Departmental Coversheet

MSc in Computer Science 2020-21

Project Dissertation

Project Dissertation title: Voting and Strategic Candidacies in Primary Systems

Term and year of submission: Trinity Term 2021

Candidate Number: 1049855

Abstract

Voting in single stage, direct systems, where voters submit scores for each candidate and a winner is selected, depending on the voting rule used, has been studied extensively in the social choice literature. However, numerous types of elections involve a multi-stage process, with primary systems being a notable example. In a recent line of research, focused on analysing the distortion of voting rules under the primary system, a model to perform such an analysis has been introduced. In this project, we use the same model to perform a quantitative instance-wise comparison between the two systems as well as an average-case analysis on the distortion of several voting rules. Furthermore, we extend the model to allow the analysis of strategic candidacy games. We adapt two existing classes of games for the direct system to the primary model and introduce a novel one, specifically adjusted to the primary system. We analyse the properties of their pure strategy Nash equilibria in one dimension, convergence of different types of best-response dynamics and prove that the associated decision problems are NP-complete.

Contents

1 Introduction **1**

1.1 Motivation 1

1.2 Contribution 2

1.3 Project Structure 3

2 Related Work **4**

2.1 Quality of Voting Rules 4

2.2 Strategic Candidacies 9

3 Preliminaries and Model **13**

4 Comparison Between Primary and Direct Systems **16**

4.1 Voting under Plurality 16

4.1.1 Primaries are not Necessarily Better if one Party has Exactly 1 Candidate 17

4.1.2 Primaries are not Necessarily Better if the Voters and Candidates are Uniformly Distributed 19

4.2 Condorcet Winners 29

4.2.1 Symmetric Case 32

4.3 Single Transferable Vote 34

4.4 Average-case Distortion 35

4.4.1 1 Dimension . . . 36
4.4.2 3 and 5 Dimensions . . . 40
4.4.3 Polarization . . . 41
4.4.4 Overall Results . . . 45

5 Strategic Candidacies **46**

5.1 Lazy Strategic Candidacy Games . . . 47
5.1.1 Nash Equilibria . . . 47
5.1.2 Best-response Dynamics . . . 53
5.1.2.1 J- and W-dynamics . . . 53
5.1.2.2 Equilibrium Dynamics . . . 57
5.2 Eager Strategic Candidacy Games . . . 60
5.2.1 Nash Equilibria . . . 61
5.2.2 Complexity of PNE . . . 66
5.3 Keen Strategic Candidacy Games . . . 69
5.3.1 Nash Equilibria . . . 70
5.3.2 Complexity of PNE . . . 77
5.3.3 Number of Equilibria . . . 84

6 Conclusions and Future Work **87**

6.1 Reflection . . . 88

A Difference in Distortion **95**

List of Figures

1 Direct elections can be better 17

2 Direct elections are better if one party has exactly 1 candidate 18

3 Direct elections are better even if voters and candidates are uniformly distributed. 19

4 Example where the primary system produces a worse outcome for an even k and $m_1 > 1$ 28

5 Example where the primary system produces a worse outcome for an even k and $m_1 = 1$ 28

6 The primary system produces a better outcome with Condorcet winners and $\pi(a_i) = 1$ 31

7 The primary system produces a better outcome with Condorcet winners and $\pi(a_i) = -1$ 31

8 The Condorcet winner may be eliminated by STV 34

9 Difference in distortion between the primary and direct system for separable, uniformly distributed election instances 38

10 Difference in distortion between the primary and direct system for random, uniformly distributed election instances 39

11 Difference in distortion between the primary and direct system for separable general one dimensional election instances 39

12 Difference in distortion between the primary and direct system for general one dimensional election instances 40

13 Two dimensional Gaussian distributions 43

14 Further experiments for one dimensional Gaussian distributions 44

15 Further experiments for two dimensional Gaussian distributions 44

16 Example of a lazy strategic candidacy game with no PNE 53

17 J-dynamics for a lazy strategic candidacy game 55

18 J-dynamics terminating in a state with all candidates running for a lazy strategic candidacy game 57

19 Example of a PNE in eager strategic candidacy games 63

20 Example of a PNE where candidate b is the winner 65

21 Example of a keen strategic candidacy game with 4 candidates, party separability and no PNE 75

22 Further example of a keen strategic candidacy game with 4 candidates, party separability and no PNE 76

23 Number of PNE in one dimensional KSCGs 85

24 Number of PNE in four dimensional KSCGs 86

25 Difference in distortion between the primary and direct system for separable three dimensional election instances 95

26 Difference in distortion between the primary and direct system for three dimensional election instances 96

27 Difference in distortion between the primary and direct system for separable five dimensional election instances 97

28 Difference in distortion between the primary and direct system for five dimensional election instances 98

List of Tables

1 Comparison of the two systems in one dimension, in terms of the utilitarian social cost of the winner produced 37

2 Comparison of the two systems in three and five dimensions, in terms of the utilitarian social cost of the winner produced 41

3 One dimensional polarization 42

4 Two dimensional polarization 44

5 Overall results 45

6 Voters' preferences in the proof of Theorem 4 79

7 Candidates' preferences in the proof of Theorem 4 80

8 Number of PNE in KSCGs 85

1 Introduction

1.1 Motivation

Voting is one of the fundamental tools in countless decision-making scenarios, as it aggregates voters' ranked preferences over a set of candidates or alternatives to produce an outcome reflecting their collective opinion. Generally, given the voters' preferences, a voting rule is used to determine the winning candidate. Perhaps, the most widely known application of voting is that of political elections. As such, it is natural to think that a voter's preferences may be influenced by the candidates' standings on matters that are important to them, which can include: taxes, voting and abortion rights, environmental issues or gun control. Such a setting can be modelled by the use of a metric space, where the candidates and voters are placed, and the associated distance function determines the likeability of candidates.

