Three Empirical Analyses of Voting

Chang Geun Song

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Economics

Nicolaus Tideman, Chair
Richard Ashley
Eric Bahel
Florenz Plassmann

May 5, 2022
Blacksburg, Virginia

Keywords: Condorcet's paradox, Social choice, Alternative vote, Collective decision making, Instant-Runoff Voting, Election, Agent-based modeling

Three Empirical Analyses of Voting

Chang Geun Song

(ABSTRACT)

To evaluate voting rules, it would be good to know what universe election outcomes are drawn from. Election theorists have postulated that elections might be drawn from various stochastic preference models, including the IC and IAC conditions, but these models induce empirically contradicted predictions. We use two distinct data sets, FairVote and German Politbarometer survey. Based on the data information, we suggest approaches that differ from those probabilistic models to better approximate the actual data in Chapter 3 and 4. Chapter 5 applies the spatial model for four-candidate in a three-dimensional setting. We also offer a significant gap between the actual and simulated data under the IAC conditions by comparing their statistical characteristics.

Three Empirical Analyses of Voting

Chang Geun Song

(GENERAL AUDIENCE ABSTRACT)

Through the 1884 Third Reform Act, the plurality rule (or first-past-the-post system) runs to elect parliament members for the first time. More than a hundred years passed after the Act, and election theorists have suggested various alternatives, the plurality rule is the second most used rule worldwide for national elections for now. One main reason is that researchers do not reach an agreement on the best alternative rule. Theorists have evaluated different voting rules under probabilistic assumptions, but real-world examples contradict the predictions of these models. In this dissertation, we suggest different approaches provide a better approximation to the actual data. In Chapter 3 and 4, we go backward: analyze how voters of each preference order are distributed in real data first, then set a model for estimating the frequency of paradox. In chapter 5, we extend an existing model with higher dimensionality. Then using the model, we offer empirical evidence showing the gap between the actual and simulated data under a popular probabilistic model.

Dedication

This dissertation is dedicated to my parents, Tae Yeol Song and Sun Sim Kim.

They have always supported and loved me unconditionally.

This work is also dedicated to my sister Hwayoung Song.

The box full of snacks she sent for my every birthday helped me deal with stress.

I dedicate this work and thank all of my Virginia Tech fellows.

Especially to Sunjin Kim and Hoang Phuong Do.

They always cheered me up when I was stuck by the challenges of graduate student life:

All academic and non-academic kinds.

Acknowledgments

First but foremost, thanks to my advisor, Dr. Nicolaus Tideman, for every support, guidance, and patience throughout all of my Ph.D experience. Especially when I was struggling as a stranger to Public Choice and Voting literature, he set me to the correct starting point and gave me great motivation to work on it. He is full of new ideas and has a unique perspective, and I always enjoy talking with him. Every meeting I had was like a lecture of his, a bit difficult to catch up sometimes, so priceless, and I learned a lot.

I also thank my committee members, Dr. Richard Ashley, Dr. Eric Bahel, and Dr. Florenz Plassmann, for their support. Without them, I would not have been able to finish this journey. Notably, I am grateful to Dr. Florenz Plassmann, who always gives me thoughtful and sincere comments about my writings, research, and the job market.

I thank Dongwoo Lee and Brian D'orazio for proofreading my work and giving comments.

Lastly, I appreciate Dr. Yong Gwan Kim and Dr. Kyung Hwan Baik. Thanks to their guidance and encouragement, I could set my path and take my very first step to become an Economist.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

It is simple to understand, easy to process, and cost-saving. The plurality rule is the globally most used system in elections for national legislatures, along with proportional representation [1].

Studies argued the weaknesses of the plurality rule and suggested other rules to cure them. They are supposed to be better than the old ones: more straightforward, cheaper, fairer, more reflecting individual's preferences, incentivizing voters to be more honest. Researchers studied for a long time to evaluate different voting systems, using theoretical criteria. Unfortunately, to check most of the criteria, complete rankings of preferences are required, not just first choices. However, since systems other than plurality were rarely adopted in the real world, finding relevant data and empirical evidence to champion their theories was rare. Instead, researchers lean on probabilistic models and simulations for their arguments. Alternatively, other researchers analyze available actual data, although the amount is not abundant. They insist that the theoretical approach of estimating voting paradoxes and criteria violations is a bit off from the actual outcomes or intimidates people by exaggerating the hypothetical situation. Eventually, there is no consensus of which one is the best electorate system yet.

Nevertheless, at some point, alternative voting rules have been adopted. For example, some countries (Australia and Ireland) adopt *Instant-Runoff Voting* (IRV) for national elec-

[1] From International Institute for Democracy and Electoral Assistance (IDEA, https://www.idea.int/data-tools/question-view/130355)

tions, and some states in the US also use the system for local elections. Additionally, in more than 50 US colleges and political party elections, IRV is used. Conflicts between supporters and critics exist, as expected.

That concern motivates this research. To find out the best voting rule, accurate evaluation of each voting systems in realistic setting is required. Through these series of study, we offer different points of views.

This field of study gets more and more essential as we are ready with advanced technology and environments enough to operate more complicated electoral systems without much expenditure: 1. High computing power allows to calculate the winner under a complicated voting system without paying heavy cost, 2. High penetration rate of smart devices grants possibility of fully online voting[2], and 3. Block-chain technology could deal with potential election manipulation problem.

Newly assign costs should be considered as well, but while counting cost-saving from the problems in representative democracy such as principal-agent problem, the benefits from reflecting group members' preferences more accurately, the cost such as instruction and adaptation from introducing new electoral system would be bearable.

Chapter 2 introduces terms and concepts frequently appearing in the following chapters. General information of the data sets we use is also provided. In Chapter 3, we suggest more data-oriented analyses for estimating the frequency of the cycle. Chapter 4 also deals with the voting cycle but with a different approach. Additionally, we analyze Condorcet inconsistency. Chapter 5 focuses on the Normal Spatial model. The model better fits actual data than other probabilistic models in three-candidate elections. In the chapter, we extend the model for four-candidate cases.

[2]In the US, the share of adults who owned a smartphone was 85% in February 2021 (https://www.statista.com/statistics/219865/percentage-of-us-adults-who-own-a-smartphone/), and the share of adults who owned a laptop was 74% in February 2019 (https://www.statista.com/statistics/756054/united-states-adults-desktop-laptop-ownership/).

Chapter 2

Preliminaries

2.1 General

2.1.1 Preference Order and Preference Schedule

For arbitrarily labeled alternatives **A**, **B**, and **C**, we refer **AB** as the preference order that strictly prefers **A** to **B**, and **ABC** is the preference that prefers **A** to two others, and **B** to **C**. When a voter is indifferent between two candidates, **A** and **B**, we describe it with parenthesis as (**AB**).

Counting only strict preference, a voter can have one of six rank orders. We call a collection of voters' preference orders a preference schedule or a preference profile. For three voters, for example, we have 216 possible preference schedules: $\{\mathbf{ABC}, \mathbf{ABC}, \mathbf{ABC}\}$, $\{\mathbf{ABC}, \mathbf{ABC}, \mathbf{ACB}\}$, ..., and $\{\mathbf{CBA}, \mathbf{CBA}, \mathbf{CBA}\}$.

We use $n[\mathbf{ABC}]$ as the number of voters with **ABC**. **ABCA** will be used for cyclical preference, preferring **A** to **B**, **B** to **C**, and **C** to **A**. For the oppositely directed cycle, we use **CBAC**.

2.1.2 Condorcet Paradox (Voting Cycle)

Collective preference may be cyclic with more than two alternatives; even when every individual preference satisfies transitivity. We call it the Condorcet paradox, and the voting cycle occurs when voters report their preferences truthfully. When the cycle exists, it is hard to tell what option is the best for the group. Specifically, no voting system properly works because there will always be people more than the majority who prefer another alternative to the chosen option under any decision-making process. Consequently, analyzing voting cycles and suggesting rules to deal with the cycle have been popular research topics.

2.1.3 Condorcet Method and Condorcet Criterion

The Condorcet method is a ranked-ballot voting rule that determines who wins an election based on pairwise majority comparisons for all possible pairs. When a candidate beats all others in pairwise comparisons, the candidate is called the Condorcet winner and considered the best candidate if such a candidate exists. When a voting system always chooses the Condorcet winner as the winner of the election, the voting system satisfies the Condorcet criterion, or it is Condorcet consistent, and this is one of the criteria for evaluating a voting system. Darlington (2017) argues that Condorcet consistency is crucial and needs to be severely considered even if it means giving up other criteria. If a voting system is not Condorcet consistent, then the problem that the voting cycle induces, the existence of an alternative favored by the majority over the winner of the system, is even worse.

2.2 Stochastic Preference Models

In most cases, each voter's complete preference order is necessary to evaluate criteria and check various voting paradoxes in different electoral systems. However, since large data sets of the relevant sort are scarce, researchers have mainly used hypothetical preference profiles generated by probabilistic assumptions.

When we have N alternatives, we have $N!$ possible strict orderings. Thus, frequency distribution over the strict orderings would be expressed by a set of $N!$ numbers, such that the sum of components is 1. Specifically, for three options, **A**, **B**, and **C**, we have six strict rankings **ABC**, **ACB**, **BAC**, **BCA**, **CBA**, and **CAB**. Probabilistic models assign a distribution to these six.

The Impartial Culture (IC) and *Impartial Anonymous Culture* (IAC) models are standard. They are widely used for analytical approaches and simulations, due to their convenience for calculations. However, they are criticized because they are not very realistic.

2.2.1 Impartial Culture (IC) and Impartial Anonymous Culture (IAC) Model

Guilbaud (1952) proposes the IC model, in which, for every voter, each preference order will be realized with equal probability. Hence, for a three-candidate case, every individual would have one of six preference orders with 1/6 probability under the IC assumption. This assumption yields equal probability on each preference schedule as well.

The IAC model is developed by Kuga and Nagatani (1974) and Gehrlein and Fishburn (1976). The IAC assumption appends anonymity to the IC, and does not take account of the order of voters in a preference schedule. It counts a voting situation as a set and does not differentiate two different preference schedule if the elements' composition is the same.

A set is called an *anonymous equivalence class* (AEC). The IAC assumption assigns an equal probability of realization to all possible sets of preference profiles, or AECs. One could relate IC to permutations with repetition and IAC to combinations with repetition.

2.2.2 Calculation of Voting Cycle Probability under IC and IAC Assumption

The frequency of the event can be calculated by (the number of cases that the event occurs)/(total number of cases).

Suppose three candidates and three voters, the simplest case. The preference profile will be a collection of three preference orderings. The number of all possible preference profiles is $6^3 = 216$. They can be categorized by $(3+6-1)!/3!(6-1)! = 56$ voting situations or AECs. Consider a cycling **ABCA** (**A** beats **B**, **B** beats **C**, and **C** beats **A**). The cycle occurs when three voters have **ABC**, **BCA**, and **CAB**. Therefore, the total number of preference profiles in which the event occurs is six, the permutations of **ABC**, **BCA**, and **CAB** under the IC assumption. On the other hand, it is 1, the combination of **ABC**, **BCA**, and **CAB** under the IAC assumption. The other cycling **CBAC** also should be considered. Hence the probability of voting cycle in three candidates and three voters cases is $12/216 = 0.0556$ and $2/56 = 0.0357$ for IC and IAC assumptions, respectively.

The following Table shows probabilities for some finite numbers of voters and that for infinite numbers of voters. Probabilities show a significant gap between odd number cases and even number cases. The reason is that semi-cycles are also counted for cases of even voters. In cases where **A** is not less preferred to **B**, **B** is not less preferred to **C**, and **C** is not less preferred to **A**, and the opposite directed ones are all counted. The probabilities converge as the number of voters increases (Gehrlein, 2002, 2006).

the number of Voters	IC	IAC
3	0.0556	0.0357
4	0.3750	0.2857
5	0.0694	0.0476
6	0.3164	0.2208
7	0.0750	0.0530
8	0.2814	0.1841
9	0.0780	0.0559
10	0.2579	0.1608
⋮	⋮	⋮
∞	0.0877	0.0625

Table 2.1: Probabilities of the cycle with three candidates by the number of the voters in IC and IAC assumption

2.2.3 The Normal Spatial Model

The normal spatial model was developed by Good and Tideman (1976). This model assumes an attribute space, and an individual's preference ordering is determined by the (Euclidean) distances from the options to his or her most desired combination of attributes, with closer alternative ranks higher. The normal spatial model assumes that the voters' most desired options are drawn from a (multivariate) normal distribution. The bisecting hyperplanes of the line segments that connect all pairs of candidates divide the space into sectors corresponding to different strict preference orderings. The voters' probability density integration over all such sectors specifies the probabilities. One needs at least two dimensions to represent all six possible strict rankings for three candidates. For four candidates, at least three dimensions are needed for the 24 possible strict rankings.

2.3 Data Description

In this dissertation, we use two different types of data sets. One is actual election data, from FairVote. The other is survey data from the German Politbarometer. In each chapter, we have different purposes and use different analysis methods. Hence data cleaning process of each chapter varies. In this section, we provide general information, and the detail of the process will be described in each chapter if needed.

2.3.1 FairVote

FairVote data is actual data of numerous local elections across the United States[1]. The elections are sporadic; in the data we have, the earliest is November 2004, and the latest is November 2020. The electorate system is Instant Run-off Voting. The data is a rank-order-ballot of voters, and it includes more specified preference data of voters than data of elections run by plurality rule.

Every election data of FairVote shows individual voters' first-choice, second-choice, and third-choice (or more than three, it varies). In some elections, voters only can choose a limited number of candidates out of the candidate pool, which candidates' number is more than the limitation. We cannot specify the voter's complete preference in these cases, because the preference for not-ranked candidates is not observable. Kilgour et al. (2020) warn against using data with incomplete preference order, since it could significantly impact voting outcomes, but we set aside the ballot truncation problem for our study.

Each choice in the ballot is supposed to be filled with a candidate's name in the candidate pool, but there are some anomalous: *skipped*, *write-in*, and *overvote*.

[1] Including San Francisco, Cambridge, Minneapolis, Berkeley, Oakland, and New York City, more than 40 cities have been used or plan to use Instant-Runoff Voting from 2020 to 2022 (`https://www.fairvote.org/data_on_rcv`)

Skipped is literally when the voter did not make his/her choice for first-choice, second-choice, or third-choice. *Write-in* occurs when a voter wants to vote for someone not on the candidate list and writes the person's name on a ballot. Finally, *Overvote* is the case in which the voter gives the same rank to more than two candidates.

Skipped, *write-in*, and *overvote* are removed and filled up by the following preferred candidate if it is possible. Specifically, **A**-[skipped]-**B**, **A**-[write-in]-**B**, or **A**-[overvote]-**B** will be considered as the voter prefer **A** the most and then **B** the second most, **A**-**B** or shortly **AB**. All weak preferences are not counted.

2.3.2 German Politbarometer

German Politbarometer data is survey data that are freely available on the official website[2]. The survey is conducted monthly (approximately); For now, the data is available starting from March 1977 and ending in December 2020. Participants answered more than one hundred questions for each survey, and we used a set of questions categorized as *skalometer*. For each question, the participant gives a score of one of the integers from -5 to 5 for each politician.

The politician pool for the questionnaire consists of influential politicians at the time of the survey. There are mainly German politicians as active candidates in the next elections. However, retired politicians, foreign presidents, or foreign prime ministers are also in some surveys[3], so it is close to a general popularity poll. There is no regular rule for forming the

[2]Anyone could access data after signing up at the gesis website (https://www.gesis.org/en/elections-home/politbarometer). They yield the data in SPSS and STATA data format. The survey contains various questions such as demographic characteristics, region of residence, and weekly working hours.

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

pool; some politicians continue to be on the list of the following survey, and others do not. Some excluded politicians are re-included in later surveys' pools.

We assume that the participants would vote to comply with their evaluations. Thus we derive voters' preference order from the score data. If the participant assigned the same score on multiple politicians, we assumed the voter equally preferred or indifferent. On the other hand, if the participant did not assign any score to some politicians, we considered he would not rank for them in the election. For instance, if a participant assigns 4, −1, **blank**, 3, −1 for **A**, **B**, **C**, **D**, and **E**, then it counts as **A**-**D**-(**B** and **E**) or shortly **AD**(**BE**). We eliminate voters assigning an identical score to all candidates as careless evaluation and providing no information about preference. For the research purpose, each survey is regarded as an election, and we assume that all surveys are independent of each other. Additionally, any impact from which party the candidate belonged is not independently handled for our study; instead, we take that it is also just one of the attributes for evaluation.

2.3.3 Statistics of the Data

In 172 observations of FairVote data, three to 35 candidates are in each election (5.372 on average), and the elections range in size from 1,557 to 298,808 ballots (40253.198 on average). On the other hand, we have 1,022 surveys as observations in Politbarometer data. Four to 21 politicians are evaluated (10.62 on average), and the number of participants ranges from 179 to 3,345 (930.05 on average).

Chapter 3

Estimating the Probability of a Voting Cycle

3.1 Introduction

In this chapter, we analyze German Politbarometer data and then suggest a way to explain the observed frequency of the Condorcet paradox and show the departure of the result from the predictions of previous models. The modeling process is based on empirical distribution, unlike previous probabilistic models. What makes our work different from the others is that we directly connect the cardinal information to the probability of winning in the pairwise election and establish a model to estimate the frequency of voting cycles. Specifically, we introduce the concept of a *fraction-valued median*, combining the standard median and the distributions of evaluations below and above the median. Then, we check the usefulness of the gap in the fraction-valued medians of scores of two candidates for predicting which candidate wins in pairwise comparisons. Finally, we consider a cycle as a result of three consecutive head-to-head elections, then calculate the probability by multiplying three estimated probabilities.

In Section 3.2, we introduce predictions of the occurrence of the Condorcet paradox in previous probabilistic models, simulations, and actual data. Next, Section 3.3 describes the methods of analysis and the results. Finally, Section 3.4 summarizes and proposes future research.

3.2 Literature Review

We will summarize the estimation of probabilistic models and simulations in previous research. There exist various models for voter's preference. In this chapter, we narrow down our scope and introduce estimation under Impartial Culture (IC), Impartial Anonymous Culture (IAC) and the normal spatial model. We also introduce articles related to how often the voting cycle happens in practice. We purposely omit research on actual elections published before 2010 (see Gehrlein and Lepelley, 2011, p. 13 and Van Deemen, 2014, Section 5).

3.2.1 Probabilities Implied by the Models

Gehrlein (2002, 2006) 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 becomes large enough, the IC model and the IAC model yield probabilities of the Condorcet paradox decrease and converge to 8.77% and 6.25%, respectively. On the other hand, the estimated frequency of a voting cycle under the normal spatial model converges to 0% as the number of voters increases. With a large enough number of voters, a candidate closer to the mode of normal distribution will undoubtedly have more votes than candidates farther from the mode unless the mode of the normal distribution locates at an equal distance to two candidates or more, which is a zero probability event.

3.2.2 Simulations Based on Actual Data

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

Darmann et al. (2019) use online survey data collected two weeks before the 2015 parliamentary elections in the Austrian federal state of Styria. Respondents were asked to assign points in the range [-20,20]. The authors report zero cycles in survey data and 13 voting cycles out of 1,000 bootstrapped profiles. Regenwetter et al. (2007) compare the collective preference of candidates under Condorcet, Borda, Plurality, and Instant-Runoff systems with simulated preference profiles. The data they use is from the American Psychological Association's presidential elections run by the Instant-Runoff system with five candidates and more than 15,000 voters. They simulate data by point estimations and bootstraps based on four elections with strict and weak order assumptions. They report no voting cycle in point estimates and 8,000 bootstrapped profiles with 100% confidence.

Another notable empirical study is Tideman and Plassmann (2012). They use actual election data from the Electoral Reform Society, and survey data from the American National Election Studies. Based on the actual preference distributions, they simulate 1,000,000 elections for each different number of voters under various models[1]. Specifically, in IC, IAC, and spatial models with 1,000 voters, the estimated voting cycles are 12.45%, 6.55%, and 0.53% of simulated elections. As the number of voters increases, they converge to 9.01%, 6.29%, and 0%, respectively.

[1] Besides IC, IAC, and spatial models, they also simulate under alternative versions of IACs, Dual culture, Borda model, Condorcet model, *Equal probability for same first*, and Unique unequal probabilities.

