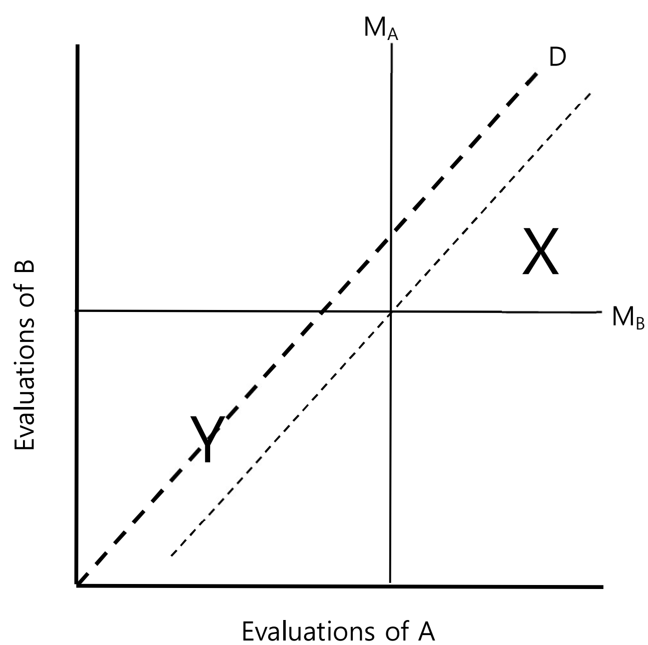
If we can predict the sign of d_{AB}, d_{BC}, and d_{AC}, then we can also predict whether there exists the voting cycle or not. Assuming the voters cast their ballots based on their evaluation, then comparing the distribution of scores will be a great predictor for the sign of majorities and existence of the voting cycle. Thus, we pay attention to compare peoples' evaluation, specifically the median evaluation of each alternative.

Why median matters?

Let the voters' cardinal evaluation of two candidates are distributed on two-dimensional space as **<figure 2>**. Voter's evaluation of A for x-axis and that of B for y-axis. M_A and M_B are the median evaluation of candidates **A** and **B**, respectively. If voters cast ballots based on their evaluation, **B** will be selected by the majority rule if and only if more than a half of voters are located above line *D*, which is a 45-degree line. Due to the half of voters are on the right of M_A, and half voters are below M_B, however, winning more votes is likely to the candidate **A**.

Another way to think about the possibility is that rotating the horizontal line M_B 45 degrees, counterclockwise and look the upper-left partition of the rotated line. For a majority in favor of candidate **B**, such a rotation, ruling out voters in area X from above the line M_B and adding voters in area Y from below the line, must include so many voters above the line. In either way, it would be more unlikely when the intersection of M_A and M_B moves farther below line *D*.

<figure 2, Median and Majority>



We check each candidate's median evaluation and which one would win more votes in all 55,321 pairs of candidates from 1,022 surveys. Among 41,512 pairs with a gap between two candidates' median of evaluation exists, a candidate with a lower median beats the other in 213 cases. It happens only when the gap is 0.5 and 1, and the portion is lesser when the gap is 1 compared to 0.5.

median gap	pairs	beat by lower candidate
0	13809	0
0.5	297	22
1	22195	191
1.5	203	0
2	11864	0
2.5	71	0
3	4639	0
3.5	28	0
4	1576	0
4.5	8	0
5	493	0
5.5	1	0
6	93	0
6.5	2	0
7	31	0
7.5	0	0
8	7	0
8.5	0	0
9	4	0
9.5	0	0
10	0	0
total	55321	213

<table2, the number of pairs and beat by lower with respect to median gap>

Comparing the median of evaluation seems a fair way to predict who wins in paired comparison. A substantial problem with using the median and the gap in our analysis is that the survey allows only 11 integers for scoring, so many ties in pairwise comparisons. Indeed, we have pairs that have the same median evaluation between two candidates in about 1/4 of the total cases. We need to identify a better measure of candidate quality that can be considered as a reasonable explanatory factor for the difference between votes that two candidates receive as the median and be easy to observe and calculate.

The Fraction-Based Median (FBM)

We construct a new measure that has the character of a median but generally does not result in ties, despite only a few discrete values that the underlying variable can take. The "fraction-based median (FBM)" takes non-integer value based on the median and the placement of the median in the set of median-valued observations.

It is derived in the following way.

-Let **M** be the integer median of the numbers in the set,

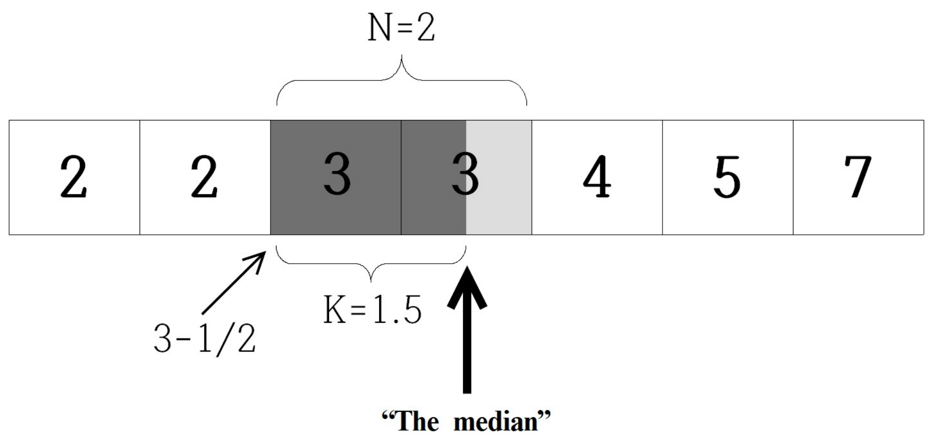
-Let **N** be the number of observations that take the standard median value,

-Let **K** be the number of observations to the left of "the median" that take the standard median value.