Multi-stage voting systems are commonly used in political settings in countries such as the United States, the United Kingdom, Russia, Hungary or Costa Rica. An example of such a system is the primary one, where voters affiliated with a party vote over candidates from their own party. The primary winners then advance to a general election, where the same set of voters submit their preferences again, without taking into account the party affiliation. Given their use, primary systems do present interest, from a theoretical point of view, especially in a comparison against the direct system.

Any political election can be susceptible to forms of manipulation, either from the side of voters, who might be incentivised to vote for a different candidate than their most preferred alternative, or from the side of candidates, who themselves, can have incentives that influence their decision on whether to participate in an election or not (examples of such incentives could be that candidates have their own preferences on who they would like to win the election, or that the costs associated to participating in the election may be insurmountable if they are not capable of winning). While strategic candidacies in direct voting systems have received significant attention, for the primary system this aspect has not yet been investigated.

1.2 Contribution

The contributions related to this project are twofold. Firstly, we extend the work of Borodin et al. [2019] by performing an instance-wise comparison between the primary and direct systems in one dimension, using several voting rules and focusing on the utilitarian social cost of the resulting winners. For the case where voters and candidates are uniformly distributed, under plurality voting, the primary system outperforms the direct one in the vast majority of cases. However, for the Condorcet-consistent Copeland rule, the direct system is almost always better. Next, we perform our own average-case analysis on the distortion of plurality, anti-plurality, plurality with run-off, Borda, Harmonic Borda, Copeland and single transferable vote (STV) voting rules under the two systems. Although we adopt a slightly different methodology in our simulations than Borodin et al. [2019] do, the results are similar: plurality voting is better used in primary systems, while Condorcet-consistent rules are more suitable for the direct system.

Secondly, we present the first analysis of strategic candidacy games under the primary system, focusing on the properties and existence of pure strategy Nash equilibria in one dimension, as well as the reachability of such states by two categories of best-response dynamics. Moreover, we investigate the computational complexity of several associated decision problems, which we prove to be NP-complete, when higher dimensions are considered for the metric space. Lastly, for keen strategic candidacy games (i.e. where candidates receive additional utility if they participate in an election), we perform an average-case analysis on the number of equilibria for a candidate set of size four, for which we conclude that such games generally have only one pure strategy Nash equilibrium, with slightly more variation for a small value of the participation bias.

1.3 Project Structure

- Section 2 contains a brief overview of the related literature.
- Section 3 describes the model that will be used throughout this project to represent election instances.
- In Section 4 we describe the results of our comparison between the two systems for several voting rules
- In Section 5 we present our analysis of strategic candidacy games under the primary system.
- Section 6 presents a summary of the work that has been conducted in this project.

2 Related Work

2.1 Quality of Voting Rules

An important area of research in social choice theory concerns selecting appropriate voting rules for different applications, including traffic applications [Dennisen and Müller, 2015], constructing meta-search engines and reducing their spam [Dwork et al., 2001], movie recommender systems [Ghosh et al., 1999] and political elections (Gehrlein et al. [2017] provide a great overview, not limited to political elections).

For this purpose, there are several main lines of research. Firstly, there is the normative approach, which describes axioms that a voting rule should satisfy. Known voting rules are then compared against these axioms. Distinguished papers in this area are those of May [1952], Arrow et al. [1963], Gibbard [1973], Satterthwaite [1975] and Young [1975]. A general conclusion would be that there is no perfect voting rule. A famous result, known as "Arrow's impossibility theorem" states that no voting rule satisfies five reasonable conditions termed nondictatorship, Pareto dominance, unrestricted domain, social ordering and independence of irrelevant alternatives.

Secondly, a noise model can be used, where it is assumed that an optimal candidate exists and voters' votes coincide to noisy estimates of the optimal candidate. A voting rule then corresponds to the maximum likelihood estimate of the optimal candidate [de Caritat de Condorcet, 1785]. Conitzer and Sandholm [2012] investigate for which widely used voting rules there exists a noise model for which the maximum likelihood estimate is the rule itself. Another approach, called distance rationalizability, describes a voting rule in terms of a class of elections with a clear winner as well as a distance function [Meskanen and Nurmi, 2008; Elkind et al., 2009, 2010].

Finally, spatial models, where voters and candidates are situated into a metric space have also been studied extensively [Plott, 1967; Enelow and Hinich, 1984, 1990; Ordeshook and McKelvey, 1990; Schofield, 2007; Skowron and Elkind, 2017; Elkind et al., 2017; Anshelevich et al., 2018] and are a very popular choice, for their intuitive interpretation: the closer a voter is to a candidate, the more they prefer that candidate. The quality of the winner produced by a specific voting rule for a given election instance is measured by the sum of distances between them and each voter, which is related to their utilitarian social welfare. As a result, candidates with a lower utilitarian social cost are more preferred.