3.2.3 In Actual Data

Mattei (2011) measures 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. The elections 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 $1-5$ (assigning half stars permitted) for how well they liked the movies they watched[23]. The author makes three subsets by randomly drawing 2,000 movies from the 17,770 movies. He then generates all possible three-alternative elections and four-alternative elections in each subset. A derived election is included in the analysis only if it has more than 350 voters. The percentages of three-alternative elections that exhibit the Condorcet paradox in the three independent subsets are 0.041%, 0.044%, and 0.056%.

Mattei et al. (2012) extend the previous work. While using the same data, the authors make ten subsets with non-overlapping sets of 1,777 movies and check for voting cycles for all possible triples and quadruples in each subset. The frequency of the cycles in the top three candidates in four alternative elections is lower than 0.11%.

Kurrild-Klitgaard (2018) provides possible evidence of a 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 the 2016 Republican party presidential primaries. The result shows a voting cycle for the top three candidates. Feizi et al. (2020) also analyze several polls before an actual election. The polls began some weeks before the 2017 Iranian presidential election. The polls included the pairwise comparison of candidates. The authors conclude that there was no voting cycle for three leading candidates in any poll.

[2] One can find detailed information at `https://www.netflixprize.com/`

[3] The rating by stars system has been abolished and replaced with ratings by *like* or *dislike* system

3.3 Analysis

3.3.1 Data Descriptions

This chapter uses German Politbarometer data, beginning in 1977 and ending in 2019. For convenience, we confine our analysis to triples of candidates, the minimum number of alternatives needed for a Condorcet paradox. We construct triples from each survey for all possible combinations of three politicians in the survey. 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 assigned scores to neither of the two politicians for each pairwise comparison. There are 206 majority cycles (including semi-cycle) in 181,579 triples, a rate of 0.113%, or about 1 out of 881. Thus, the actual data shows a significant departure from the limit probability of the IC and IAC models, 8.77% and 6.25%.

3.3.2 Framework for Analysis

To check whether there is a voting cycle, we need to know the majorities in three paired comparisons. For arbitrarily labeled alternatives, **A**, **B**, and **C**, we define $d_{\mathbf{AB}}$ as the vote difference between **A** and **B**, or $n[\mathbf{AB}] - n[\mathbf{BA}]$.

As Table 3.1 shows, there are eight possible combinations of the signs 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 preference of the three options.

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

Table 3.1: Eight possible profiles of signs of d_{ij} and implied preference orderings

If we can accurately predict the sign of $d_{\mathbf{AB}}$, $d_{\mathbf{BC}}$, and $d_{\mathbf{AC}}$, then we can also predict whether or not there is a voting cycle. Furthermore, in the following subsection, we argue that the distributions of *skalometer* scores are good predictors of which option is preferred by more voters in a head-to-head comparison, and thus the signs of majorities and the existence of voting cycles. The median evaluation of each candidate is the fundamental concept of this study.

3.3.3 Why Does the Median Matters

Let the voters' cardinal evaluation of two alternatives be distributed on a two-dimensional space as in Figure 3.1. A voter's evaluation of **A** is on the x-axis, and that of **B** is on the y-axis. $Med_{\mathbf{A}}$ and $Med_{\mathbf{B}}$ are the median evaluation for **A** and **B**, respectively. While assuming that 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. However, due to half of the voters being on the right of $Med_{\mathbf{A}}$, and half voters being below $Med_{\mathbf{B}}$, option **B** is less likely to have a majority of the votes.

Another way to think about the possibilities is to rotate the horizontal line $Med_{\mathbf{B}}$

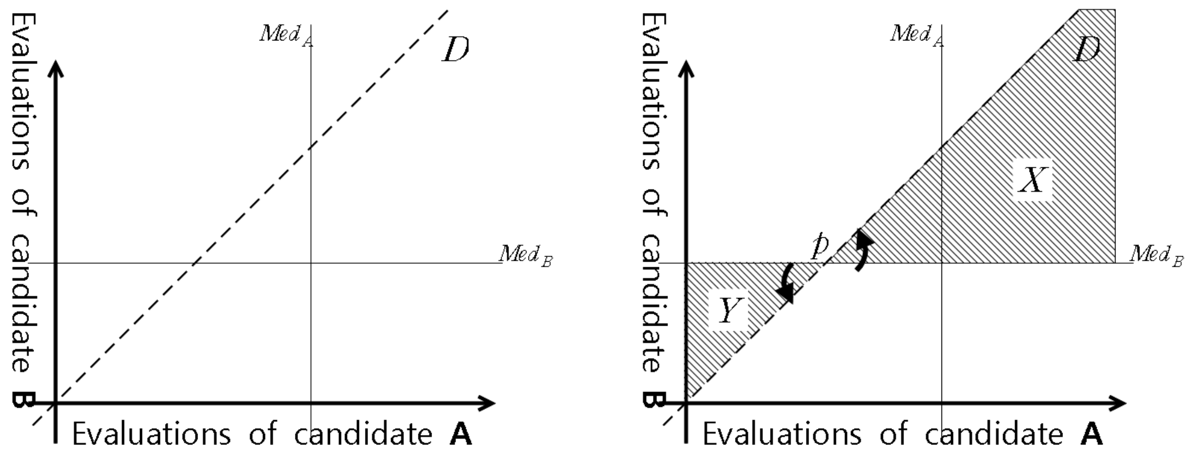


Figure 3.1: Median and majority

around a point p, 45 degrees, counterclockwise; The line will coincide with the dashed line D. The voters who rank **A** ahead of **B** are distributed in the area below the dashed line D. Line $Med_{\mathbf{B}}$ represents an even split of the voters. Therefore, for a majority in favor of **B**, the number of voters in area X should be greater than in area Y. If we suppose a **well-behaved** distribution of voters evaluations for two alternatives, it is more plausible to have more voters in X than Y.

To measure the impact of relative median evaluations on the probability of winning, 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 evaluations, the candidate with the lower median beats the other (we will call this an *upset* throughout this chapter) in 213 pairs. As Table 3.2 shows, it happens only when the gap is 0.5 or 1.

median gap	the number of pairs	the number of upset		median gap	the number of pairs	the number of upset
0	13,809	N/A		5.5	1	0
0.5	297	22 (7.41%)		6	93	0
1	22,195	191 (0.86%)		6.5	2	0
1.5	203	0		7	31	0
2	11,864	0		7.5	0	0
2.5	71	0		8	7	0
3	4,639	0		8.5	0	0
3.5	28	0		9	4	0
4	1,576	0		9.5	0	0
4.5	8	0		10	0	0
5	493	0		Total	55,321	213(0.39%)

Table 3.2: The number of pairs and upset by median gap in all possible pairwise comparisons

The upset rate by the median evaluation gap (Table 3.2) shows that the statistics have desirable characteristics as an indicator of the occurrence of an upset: First, upsets are not detected after a particular value, 1; Second, the rate is a decreasing function of the median gap; Third, the rate is easy to calculate.

Comparing the median evaluations seems a good way to predict which candidate will win in paired comparison. However, a substantial problem with using the median gap in our analysis is that the survey allows only 11 integers for scoring, so there are many ties in pairwise comparisons of medians. Indeed, we have pairs that have the same median evaluation between two candidates in about 1/4 of the total cases. Therefore, we need to identify another better measure of candidate quality, which would be a better explanatory factor of the difference between votes that two candidates receive and would be not difficult to observe and calculate.

3.3.4 The Fraction-Valued Median (FVM)

We offer a measure that has the character of a median but generally does not result in ties, despite there being only a few discrete values that the underlying variable can take. The *fraction-valued median* (FVM) takes a non-integer value based on the standard median and the placement of the median in the set of standard median-valued observations. It is derived in the following way.

-Calculate **M**, the standard median of the set,

-Count **N**, the number of observations that takes the standard median value,

-Count **K**, the number of observations that takes value lower than standard median.

Then FVM of the set is $\mathbf{M} - 1/2 + \mathbf{K}/\mathbf{N}$. By its definition, FVM can only have the value from $\mathbf{M} - 1/2$ to $\mathbf{M} + 1/2$, and even if two sets have the same standard median, it will generally be true that one will have a higher FVM than the other. Note that if $\mathbf{N} = 0$ or $\mathbf{N} = 1$, we define **M** as the FVM.

Intuition of FVM is following: Consider a ruler that only has major grids. The best way to guess something's length that does not precisely fit a grid is to measure how it is relatively far from the two closest grids. $\mathbf{M} - 1/2$ applies to the closest previous number, and $\mathbf{K}/\mathbf{N}$ is a normalized distance from it.

As an example of a calculation of an FVM, the set $\{2, 2, 3, 3, 4, 5, 7\}$ has an FVM of 3–0.5 + 0.75 = 3.25 (Figure 3.2). The value is between 2.5 and 3.5, which indicates that the standard median is 3, and having a greater FVM than the standard median indicates that the number of observations in the given set that take a value 3 is located more to the "left." In other words, the share of scores above 3 is larger than that of below 3. That is how the *usual majority* (Fabre, 2021) deals with tie-breaking under the highest median rule, checking normalized differences between scores above and below the standard median. FVM and the usual majority are started from different ideas but have the same essence and are mathematically identical[4].

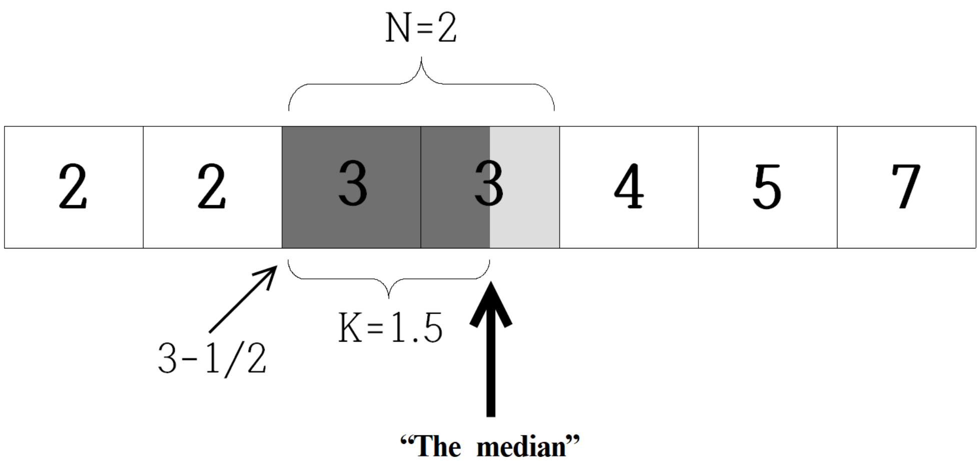


Figure 3.2: Idea of Fraction-Valued Median

The cumulative density function (CDF) of a set of discrete random variables is a step function. The standard median is the score where the CDF intersects a horizontal line at 0.5. Contrarily, suppose a series of line segments pass through the midpoints of all of the horizontal segments of the CDF (Figure 3.3). Then, the FVM will be at the fractional score where that series of line segments intersects the horizontal line at 0.5.

[4]For a tie-breaking in the highest median rule, Fabre (2021) suggests comparing *usual majority*, $\frac{1}{2}\frac{p-q}{1-p-q}$, where p and q are shares of scores above and below the standard median. In their terms, $\mathbf{K}$ and $\mathbf{N}$ match $0.5-q$ and $1-p-q$, respectively. Substituting those to $-1/2+\mathbf{K}/\mathbf{N}$ (second parts of FVM) would have the exact mathematical representation.

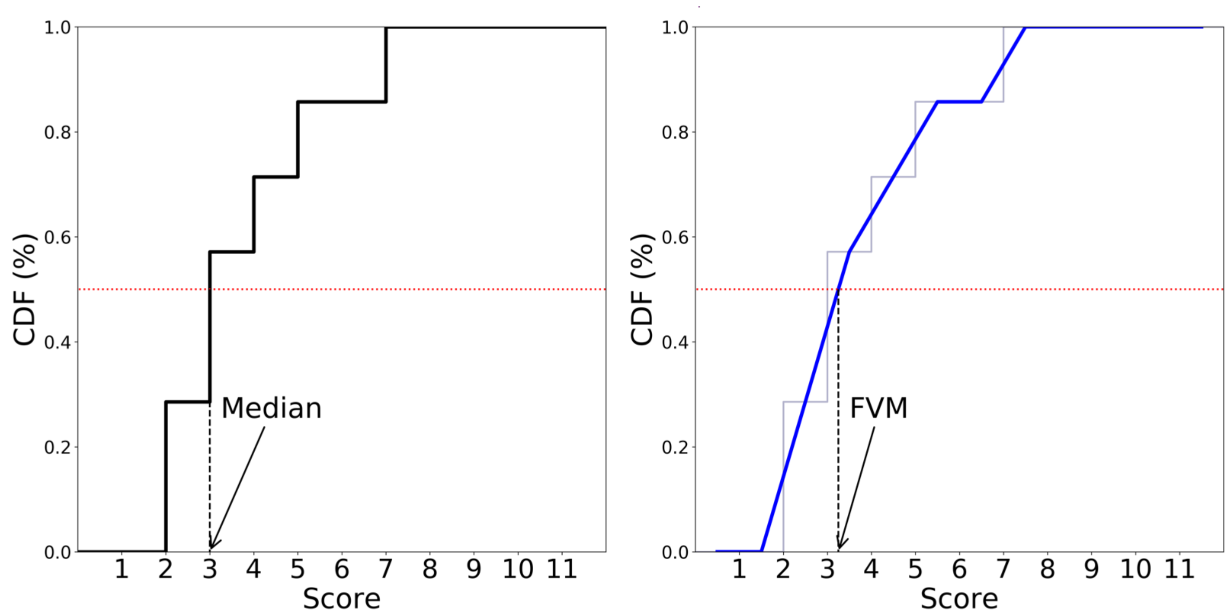


Figure 3.3: CDF, Median and Fraction-Valued Median

3.3.5 Frequency of upset

We restrict our interest to the top triple of each survey, the three politicians with the highest FVMs. The main reason for this restriction is that the frequency of voting cycles resulting in no Condorcet winner because of upsets by a fourth or lower evaluated politician can be expected to be very much lower than the frequency for cycles among the top three.

If scores assigned by participants for two candidates are not correlated, then in any paired comparison, the candidate with the higher FVM score is more likely to win, but it is not necessary; we will call the expected event a ***straight*** while using ***upset*** for the case that the candidate with the lower FVM beats the opponent as it is defined previously.

We checked how often upsets happen in all possible pairwise comparisons by the difference in fraction-valued medians in 3,066 pairs from the top three candidates of 1,022 surveys. Then, 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 FVM difference between two candidates, as shown in Figure 3.4[5]

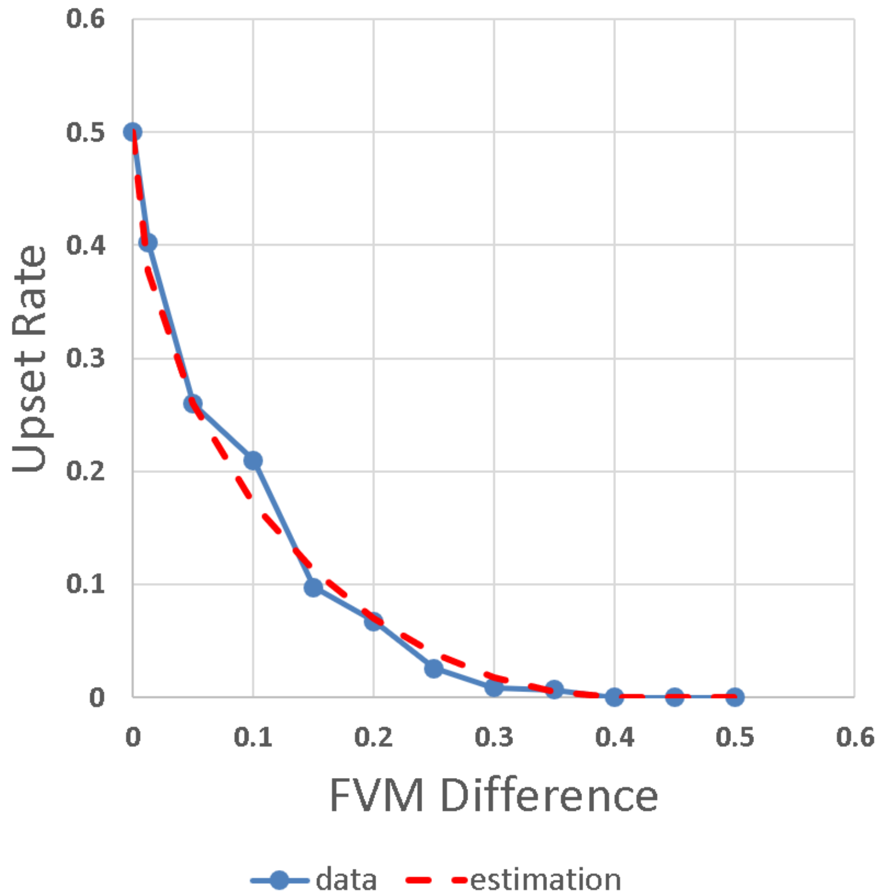


Figure 3.4: Upset rate by FVM gap

$$U(D) = \frac{1}{2}\left[1 - \sqrt{1 - \left(\frac{D - 0.4081}{0.4081}\right)^2}\right] \tag{3.1}$$

where D is the difference in FVM. Technically, the upset rate at a 0 FVM gap is not defined, but setting the value to 0.5 is reasonable because each candidate should win half the time when they have the same FVM. The number 0.4081 is the horizontal intercept that makes each area under the ellipse curve and the line from the data equal. In other words, the model says that the probability of an upset is 0 when the FVM gap is greater than 0.4081. In the

[5] We attach graphs and notes about the estimation of the upset rates by the mean score in Appendix A.

data, the largest FVM gap leading to an upset was 0.3352.

For a given FVM difference, D, $U(\cdot)$ generates the probability estimation of an upset, and $1-U(\cdot)$ estimates the probability of a straight. A voting cycle requires either (*straight*, *straight*, *upset*) or (*upset*, *upset*, *straight*) in pairwise comparisons, and the probability of a voting cycle can be calculated as follows

$$p\left(d_{\mathbf{12}}, d_{\mathbf{23}}, d_{\mathbf{13}}\right) = \left[1-U\left(d_{\mathbf{12}}\right)\right]\left[1-U\left(d_{\mathbf{23}}\right)\right]U\left(d_{\mathbf{13}}\right) + U\left(d_{\mathbf{12}}\right)U\left(d_{\mathbf{23}}\right)\left[1-U\left(d_{\mathbf{13}}\right)\right]$$

where d_{ij} is FVM difference between ith and jth candidates. We label **1**, **2**, and **3** for candidates by FVM: **1** for the highest, **2** for the second-highest, and **3** for the third-highest.

We extract ${d_{ij}}^n$s (FVM differences in the nth survey) from 1,022 surveys, and calculate $\sum_{n=1}^{1022} p^n\left({d_{\mathbf{12}}}^n, {d_{\mathbf{23}}}^n, {d_{\mathbf{13}}}^n\right)$ as a simple estimate of voting cycles, and we have 9.1837 estimated cycles (0.899%).

3.3.6 Distributions of $d_{\mathbf{12}}$, $d_{\mathbf{23}}$, $d_{\mathbf{13}}$

Besides the function for estimating upset rate as a function of D, the question of interest is the distribution of D. The empirical frequency distribution of FVM differences between candidates **1** and **3** is shown in Figure 3.5. We fit a generalized gamma distribution (three-parameter gamma distribution) in the data by its shape. 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, must be **2**.