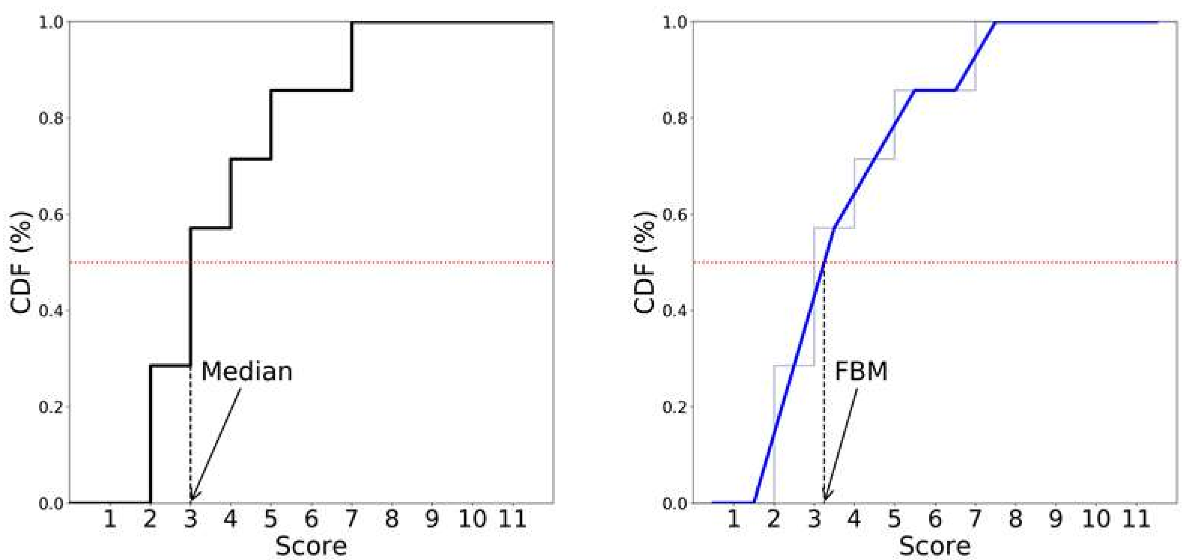
Then FBM of the set is M – 1/2 + K/N. If the median occurs precisely in the middle of the observations that have the median value, then the FBM is M. By its definition, FBM can only have the value from M-1/2 to M+1/2, and even two sets have the same standard median, it will generally be true that one will have a higher FBM than the other. Note that if the standard median is not an integer, then define FBM as the median.

As an example of a calculation of an FBM, the set {2,2,3,3,4,5,7} has an FBM of 3 – 0.5 + 0.75 = 3.25 **(figure 3a).** The value 3.25 indicates that the median is 3 and, being greater than the median, indicates that the number of observations in the given set that take a value higher than 3 is greater than the number that takes a value less than 3.

Graphically, FBM is described as follows.



<figure 3a, Idea of Fraction-Based Median>



<figure 3b, Median and Fraction-Based Median by CDF>

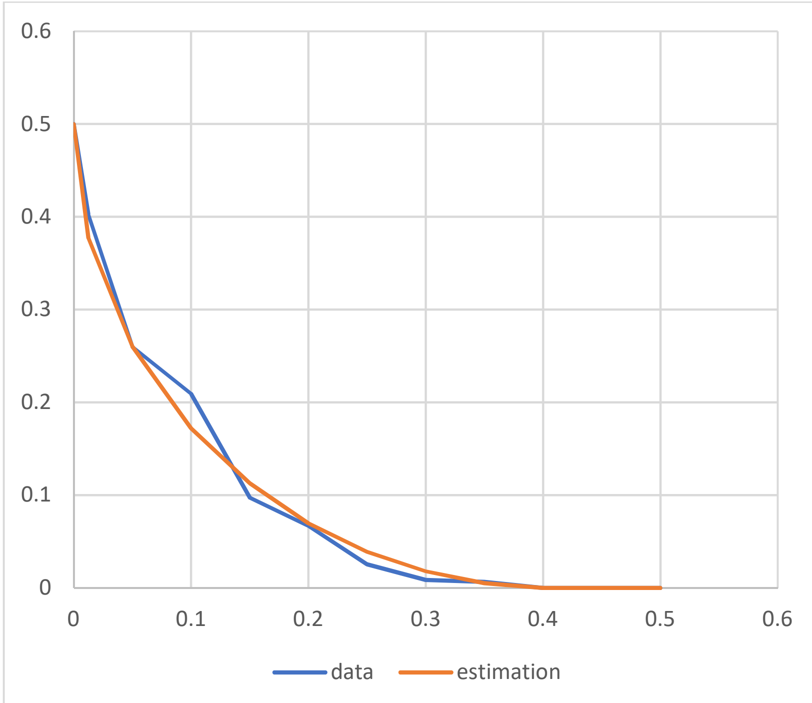
The CDF of a set of discrete random variables is a step function. The standard median is the score of the intersects of a horizontal line at the height 0.5 and the CDF. Suppose a series of line segments, passing through the midpoints of all of the horizontal segments of the CDF **(the second graph on figure 3b)**. The FBM will be at the fractional score where that series of line segments intersects the horizontal line at the height of 0.5.

Upset rate

Now we restrict our interest to the top triple of each survey, the three politicians with the highest FBMs. The main reason for this restriction is that the frequency of voting cycles being created by a fourth or other politician and resulting in no Condorcet winner can be expected to be much lower than the frequency for cycles among the top three.

We assign labels **1**, **2**, and **3** according to FBMs. That is, **1** is the politician who has the highest FBM among all candidates in a survey, **2** is the second highest and **3** is the third highest.

If scores assigned by participants are not correlated, then in any paired comparison, the candidate with the higher FBM score is more likely to win the paired comparison. While the candidate with the lower FBM might beat the candidate with the higher FBM, we will call such an event an "***upset***" and the usual result a "***straight***." We checked how often upsets happen in all possible pairwise comparisons to the difference in fraction-based medians. We have 3,066 pairs from the top three candidates of 1,022 surveys. We fit the data for the upset rate, U, by the arc of an ellipse. The data shows a clear downward trend for the frequency of upsets as a function of the FBM difference between two candidates, as shown in the **<figure 4>**.