In this setting, it is natural to investigate the distortion of specific voting rules, a term initially defined by Procaccia and Rosenschein [2006] and subsequently investigated in the works of Feldman et al. [2016]; Goel et al. [2017]; Anshelevich et al. [2018]; Munagala and Wang [2019]; Gkatzelis et al. [2020]; Kempe [2020a,b], who all consider spatial models, as well as Caragiannis and Procaccia [2011]; Boutilier et al. [2015]; Caragiannis et al. [2017], who follow an older approach with normalised utilities. The distortion of a voting rule represents a worst-case notion; more precisely, it is the supremum of the ratio between the utilitarian social cost of the winning and optimal candidates of an election.

Notably, Anshelevich et al. [2018] (we note that this paper combines the results of two other papers, presented in 2015 and 2017, respectively) are the first to measure the distortion of widely known voting rules, including plurality, Borda, Copeland, STV. The Harmonic Borda voting rule, introduced by Boutilier et al. [2015] is also included in their analysis. The distortion of plurality and Borda is shown to be linear in the number of candidates; more precisely, $2m-1$, where m is the number of candidates, while the distortion of STV is upper bounded by $O(\ln m)$ and lower bounded by $O(\sqrt{\ln m})$. This means that STV yields much better results in the worst case. Moreover, the distortion of the Harmonic Borda voting rule is proved to be asymptotically better than those of plurality or Borda, although it is very close to being linear. However, all these voting rules are outperformed by Copeland, which has a distortion of 5. Lastly, the authors provide a lower bound of 3 for the distortion of any deterministic voting rule (plurality, Borda, Copeland and Harmonic Bords are all deterministic voting rules, while STV, in this paper, is presented as non-deterministic).

Initially, the Ranked Pairs voting rule was conjectured to achieve a distortion of 3. Goel et al. [2017] refute this conjecture about the distortion of the Ranked Pairs voting rule, which is shown to have a lower bound of 5 - in fact, Kempe [2020a] proved that its distortion is $\Theta(\sqrt{m})$. Before the findings of Munagala and Wang [2019] were presented, no deterministic voting rule with a distortion lower than 5 was known. However, in their paper, the authors design a new voting rule, which has a distortion of $2 + \sqrt{5}$, or approximately 4.236, making considerable progress in constructing a deterministic voting rule with a distortion of 3. With the help of the properties that a voting rule needs to satisfy in order to achieve a distortion of 3, outlined by Munagala and Wang [2019], Gkatzelis et al. [2020] successfully described a voting rule with a distortion of 3, termed "plurality matching".

As for randomised voting rules, Feldman et al. [2016] evaluate the distortion for randomised truthful voting mechanisms and come up with a tight bound of 2 on the line, as well as other lower bounds for more general metric spaces. Goel et al. [2017] also consider randomised tournament rules and provide lower bounds for their distortion. Anshelevich and Postl [2017] slightly improve the lowest value of the distortion of a voting rule, by showing that the Randomised Dictatorship voting rule, where a voter is randomly selected and their most preferred candidate is declared the winner, has a distortion of $3 - \frac{2}{n}$, where n is the number of voters. Because it is often the case that the number of candidates is smaller than the number of voters, the result of Kempe [2020b] also needs to be mentioned, as the author describes a randomised voting rule with a distortion of $3 - \frac{2}{m}$, where m is the number of candidates.

However, a different direction of work on the notion on distortion was firstly introduced by Procaccia and Rosenschein [2006] and subsequently considered in Caragiannis and Procaccia [2011]; Boutilier et al. [2015]; Caragiannis et al. [2017]. The approach in the mentioned papers is that, for each voter, utilities are assigned to candidates, depending on their preference. As a result, the notion of distortion can be analogously defined. Procaccia and Rosenschein [2006] proved that, under this setting, the Borda voting rule has an unbounded distortion. The distortion of the plurality voting rule was shown to be $\Omega(m^2)$ [Caragiannis and Procaccia, 2011] and in fact, it was also proven, by Caragiannis et al. [2017], that no deterministic voting rule can have a distortion that improves the bound of plurality, i.e. $\Omega(m^2)$. Lastly, Boutilier et al. [2015] investigated randomised voting mechanisms and described lower and upper bounds for their distortion, as well as an average-case model, for which voters' utilities for each candidate are drawn from a known distribution. They also showed that when the distribution is uniform over an interval, the optimal voting mechanism is precisely the Borda voting rule.

Coming back to spatial models, Elkind et al. [2017] analysed several multiwinner voting rules by performing various experiments in a two dimensional Euclidean model, with voters and candidates generated using four distributions. They looked at which voting rule would be best to be used, for the following scenarios : parliamentary elections, portfolio/movie selection or shortlisting.

For additional information on the notion of distortion, under both types of models described, we refer the reader to Anshelevich et al. [2021], who offer a very concise summary of the work that has taken place in this research area over the years.

All the above mentioned papers have in common the fact that they analyse direct elections, where voters submit their preferences and a candidate is declared the winner, depending on the voting rule used. A recent line of research focuses on primary systems [Borodin et al., 2019], a type of multi-stage elections, where voters and candidates are affiliated to one party and each party hosts its own primary election, where voters affiliated to that party vote over candidates from the same party and a primary winner is designated. Each primary winner then advances to the general election, where the party affiliation is no longer relevant: a voter will vote for the candidate they prefer the most.