The probability density function of the generalized gamma distribution is

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

Figure 3.5 shows the estimated and the actual frequency of the FVM gap between **1** and **3**. We apply the maximum likelihood method to find the parameters.The likelihood maximized when $(a, d, p) = (1.1531, 1.7766, 2.0646)$. We would like to estimate the dis-

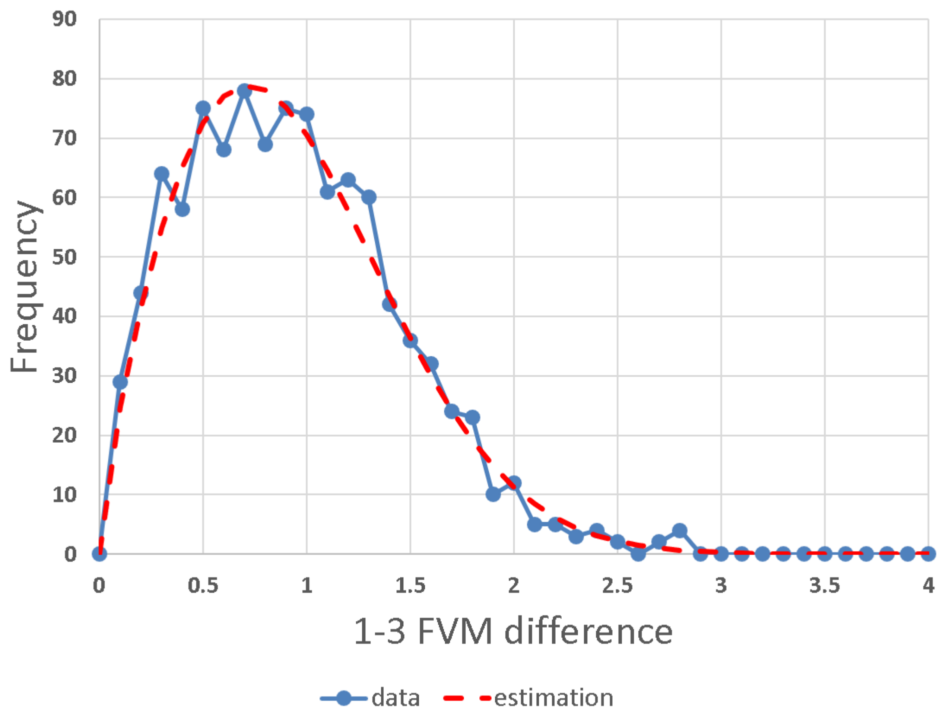


Figure 3.5: Empirical distribution of FVM gap between 1 and 3, and estimation with three-parameter gamma distribution, $(a, d, p) = (1.1531, 1.7766, 2.0646)$

tributions of the FVM difference between candidates **1** and **2**, and **2** and **3**. These FVM differences depend on the FVM difference between candidates **1** and **3**. Hence, we do not assume each FVM difference is drawn independently: we assume an FVM difference between candidates **1** and **3** is known first, then divided into $\mathbf{1} - \mathbf{2}$ difference and $\mathbf{2} - \mathbf{3}$ difference with a specific share rate. Thus, if we define the share of FVM difference between **1** and **2** in the FVM difference between **1** and **3** as $z = \frac{FVM_1 - FVM_2}{FVM_1 - FVM_3}$, and examine the empirical distribution of this share. Figure 3.6 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 FVM difference between **1** and **2** is likely to be greater than **2** and **3**.

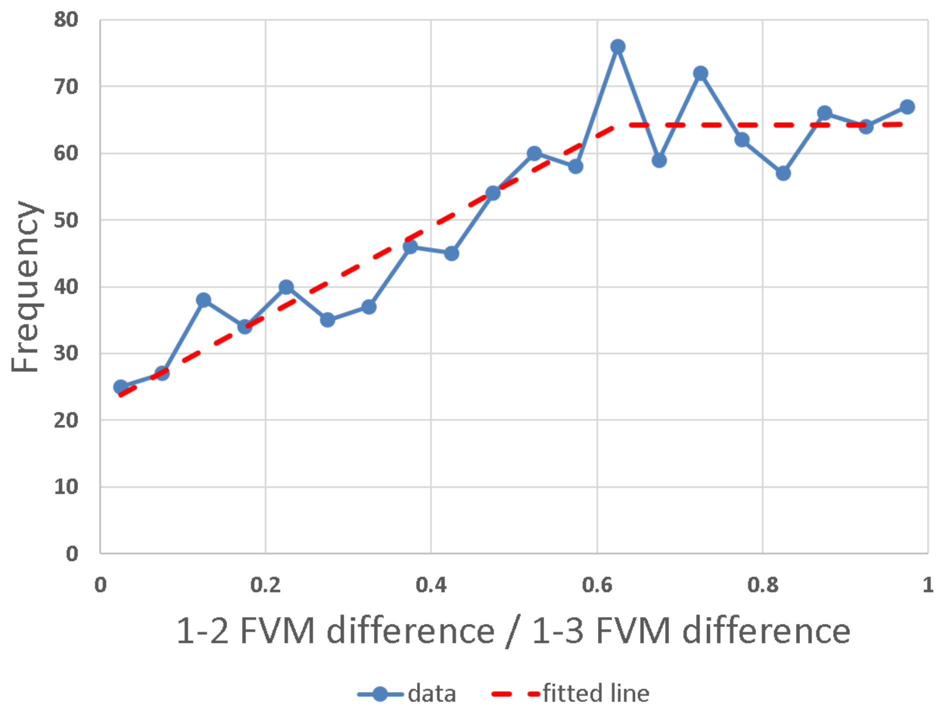


Figure 3.6: Empirical distribution of the ratio z and estimation with two line segments

$$g(z) = \begin{cases} 64.206 + 67.327(z - 0.625), & if\ z \leq 0.625 \\ 64.206 + 0.320(z - 0.625), & if\ z > 0.625 \end{cases} \tag{3.2}$$

The generating process of FVM difference between 1 and 2, 2 and 3, and 1 and 3 can be described by

1. Draw D from a generalized gamma distribution with parameters $a = 1.1531$, $d = 1.7766$, and $p = 2.0646$,

2. Draw z from the normalized version of the distribution specified by Equation (3.2),

3. Divide D into a difference between the FVMs of **1** and **2** and **2** and **3** by z.

The probability estimation is an integration of the probability of two "types" of the cycle with a given **1-3** FVM difference and proportions of **1-2** FVM difference and **2-3** FVM difference. The expectation of the probability of a cycle can be calculated as

$$\begin{aligned} P \;=\; \int_0^{10} f(D)\bigg(\int_0^1 g(z)\Big[&[1-U(zD)][1-U((1-z)D)]U(D) \\ &+U(zD)U((1-z)D)[1-U(D)]\Big]dz\bigg)dD \end{aligned} \tag{3.3}$$

The estimation is 0.0088 (0.88%) or one cycle every 113.6 elections. In the actual data, on the other hand, there are 0 cycles in 1,022 triples of top candidates from all surveys.

3.3.7 An extended model

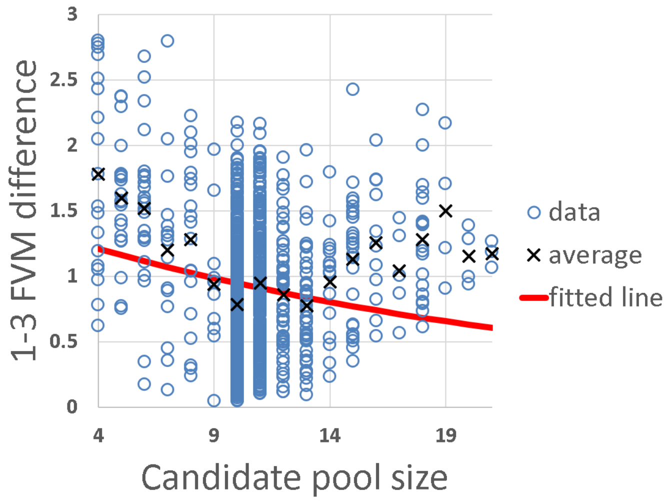
We choose only the top three candidates among the candidate pool. In each survey, the candidate pool size varies from 4 to 21. It is reasonable to expect that the average values of differences in FVMs within the top triple will shrink when the top triple is chosen from a survey that has more politicians. Intuitively, when all politicians have FVMs in the fixed interval $[-5, 5]$, having more politicians on the list can be expected, on average, to lead to a reduction in every gap between candidate to candidate, and the gap will be 0 in the limit as the number of candidates increases[6].

The data display a weak negative correlation between the number of candidates in a survey and the FVM difference between candidates **1** and **3**, or **D**. The standard deviation somewhat gets smaller (Table 3.3 and Figure 3.7).

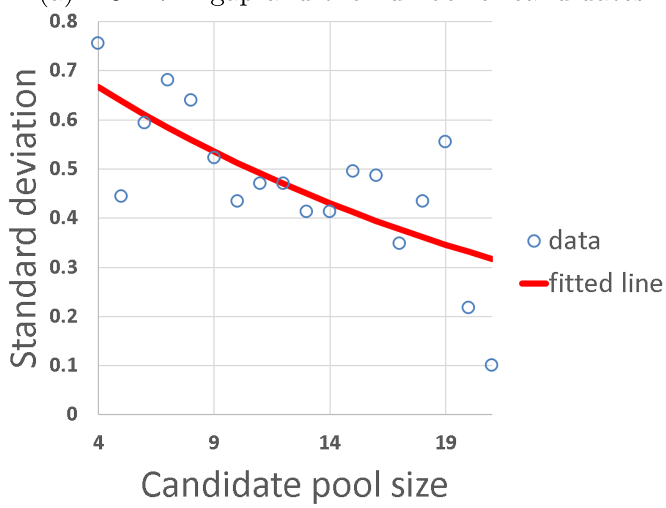
[6]We attach the table of 95% confidence interval of the gap between the first and the third order statistics, $X_{(1)} - X_{(3)}$, in Appendix

Number of candidates	Number of instances	Mean difference in 1-3 FVM	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

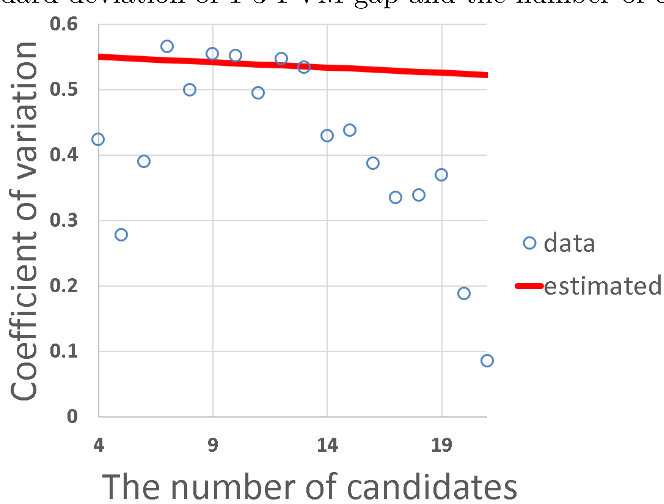
Table 3.3: Statistics of the data



(a) 1-3 FVM gap and the number of candidates



(b) Standard deviation of 1-3 FVM gap and the number of candidates



(c) Coefficient of variation and the number of candidates

Figure 3.7: Empirical distribution and estimation of 1-3 FVM gap

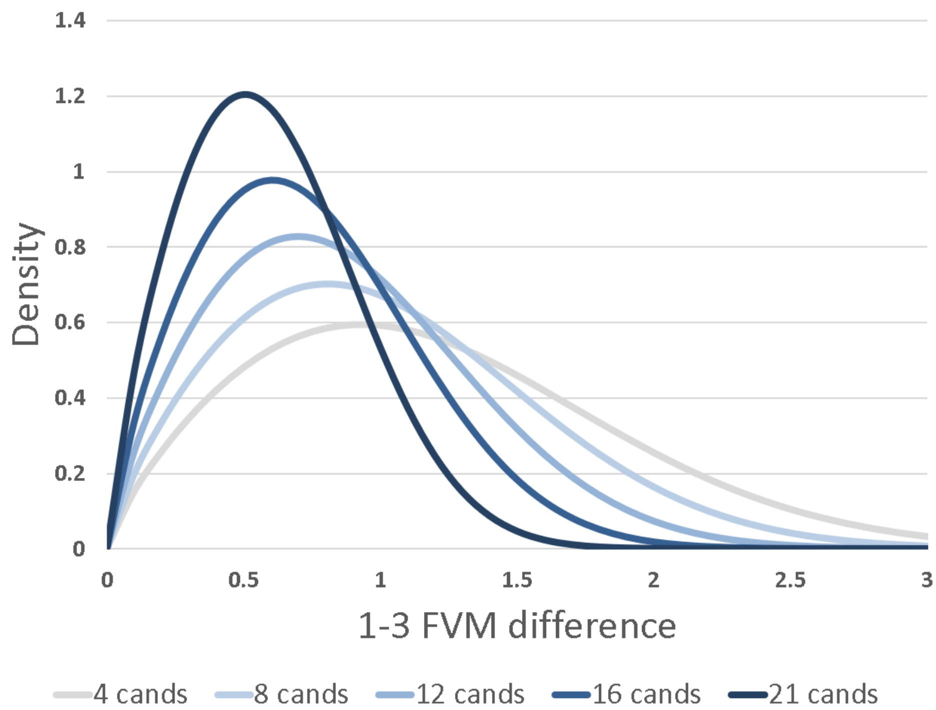


Figure 3.8: Estimated distribution of 1-3 FVM gap and the number of candidates

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, are supposed to be differed by the number of candidates.

We use the method of moments. To specify two parameters, mean and standard deviation are required; we use an exponential function and the least square method for each parameter estimation (Figure 3.7).

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

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

where μ and σ are the mean and the standard deviation of FVM difference between **1** and **3**, and 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

$$\frac{\sigma}{\mu} = \frac{\sqrt{\left(\frac{\Gamma((d+2)/p)}{\Gamma(d/p)}\right) - \left(\frac{\Gamma((d+1)/p)}{\Gamma(d/p)}\right)^2}}{\frac{\Gamma((d+1)/p)}{\Gamma(d/p)}}$$

With $\mu(C)$, $\sigma(C)$, d, and given C, the equation is solvable for p. Moreover, from the mean equation of the distribution, one could also get the parameter a, which is corresponding to C.

$$\mu = a\frac{\Gamma((d+1)/p)}{\Gamma(d/p)}$$

Let $f_C(x)$ be a generalized gamma distribution with parameters $a(C)$, $d = 1.7766$, and $p(C)$. The shape is of the same general form, but with a squeezing. The distribution is pushed toward the vertical axis as the number of candidates increases as in Figure 3.8: The model applies our assumption that differences in FVMs within the top triple tend to shrink when a survey has many candidates.

Maintaining all other assumptions from the base model, the probability of a cycle can be calculated as

$$\begin{aligned} P(C) \;=\; \int_0^{10} f_C(D)\Bigg(\int_0^1 g(z)\Big[&[1-U(zD)][1-U((1-z)D)]U(D) \\ &+U(zD)U((1-z)D)[1-U(D)]\Big]dz\Bigg)dD \end{aligned} \tag{3.4}$$

The only difference between the new Equation (3.4) and the Equation (3.3) is that the distribution of D varies by C. The probability calculation results for each number of candidates are shown in Table 3.4 and Figure 3.9.

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

Table 3.4: The number of candidates and corresponding estimated probability

As the number of candidates increases, the FVM difference among candidates decreases, and the possibility of upsets increases, so a voting cycle is more likely to exist.

3.4 Discussion

Theories and models have warned of the risk of the Condorcet paradox and have sought to estimate the frequency of cycles in actual elections. However, voting cycles rarely happen in surveys and actual elections, and our primary objective is to find a better method for bridging the gap between theoretical models and reality.

In the present chapter, We propose an approach based on the distribution of cardinal

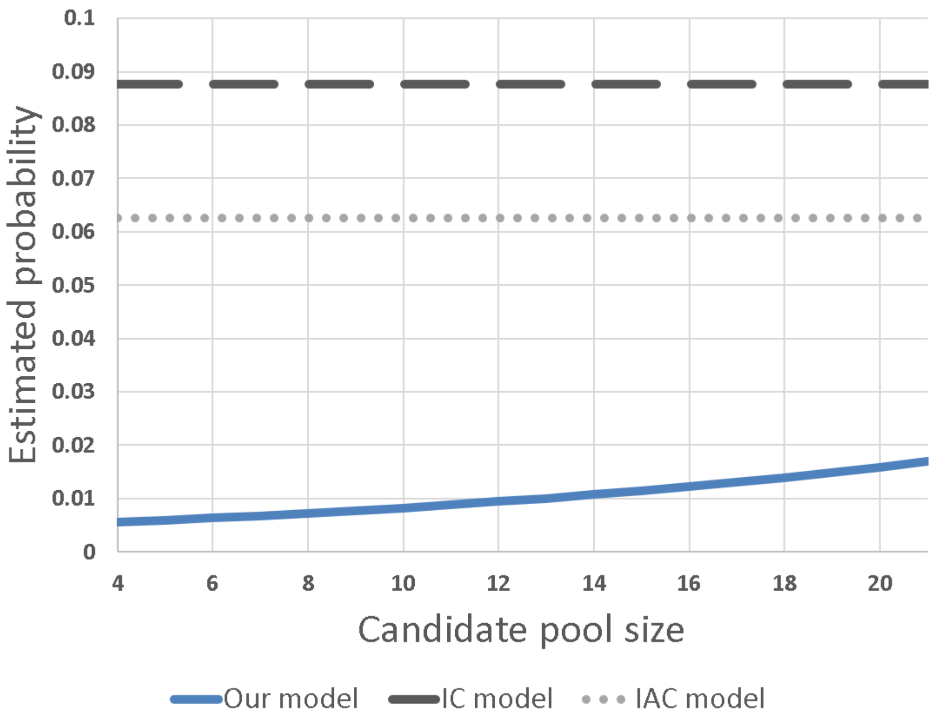


Figure 3.9: Estimated probability of a cycle

evaluation data. Using unusual techniques and data set with the cardinality of preferences distinguishes our model from others. However, because of the uniqueness of the data, it is hard to get another data set to apply our methodology and compare.

The data shows 0 cycles in 1,022 top triples and 206 cycles in all 181,579 all possible triples. On the other hand, in our base model, the estimated probability of a voting cycle among the top three candidates in a survey is 0.0088, or 0.88%. For a binomial distribution with $p = 0.0088$ with 1,022 trials, the expected value of the number of successes is 8.9836, and the standard deviation is 2.653. With the additional assumption of dependence on the number of candidates, the probability ranges from 0.0056 to 0.0169 as the number of candidates increases. The expected cycle is 8.9224, and the standard deviation is 2.9735.

Our approach is at least one step closer to an accurate model, since our results suggest a much lower probability of a Condorcet paradox than the 8.77% and 6.25% that the IC and IAC models offer. Moreover, there are several aspects of the modeling process that we have ignored, which can be examined in 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 that Politbarometer provides. Regarding the weighting factor, since we split the observations by region and questionnaire, it is tricky to adjust the factor appropriately for each set.

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 correlated from one month to the next, reducing the information of observations and increasing the standard error of the number of cycles.

Third, participants have too many candidates to assess and judge in some surveys. For instance, when a participant had 20 or 21 candidates to evaluate, he could feel fatigued and slide over some "unpopular" candidates without serious consideration and evaluation. Moreover, even if a participant has a strict preference for candidates and was willing to give a sincere response, he has only 11 numbers for evaluation and cannot describe a complete rank order. If either is the case, there will be better results when we set aside or independently analyze surveys with more than a reasonable number of candidates.

Fourth, we did not handle blanks in the survey. In pairwise comparison, we only counted the participants who evaluated both candidates and got rid of those who did not. 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. Additionally, we do not use any random process, and this will allow others to replicate our work without any differences. One way to deal with the blanks is to consider them a voter's choices under uncertainty or ambiguity. Still, it is challenging to fill in the blanks with reasonable numbers.

Chapter 4

The Frequency of Cycles and Condorcet Inconsistency under IRV:in FairVote and Politbarometer Data

4.1 Introduction

We determine the frequency of the Condorcet paradox and Condorcet inconsistency in simulated elections; US local elections and survey data. Then we estimate frequencies of the two anomalies by empirical approaches. Finding the gap between the theory and the data is our primary goal.

We focus on the Condorcet paradox and Condorcet inconsistency because we believe the Condorcet system is the best option for group decision. However, with the anomalies, the winner of the group decision is less favored than an alternative by the majority. Thus it undermines the coherence of the voting system.

Consider a three-candidate election with a Condorcet winner and Condorcet loser, and think of the vote gap between the two as an indicator how likely the voting cycle is to happen. Consider two extreme cases: the winner beats the loser in a **pairwise majority** by 51 : 49 and by 99 : 1. Only two voters need to change their ballots to make a turnover in the first case, and the voting cycle would occur. On the other hand, 50 voters are needed for the second case.

Condorcet inconsistent outcomes under IRV happen when the Condorcet winner wins the fewest votes under **plurality**. Therefore, we consider the first-choice vote gap between the Condorcet winner and a candidate with fewer first-choice votes in two non-Condorcet winners as an indicator of the likelihood of Condorcet inconsistency under IRV.