<figure 4, upset rate>

$$\mathrm{U(D)} = \frac{1}{2}\left[1 - \sqrt{1 - \left(\frac{D - 4081}{0.4081}\right)^2}\right]$$

Where D is the difference in FBMs, and the number 0.4081 matches the integral under the ellipse to the integral of bars from data.

For a given FBM difference, D, U(.) generates the probability estimation of an upset, and 1-U(.) estimates the probability of a straight.

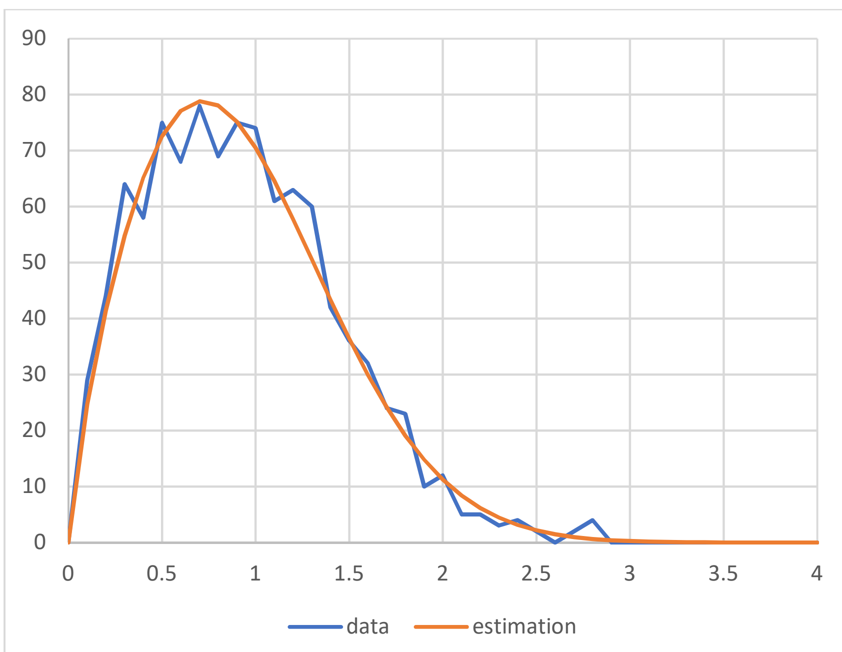
Frequency of d_12, d_23, d_13

Besides the function for estimating upset rate by given D, the other should be concerned is that the distribution of D. The empirical frequency distribution of FBM differences between candidates 1 and 3 is shown in **<figure 5>**. By its' shape, we fit a generalized gamma distribution (three-parameter gamma distribution) in the data. Specifically, for the distribution function to emerge from the origin with a finite, positive slope and no curvature, as seems to be true of the data, the second parameter of the generalized gamma distribution, d, is around 2.

The probability density function of generalized gamma distribution is as follows.

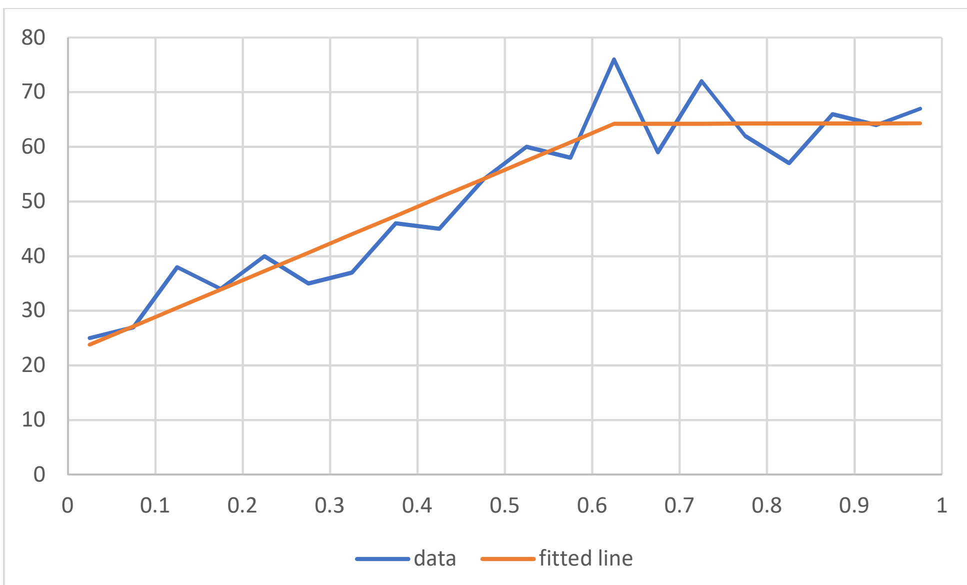
$$\mathrm{f}(\mathrm{x}: \mathrm{a}, \mathrm{d}, \mathrm{p}) = \frac{p/a^{d}}{\Gamma(\mathrm{d}/\mathrm{p})} x^{d-1} e^{-(x/a)^{p}}$$

The likelihood maximized when (a,d,p) = (1.1531, 1.7766, 2.0646), <figure 5> shows that the estimated with the real frequency of the data.



<figure 5. real distribution of D and estimation with three-parameter gamma distribution when (a,d,p) = (1.1531, 1.7766, 2.0646), >

After that, we also need to care about the FBM difference between candidates 1 and 2, and 2 and 3. Those FBM differences depend on the FBM difference between candidates 1 and 3. Hence, we do not assume each FBM difference is drawn independently; we assume that FBM difference between candidates 1 and 3 is drawn and the proportions in which this division is drawn instead. Thus, if we define the share of FBM difference between 1 and 2 in the FBM difference between 1 and 3 as $z = \frac{FBM_1 - FBM_2}{FBM_1 - FB\ _3}$, we can examine the empirical distribution of this share. **<figure 6>** shows that it seems to have a relatively steep upward trend until a certain point, and then the degree gets mild. For simplicity, we fit two line segments. The empirical distribution shows that the FBM difference between 1 and 2 is more likely to be greater than 2 and 3.