Borodin et al. [2019] use a spatial model to define election instances, where voters and candidates are situated in a metric space. Three types of families of instances are considered, depending on the metric space. The most general type is an arbitrary metric space, followed by the metric space $\mathbb{R}^k$ with the Euclidean distance. The last type of family contains separable instances in $\mathbb{R}^k$, where both voters and candidates from opposing parties must be separated by an hyperplane. To perform a quantitative analysis of the direct and primary systems, the authors rely on the notion of distortion, over the three types of families of instances. What enables them to quantitatively compare the distortion of a voting rule is that once the set of voters and candidates, as well as their party affiliations are fixed, the voting rule is used to determine the winner in both the direct and the primary systems. Resulting winners can then be compared in terms of their utilitarian social cost. Important results are obtained in their theoretical analysis, which focuses on primaries with two parties.

Firstly, they show that any voting rule can have an unbounded distortion, for families of separable instances in $\mathbb{R}$. However, their result is restricted to the case where one of the parties has disproportionally less voters affiliated with it, i.e. the ratio between the number of voters associated with that party and the total number of voters tends to 0, as the number of voters tends to infinity. Next, a stronger result is related to individual instances, rather than families of instances. More specifically, if we set an election instance, an upper bound for the ratio between the utilitarian social cost of the winner and that of an optimal candidate in the primary system can be obtained. The upper bound depends on the ratio between the utilitarian social cost of the winning candidate and that of an optimal candidate in the primary of each party, which can be viewed as a direct election. This implies that the distortion of a voting rule under the primary system is also upper bounded by its distortion under the direct system, multiplied by a constant. This result holds for any of the three types of families of instances.

Secondly, while the previous results imply that the distortion under the primary system can only be a constant times higher than the distortion under the direct system, the authors also prove that the distortion under the direct system is upper bounded by the distortion under the primary system, for the first two types of families of instances, i.e. excluding separable instances. In other words, in terms of distortion of the voting rule used, the primary system will never be outperformed by a large margin by the direct system.

Lastly, the advantages of party separability in $\mathbb{R}$ are analysed: a voting rule is constructed such that its distortion under the primary system is upper bounded by a constant, while the distortion under the direct system is unbounded. This shows that primaries have potential to perform significantly better than direct elections, while the converse has already been shown not to be true.

Several experiments are carried out to measure the distortion of the plurality, Borda, STV, Copeland and maximin voting rules, by varying the value of k for the metric space $\mathbb{R}^k$ with the Euclidean distance, and for a fixed number of voters and candidates, with the voters uniformly distributed in $[0,1]^k$. In their average-case analysis, the primary system, for both separable and non-separable instances, produces a winner with a lower utilitarian social cost than the direct system, the only exception being plurality, for non-separable one dimensional instances, for which the direct system is better, by a very small margin. The results of the experiments suggest that a quantitative, instance-wise, analysis of the two systems, for separable instances, might yield interesting results. We show that even for very restricted elections, with both voters and candidates uniformly distributed, the direct system can still produce a winner with a lower utilitarian social cost than that of the winner produced by the primary system.

2.2 Strategic Candidacies

Another widely researched topic is that of strategic candidacies for direct elections, where candidates also have preferences over who can win the election. As a result, a candidate may prefer to join or withdraw from an election, if they are capable of manipulating the outcome of an election, such that the resulting winner is more preferred. Naturally, this sets the foundations of a non-cooperative strategic candidacy game, for which it is of high interest to characterize its pure strategy Nash equilibria (PNE).

Strategic candidacy games were introduced by Dutta et al. [2001], where they investigated whether the strategy profile with all candidates running is a pure strategy Nash equilibrium for a strategic candidacy game where candidates have self-supporting preferences (i.e. a candidate's most preferred outcome of an election is them winning). Their response is in the negative, for any unanimous (i.e. if all voters have the same most preferred candidate, then that candidate wins the election) and non-dictatorial (a voting rule is a dictatorship if the most preferred candidate of a voter, selected before looking at other voters' preferences, is declared the winner) voting rule.

Lang et al. [2013] build on the results of Dutta et al. [2001] and prove that a game with four candidates and an odd number of voters has at least one PNE, if a Condorcet-consistent voting rule (e.g. Copeland, maximin) or Borda is used, whereas for others voting rules, such as plurality, plurality with run-off or STV, the result no longer holds. The positive results for Borda and maximin do not generalise for games with more than four candidates, however, provided the number of voters is odd and the voting rule used is Copeland, any strategic candidacy game admits at least one PNE.

Polukarov et al. [2015] focus on equilibrium dynamics under plurality voting. In dynamic candidacies, given a strategy profile, which can be seen as the initial state of the dynamics, unless that profile is a PNE, there is at least one agent who would prefer to deviate from the profile, i.e. either join or withdraw from the election. This is defined as an improving move and a state is an equilibrium state if no agent has an improving move. An improving path is then a sequence of states, such that each state is obtained by an improving move from the previous one and they raise the question whether an improvement path can reach an equilibrium state, which is answered in the positive, as every improvement path is finite with probability 1. This result continues to hold for a more general setting, with refusing voters, i.e. voters that block their most preferred candidate if they withdraw from the election, and may only unblock them if all other running candidates are also blocked. Moreover, several decision problems are introduced, related to reachability properties of equilibrium dynamics, which are shown to be either NP-complete or NP-hard.