This chapter is a collection of small studies. The next section briefly provides basic information on the data used and Instant-Runoff Voting. Then, in the following subsections, we deal with: First, the frequency of a voting cycle; Second, the frequency of Condorcet inconsistency outcomes; Third, Three candidates in a one-dimensional spatial model. Finally, in Section 4.3, we summarize our work.

4.2 Analysis

4.2.1 Data Descriptions

In this chapter, we use FairVote and German Politbarometer data. For FairVote, our data consists of ranked ballots for 172 elections in cities of the USA between 2004 and 2020. For Politbarometer, we use 811 surveys from 1977 to 2017.

Compared to the previous chapter, one difference is the way of dealing with blanks. In the present chapter, leaving a question without evaluation is the case where a voter does not have enough information to judge the candidate. Additionally, voters are extraordinarily risk-averse, so they are afraid of the possibility that unknown politicians will turn out to be the worst of worst ones. Thus we consider a blank as a lower evaluation than -5, namely -6. Interestingly, this change in the data handling process makes a “big” difference.

The Condorcet winner exists in all elections of FairVote, but there is no Condorcet winner in 5 elections of Politbarometer (0.62%). Among the top three candidates (Condorcet ranking), a **voting cycle happens** only in the Politbarometer data, but just in one election (0.12%). The other four cases are tow-way ties at the top. We also count how often the IRV does NOT choose the Condorcet winner in each data set; The numbers are 1 out of 172 (0.58%) and 19 out of 806 (2.36%), respectively. If one applies simple plurality, the Condorcet winner does not win in 9 out of 172 (5.23%) and 75 out of 806 (9.31%) times in the two data sets.

For the cases that are impossible to define the top three candidates because of the cycles or ties, break the cycle and tie by breaking the weakest link (comply on Ranked Pairs, Tideman, 1987) and pick one who is winning in their head-to-head election. Then, we refer to the first candidates from above as the Condorcet winner for later on.

In some elections of FairVote, the top three candidates are not similarly competitive. For example, in an election for the city council member, district1, San Leandro, in 2010, the third candidate was ranked by only 0.757% of the total votes. In contrast, the first and second candidates were ranked by 84.088% and 72.278% of voters. A couple of elections had one dominating candidate and two "weak" candidates. In an election for the mayor of San Francisco in 2007, three top candidates were ranked by 89.312%, 20.113%, and 16.339% of voters, respectively. They are not suitable for our research; we exclude eight elections from our observations of FairVote.

4.2.2 Instant-Runoff Voting (IRV)

Instant-Runoff Voting (IRV) is a voting systems in which voters cast a vote with their ranking for some or all candidates, also called Ranked-Choice Voting. The detailed process for counting the votes is as follows.

1. Count the first-choice votes and find a candidate who wins the most.
2. If the candidate wins more than a half of the total votes, then the candidate is the winner, and the process is done.
3. Otherwise, eliminate a candidate with the fewest first-choice votes. In each ballot, if the eliminated candidate is in the ranking, the ranking of the ballot will be re-organized without the candidate, as considering other lower-ranked candidates will be ranked up by one.
4. Repeat the process until one ends up with the winner of a majority.

Under the plurality rule, especially under a two-party system, voters who prefer a third option are incentivized to choose the lesser evil of the two huge parties rather than revealing their honest first pick. Additionally, when the third party is politically close or similar to one of the two dominant parties, the third party might "steal" enough votes to affect the winner between the two parties. By its design, IRV is supposed to make up for those shortcomings, but other new, different disadvantages pop up, such as and monotonicity failure.

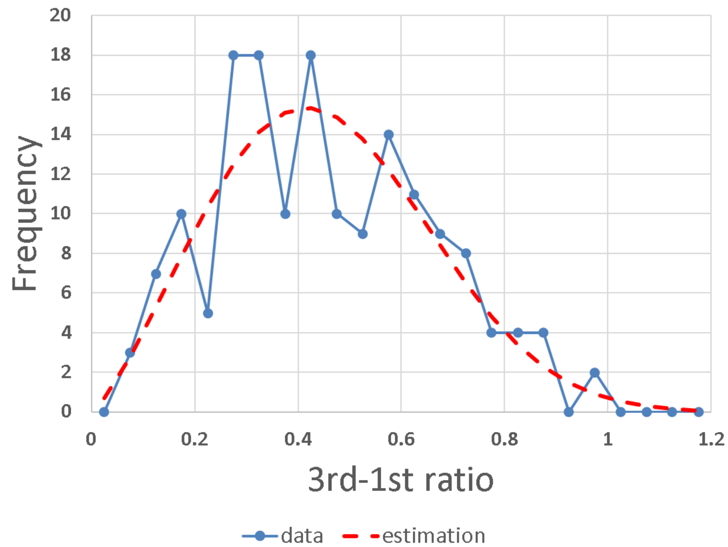
4.2.3 Estimating the Probability of the Top Cycle

We count the votes that the first candidate (the Condorcet winner) wins and the votes that the third candidate wins in their head-to-head election and define the *3rd-1st ratio*, which is the ratio of the latter to the former. If the ratio gets closer to 1, only a few votes are needed by the third candidate to beat the first candidate and closer to having a voting cycle.

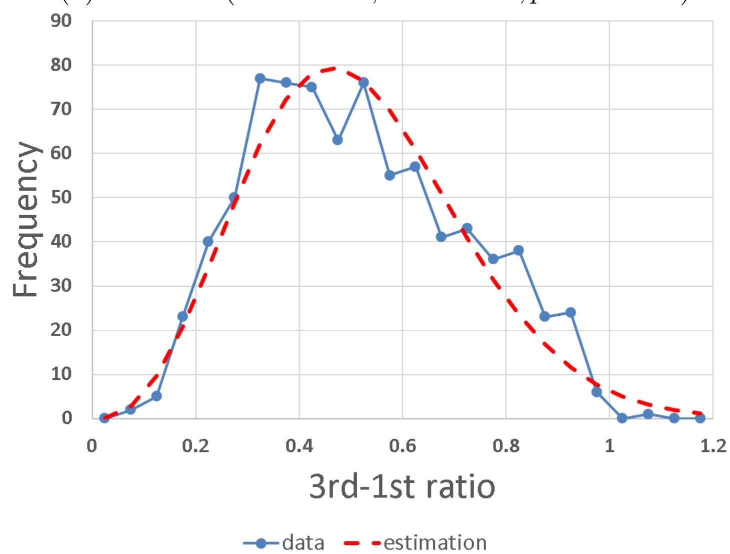
Based on the shape of the empirical distribution of 3rd-1st ratio from the two data sets, we decided to use a generalized (three-parameter) gamma distribution and take three parameters by maximum likelihood estimation to fit the data. The density function of the distribution is as follows.

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

With the estimated distribution, the ratio will be greater than 1 with the probability of 0.0068 and 0.0156 in the two data sets, respectively. As we mentioned above, on the other hand, the data show 0 and 1 out of 806 (0.0012).



(a) FairVote ($a = 0.5463, d = 2.2817, p = 2.56659$)



(b) Politbarometer ($a = 0.4058, d = 3.5155, p = 1.9438$)

Figure 4.1: Empirical distribution and estimation of 3rd-1st ratio

The 3rd-1st ratio indicates how competitive the third candidate is compared to the Condorcet winner. Since having more candidates in the election means probabilistically having a more competitive third candidate, we expect increases in the 3rd-1st ratio (equivalently, decreasing the vote gaps) as the number of candidates increases. While observing the FairVote data, we find that the mean of the 3rd-1st ratio increases gradually as the number of candidates increases, but the variance does not change significantly. Hence, we suppose the distribution of the 3rd-1st ratio has a fixed shape, but the support shifts with the number of candidates.

We use the four-parameter beta distribution for convenience. Because the distribution could be fitted with various shapes and support, we get a good approximation of parameters by the method of moments. The distribution has two shape parameters, (α, β), and the other two, (L, U), for the lower and upper bound. The probability density function is as follows.

$$f(x;\alpha,\beta,L,U) = \frac{1}{(U-L)^{\alpha+\beta-1}} \frac{(x-L)^{\alpha-1}(U-x)^{\beta-1}}{\mathbf{B}(\alpha,\beta)}$$

where

$$\mathbf{B}(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

In our model, we need to define L (or U) as a function of the number of candidates. To capture the diminishing shifting of the support, we use the logarithm function.

$$L(N;A,C,D) = A + C \cdot \ln\left(\frac{N+D}{N_0+D}\right)$$

or

$$U(N;A,B,C,D) = A + B + C \cdot \ln\left(\frac{N+D}{N_0+D}\right)$$

where N is the number of candidates, N_0 is the minimum number of candidates in the data set, A is the initial lower bound, B is the support length, and C and D are constant terms that capture support shifting.

By maximum likelihood estimation, we take six parameters, $(\alpha, \beta, A, B, C, D)$. The sets of parameters we got are $(1.537, 2.744, 0.050, 0.933, 0.678, -2.344)$ and $(2.263, 3.065, 0.044, 1.022, 0.002, -3.999)$ for the data sets.

A Four-parameter beta distribution with a support-shifting assumption provides a good fit for FairVote data, but it barely fits Politbarometer data.

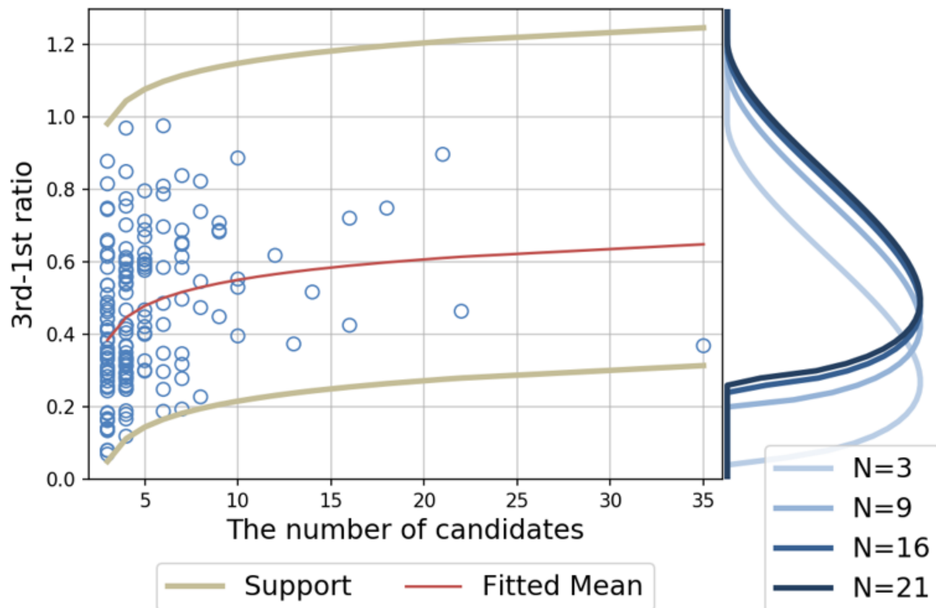


Figure 4.2: Scatter plot of 3rd-1st ratio and estimated marginal distributions, FairVote

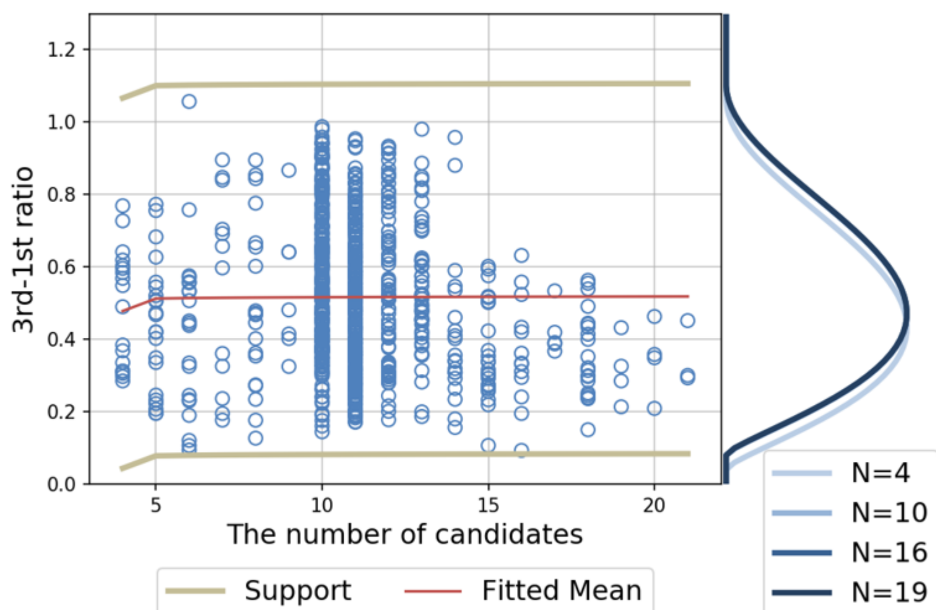


Figure 4.3: Scatter plot of 3rd-1st ratio and estimated marginal distributions, Politbarometer

By integrating the probability density of the 3rd-1st ratio greater than 1, we have probability calculations for a voting cycle for each number of candidates. As shown in Table 4.1 and 4.3, the numbers are smaller than models under Impartial Culture (IC) and Impartial Anonymous Culture (IAC), where, for three-candidate elections, the numbers are 0.0877 and 0.0625, respectively (Gehrlein, 2002, 2006).

N	FairVote	Polit-barometer		N	FairVote	Polit-barometer
3	0	(N/A)		13	0.0196	0.0046
4	0.0005	0.0012		14	0.0215	0.0046
5	0.0023	0.0040		15	0.0233	0.0046
6	0.0045	0.0042		16	0.0251	0.0047
7	0.0068	0.0043		17	0.0269	0.0047
8	0.0091	0.0044		18	0.0285	0.0047
9	0.0113	0.0044		19	0.0302	0.0047
10	0.0135	0.0045		20	0.0317	0.0047
11	0.0156	0.0045		21	0.0333	0.0048
12	0.0176	0.0046				

Table 4.1: Estimated probability of the cycle by the number of the candidates

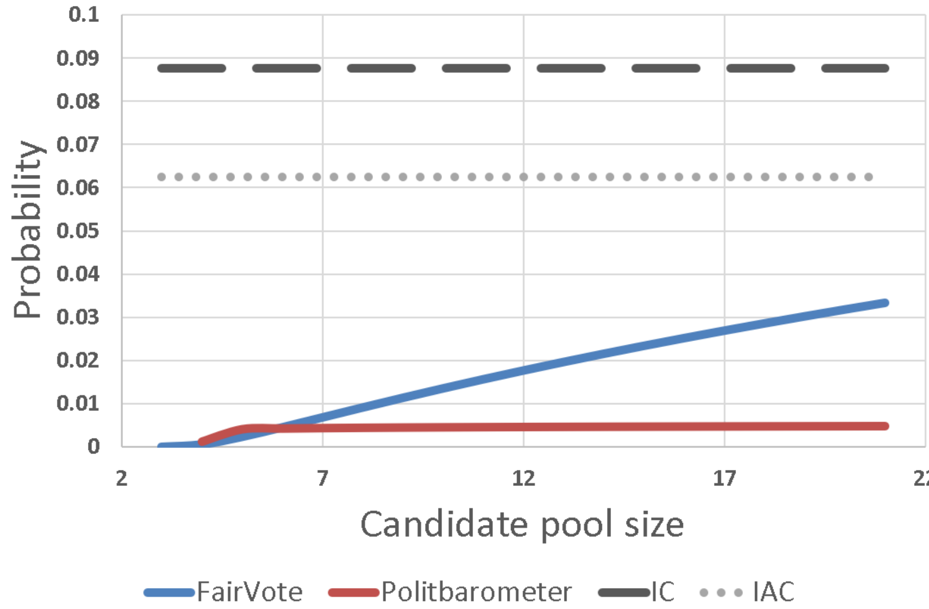


Figure 4.4: Estimated probability of a voting cycle

4.2.4 Estimating the Frequency of Condorcet Inconsistent Outcomes under IRV

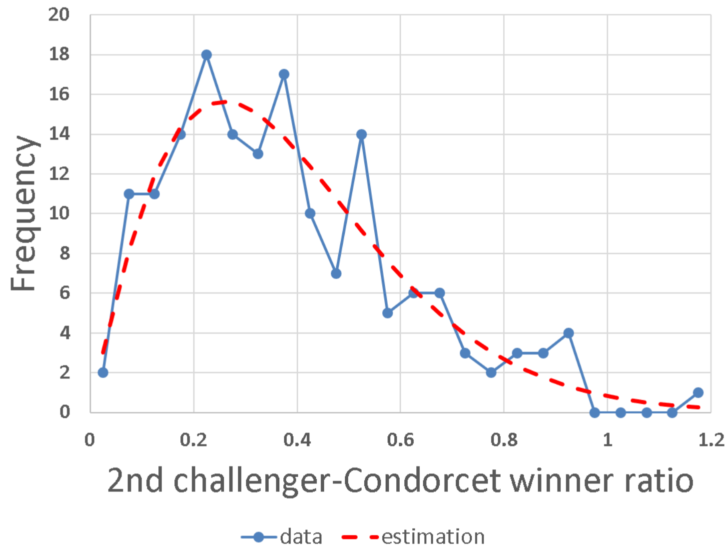
IRV is a voting system that does not satisfy the Condorcet criterion. In this subsection, we suggest a way of predicting the frequency of the Condorcet inconsistent outcomes, based on an empirical distribution.

The Condorcet winner differs from the IRV winner in only one case out of 164 elections (0.61%) from FairVote (mayoral election of Burlington in 2009) and for Politbarometer data in 19 cases out of 806 elections (2.36%). Eighteen cases (1 from FairVote 17 from Politbarometer) are ones where the IRV winner is ranked the second placement by the Condorcet method, and the Condorcet winner is ranked the third placement by IRV.

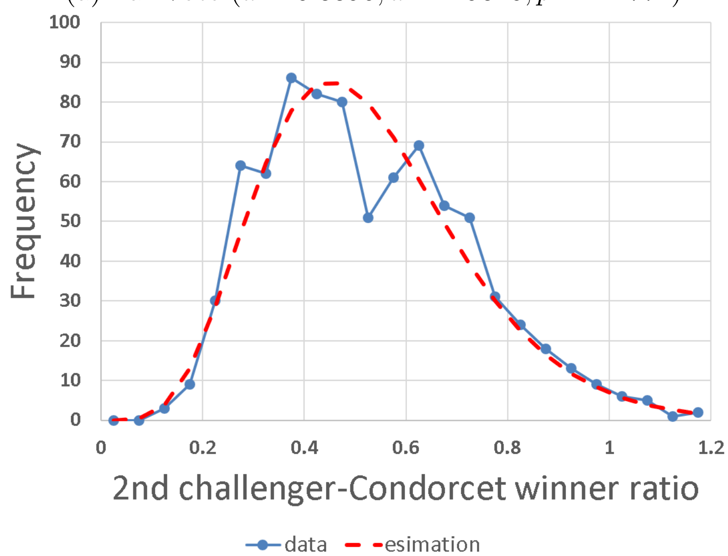
In the Politbarometer survey of April 1987, the IRV winner is ranked in second place by the Condorcet method, as formerly, but the Condorcet winner becomes the fourth candidate by IRV. In the survey of June 1996, the Condorcet winner is ranked third by IRV, and three candidates (including the IRV winner) are in a cycle for second, third, and fourth place by the Condorcet method.

To have a Condorcet consistent outcome under IRV, the Condorcet winner must not be the plurality loser. Hence, to estimate the susceptibility of IRV to Condorcet inconsistency, we count the votes that the Condorcet winner wins and the second challenger wins in the plurality voting among the top three candidates. Then we define *2nd challenger-Condorcet winner ratio*, the ratio of the first-choice votes the second challenger votes to the Condorcet winner votes. The second challenger is the candidate who wins fewer votes between the two candidates other than the Condorcet winner. When the ratio is close to 1, fewer votes are needed to defeat the Condorcet winner by IRV. When it is greater than 1, the Condorcet winner will be eliminated under IRV with the three candidates, or on the round with the three candidates remaining.

We deploy a generalized (three-parameter) gamma distribution and take three parameters by maximum likelihood estimation to fit the data.