<figure 6, distribution of the ratio z>

$$g(z) = \begin{cases} 64.206 + 67.327(z - 0.625) & if\ z \leq 0.625 \\ 64.206 + 0.320(z - 0.625) & if\ z > 0.625 \end{cases}$$ **<equation 1>**

The joint distribution of the FBM difference between 1 and 2, 2 and 3, and 1 and 3 can be represented by a draw of D from a generalized gamma distribution with parameters a = 1.1531, d = 1.7766, p = 2.0646. Followed by a draw from the normalized version of the distribution specified by <equation 1>, divide the difference between the FBMs of 1 and 3 into a difference between the FBMs of 1 and 2 and between those of 2 and 3.

The probability estimation is an integration of the probability of two "types" of cycle given 1-3 FBM difference and proportions of 1-2 FBM difference and 2-3 FBM difference. A voting cycle requires either (straight, straight, upset) or (upset, upset, straight) in pairwise comparisons. The expectation of the probability of a cycle can be calculated as

$$P = \int_0^{10} f(D) \left(\int_0^1 g(z) \, [[1 - U(zD)][1 - U((1-z)D)]U(D) + U(zD)U((1-z)D)[1 - U(D)]] \mathrm{dz} \right) \mathrm{dD}$$

<equation 2>

The estimation is 0.0088 or 0.88%. The estimation is a bit off from what data shows; in 1,022 triples of top candidates of each survey, there is 0 cycle.

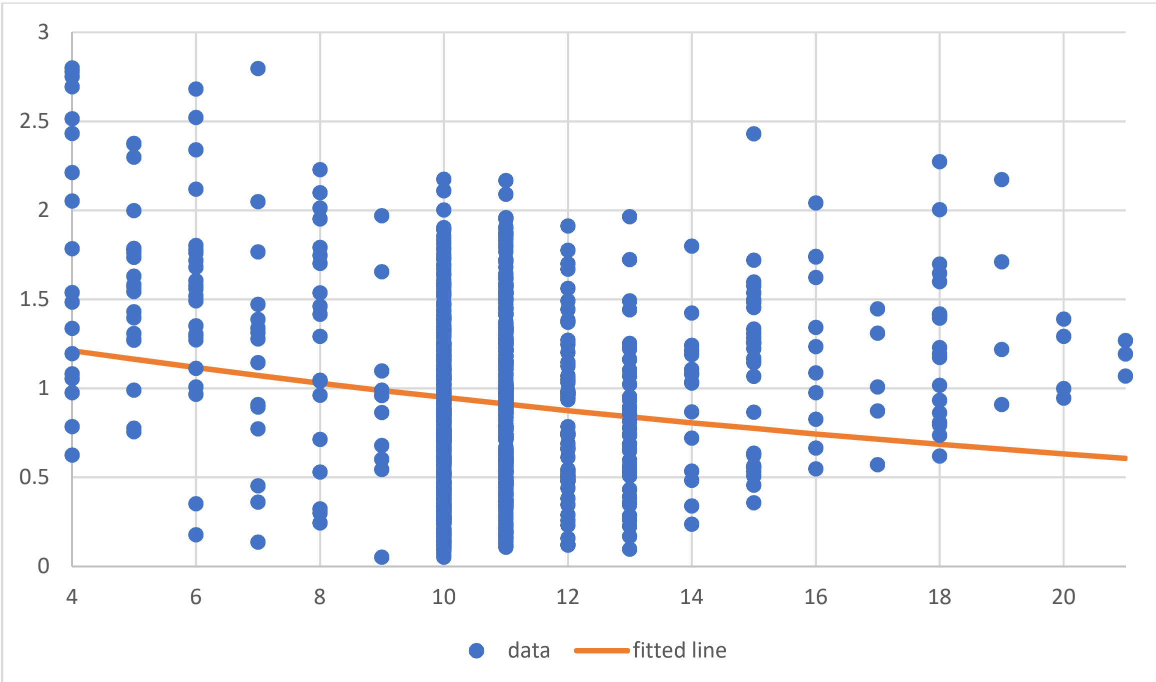
An extended model

For our research, we choose only the top three candidates among the candidate pool in each survey, and the number of politicians in the pool varies in each survey from 4 to 21. It is reasonable to expect that the average values of differences in FBMs within the top triple will shrink when the top triple is chosen from a survey that has more politicians. Intuitively, when all politicians have FBMs in the fixed interval [-5,5], more politicians can be expected, on average, to lead to a reduction in every gap between a candidate to a candidate, and it will be 0 in the limit.