Obraztsova et al. [2015] consider lazy candidacies under plurality voting, where candidates prefer to withdraw if their participation cannot influence the winner of the election. They show that if a candidate is Pareto dominated by another candidate, who is preferred by a tie-breaking rule, then that candidate cannot be the winner in a PNE of the game. This highlights the importance of the tie-breaking rule and its possible impact over the outcome of an election. Other results are related to Condorcet winners: if the game admits a Condorcet winner, then the strategy profile with the Condorcet winner being the only candidate running is a PNE, if they have self-supporting preferences. However, they also give an example of a game that has a PNE where the Condorcet winner does not participate in the election. Two important decision problems, related to deciding whether a game has a PNE and a PNE with a specific candidate as the winner are shown to be NP-complete. Lastly, a weaker type of best-response dynamics than those considered in Polukarov et al. [2015] are studied, called J-dynamics and W-dynamics. In this setting, candidates are not allowed to join an election if they had previously withdrawn. As a result, the convergence of such dynamics is guaranteed in at most m steps, where m is the number of candidates. Examples of games where J-dynamics terminate in a state with all candidates running, or for which no W-dynamics converge to a state with the Condorcet winner running are presented. Lastly, four decision problems, related to these types of dynamics and their convergence to specific states of the game, are proven to be NP-complete.

Lang et al. [2019], motivated by the 2000 presidential elections in the United States between Bush, Al Gore and Nader, propose a different type of strategic candidates which they term keen candidacies. Candidates' preferences are defined in terms of utilities: the more a candidate is preferred, the higher their utility. Opposite to lazy candidacies, there is a participation bias which contributes to the overall utility a candidate receives from an election, so that a candidate will always prefer to run in the election, if their participation cannot influence the election winner. As a result, even for a small value of the participation bias, several results that are true for standard candidacy games no longer hold. Firstly, for games that admit a Condorcet winner, the strategy profile with the Condorcet winner being the only candidate running, is not a PNE, because, as the authors prove, in each PNE of keen candidacy games there must be at least two agents running. Secondly, there are games with four candidates that do not have a PNE, even if the voting rule used is Copeland, in contrast with the results obtained for standard candidacy games by Lang et al. [2013]. The same is also true if plurality is used instead of Copeland, in line with the properties of standard candidacy games. Next, multi-party elections with medium participation bias are considered, for which there are three properties that each PNE must satisfy. Moreover, an exponential upper bound for finding all equilibria is given. Intuitively, for large values of the participation, there is a unique PNE; more specifically, the strategy profile with all candidates running. In the last part of their paper, the authors look at the number of equilibria of keen strategic candidacy games, which decreases significantly, compared to the upper bound, even for small values of the participation bias. Lastly, an average case analysis is performed, to identify the number of equilibria for this type of games in practice, for various values of the participation bias. The two voting rules considered are plurality and Copeland and results are presented for games with five candidates, although for three to eight candidates, the authors claim to have very similar results. The trend of both voting rules is that, as the participation bias increases, the number of equilibria is almost always one. In fact, intuitively, in most of those cases, the strategy profile with all candidates running is the unique PNE. For plurality, there are a few more instances, for smaller values of the participation bias, that admit up to four PNE, however, for Copeland, irrespective of the value of the participation bias, almost all instances admit only one PNE.

Brill and Conitzer [2015] focus on single-peaked preferences, i.e. where voters and candidates can be positioned on a line, but in their analysis, strategic candidacies and strategic voting are combined. Importantly, without enforcing single-peaked preferences, they prove that computing the candidate stable set is NP-complete, when voting by succesive elimination is used.

Sabato et al. [2017] introduce a new model of strategic candidacies, called real candidacy games, where candidates have a continuous range of positions on a line, which can influence the preference of the voters. They analyse the existence of PNE, for different types of voting rules, including plurality and Condorcet-consistent ones, under various scenarios, as well as the impact of lexicographic and random tie-breaking rules.

3 Preliminaries and Model

We use similar notations to those in the related literature and, following the works of Borodin et al. [2019], we also use a spatial model. A useful observation is that such models are not restrictive by any means, as arbitrary preference orders over candidates can be expressed if the dimension of the metric space is large enough. We denote the set of voters by V and the set of candidates by A. Both voters and candidates are situated in the metric space (M, d), where M is a set and d is a distance function, $d : M \times M \to \mathbb{R}$, satisfying the following properties:

- Non-negativity: $\forall a, b \in M, d(a, b) \geq 0$
- Identity of indiscernibles: $\forall a, b \in M, d(a, b) = 0 \Leftrightarrow a = b$
- Symmetry: $\forall a, b \in M, d(a, b) = d(b, a)$
- Triangle inequality: $\forall a, b, c \in M, d(a, b) + d(b, c) \geq d(a, c)$

$\rho : V \cup A \to M$ positions voters and candidates in M.

The candidate preferences of a voter are entirely dependent on the distances between their position in the metric space and those of the candidates. Formally, a voter v prefers candidate a to candidate b, written $a \succ_v b$, if and only if $d(\rho(v), \rho(a)) < d(\rho(v), \rho(b))$ (we sometimes use $d(v, a)$ for the distance between the positions of v and a, for readability purposes). To address the existence of equidistant candidates for a voter, we use a tie-breaking rule $\triangleleft$, so that if $d(\rho(v), \rho(a)) = d(\rho(v), \rho(b))$ and $a \triangleleft b$, then $a \succ_v b$ and we say that "a is preferred to b by the tie-breaking rule". In other words, $\triangleleft$ is a priority order over the candidates set, A.