(a) FairVote ($a = 0.3390, d = 1.9840, p = 1.4772$)



(b) Politbarometer ($a = 0.1170, d = 6.1376, p = 1.1247$)

Figure 4.5: Empirical distribution and estimation of 2nd challenger-Condorcet winner ratio

With the estimated distribution, the ratio will be greater than 1 with the probability of 0.0213 and 0.0223 in the two data sets, respectively. On the other hand, the two data sets show 1 out of 164 (0.0061) and 19 out of 806 (0.02357). Merrill III (1984) is comparable to our result. That study checks Condorcet consistency under various candidates, voters, and voting rules settings. Especially for the runoff rule, there are Condorcet inconsistent outcomes in 3.8% of 10,000 elections. The simulation is based on the spatial model, with "similar" three candidates and 25 voters. He notes, however, that the inconsistency rate goes up as the number of voters increases.

Additionally, like in the previous subsection, we apply an assumption to the base model that a bigger the candidate pool likely to yield more competitive contenders. The data fitting and estimated probability of extended model from the two data sets are following.

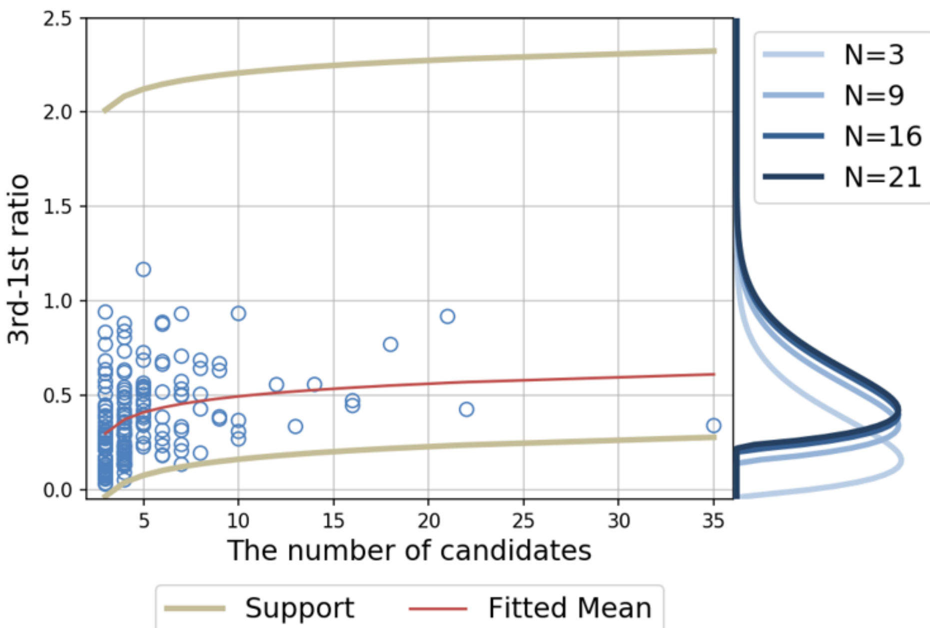


Figure 4.6: Scatter plot of 2nd challenger-Condorcet winner ratio and estimated marginal distributions, FairVote

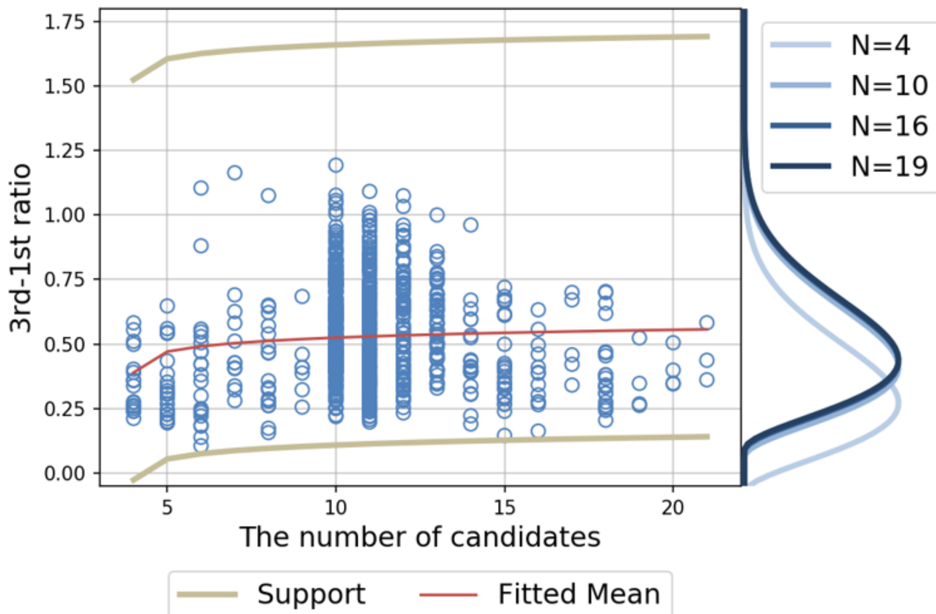


Figure 4.7: Scatter plot of 2nd challenger-Condorcet winner ratio and estimated marginal distributions, Politbarometer

N	FairVote	Polit-barometer		N	FairVote	Polit-barometer
3	0.0054	(N/A)		13	0.0309	0.0182
4	0.0099	0.0030		14	0.0325	0.0188
5	0.0135	0.0086		15	0.0340	0.0194
6	0.0164	0.0110		16	0.0355	0.0199
7	0.0190	0.0126		17	0.0369	0.0204
8	0.0214	0.0139		18	0.0383	0.0209
9	0.0236	0.0150		19	0.0396	0.0214
10	0.0256	0.0159		20	0.0408	0.0218
11	0.0274	0.0167		21	0.0421	0.0222
12	0.0292	0.0175				

Table 4.2: Estimated probability of Condorcet inconsistent outcome by the number of the candidates

4.2.5 One-dimensional Spatial Model for Three Candidates: Empirical Evidence

Spatial Model for Three Candidates in One Dimension

Suppose three candidates, each positioned on the line segment $[0, 1]$, depending on political attribute. The voters' ideal political attributes are uniformly distributed on the line segment, and each voter ranks the candidate by closeness to their ideal. The setting is intuitive and straightforward, thus widely used. However, when we count the preference order of voters, the model reveals a flaw; The model is only capable of accommodating four preference orders.

Thus, as shown in the figure below, when candidate B locates between **A** and **C**, the voters who have preference order **ACB** and **CAB** cannot be represented with the model. However, the model will fit data perfectly fit if all voters have single-peaked preferences.

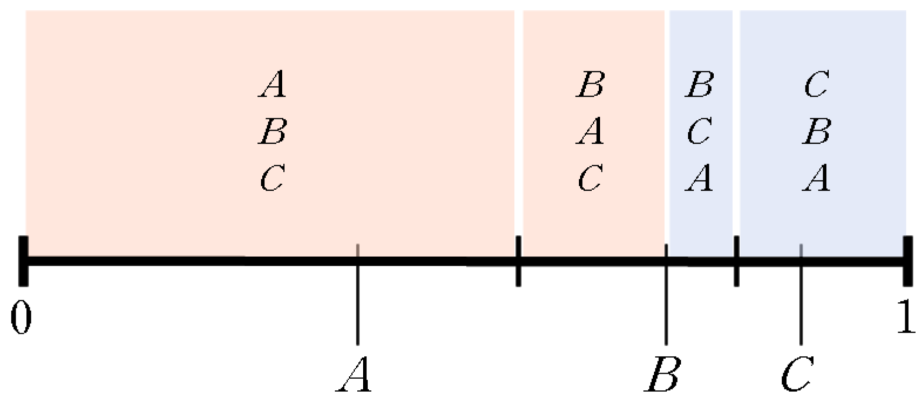


Figure 4.8: One-dimensional spatial model and preference order representation

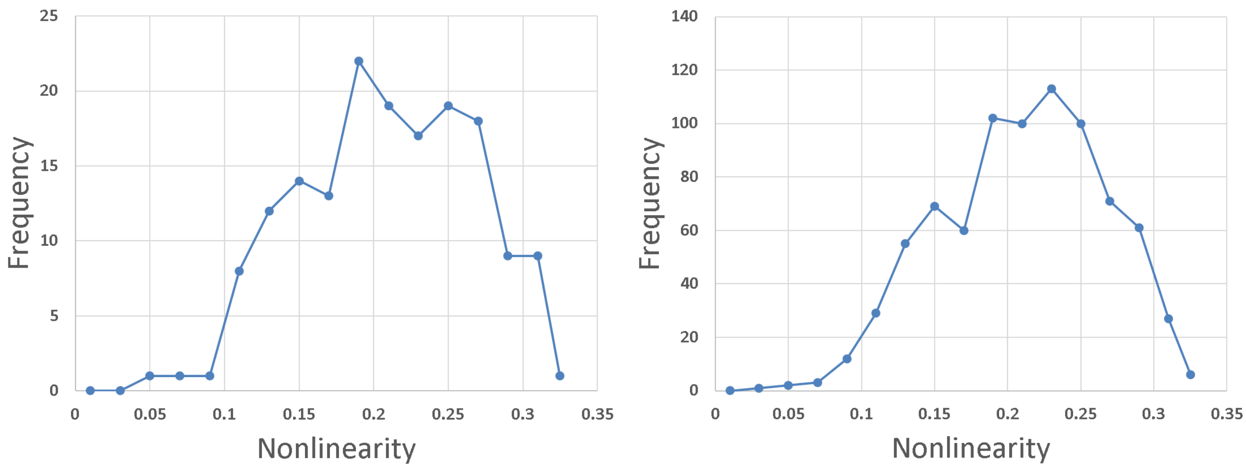
Estimation and Result

Using the data, we determined the preference order of each voter and counted the numbers for each type. Then, by checking the share of each preference order, we see how well a one-dimensional spatial model reflects the data.

If an election with three candidates can be described perfectly by a one-dimensional model, then one (and only one) candidate has no third-choice votes, like candidate **B** in the previous example. Therefore, voters' least preferred candidates matter, so we consider only the votes with a fully specified preference for the top three candidates. Next, we count the third-choice votes for each candidate and define *nonlinearity*, the smallest share of third-choice votes, to measure how close the data is to the one-dimension spatial model. The nonlinearity is necessarily less than 1/3, and closeness to 0 means a one-dimensional spatial model well captures it.

The minimum of the smallest share of third-choice is 0.049 in FairVote and 0.038 in

Politbarometer. 98.17% and 97.78% of the two data sets range from 0.1 to 0.32, and thus we claim that the one-dimensional spatial model does not fit our data well. Thus, using the one-dimension spatial model cannot result in reasonable estimations. Another finding is that distributions of nonlinearity from two data sets seem not significantly different from each other. See Figure 4.7.



(a) FairVote (b) Politbarometer

Figure 4.9: Empirical distribution of nonlinearity

4.2.6 Nonlinearity and Failure of the Condorcet Criterion under IRV

As the figure previously described, the candidate positioned between two other candidates or who gets the least third-choice vote is not necessarily the Condorcet winner and vice versa. However, according to the data we use, the Condorcet winner is the one who has the fewest third-choice votes in more than 85% of elections (FairVote:142 out of 164, Politbarometer:692 out of 811).

We use nonlinearity and 2nd challenger-Condorcet winner ratio from earlier sections, and Figure 4.8 and Table 4.3 show a non-linear positive relationship between the two. However,

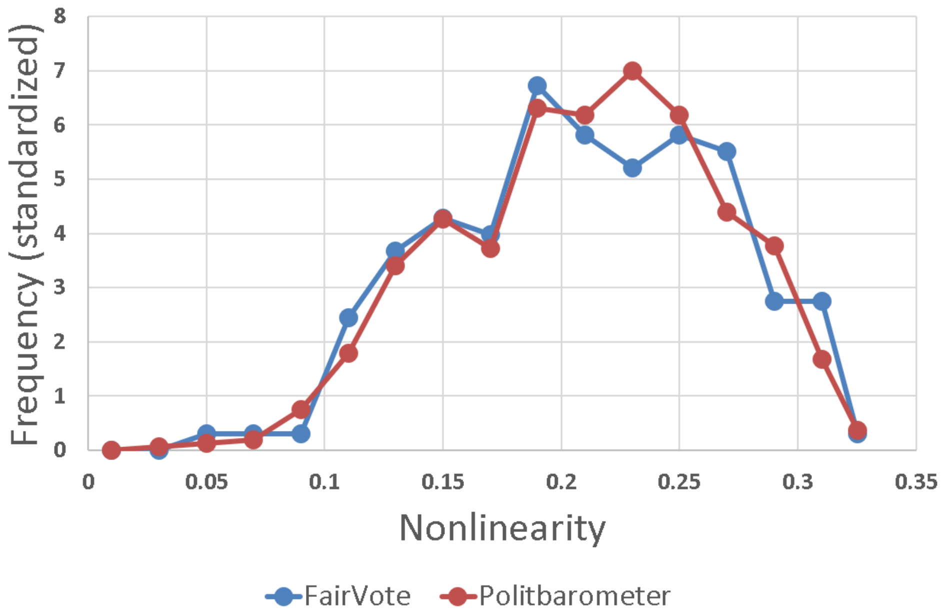
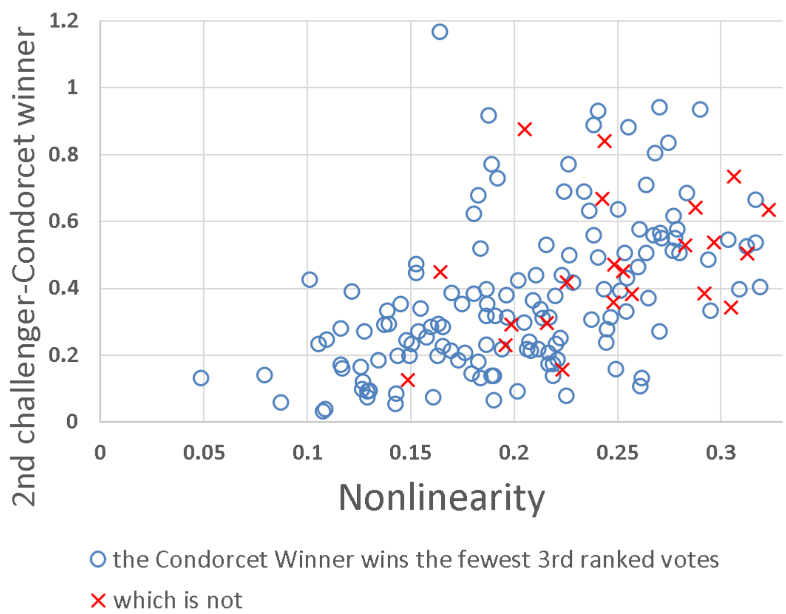
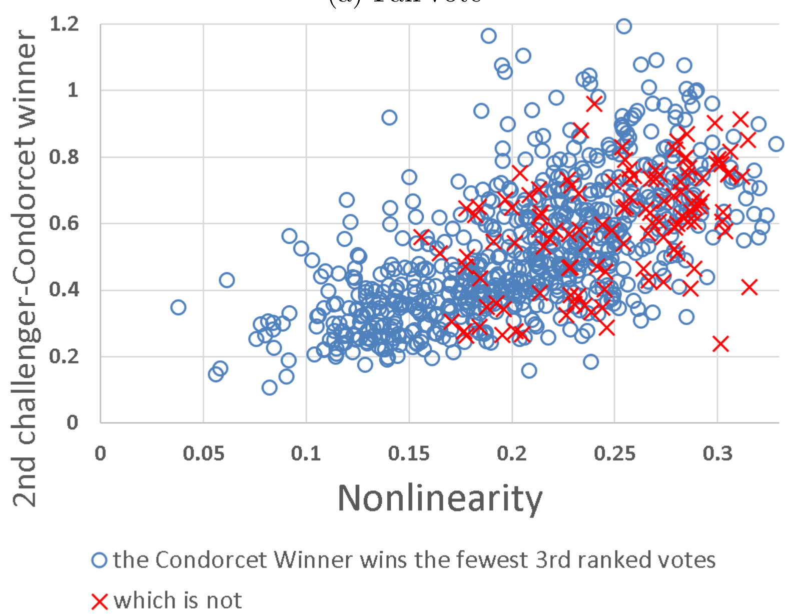


Figure 4.10: Standardized distribution of nonlinearity

we reach no conclusion as to which one causes the other, or if they move together as a result of some other causes.



(a) FairVote



(b) Politbarometer

Figure 4.11: Nonlinearity and Condorcet inconsistency in each data set

When we adopt the idea of the spatial model for three candidates in one dimension, the Condorcet winner is the candidate located closest to the median voter. If the Condorcet winner loses in the first round of IRV (or plurality rule) it happens because he is squeezed

	FairVote	Politbarometer
The Condorcet winner wins the fewest 3rd-choice votes	0.6164	0.6821
Otherwise	0.6633	0.5287
All	0.6268	0.6696

Table 4.3: Linear correlations in each data set

from two different sides by the other two candidates.

For the one-dimension spatial model, the positions of three candidates for a Condorcet inconsistent outcome is relatively simple. Put one candidate, the Condorcet winner, close to the median voter, another candidate "close enough" to the left, and the other candidate "close enough" to the right of the Condorcet winner.

However, the conditions will be more complicated for the two-dimension setting than in one-dimension. Thus, it seems to make sense that when election data fit well with a one-dimensional model, it is more likely that Condorcet inconsistency occurs. Nevertheless, our data shows the opposite. There could be another driving force that is counter. **The Condorcet winner has the fewest third-choice votes in the data** mean the Condorcet winner has many first and second-choice votes, and the probability of being beaten by the second challenger is lower. In other words, two other candidates are not located close enough to squeeze the Condorcet winner.

Let us consider this from a different point of view. For convenience sake, assume two-dimensional attribute space, voters' ideal point distribution follows a bivariate normal distribution, and there are three points describing candidates' attributes. A lower 2nd challenger-Condorcet winner ratio means the second challenger wins relatively fewer votes.

Geometrically, this implies the second challenger is located further from the mode of bivariate normal distribution. For example, consider the triangle formed by the three candidates' locations and stretch one vertex to the other two vertexes in the opposite direction.

The new triangle would be flatter and lower compared to the original one. Therefore, the lower 2nd challenger-Condorcet Winner ratio means that the three points of candidates are located in more "linear" positions.

4.3 Discussion

Although there are several defects of the plurality rule, it is the most or the second most widely used system in history. While many other systems have been suggested, it is still controversial which voting rule should be the next standard. IRV seems to be a strong contender, and it is already used in many elections. The anomaly of Condorcet inconsistency exists, based on mathematics and theoretical modeling, and the predicted possibility is so high that it cannot be ignored. However, real-world data says the frequency is much lower than the expectation.

With actual election data under IRV and surveys with scoring data, we check the frequency of the voting cycle and Condorcet inconsistent outcomes. We introduce empirical-based modeling using vote gaps in three top candidates as indicators of measuring how the voting profile is close to each anomaly and find any statistical regularities of the factors' empirical distributions.

From 164 simulated elections from FairVote, we have zero voting cycle and one Condorcet inconsistent outcome under IRV (0.0061). Our base model predicts 0.0068 and 0.0147 for each anomaly. On the other hand, from 811 simulated elections from Politbarometer, there exist one voting cycle (0.0061) and 19 Condorcet inconsistent outcome under IRV (0.0236, 5 elections are excluded because the Condorcet winner does not exist). Our model yields 0.0156 and 0.0211. The predictions are not perfect but it is closer to the actual data.

Moreover, we offer empirical evidence that the one-dimensional spatial model, which is easy to understand and set, does not depict actual data well. We also discovered that the one-dimensional setting is less capable when the outcome is Condorcet inconsistent or closer to it. It is different from our conjecture, so this would be our next project.

Analyzing another paradox that IRV is vulnerable to is also intriguing. The monotonicity failure is related to our research: upward monotonicity failures can occur in IRV only when there is a voting cycle or Condorcet inconsistency (Ornstein and Norman, 2014). Miller (2017) specifies the necessary and sufficient conditions of voting profile for monotonicity failure in IRV with three candidates. Since the anomaly is counter-intuitive and regarded as a serious one, many studies deal with it (Lepelley et al., 1996; Felsenthal and Tideman, 2013, 2014; Plassmann and Tideman, 2014; Graham-Squire and Zayatz, 2021). Like voting cycles and Condorcet inconsistency, frequency of monotonicity failure should also be not very rare according to theoretical models, but some argue that it does not occur in actual elections and is exaggerated. When we combine these conditions with the approaches used in this chapter, developing a model for monotonicity failure seems possible.