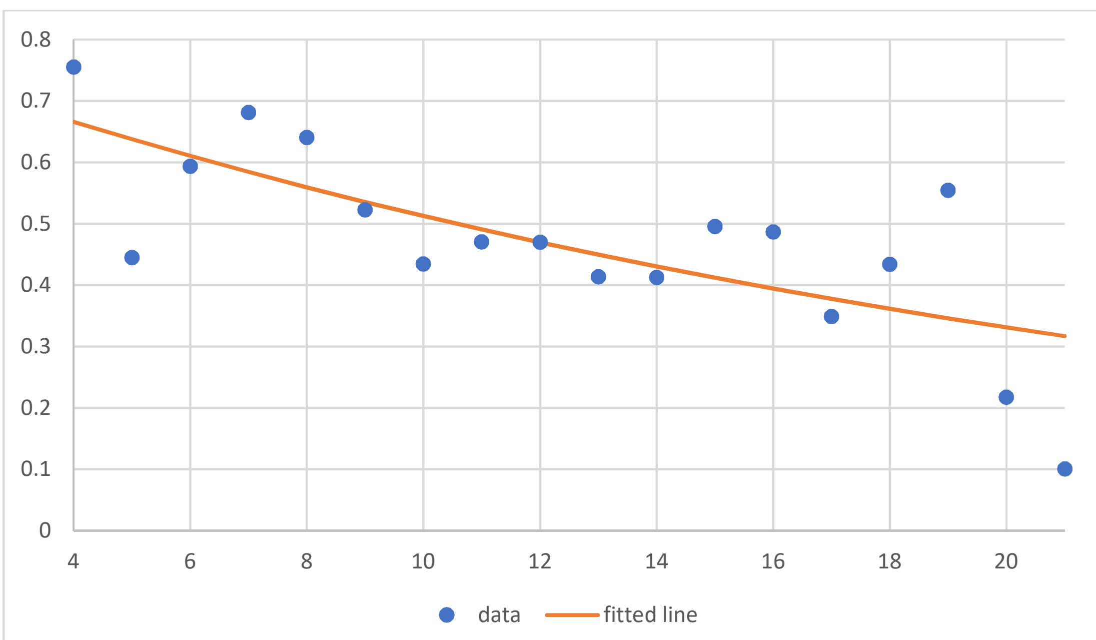
The data display a weak negative correlation between the number of candidates in a survey and the FBM difference between candidates 1 and 3, or D. The standard deviation somewhat gets smaller (**<table 3>**, **<figure 7>**, **<figure 8>**).

Number of candidates	Number of instances	Mean difference in 1-3 FBM	Standard deviation	Coefficient of variation	Standard error of estimate
4	18	1.7825	0.7558	0.4240	0.1781
5	22	1.5989	0.4449	0.2782	0.0949
6	23	1.5207	0.5936	0.3904	0.1238
7	15	1.2041	0.6812	0.5658	0.1759
8	19	1.2831	0.6405	0.4992	0.1469
9	11	0.9426	0.5228	0.5546	0.1576
10	483	0.7876	0.4345	0.5516	0.0198
11	250	0.9515	0.4705	0.4944	0.0298
12	48	0.8593	0.4704	0.5474	0.0679
13	44	0.7759	0.4139	0.5335	0.0624
14	16	0.9608	0.4127	0.4296	0.1032
15	26	1.1314	0.4957	0.4382	0.0972
16	11	1.2559	0.4868	0.3876	0.1468
17	5	1.0412	0.3490	0.3352	0.1561
18	20	1.2805	0.4340	0.3389	0.0970
19	4	1.5021	0.5549	0.3694	0.2774
20	4	1.1556	0.2175	0.1882	0.1088
21	3	1.1758	0.1008	0.0857	0.0582

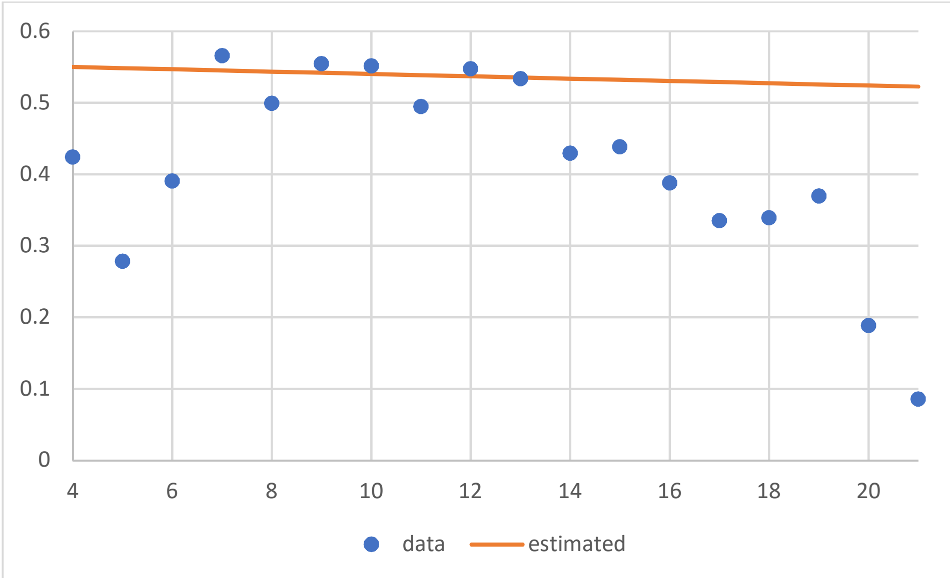
<table 3, relative statistics of the data>



<figure 7, 1-3 FBM difference and the number of candidates>



<figure 8, 1-3 standard deviation and the number of candidates >



<figure 9, The coefficient of variation $\frac{\sigma}{\mu}$, data and estimated>

For convenience, we keep our assumption that the data follow a generalized gamma distribution with d = 1.7766 that we have found earlier. Only the other two parameters, a and p, for the distribution differ by the number of the candidates. We use the method of moments.

To specify two parameters, mean and standard deviation are required, we use exponential function capture the trend, and least square method for the parameters **(<figure 7> and <figure 8>).**

$$\mu(C) = 1.42487\exp(-0.04064 \times C)$$

$$\sigma(C) = 0.79343\exp(-0.04367 \times C)$$

where μ and σ are the mean and the standard deviation of FBM difference between 1 and 3. C is the number of candidates in the survey.