As mentioned, when dealing with primary systems, voters and candidates need to be affiliated to a party. We focus on the setting with two parties, denoted 1 and -1 and we use $\pi : V \cup A \to \{1, -1\}$ to assign parties. As a result, in the primary of each party, only voters affiliated with that party are able to vote for candidates from the same party.

Scores are assigned to each candidate and, depending on the voting rule used, a candidate is declared the primary winner. The winners from the primary of each party then advance to a general election, where the majority rule is used, as it is known that for two candidates, most voting rules are equivalent to the majority rule.

An instance of an election is then a tuple, $I = (V, A, M, d, \rho, \pi, \triangleleft)$ for which, irrespective of which of the two systems is used (primary or direct), a candidate will be declared the winner. If the direct system is used, we note that the party affiliation function π has no effect on the outcome of the election. The utilitarian social cost of a candidate $c \in A$ is defined as $C^I(a) = \sum_{v \in V} d(\rho(v), \rho(c))$. Moreover, for each instance I, there must exist at least one optimal candidate $a_{\text{OPT}} \in \arg\min_{a \in A} C^I(a)$, who has the lowest utilitarian social cost, out of all candidates in A. For simplicity, we omit the superscript I when describing the utilitarian social cost of a candidate if it is clear from the context to which election instance we are referring.

Since we will be focusing on an instance-wise comparison of the two systems, it suffices to define the distortion of a voting rule f as the ratio between the utilitarian social cost of the winner produced by f for an election instance I, and that of the optimal candidate, $\phi(f, I) = \frac{C^I(f(I))}{\min_{a \in A} C^I(a)}$, however, we once again remind the readers that distortion is usually a worst-case notion.

Lastly, we describe the voting rules that will be mentioned in the following sections:

- **Plurality:** Each voter gives one point to their most preferred candidate and the candidate who has received the highest number of votes in declared the winner. In the case that there are several candidates tied with the highest score, the tie-breaking rule $\triangleleft$ is used to declare a winner.
- **Anti-plurality:** Each voter gives one point to their least preferred candidate and the candidate with the lowest number of points received is declared the winner. In case of ties, the tie-breaking rule $\triangleleft$ is used to declare a winner.
- **Plurality with run-off:** Each voter gives one point to their most preferred candidate and the first two candidates, in terms of the number of points received advance

to the run-off, unless there is one candidate who has received more than 50% of the points, who is declared the winner. In the run-off, the voters vote for one of the two candidates and the candidate with the highest number of votes is declared the winner. Any ties are resolved according to the tie-breaking rule $\lhd$.

- **Borda:** Each voter gives $m - i$ points to the candidate ranked the i^{th} in their preferences, for $1 \leq i \leq m$, where m is the number of candidates. The candidate with the highest number of points is declared the winner. Ties are resolved according to the tie-breaking rule $\lhd$.

- **Harmonic Borda:** Similarly to Borda, each voter gives $\frac{1}{i}$ points to the candidate ranked the i^{th} in their preferences, for $1 \leq i \leq m$, where m is the number of candidates. The candidate with the highest number of points is declared the winner. Ties are resolved according to the tie-breaking rule $\lhd$.

- **Copeland:** The score of each candidate is the difference between the number of pairwise elections they win and the number of pairwise elections they lose. The candidate with the highest score is declared the winner. Any ties are resolved according to the tie-breaking rule $\lhd$.

- **Single Transferable Vote (STV):** Each voter gives one point to their most preferred candidate and if no candidate has received more than 50% of the points, the candidate with the lowest number of points is eliminated, with ties being resolved according to the tie-breaking rule $\lhd$. The voters who had previously given points to the eliminated candidate, give their point to their most preferred candidate who is not yet eliminated. The procedure finishes when there exists a candidate who has received more than 50% of the points.

An important concept in social choice theory is that of a Condorcet winner. A Condorcet winner is a candidate that wins against any other candidate in a pairwise election. We note that out of the seven voting rules we described, Copeland is the only one that is Condorcet-consistent, i.e. it always selects the Condorcet winner, if one exists. As a result, the Condorcet winner would obtain a score of $m - 1$, if the Copeland voting rule were to be used.

4 Comparison Between Primary and Direct Systems

In this section, we perform an instance-wise comparison between the two systems in one dimension, based on the utilitarian social cost of the winning candidates. We continue with an average-case analysis of the distortion of the described voting rules under the two systems, for which we also consider higher dimensions. More specifically, for the instance-wise comparison, the metric space used is $(\mathbb{R}, d)$, where d is the standard Euclidean distance. Most of the results are related to voting under plurality, however, we also consider Condorcet-consistent voting rules and make an interesting observation for STV.

To help with readability, we use $sc_g(a, \{a, b\})$ to denote the score of candidate a in a general election against candidate b. In the primary system, the general election will always be between at most two candidates. Similarly, we use $sc_p^{\pi(a)}(a)$, with $\pi(a) \in \{-1, 1\}$, to denote the score of candidate a in the primary of their party. For the direct system, we use $sc(a)$ to denote the score of candidate a. Moreover, when we write that one of the systems is better than the other, we mean that the utilitarian social cost of the winner produced by the former is lower than the utilitarian social cost of the winner of the latter.