Chapter 5

The Normal Spatial Model for Four Candidates in Three Dimensions: Parameterization and Goodness-of-Fit

5.1 Introduction

While many diverse voting systems have been proposed, researchers have been trying to identify the best system. Many studies have suggested a variety of criteria to evaluate voting systems. Most criteria are most relevant when there are more than two options, and they require voters' rankings of the candidates. However, because of the paucity of such data in practice, it is difficult for any opinion to be supported by empirical evidence, and researchers have been forced to use statistical modeling that assumes a distribution of preference ordering.

Widely used preference distribution assumptions are Impartial Culture (IC) and Impartial Anonymous Culture (IAC). IC assumes that each voter can have each of the possible strict preference orderings with equal probability, so all possible preference schedules will have the same probability of realization. In the IAC model, one does not keep track of voters' names: preference schedules with the same elements are treated as the same outcome. Then IAC assigns equal probability to each *anonymous equivalence class* realization.

The normal spatial model (Good and Tideman, 1976) has a somewhat different approach than IC and IAC. Hotelling (1990) and Downs et al. (1957) initiate and develop the idea of a spatial model of politics. The Hotelling-Downs model assumes that voters' ideal points are distributed in a one-dimensional policy space, the voters prefer the candidate whose political attribute is closest to their ideal, and two or more candidates choose their locations to win the most votes.

Simulations based on the spatial model forr three-candidate in two-dimensional space imitate actual data better than other probabilistic models (Plassmann and Tideman, 2012). We are interested in whether a more generalized version also works well, namely four candidates' political attributes and voters' ideal points distributed in a three-dimensional space.

This study is meant neither to introduce a new theory nor to address empirical work such as the frequency of paradoxes. Instead, we focus on the middle process: measuring how well a model fit a data. We also show that simulated data with the IAC assumption is statistically far from actual voting outcomes. We assert the significance of our work and differentiation from previous studies in to two aspects.

First, we adds fourth candidate. Previous studies mainly consider three-candidate cases. Presuming three candidates is realistic enough and may be fair for analyzing most of the voting criteria and paradoxes, including the infamous Condorcet paradox. However, there is a point to analyzing four-candidate cases. De Sinopoli et al. (2006) examine four-candidate elections under approval voting and provided fairly engaging examples: the Condorcet winner can receive NO votes, and strategic stability does not imply sincerity. Niemi and Riker (1976) note that "And a third-place or even a fourth-candidate could exercise considerable influence over the choice of the president, as Henry Clay did in 1824 and George Wallace hoped to do in 1968." Kurrild-Klitgaard (2001) mentions that the introduction of the fourth candidate could make a significantly different outcome in an actual election: Prime ministerial candidates in the Danish election of 1994.

Second, we set the model with three dimensions. Good and Tideman (1976) provided all theoretical backgrounds of the normal spatial model for the general n-candidate case. We do give expression to the theory in the data fitting process. However, as there is a big jump from a one-dimensional to a two-dimensional model, dealing with a three-dimensional model is yet more elaborate. We need to estimate eight parameters to define the locations of candidates relative to the distribution of voters', and calculating the probability density is another problem: With three-candidate cases in a two-dimensional setting, estimating parameters involves only calculating integral of the density function of the bivariate normal distribution on six sectors. In a four-candidate and three-dimensional setting, one needs to calculate the density of the trivariate normal distribution on twenty-four partitions defined by six bisecting planes, each of which bisects the line segment between the locations of two candidates.

The rest of the chapter will be organized as follows: In Section 5.2, we discuss relevant literature and point out the main differences between previous work and our work. In Section 5.3, we briefly explain how we handle data, describe the framework step by step, and show statistical evidence of the invalidity of IAC simulation. In the last section, Section 5.4, we summarize our work.

5.2 Literature Review

Chamberlin and Cohen (1978), Merrill III (1984), Green-Armytage (2014), and Green-Armytage et al. (2016) assume multiple candidates and multiple dimensions, [1] then measure various things under different voting rules. These studies simulate the model by randomly

[1] Chamberlin and Cohen (1978) assume four candidates in four dimensions, Merrill III (1984) assumes 3, 4, 5, 7 candidates in 1∼4 dimensions, Green-Armytage (2014) assumes 3∼6 candidates in 1,2,4,8,16 dimensions, and Green-Armytage et al. (2016) assume three candidates in 1 and 8 dimensions.

drawing the candidates' and voters' locations from a multivariate normal distribution. On the other hand, we focus more on the parameterization of four-candidate election data to fit the spatial model. Moreover, by fitting the actual data and simulated data under the IAC assumption, we measure the gap between the real data and the IAC model.

Plassmann and Tideman (2012), Tideman and Plassmann (2012), and Plassmann and Tideman (2014) are more aligned with our work. Plassmann and Tideman (2012) and Tideman and Plassmann (2012) offer statistical evidence that the simulated preference profiles under the normal spatial model are very close to actual data with three candidates in a two-dimensional setting. With simulations of the distribution of voters' preferences under the spatial model for three candidates in two dimensions, Plassmann and Tideman (2014) evaluate various voting rules, check criteria, and the frequency of the Condorcet paradox with three-candidate cases.

Matje (2016) also deals with the spatial model. He demonstrates by a statistical approach that the spatial model in a two-dimensional setting is invalid for four-candidate cases. The main argument is that only 18 strict preference orders can be represented in a two-dimensional setting, which does not correspond to real preference distributions.

5.3 Analysis

5.3.1 Data Description

In this chapter, we use German Politbarometer from March 1977 to December 2017. We use 811 surveys by month and region (east/west). We analyze four-candidate elections, and we want to avoid dealing with blanks in evaluations as much as possible, so we use the most evaluated four candidates from each survey.

We derive voters' preferences for four candidates based on their evaluations. If a voter's evaluations of two or more candidates are the same, then we refer the voter is indifferent among them. We exclude voters who assign the same score to all four candidates.

For convenience, we only model strict preference orders and assume that voters who do not have strict preference orders would make up their minds relying on how others do. Specifically, for four candidates, **A**, **B**, **C** and **D**, if **A** and **B** are preferred by 6 : 4 by decisive voters(voters who have strict preference order on A and B), we do binomial sampling with a probability of 0.6 to assign voters of (**AB**)**CD** to **ABCD** and **BACD**.

5.3.2 Creating Data under IAC Assumption

To check how IAC assumption is close to the real data, we simulate voting outcomes based on IAC assumption. We follow a simulation method Lepelley et al. (2000) suggest, with omitting the detail (see e.g. Hammersly and Handscomb, 1964). The method is applying Monte-Carlo method for the IAC model: Relative frequencies of voters of each 24 strict preference order "are considered as realizations of a uniform random variable on [0,1]." We draw 23 numbers from uniform distribution on $[0, 1]$. The drawn 23 numbers partition the support to 24 parts, and the each part's length corresponds to the share of voters for each strict preference order. We draw repeatedly to simulate 811 voting outcomes For each simulation, we apply the number of voters of surveys in order. For instance, for the first draw, we multiply 860 to 24 lengths which is the number of valid voters of the first survey, March 1977. Basically, the result values of multiplication are most likely non-integer. Thus the numbers are rounded, so the total number of voters in simulated data could be different from that of the survey by a few. In our simulations, the maximum gap is four, and the sum of the absolute values of gaps is 884.

5.3.3 Fitting the Data to the Normal Spatial Model

At least three dimensions are required for four-candidate cases (Good and Tideman, 1976; Matje, 2016). Thus, we set three dimensions with normally distributed voters' ideal points and four points for candidates located by their (political) attributes. The candidates' locations are assumed to be commonly known to the voters, who prefer closer candidates by a (Euclidean) distance from their ideal.

From the four points of candidates, by connecting them, we have a tetrahedron. There are six bisecting planes, one for each pair of points, and those planes all pass through one point, the circumcenter of the tetrahedron. The planes cut the space into 24 partitions, and the ideal points of voters located in each partition correspond to voters having one of the possible 24 strict preference orders.

To define the spatial model for four candidates in three-dimensional space, one needs to define five points: four for candidates' attributes, and the other for the mode of the trivariate normal distribution. Instead of 15 numbers, eight parameters are sufficient to define a model. Thus, fitting the data to the spatial model is finding the best eight parameters that induce the closest 24 numbers to the actual voters' share of 24 preference orderings. We use maximum likelihood estimation to find the best estimates.

To calculate the likelihood, given eight parameters, once construct the six bisecting planes, determines the mode of the trivariate normal distribution, and measures the integral of the trivariate normal distribution on the 24 three-dimensional polygons, but the processes are costly. Because the function is a mixture of complicated functions such as trigonometric, min, and chi-square density, it does not have a closed form. Thus, we use the numerical optimization method instead of solving the maximization problem analytically. For the same reason, numerical differentiation is used to get first-order, second-order, and cross derivatives for the optimization method. We employ the gradient method and the Gauss-Newton method.

5.3.4 Comparison of the real data and the simulated data

We offer a partial answer to whether the spatial model captures the actual data better than other stochastic preference assumptions for more than three candidates cases: We compare the real data and simulated data under the IAC assumption by the spatial model for four candidates.

The IAC assumption covers all possible voting situations with uniform probability. If the real preference distributions have specific regularities or tendencies, some of the (technically) possible voting situations would happen with very low probability or not happen. In other words, if one simulates preference distributions under IAC assumptions, some of them may not be matched to the real cases. On the other hand, if two data sets are sampled from an identical population, they should be similar statistically. We compare the real and simulated data by comparing the gap between each and the fitted probabilities by the spatial model.

Matje (2016) uses the likelihood ratio to compare the actual outcomes, simulations under the IAC, and simulations under a modified version of the spatial model. The result shows a significant difference in distributions of the likelihood ratio of IAC and the actual data. On the other hand, we will measure the distance between two sets of 24 numbers for each data set.

$$d = \sqrt{\sum_{i=1}^{24}(q_i - p_i)^2}$$

where $(q_1, q_2, \cdots, q_{24})$, is the share of voters for each preference order in the data. Thus, each q_i is calculated by (the number of voters for corresponding preference order/total number of voters). $(p_1, p_2, \cdots, p_{24})$ is generated probabilities by the spatial model: We fit a spatial model onto data, estimate eight parameters define 24 partitions in three dimensions, then p_i is the density on the corresponding partition.

Figure 5.1 shows two very distinguished distributions of each data. Therefore, we could

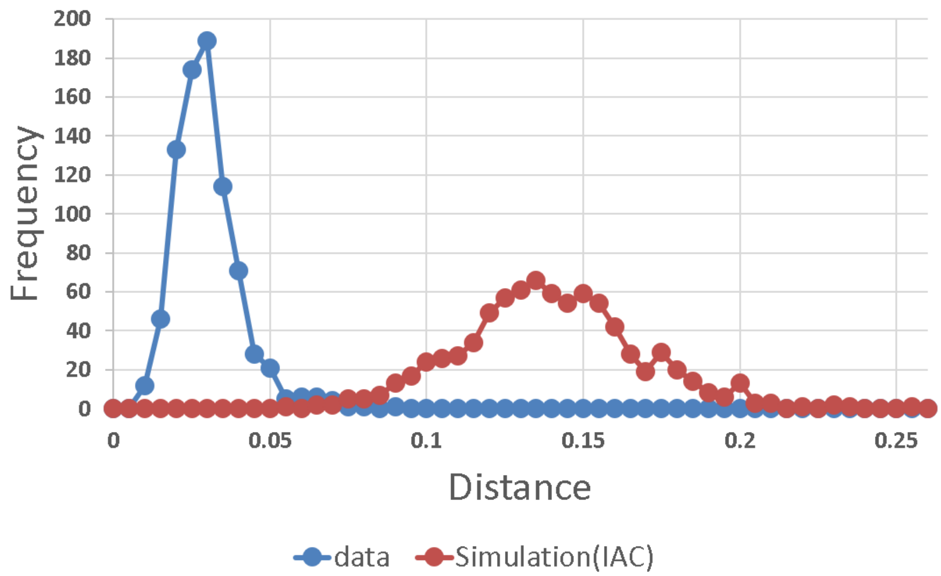


Figure 5.1: Distance to a fitted spatial model of each data

tell that simulation under the IAC assumption is not reflecting the actual outcomes well for four-candidate cases. Specifically, the actual outcomes are nearer to the closest outcomes that the spatial model can depict; however, simulated data are relatively farther from it.

5.4 Discussion

The probabilistic models have been widely used over periods for their advantages. However, they begin with unrealistic assumptions and show significant gaps compared to actual election results. Because of that, it is doubtful that the analyses or expectations based on probabilistic models are realistic. On the other hand, simulations based on the spatial model greatly mimic actual voting outcomes in three-candidate cases. We offer potential evidence supporting that the model also works in four-candidate cases.

We use the spatial model for four candidates in three-dimensional space. The main framework is identical to the one-dimensional or two-dimensional setup, but higher dimensionality makes the work challenging and costly. Eight parameters are required to define a three-dimensional spatial model. Parameter estimation costs heavily because of the high dimensionality and the complexity of the functional form.

We fit the actual number of participants for each preference ordering in survey data and the simulated preference distribution based IAC assumption to the spatial model. Then we measure the distance from the data and to the model fitted. By comparing two distributions of the distance, we can tell that the two data sets are hardly coming from the same population, so we insist that the IAC assumption does not reflect the real voter's preference in the four-candidate election.

However, the work done in this chapter is not sufficient to show that the normal spatial model is the best model for analyzing four-candidate cases. Thus, our following project should offer a proper method to show it and empirical evidence to complete the goal. The prerequisite step is knowing about the joint distribution of eight parameters from the data. Then we can generate data based on the spatial model and compare it to the actual outcome.

Another future research topic is finding evidence of Political polarization with the spatial model framework. The polarization is not a singular issue, mainly where two main political parties exist. Thus, applying bimodal distributions instead of unimodal normal distribution might induce better figures for explaining voting outcomes in some elections, such as the recent US or South Korea elections. We have done a small extra study: from each fitted spatial model of data, we put two modes of the trivariate normal distribution by arbitrary slight perturbation of x, y, z-coordinates from the initially estimated mode and compared the likelihoods. The study results show that the perturbations do not improve the likelihood of any data.

Not a main point in terms of the research, but the optimization algorithm and programming part is essential. The parameter estimation process we programmed has lots of room to be more efficient: To reduce time spent and errors from numerical differentiation, applying a first-degree optimization algorithm could be important.

Bibliography

J. R. Chamberlin and M. D. Cohen. Toward applicable social choice theory: A comparison of social choice functions under spatial model assumptions. *American Political Science Review*, 72(4):1341–1356, 1978.

R. B. Darlington. Why condorcet consistency is essential. *arXiv preprint arXiv:1706.01841*, 2017.

A. Darmann, J. Grundner, and C. Klamler. 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, 2019.

F. De Sinopoli, B. Dutta, and J.-F. Laslier. Approval voting: three examples. *International Journal of Game Theory*, 35(1):27–38, 2006.

A. Downs et al. An economic theory of democracy. 1957.

A. Fabre. Tie-breaking the highest median: alternatives to the majority judgment. *Social Choice and Welfare*, 56(1):101–124, 2021.

M. Feizi, R. Ramezanian, and S. Malek Sadati. Borda paradox in the 2017 iranian presidential election: empirical evidence from opinion polls. *Economics of Governance*, 21(2):101–113, 2020.

D. S. Felsenthal and N. Tideman. Varieties of failure of monotonicity and participation under five voting methods. *Theory and Decision*, 75(1):59–77, 2013.

D. S. Felsenthal and T. N. Tideman. Interacting double monotonicity failure with direction of impact under five voting methods. *Math. Soc. Sci.*, 67:57–66, 2014.

M. Forschungsgruppe Wahlen, M. Berger, W. G. Gibowski, D. Fuchs, M. Kaase, H.-D. Klingemann, D. Roth, U. Schleth, W. Schulte, B. EMNID, and More. Politbarometer 1977-2019 (cumulated data set). GESIS Data Archive, Cologne. URL https://www.gesis.org/en/elections-home/politbarometer.

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

W. V. Gehrlein. *Condorcet's paradox*. Springer, 2006.

W. V. Gehrlein and P. C. Fishburn. The probability of the paradox of voting: A computable solution. *Journal of Economic Theory*, 13(1):14–25, 1976.

W. V. Gehrlein and D. Lepelley. *Voting Paradoxes and Their Probabilities*, pages 1–47. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011. ISBN 978-3-642-03107-6. doi: 10.1007/978-3-642-03107-6_1. URL https://doi.org/10.1007/978-3-642-03107-6_1.

I. J. Good and T. N. Tideman. 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, 1976.

A. Graham-Squire and N. Zayatz. Lack of monotonicity anomalies in empirical data of instant-runoff elections. *Representation*, 57(4):565–573, 2021.

J. Green-Armytage. Strategic voting and nomination. *Social Choice and Welfare*, 42(1): 111–138, 2014.

J. Green-Armytage, T. N. Tideman, and R. Cosman. Statistical evaluation of voting rules. *Social Choice and Welfare*, 46(1):183–212, 2016.

G. T. Guilbaud. Les theories de pinteret general et le probleme logique de fagregation. *Economic appliquee*, 5:501–584, 1952.

J. Hammersly and D. Handscomb. Monte carlo methods (methuen and company, london). 1964.

H. Hotelling. Stability in competition. In *The collected economics articles of Harold Hotelling*, pages 50–63. Springer, 1990.

D. M. Kilgour, J.-C. Grégoire, and A. M. Foley. The prevalence and consequences of ballot truncation in ranked-choice elections. *Public Choice*, 184(1):197–218, 2020.

K. Kuga and H. Nagatani. Voter antagonism and the paradox of voting. *Econometrica: Journal of the Econometric Society*, pages 1045–1067, 1974.

P. Kurrild-Klitgaard. An empirical example of the condorcet paradox of voting in a large electorate. *Public Choice*, 107(1):135–145, 2001.

P. Kurrild-Klitgaard. Trump, condorcet and borda: Voting paradoxes in the 2016 republican presidential primaries. *European Journal of Political Economy*, 55:29–35, 2018.

D. Lepelley, F. Chantreuil, and S. Berg. The likelihood of monotonicity paradoxes in run-off elections. *Mathematical Social Sciences*, 31(3):133–146, 1996.

D. Lepelley, A. Louichi, and F. Valognes. Computer simulations of voting systems. *Advances in Complex Systems*, 3(01n04):181–194, 2000.

T. Matje. *Empirical Analyses of a Spatial Model of Voter Preferences*. PhD thesis, Virginia Tech, 2016.

N. Mattei. Empirical evaluation of voting rules with strictly ordered preference data. In *International Conference on Algorithmic Decision Theory*, pages 165–177. Springer, 2011.

N. Mattei, J. Forshee, and J. Goldsmith. An empirical study of voting rules and manipulation with large datasets. *Proceedings of COMSOC*, 59, 2012.

S. Merrill III. A comparison of efficiency of multicandidate electoral systems. *American Journal of Political Science*, pages 23–48, 1984.

N. R. Miller. Closeness matters: monotonicity failure in irv elections with three candidates. *Public Choice*, 173(1):91–108, 2017.

R. G. Niemi and W. H. Riker. The choice of voting systems. *Scientific American*, 234(6): 21–27, 1976.

J. T. Ornstein and R. Z. Norman. Frequency of monotonicity failure under instant runoff voting: estimates based on a spatial model of elections. *Public Choice*, 161(1):1–9, 2014.

F. Plassmann and T. N. Tideman. How to predict the frequency of voting events in actual elections. *Available at SSRN 1911286*, 2012.

F. Plassmann and T. N. Tideman. How frequently do different voting rules encounter voting paradoxes in three-candidate elections? *Social Choice and Welfare*, 42(1):31–75, 2014.

S. V. Popov, A. Popova, and M. Regenwetter. Consensus in organizations: Hunting for the social choice conundrum in apa elections. *Decision*, 1(2):123, 2014.

M. Regenwetter, A. Kim, A. Kantor, and M.-H. R. Ho. The unexpected empirical consensus among consensus methods. *Psychological Science*, 18(7):629–635, 2007.

T. N. Tideman. Independence of clones as a criterion for voting rules. *Social Choice and Welfare*, 4(3):185–206, 1987.

T. N. Tideman and F. Plassmann. Modeling the outcomes of vote-casting in actual elections. In *Electoral Systems*, pages 217–251. Springer, 2012.

A. Van Deemen. On the empirical relevance of condorcet's paradox. *Public Choice*, 158(3): 311–330, 2014.

Appendices

Appendix A

Appendix for Chapter 3

We attach simple empirical data that fortify our assumptions but are omitted in Chapter 3.