The coefficient of variation, the ratio of the standard deviation to the mean, of the generalized gamma distribution is

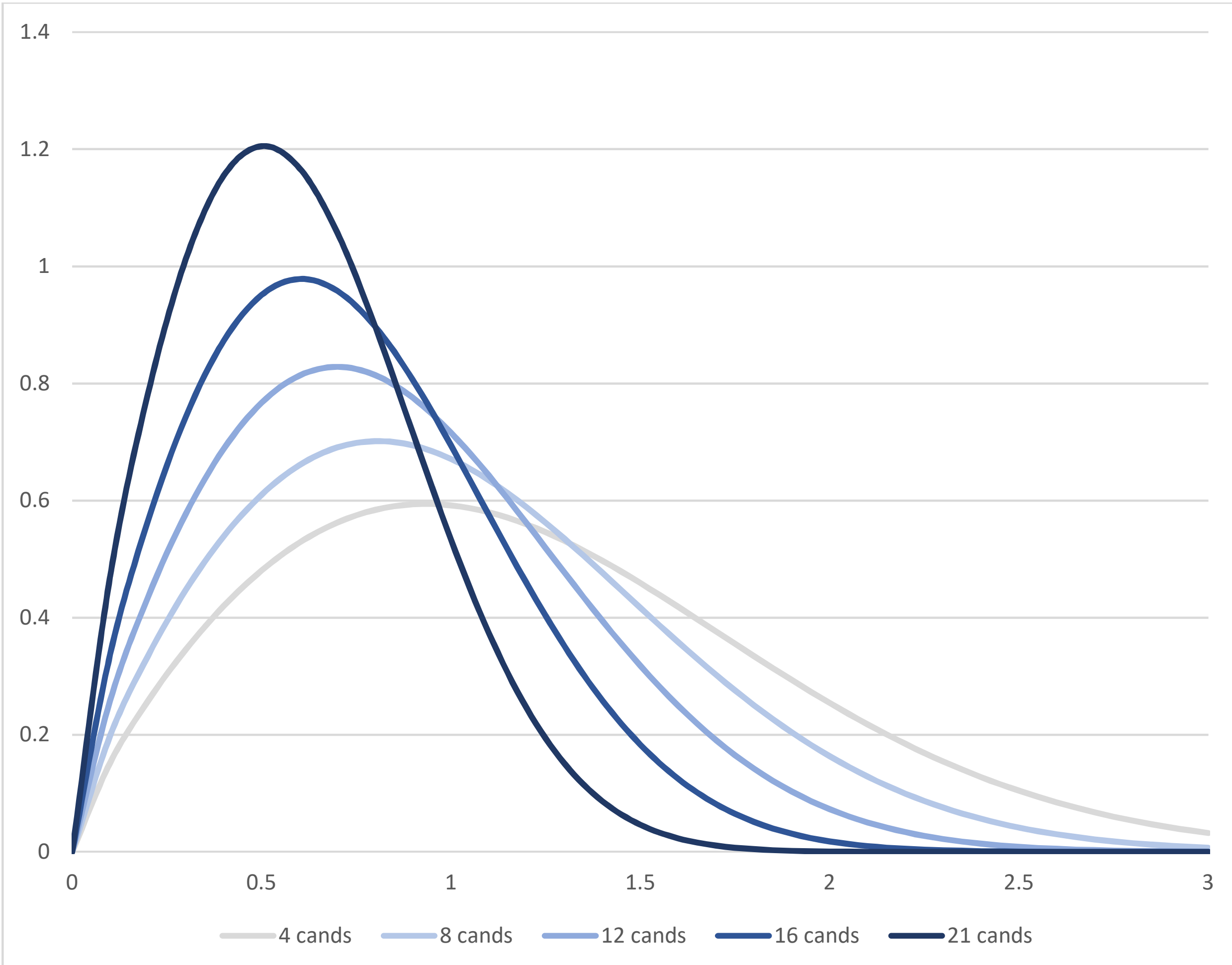
$$\frac{\sigma}{\mu} = \frac{\sqrt{\left(\frac{\Gamma((\mathrm{d}+2)/\mathrm{p})}{\Gamma(d/p)}\right) - \left(\frac{\Gamma((\mathrm{d}+1)/\mathrm{p})}{\Gamma(d/p)}\right)^2}}{\frac{\Gamma((\mathrm{d}+1)/\mathrm{p})}{\Gamma(d/p)}}$$ **<equation 3>**

The mean of the distribution can be found by

$$\mu = \mathrm{a}\,\frac{\Gamma((\mathrm{d}+1)/\mathrm{p})}{\Gamma(d/p)}$$ **<equation 4>**

We have with $\mu(\mathrm{C})$, $\sigma(\mathrm{C})$, and d, and solving **<equation 3 and 4>** to p and c, one could get p and then a that corresponds to each C.

Let $\mathrm{f}_C(x)$ is a generalized gamma distribution with a(C), d = 1.7766, p(C). The shape is of the same general form as in **<figure 9>**, but with a squeezing of the distribution toward the vertical axis as the number of candidates increases. This is applying our assumption that differences in FBMs within the top triple tend to shrink when a survey has many candidates.



<figure 9, estimated distribution of 1-3 FBM differences for the number of candidates>

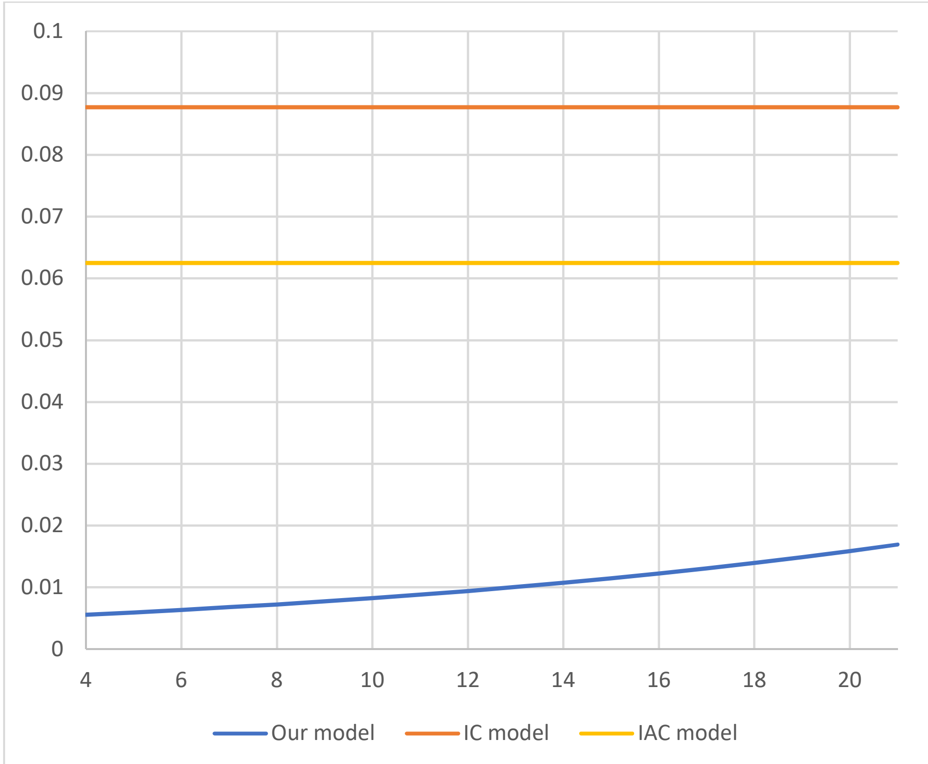
Reserve all other assumptions, the probability of a cycle can be calculated as