4.1 Voting under Plurality

It is relatively easy to see that the primary system can produce a winner with a higher utilitarian social cost than the direct system, and the instance from Figure 1 is a good example, with $A = \{a_1, a_2, a_3, a_4\}, V = \{v_1, v_2, v_3, v_{-1}, ..., v_{-k-1}\}, \pi(v_1) = \pi(v_2) = \pi(v_3) = \pi(a_1) = \pi(a_2) = 1, \pi(v_{-i}) = \pi(a_3) = \pi(a_4) = -1$, for $1 \leq i \leq k+1$. We let $\rho(v_1) = 0$ and we omit to write down the remaining positions, as they can be clearly deduced from the annotations in Figure 1.

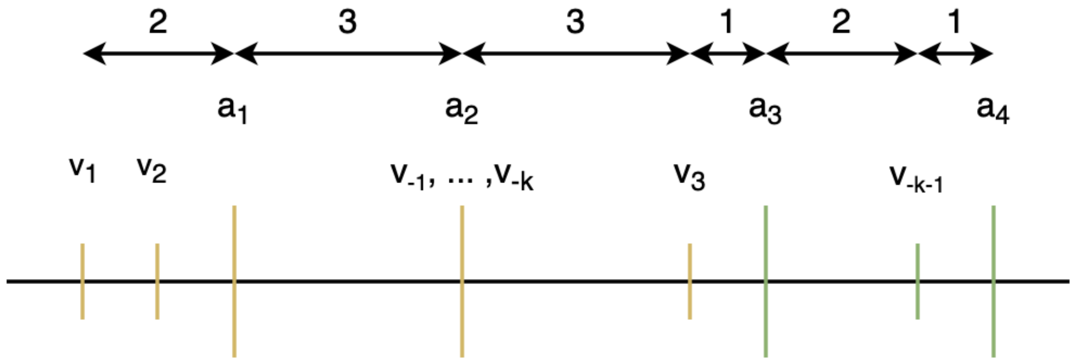


Figure 1: Direct elections can be better

In the primary system, $sc_p^1(a_1) = 2$, $sc_p^1(a_2) = 1$, so a_1 wins the primary for party 1 and $sc_p^{-1}(a_3) = k$, $sc_p^{-1}(a_4) = 1$ and a_3 advances to the general election from party -1. Because $sc_g(a_1, \{a_1, a_3\}) = k + 2$ and $sc_g(a_3, \{a_1, a_3\}) = 2$, a_1 becomes the winner of the general election. In the direct system, a_2 clearly wins, as they receive k out of the $k + 4$ possible votes. However, $C(a_1) = 2+1+3k+6+9 = 3k+18$ and $C(a_2) = 5+4+3+6 = 18$.

Therefore, we consider the following, more restrictive, settings for which we also require party separability:

- one of the parties is represented by only one candidate,
- voters and candidates are uniformly distributed,
- a particular case of the uniformly distributed scenario.

4.1.1 Primaries are not Necessarily Better if one Party has Exactly 1 Candidate

In the setting where party separability is enforced and one of the parties is represented by only one candidate, we are interested in evaluating whether the primary system always produces a winner with a lower utilitarian social cost than the direct system. For this purpose, we formulate the following conjecture, which we refute by the means of a counter-example:

Conjecture 1. *For separable election instances, if there exists only one candidate associated to one of the parties, the winner in the primary system always has a lower utilitarian social cost than that of the winner in the direct system.*

Refutation of Conjecture 1. Consider the counter-example from Figure 2, for which $V = \{v_1, ..., v_{k+1}, v_{-1}, ..., v_{-k}\}, A = \{a_1, a_2, a_{-1}\}$, with $\pi(a_2) = \pi(a_1) = 1, \pi(a_{-1}) = -1; \pi(v_1) = ... = \pi(v_{k+1}) = 1, \pi(v_{-1}) = ... = \pi(v_{-k}) = -1$ and $\rho(a_2) = 0, \rho(v_{k+1}) = -\epsilon, \rho(v_{-1}) = ... = \rho(v_{-k}) = \epsilon, \rho(v_1) = ... = \rho(v_k) = -1 - \epsilon, \rho(a_1) = -1 - 2\epsilon, \rho(a_{-1}) = 1 + 2\epsilon$:

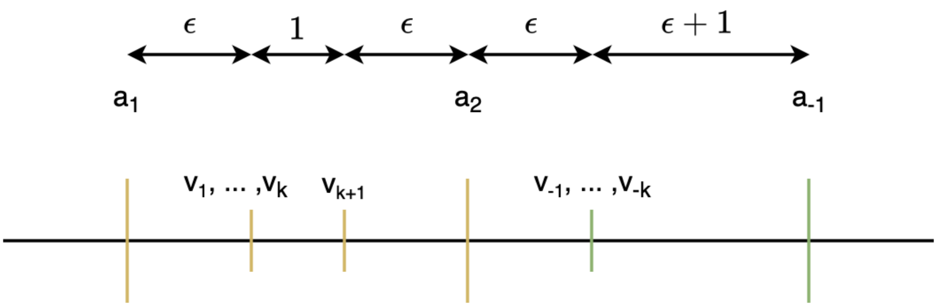


Figure 2: Direct elections are better if one party has exactly 1 candidate

In the primary system, candidate a_{-1} is the only candidate from party -1, so they win the primary. For the primary of party 1, $sc_p^1(a_1) = k, sc_p^1(a_2) = 1$, so candidate a_1 wins the primary. In the general election between a_1 and $a_{-1}, sc_g(a_1, \{a_1, a_{-1}\}) = k + 1, sc_g(a_{-1}, \{a_1, a_{-1}\}) = k$ and a_1 wins the general election.