A.1 Mean and FVM as measure for upset

A.1.1 Upset rate and mean difference / FVM difference in 55,321 pairwise comparisons (all possible pairs)

For all 1,022 surveys, there are 55,321 pairs of candidates. We check the frequency of upset by the difference in mean score and FVM score (Figure A.1). Two measurements are almost equivalently sensitive to upset. Thus, one cannot tell which one is a better indicator to estimate upset in pairwise comparisons.

A.1.2 Upset rate and mean difference / FVM difference in 3,066 pairwise comparisons (top three candidates)

For 1,022 surveys, we pick the top three candidates by mean and FVM score. For each, there are 3,066 pairs of candidates. Then we check the frequency of upset in each case by the difference in mean and FVM scores (Figure A.2). In either case, the curve of the upset rate by FVM is below the other curve.

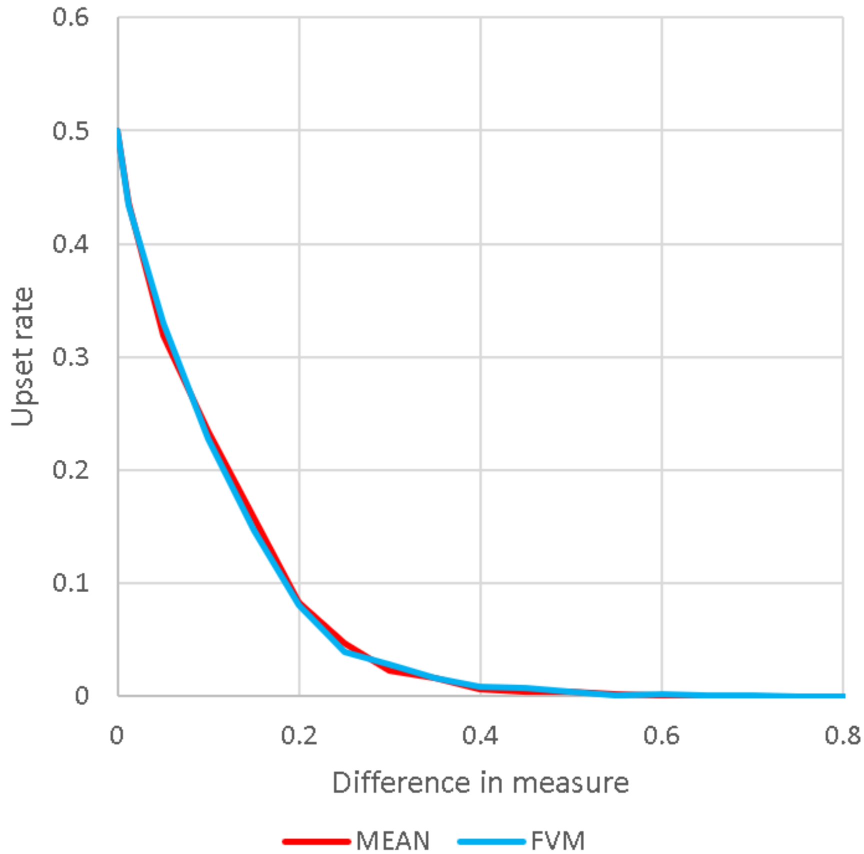
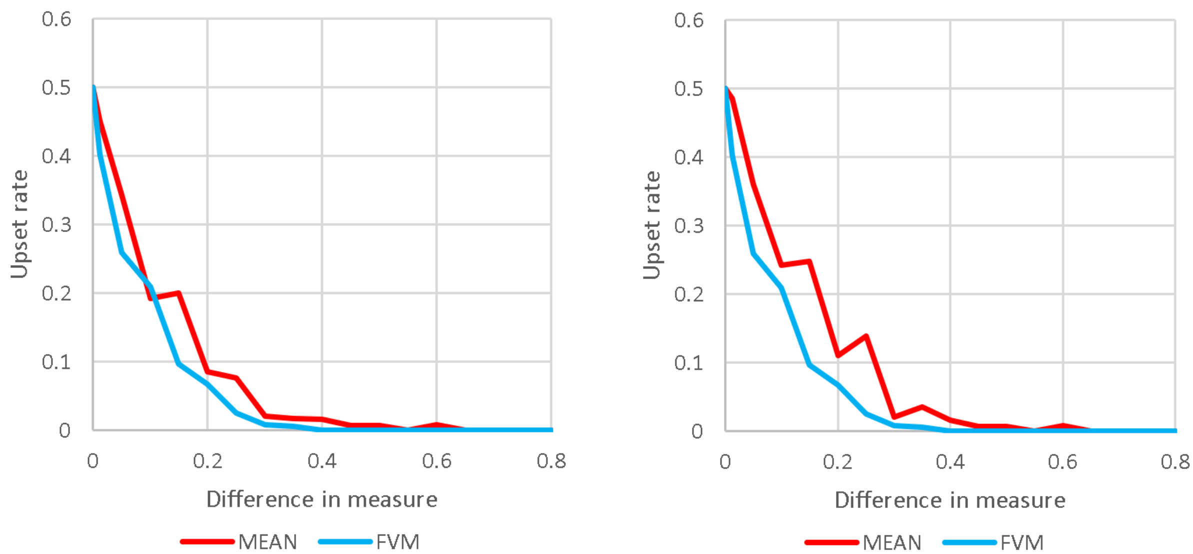


Figure A.1: Empirical distribution of upset by difference in mean and FVM, all pairs

A.1.3 Why do we NOT use the mean score of the evaluations to predict upset?

One of the essentials in our modeling is estimating the probability of the upset, and we construct a statistics, FVM, for candidate evaluation. We suggest a reason that we stick to the median in Chapter 3. 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 between the mean and FVM scores in all pairs. However, when it comes to pairs generated from the top triples, with the same numeric difference between



(a) Pairs in top three candidates by mean (b) Pairs in top three candidates by FVM

Figure A.2: Empirical distribution of upset by difference in mean and FVM, pairs in top triples

two candidates in different statistics, upset is less likely to happen in FVM. One could tell a higher FVM candidate would win in the pairwise election with more certainty, which indicates that the FVM score is a more reliable measurement to upset prediction than the mean score.

A.2 Evaluation gap between top candidates by the size of candidate pool

To support our conjecture for the extended model (Subsection 3.3.7), "the average values of differences in FVMs within the top triple will shrink when the top triple is chosen from a survey that has more politicians," we attach a result of simple simulations.

Suppose that a candidate's evaluation follows 1) standard normal distribution and 2) uniform distribution [0,1]. Then for different sizes of sampling pool, we calculate the 95%

	confidence interval (uniform)		confidence interval (normal)	
Number of candidates	lower bound	upper bound	lower bound	upper bound
3	0.49986	0.50183	1.67772	1.70851
4	0.39874	0.40032	1.32073	1.34153
5	0.33192	0.33316	1.15662	1.17343
6	0.28447	0.28546	1.05651	1.0719
7	0.24832	0.24913	0.98725	1.0002
8	0.22197	0.22265	0.95017	0.96291
9	0.19901	0.19958	0.91365	0.9259
10	0.18115	0.18164	0.88363	0.89491
20	0.09497	0.09512	0.73558	0.74427

Table A.1: 95% confidence interval of difference between $X_{(1)}$ and $X_{(3)}$

confidence interval of the mean of $X_{(1)} - X_{(3)}$, where $X_{(i)}$ is i'th biggest number among the samples. The result hints at how the evaluation gap differs by the size of the candidate pool. For each number of samples, we simulated 10,000 times. The result shows that the gap between two order statistics tends to decrease as the number of samples increases, and this result would be no different from the result qualitatively with a different support setup $[-5, 5]$ or $[1, 11]$.

Appendix B

Appendix for Chapter 5

In more detail, we introduce the parameter estimation used in Chapter 5 for the normal spatial model for three candidates in two-dimensional space and four candidates in three-dimensional space.

B.1 The Normal Spatial Model for three candidates in two-dimensional space (NSM-3C2D)

The idea of the normal spatial model is that there are locations of candidates and the mode of a normal distribution that well describe the given distribution of people's preferences.

Consider two-dimensional space and a bivariate normal distribution with variance as 1 and no covariance. We could draw a triangle by connecting points for arbitrary three points **A**,**B**, and **C**. From the triangle, we have three bisecting lines of each edge that pass through the circumcenter of the triangle, and they divide the space into six partitions. Assuming that the bivariate normal distribution represents voters' ideal points and they prefer candidates by distance from their ideal. Each area corresponds to strict preference orderings, and the density on each area is an estimated share of each strict preference ordering.

Idea of the normal spatial model is that there are locations of candidates and the mode of normal distribution that well describe the given distribution of people's preference.

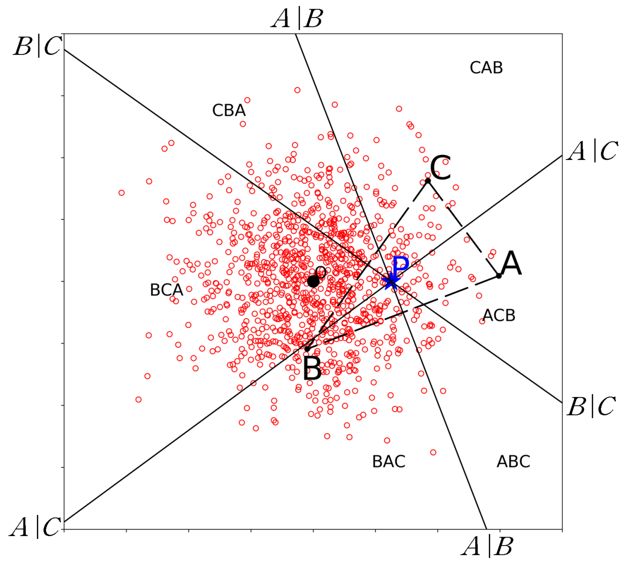


Figure B.1: Normal spatial model for three candidates in two-dimensional space and preference order representation

Technically, eight numbers are required to define four points on two-dimensional space, but we decrease its degree to four. First, defining three bisecting lines is mathematically equivalent to defining three points. Second, Rotate three lines (three points), and the mode of normal distribution against an arbitrary point induces the same partitions with the same density.

First, put the origin onto the triangle's circumcenter, **P**. Next, rotate all figures about the origin; until the line bisecting **A** and **B** (call it **A**|**B**) aligns to the x-axis. Then, we can define the other two bisect lines **B**|**C** and **A**|**C** by the angle between each of them and **A**|**B**, α and β, respectively ($0 < \alpha < \beta < \pi$). We need two more numbers to define the coordinate of the mode of the normal distribution $(x, y) \in \mathbb{R}^2$, as Figure B.2.

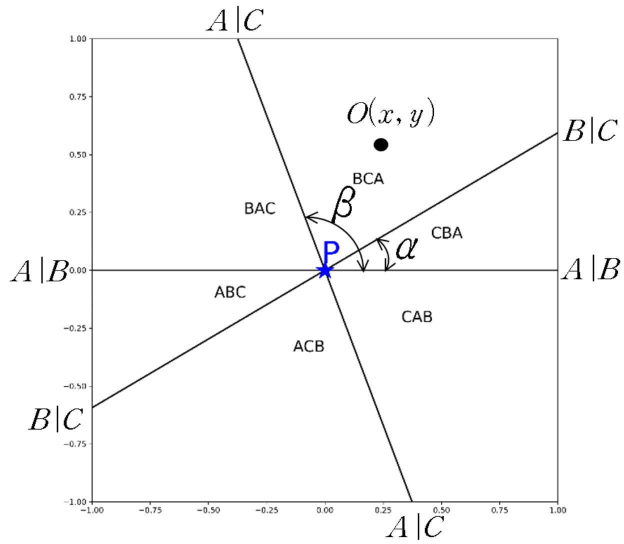


Figure B.2: Spatial model for three candidates in two-dimensional space, defined by four parameters

B.1.1 Initial guess

For oppositely locating pairs of partitions **ABC**-**CBA**, **ACB**-**BCA**, **BAC**-**CAB**, the vertically opposite angle should be the same. From the data, sum up the number of voters with **ABC** and **CBA**, and set the angle proportional to the share of the voters. In other words, $\theta_{\mathbf{ABC}} + \theta_{\mathbf{CBA}} = 360 * ((n[\mathbf{ABC}] + n[\mathbf{CBA}])/\text{total number of voters})$. Thus, we begin by setting the mode of the normal distribution (P) is at the origin, then the density of each partition has one-to-one corresponds to the angle. After getting the other four angles, we have three bisecting lines of initial guess, and so angles α_0 and β_0.

Each bisecting line divides the space into two partitions. The two partitions correspond to the share of votes two candidates win in the pairwise comparison, thus indicating where the mode of normal distribution should be. For instance, we can draw a virtual line parallel to $A|B$, shifted up by 0.45, as shown in Figure B.3. Since CDF of standard normal distribution at 0.45 is 0.674, the bisecting line cuts the distribution by 0.674 and 0.326 when the mode is on the virtual line. In other words, **B** wins 67.4% and **A** wins 32.6% in pairwise election. While applying the same to the other two bisecting lines, we have three virtual lines and a triangular part constructed by them. We pick the center of gravity (for convenient calculation) as the mode of the normal distribution of the initial guess (Figure B.4).

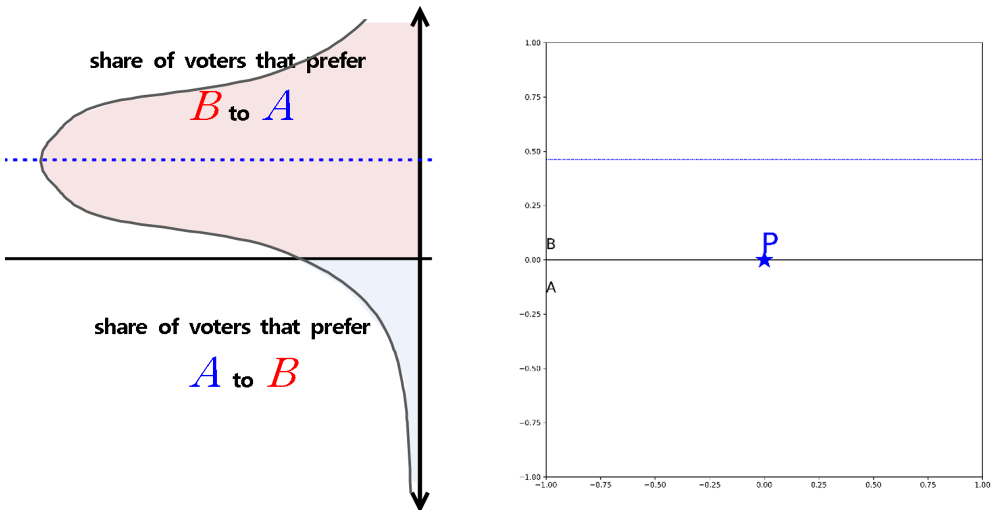


Figure B.3: Guessing location of the mode by share of A and B

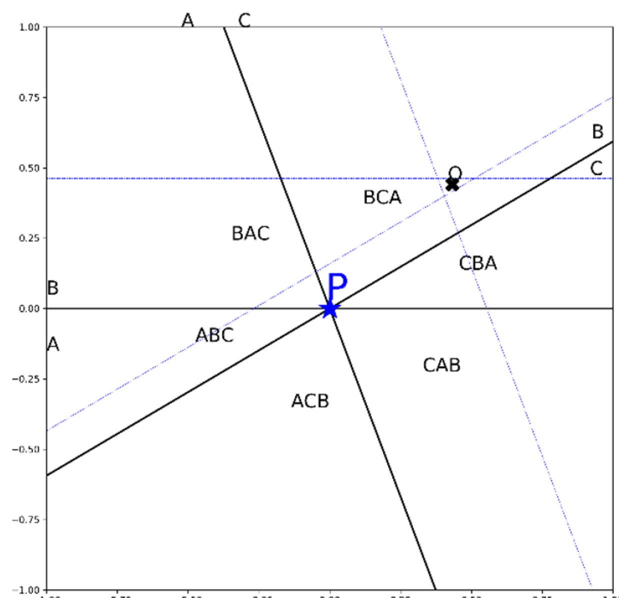
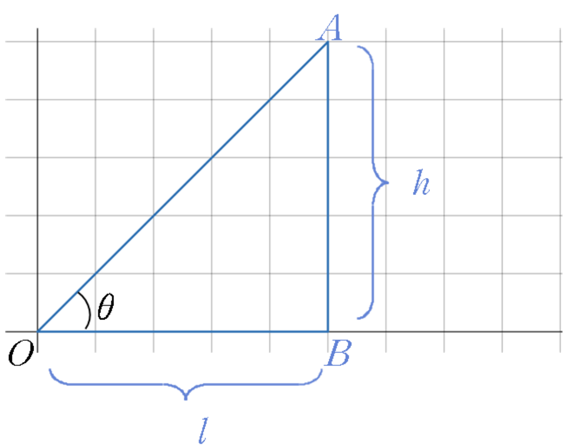


Figure B.4: Initial guess of the mode of the bivariate normal distribution

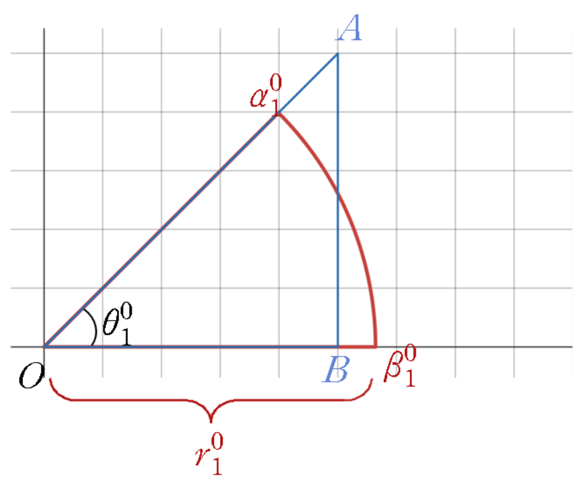
B.1.2 Ground work

Directly calculating the density is not available for each partition. Thus we cut partitions into pieces that are available to be calculated. To begin with, We introduce how to approximate the density of bivariate normal distribution on a right triangle when the mode of normal distribution at a vertex.

Suppose a right triangle OAB with l for length and h for height. O is the mode of bivariate normal distribution, and the angle of vertex O is θ (Figure B.5a).



(a) A right triangle OAB



(b) A right triangle and a sector form for approximation

Figure B.5: Density approximation of normal distribution on a right triangular area

Since the formula for calculating the density on a sector form is known, we calculate the density of a sector form $O\alpha\beta$ which has the same area as the triangle, and use it to approximate the density on the right triangle OAB (Figure B.5b).

Assume the bivariate normal distribution is standard normal with 0 covariances or,

$$\boldsymbol{\mu} = [0, 0], \qquad \boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Then the density on the sector with angle θ in radian and radius r is as follows.

$$Pr(R \leq r;\theta) = \frac{\theta}{2\pi}\left(1 - \exp\left(-\frac{r^2}{2}\right)\right)$$

The radius that makes the area of the triangle and the sector equal can be found by $r^2 = \frac{lh}{\theta}$. Hence we have an alternative representation, the function of l and h.

$$P\left(l, h\right) = \frac{\theta}{2\pi}\left(1 - \exp\left(-\frac{1}{2}\frac{lh}{\theta}\right)\right)$$

where $\theta = \tan^{-1}\left(\frac{l}{h}\right)$.

When we divide the triangle into two to let the height equal $h/2$, we get a better

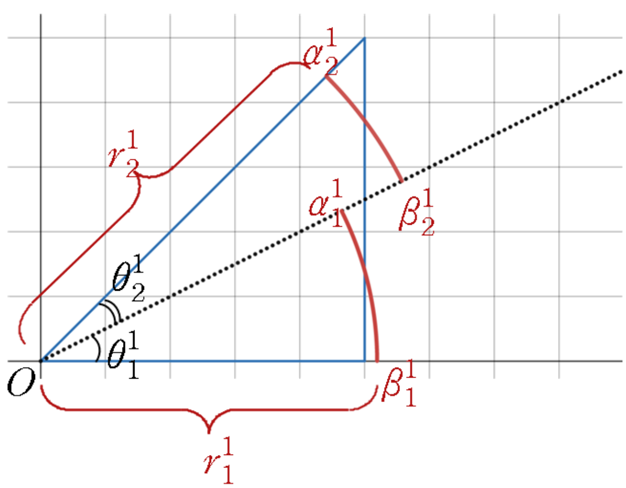


Figure B.6: A right triangle and two sector forms for approximation

approximation (Figure B.6). Then we use the density of each sector that has an equal area to each triangle. The representation of the generalized version, approximating by N sector forms, is

$$P\left(l, h; N\right) = \sum_{n=1}^{N} \left(\frac{\theta_n}{2\pi} \left(1 - exp \left(-\frac{1}{2N} \frac{lh}{\theta} \right) \right) \right)$$

where $\theta_n = \tan^{-1}\left(\frac{n}{N}\frac{l}{h}\right) - \tan^{-1}\left(\frac{n-1}{N}\frac{l}{h}\right)$. In chapter 5, we cut the triangle by 10, $N = 11$.