$$P(C) = \int_0^{10} f_C(D)\left(\int_0^1 g(z)\,[[1 - U(zD)][1 - U((1-z)D)]U(D) + U(zD)U((1-z)D)[1 - U(D)]]\mathrm{d}z\right)\mathrm{dD}$$

The only difference of the new equation from the **<equation 2>** is that the number of the candidates varies the distribution of D. The probability calculation results for each number of candidates are shown in **<table 4>** and **<figure 10>.**

Number of candidates	Estimated probability of a voting cycle
4	0.005581
5	0.005959
6	0.006363
7	0.006794
8	0.007254
9	0.007745
10	0.008269
11	0.008829
12	0.009426
13	0.010064
14	0.010744
15	0.011469
16	0.012243
17	0.013068
18	0.013947
19	0.014884
20	0.015883
21	0.016947

<table 4, estimated probability of cycle>



<figure 10, estimated probability of cycle >

As the number of the candidates increases, FBM difference between the first and the third candidates decreases, and voting cycle is more likely to exist.

4. Discussion and future work

Theories and models warned of the risk of the Condorcet paradox and tried to reveal the frequency of the cycle in the actual voting. However, voting cycle rarely happens in real cases, and finding a better method and bridge the gap between theoretical models and reality is our primary objective.

We propose a distinct approach from the distribution of the data. Using basic techniques and a unique data-set with cardinality in preference distinguishes our model from the others. At the same time, this is a double-edged sword, because it is difficult to find additional data with voters' scoring information.

The data shows 0 cycles in 1,022 top triples, and 206 majority cycles exist out of 181,579 all possible triples. On the other hand, in the base model, the estimated probability of a voting cycle among the top three candidates in a survey is 0.0088, or 0.88%. In the modified model with the additional assumption, the estimation ranges from 0.0056 to 0.0169 as the number of candidates increases. Consider a binomial distribution with p = 0.0088 with 1,022 trials. The expectation is 8.9836, and the standard deviation is 2.653.

While critics might not agree with our approach, one step closer to an accurate reflection since our results suggest a much lower probability of Condorcet paradox than the 8.77% and 6.25% that the IC and IAC models suggest. Moreover, we specify several aspects ignored for the modeling process, which can be checked out for future work.

First, we ignored some statistics. We did not count the number of voters for each survey and the representativeness of each observation, the weighting factor. Regarding the weighting factor, since we split the observations by region and by the questionnaire, it is tricky to adjust the factor correctly.

Second, for all separated surveys, we assumed that they were independent of each other. Sometimes the candidate pool changes insignificantly from a month to the following month, and aggregated evaluation of a candidate is related.

In some surveys, participants need to assess and judge many candidates. For instance, when the participant had 20 or 21 questions for 20 or 21 candidates, he could feel fatigued and move over some "unpopular" candidates without serious consideration and evaluation. If this is the case, it will provide better results when we group surveys with more than a reasonable number of candidates.

Fourth, we did not handle blanks. In pairwise comparison, we only counted the participants who evaluated both candidates and got rid of those who did not. This might be problematic. However, we do not expect the distortion from this process to be severe since this research focused on “strong″ candidates, those who are highly evaluated by most of the participants. One possibility is to consider comparing a score and the blank as a choice under uncertainty or ambiguity. It would not be tough to make a preference order. Still, it is challenging to fill the blank with numbers to get a convincing aggregated evaluation or FBM.

References

[1] Darmann, A., Grundner, J., & Klamler, C. (2019). Evaluative voting or classical voting rules: Does it make a difference? Empirical evidence for consensus among voting rules. European Journal of Political Economy, 59, 345-353.

[2] Feizi, M., Ramezanian, R., & Malek Sadati, S. (2020). Borda paradox in the 2017 Iranian presidential election: empirical evidence from opinion polls. Economics of Governance, 21(2), 101-113.

[3] Gehrlein, W. V., & Fishburn, P. C. (1976). The probability of the paradox of voting: A computable solution. Journal of Economic Theory, 13(1), 14-25.

[4] Gehrlein, W. V. (2002). Condorcet's paradox and the likelihood of its occurrence: different perspectives on balanced preferences. Theory and decision, 52(2), 171-199.

[5] Gehrlein, W. V. (2006). Condorcet's paradox (pp. 31-58). Springer Berlin Heidelberg.

[6] Gehrlein, W. V., & Lepelley, D. (2011). Condorcet efficiency and social homogeneity. In Voting Paradoxes and Group Coherence (pp. 157-198). Springer, Berlin, Heidelberg.

[7] Good, I. J., & Tideman, T. N. (1976). From individual to collective ordering through multidimensional attribute space. Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences, 347(1650), 371-385.

[8] Guilbaud, G. T. (1952). Les theories de Pinteret general et le probleme logique de Fagregation. Economic appliquee, 5, 501-584.