B.1.3 Density approximation

We can approximate six partitions using the approximating distribution on a right triangle.

1. The partition with the mode of the normal distribution (partition BCA)

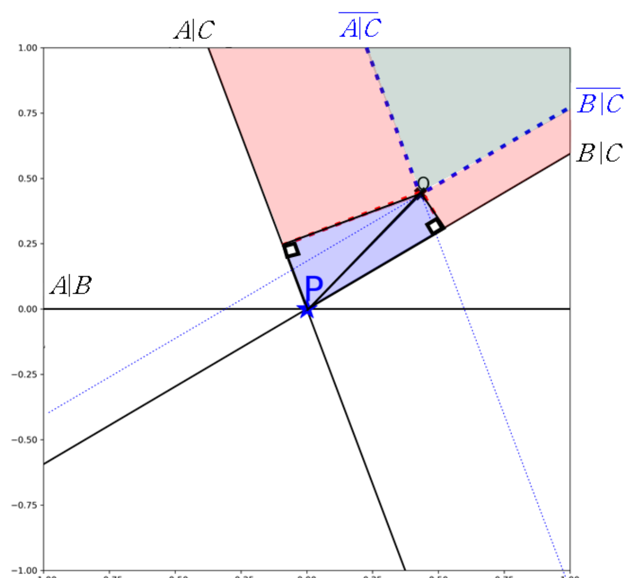


Figure B.7: Calculating density on the partition with the mode

The partition that has the mode of the normal distribution can be divided into a sector with an infinite radius (green), two half-infinite bars (red), and two right triangles (blue). The density of an infinite-radius sector is a share of the angle between $A|C$ and $B|C$ to 2π. For two half-infinite bars, the density can be found by the density function of the normal distribution with the width of the bar. Using the approximation, we get the density of triangles. By adding all, one could get the density of partition **BCA**.

2. The partition next to the partition with the mode (partition CBA)

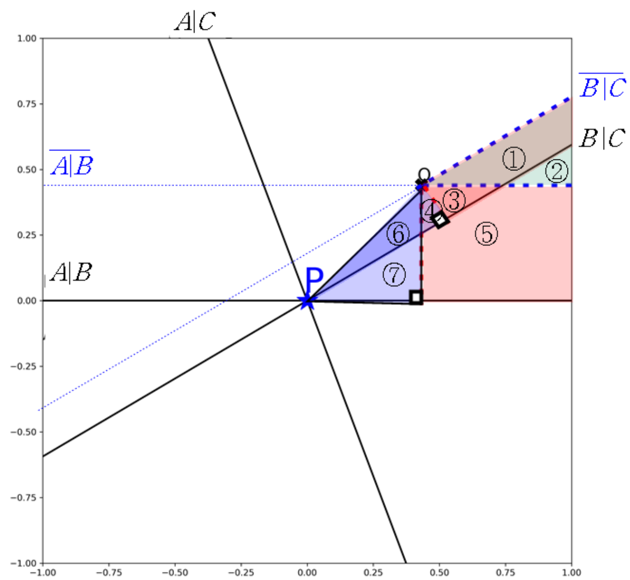


Figure B.8: Calculating density on a partition next to the partition with the mode

Graphically, the partition is ②+⑤+⑦. We calculate one circular sector with an infinite radius (①), two half-infinite bars, (①+③, ③+④+⑤), and two right triangles (④+⑥, ⑥+⑦). The following recipe provides density on the partition CBA.

(① + ②) + (③ + ④ + ⑤) − (① + ③) − (④ + ⑥) + (⑥ + ⑦) = ② + ⑤ + ⑦

3. Remainder

With the density of two partitions, we can calculate other partitions. For instance, to calculate the density of partition **BAC**, we subtract the density of the above two (partition **BCA** and **CBA**) from the density of space above the x-axis, which equals $\Phi(d)$, where $\Phi(\cdot)$ is the cumulative density function of standard normal distribution. Likewise, all other partitions (partition **ABC**, **ACB**, and **CAB**) can be derived. For six densities on six partitions, we have a discrete probability distribution, $(p_1, p_2, p_3, p_4, p_5, p_6)$. Each corresponds to the preference in lexicographic order: **ABC**, **ACB**, **BAC**, **BCA**, **CAB**, and **CBA**.

B.1.4 Likelihood maximization

With $(p_1, p_2, p_3, p_4, p_5, p_6)$, the probability mass function of the multinomial distribution and the log-likelihood function are

$$f(p_1, p_2, p_3, p_4, p_5, p_6) = \frac{N!}{n_1! \cdots n_6!} \prod_{i=1}^{6} p_i^{n_i}$$

$$llh(p_1, \cdots, p_6; n_1, \cdots, n_6) = ln(N!) - \sum_{i=1}^{6} ln(n_i!) + \sum_{i=1}^{6} n_i ln(p_i)$$

where $(n_1, n_2, n_3, n_4, n_5, n_6)$ is an observation, the number of voters have each of 6 preference order in data and $\sum_{i=1}^{6} n_i = N$. In our study, they are the number of voters have each strict preference order in a survey.

Our model has four parameters (x, y, α, β), that decide $(p_1, p_2, p_3, p_4, p_5, p_6)$. Since $(n_1, n_2, n_3, n_4, n_5, n_6)$ is given, $\ln(N!)$ and $\sum_{i=1}^{6} ln(n_i!)$ are constants, and the maximization problem is

$$\max_{x, y, \alpha, \beta} L(x, y, \alpha, \beta; n_1, \cdots, n_6) = \sum_{i=1}^{6} n_i ln(p_i(x, y, \alpha, \beta))$$

When we consider the function that induces $(p_1, p_2, p_3, p_4, p_5, p_6)$ from (x, y, α, β), it is reasonably complex, including unusual formulas and approximations. Hence we do not solve the maximization problem analytically.

B.2 The Normal Spatial Model for four candidates in three-dimensional space (NSM-4C3D)

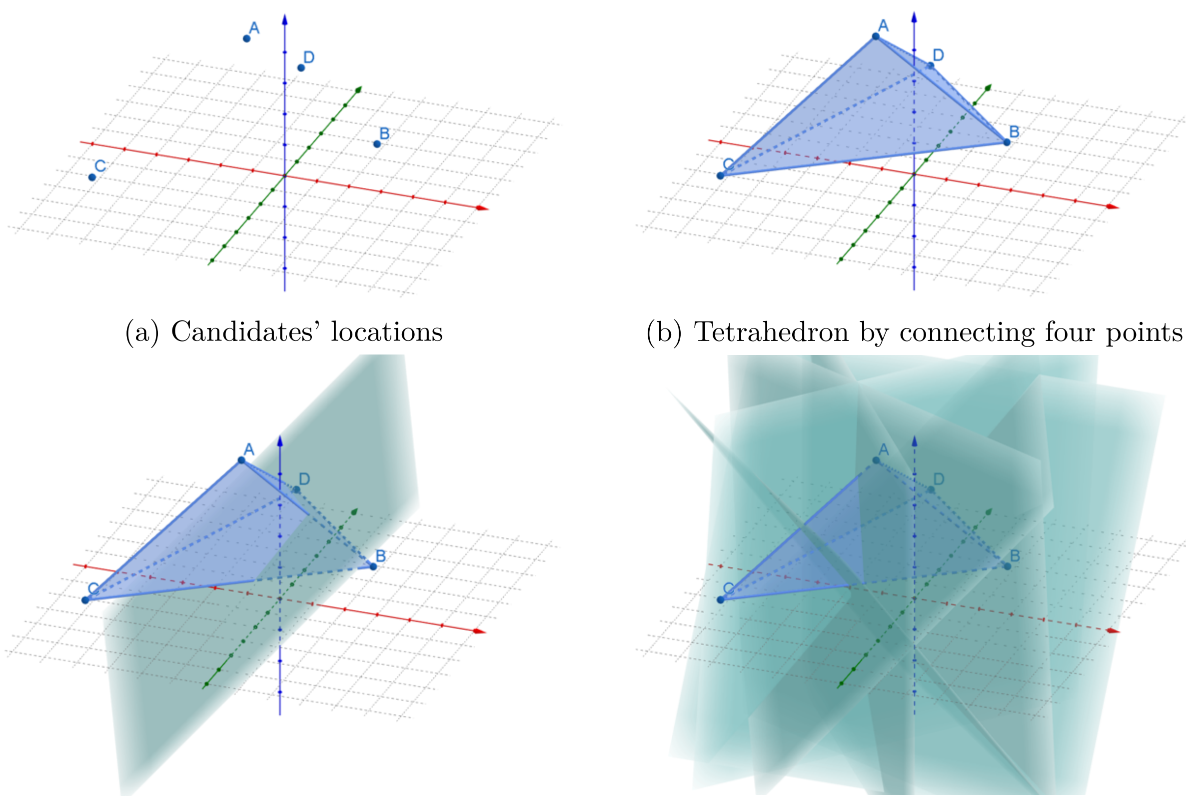
The objective is the same, finding the best locations of candidates and the mode of multivariate normal distribution. However, in three-dimensional space this time.

Adding one extra dimensionality induces substantial changes. Now, we need to consider the trivariate normal distribution. Connecting arbitrary four points in three dimensions, $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$, will form a tetrahedron. Six bisecting planes of six tetrahedron edges pass through the circumcenter of the tetrahedron, but we have 24 partitions for 24 preference orderings.

As we find parameters that define three bisecting lines instead of coordinates of three points in three-candidate and two-dimension settings, we define six bisecting planes, not four points' coordinates in xyz-space: Knowing five angles is enough to define six bisecting planes.

Rotate three dimensions to put three points $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ on a horizontal plane (call $\mathbf{ABC}$ plane), to put $A|B$ plane aligned to $y = 0$, parallel to xy-plane, and by parallel translation to adjust the circumcenter of the tetrahedron to the origin. Then it exactly corresponds to the model with three candidates in two-dimensional space.

Not three bisecting lines and the origin as intersection point, we have three bisecting planes of $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ and z-axis as the line of intersection. Without loss of generality, suppose the location of $\mathbf{D}$ is above xy-plane. Then take a vector $(0, 0, 1)$ initiated from the origin, on the line of intersection of three bisecting planes of $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$, and coming toward $\mathbf{D}$. Refer to it as the $\mathbf{D}_+$ vector. We also can take unique $\mathbf{A}_+$, $\mathbf{B}_+$, and $\mathbf{C}_+$ vectors by



(a) Candidates' locations

(b) Tetrahedron by connecting four points

(c) Bisecting hyperplane of A and B

(d) Six bisecting hyperplanes

Figure B.9: The normal spatial model for four candidates in three-dimensional space

finding each line of intersection of three bisecting planes of **BCD**, **ACD**, and **ABD** planes. By definition, two different + vectors span a space which is the bisecting plane formed by the other two. Specifically, $\mathbf{A}_+$ and $\mathbf{B}_+$ span $C|D$.

While fixing the $A|B$ plane as $y = 0$, $B|C$ and $A|C$ can be defined by two between angles (α, β) on xy-plane exactly as a two-dimensional setup. Since we have $B|C$, we know x and y component of vector $\mathbf{A}_+$, and z component of $\mathbf{A}_+$ vector can be defined by between angle of $\mathbf{A}_+$ and $\mathbf{D}_+$ on $B|C$. Therefore, we need five angles, $(\alpha, \beta, \gamma, \delta, \varepsilon)$, to define $\mathbf{A}_+$, $\mathbf{B}_+$, and $\mathbf{C}_+$, and thus six bisecting planes. As two-dimensional setting, $0 < \alpha < \beta < \pi$ should be satisfied. By definition, γ, δ, and ε should be greater than 0 and less than π. For coordinates of the mode of the trivariate normal distribution, we need three more parameters, $(x, y, z) \in \mathbb{R}^3$.

B.2.1 Initial guess

To begin with three candidates in two-dimensional space, we get (α, β, x, y), and they will be our initial guess for four parameters. For the other three angles $(\gamma, \delta, \epsilon)$, we set $2\pi/3$ for the initial guess. For the z coordinate of the mode, we calculate average share of **D** against **A**, **B**, and **C** in the pairwise elections and then set the initial value by applying the value to the inverse of CDF of standard normal distribution.

B.2.2 Density approximation

For the multivariate normal distribution with $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ where $\boldsymbol{\mu}$ is $p \times 1$ mean vector and $\boldsymbol{\Sigma}$ is $p \times p$ variance-covariance matrix,

$$(\mathbf{x} - \boldsymbol{\mu})^{\mathbf{T}} \boldsymbol{\Sigma}^{-\mathbf{1}} (\mathbf{x} - \boldsymbol{\mu}) \sim \boldsymbol{\chi}^{\mathbf{2}}_{\mathbf{p}}$$

While assuming the standard trivariate normal distribution,

$$\mathbf{x^2 + y^2 + z^2} \sim \boldsymbol{\chi}^{\mathbf{2}}_{\mathbf{3}}$$

For a given radius r the equation of the sphere with the center of the sphere as O is

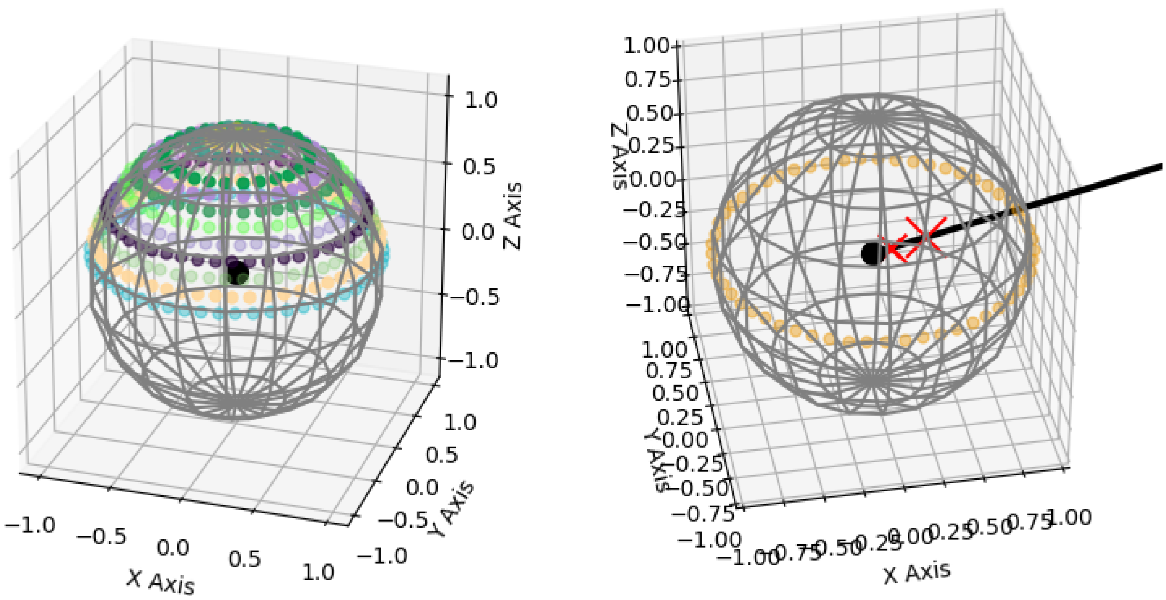
$$\mathbf{x^2 + y^2 + z^2 = r^2}$$

Therefore, the density of within-sphere, $Pr[\mathbf{x^2 + y^2 + z^2 \leq r^2}]$, is $Pr\left[\boldsymbol{\chi}^{\mathbf{2}}_{\mathbf{3}} \leq \mathbf{r^2}\right]$.

Tiling the sphere

Consider a sphere with the center at O, the mode of the trivariate standard normal distribution. Put N number of points equally distributed on the surface of the sphere. Each point is the center of an identical square-liked shape of tiles are covering the whole sphere.

Imagine a three-dimensional structure: an infinitely tall pyramid-like, made of lines



(a) Centers of 500 tiles on upper hemi-sphere

(b) A ray from the mode to the center of a tile and intersections with some bisecting planes

Figure B.10: Tiling the sphere

originating from the O and passing through each tile's vertex. Some of the six planes might cut the pyramid, and each partition will be a part of some strict preference ordering.

The idea is to approximate each cut part of the pyramid with $1/N$ part of spheres with different radiuses (Figure B.10). For example, consider a ray that starts from O,and passes through the tile's center. Let the distance from O to where the ray meets the first cut edge as l_1. If applicable, for other cuts, define the distances from O to where ray meets the edges as l_2, and so forth.

We approximate the density of the first chunk as $\frac{1}{N}Pr\left[\chi^2_{\mathbf{3}} \leq \mathbf{l_1}^{\mathbf{2}}\right]$. The second chunk is $\frac{1}{N}\left(Pr\left[\chi^2_{\mathbf{3}} \leq \mathbf{l_2}^{\mathbf{2}}\right] - Pr\left[\chi^2_{\mathbf{3}} \leq \mathbf{l_1}^{\mathbf{2}}\right]\right)$ and $\frac{1}{N}\left(Pr\left[\chi^2_{\mathbf{3}} \leq \mathbf{l_3}^{\mathbf{2}}\right] - Pr\left[\chi^2_{\mathbf{3}} \leq \mathbf{l_2}^{\mathbf{2}}\right]\right)$ for the third. When there is no more cut, then for the remaining giant chunk, the density is approximated to $\frac{1}{N}\left(1 - Pr\left[\chi^2_{\mathbf{3}} \leq \mathbf{l_3}^{\mathbf{2}}\right]\right)$.

We divided the sphere by 2,000 tiles, so with 2,000 straight rays. Then, for each line,

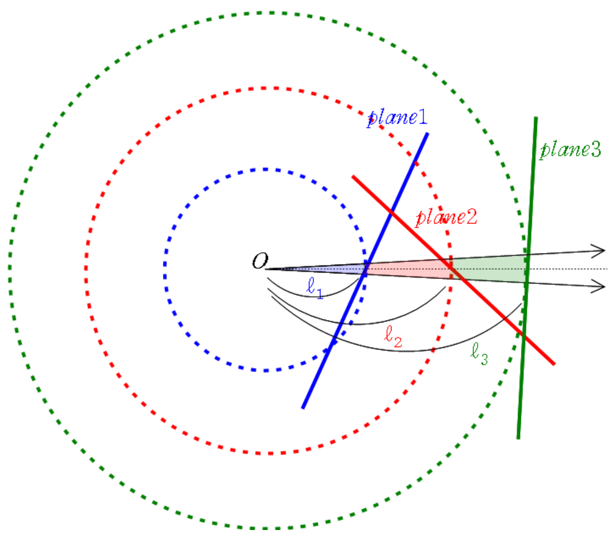


Figure B.11: A ray from the mode to the center of a tile and intersections with some bisecting planes, top view

we check the intersection with six-plane, approximate the density of chunks, define which strict preference ordering is corresponding, and sum up by the preference. It ends up with 24 numbers.

B.2.3 Likelihood maximization

Like three candidates in a two-dimensional setup, eight parameters decide the probability profile with 24 numbers. Thus, with a given probability profile $(p_1, p_2, \cdots, p_{24})$ and an observation $(n_1, n_2, \cdots, n_{24})$, the likelihood maximization problem can be set by

$$\max_{x,y,z,\alpha,\beta,\gamma,\delta,\varepsilon} L(x,y,z,\alpha,\beta,\gamma,\delta,\varepsilon; n_1,\cdots,n_{24}) = \sum_{i=1}^{24} n_i ln\left(p_i(x,y,z,\alpha,\beta,\gamma)\right)$$

The function that induces $(p_1, \cdots, p_{24})$ from $(x, y, z, \alpha, \beta, \gamma, \delta, \varepsilon)$ is highly complicated: It is undoubtedly a not-closed form function, primarily because of tiling the sphere and approximating the probability on chunks of pyramids. Hence we do not solve the maximization problem analytically.