[9] Kuga, K., & Nagatani, H. (1974). Voter antagonism and the paradox of voting. Econometrica:

Journal of the Econometric Society, 1045-1067.

[10] Kurrild-Klitgaard, P. (2018). Trump, Condorcet and Borda: Voting paradoxes in the 2016 Republican presidential primaries. European Journal of Political Economy, 55, 29-35.

[11] Mattei, N. (2011, October). Empirical evaluation of voting rules with strictly ordered preference data. In International Conference on Algorithmic Decision Theory (pp. 165-177). Springer, Berlin, Heidelberg.

[12] Mattei, N., Forshee, J., & Goldsmith, J. (2012). An empirical study of voting rules and manipulation with large datasets. Proceedings of COMSOC, 59.

[13] Popov, S. V., Popova, A., & Regenwetter, M. (2014). Consensus in organizations: Hunting for the social choice conundrum in APA elections. Decision, 1(2), 123–146. https://doi.org/10.1037/dec0000010

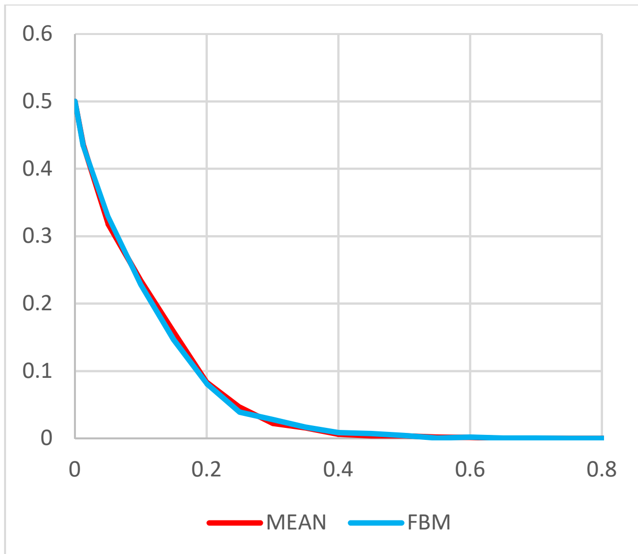
[14] Regenwetter, M., Kim, A., Kantor, A., & Ho, M. H. R. (2007). The unexpected empirical consensus among consensus methods. Psychological Science, 18(7), 629-635.

[15] Tideman, T. N., & Plassmann, F. (2012). Modeling the outcomes of vote-casting in actual elections. In Electoral Systems (pp. 217-251). Springer, Berlin, Heidelberg.

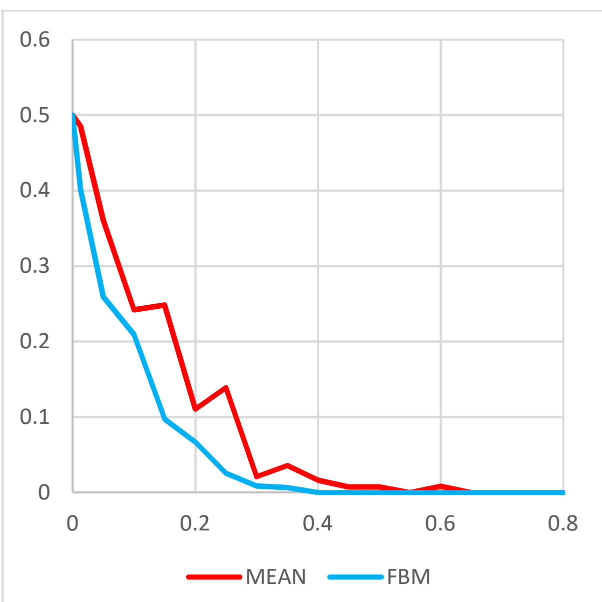
[16] Van Deemen, A. (2014). On the empirical relevance of Condorcet's paradox. Public Choice, 158(3), 311-330.

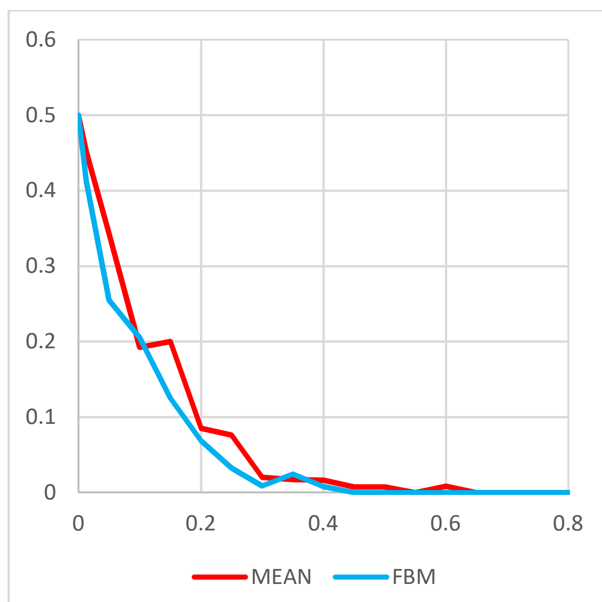
Appendix.

A. Upset rate and mean difference / FBM difference in 55,321 pairwise comparisons (all possible pairs).



B.Upset rate and mean difference / FBM difference in 3,066 pairwise comparisons (top 3 candidates).





From top 3 candidates by FBM

From top 3 candidates by MEAN

Why do we NOT use the mean of the evaluations to predict upset?

One of the essentials in our modeling is estimating the probability of the upset, and we suggest innovative statistics, FBM, for prediction. We suggest a reason that we stick to the median in section 2. It is, by definition, the median is directly linked to the distribution of the majority. In addition, it is well-known that the mean is vulnerable to extreme values. We attach what our data tell for another reason.

There is no difference in all pairs. However, when it comes to pairs generated from the 1,022 top triples, with the same numeric difference between two candidates in different statistics, upset is less likely to happen in FBM, which indicates that FBM works better than the mean to predict the upset in pairwise comparisons.