If we consider the direct election, $sc(a_1) = k, sc(a_2) = 1 + k, sc(a_{-1}) = 0$ and a_2 would be the winner.

Lastly, let's compare the utilitarian social costs of the two winners:

- $C(a_1) = k\epsilon + 1 + \epsilon + k(1 + 3\epsilon) = (4k + 1)\epsilon + k + 1$
- $C(a_2) = (k + 1)\epsilon + k(1 + \epsilon) = (2k + 1)\epsilon + k$

Clearly $C(a_2) < C(a_1)$, so the direct system would be better in this case. □

4.1.2 Primaries are not Necessarily Better if the Voters and Candidates are Uniformly Distributed

Just as in the previous case, the aim of this section is to investigate whether the primary system is always better than the direct system, if the voters and candidates are uniformly distributed and the two parties are separable. For $V = \{v_1, v_2, ..., v_n\}$, we say that voters are uniformly distributed if there exists $\delta > 0$, such that $\rho(v_i) = i\delta, \forall i \in \{1, 2, ..., n\}$. Similarly, for $A = \{a_1, a_2, ..., a_m\}$, candidates are uniformly distributed if there exists $\epsilon > 0$, such that $\rho(a_i) = i\epsilon, \forall i \in \{1, 2, ..., m\}$. Firstly, we formulate a conjecture, which we, once again, refute. Then, we further restrict the setting and perform a quantitative instance-wise comparison between the utilitarian social cost of the winners in the two systems.

Conjecture 2. *For separable election instances, if the voters and candidates are uniformly distributed, the utilitarian social cost of the winner in the primary system is lower than that of the winner in the direct system.*

Refutation of Conjecture 2. We provide a counter-example, illustrated in Figure 3, for which $V = \{v_1, v_2, ..., v_8\}, A = \{a_1, a_2, a_3\}$, with $\pi(a_1) = \pi(a_2) = \pi(v_1) = ... = \pi(v_5) = 1, \pi(a_3) = \pi(v_6) = \pi(v_7) = \pi(v_8) = -1$. Because voters are uniformly distributed, we have $\rho(v_i) = i\delta$, for some $\delta > 0$ and for the candidates we can have $\rho(a_1) = 2.5\delta, \rho(a_2) = 5\delta, \rho(a_3) = 7.5\delta$. Clearly, we can choose $\epsilon = 2.5\delta$ such that $\rho(a_i) = i\epsilon$, so the candidates are also uniformly distributed. Lastly, assume $a_2 \lhd a_1 \lhd a_3$.

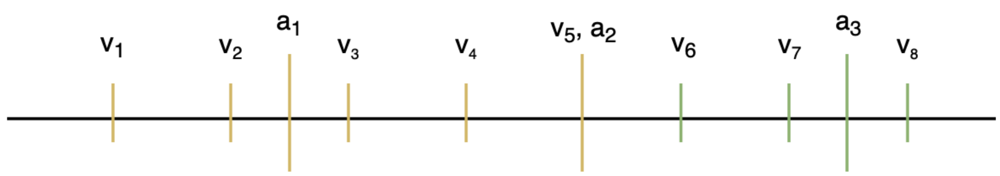


Figure 3: Direct elections are better even if voters and candidates are uniformly distributed.

In the primary system, a_3 is the only candidate from party -1, so they advance to the general election. For the primary of party 1, $sc_p^1(a_1) = 3, sc_p^1(a_2) = 2$, so a_1 wins the primary. In the general election, $sc_g(a_1, \{a_1, a_3\}) = 5, sc_g(a_3, \{a_1, a_3\}) = 3$ and a_1 becomes the winner.

In a direct election, $sc(a_1) = 3, sc(a_2) = 3, sc(a_3) = 2$ and because $a_2 \lhd a_1$, a_2 would win. Let's compare the utilitarian social costs of the two winners:

- $C(a_1) = 1.5\delta + 0.5\delta + 0.5\delta + 1.5\delta + 2.5\delta + 3.5\delta + 4.5\delta + 5.5\delta = 20\delta$
- $C(a_2) = 0 + 2\delta + 4\delta + 6\delta + 4\delta = 16\delta$

Clearly, $C(a_2) < C(a_1)$ and again, the direct system would be better.

$\square$

We further restrict this setting, so that for each candidate, there is a co-located voter. This is a plausible scenario, as in most real-life elections, candidates are able to vote for themselves. Formally, let $V = \{v_1, ..., v_n\}$ be the set of voters, with $v_1, ..., v_{n_1}$ being affiliated with party 1 and the others with party -1 and $A = \{a_1, ..., a_m\}$ be the set of candidates, with $a_1, ...a_{m_1}$ being affiliated with party 1 and the others with party -1. Assume candidates and voters are uniformly distributed, with $\rho(v_i) = i \times \delta$ and $\rho(a_i) = i \times k \times \delta$, for some $\delta > 0$ and $k \in \mathbb{N}$. We can also assume that candidate a_m would receive at least 1 vote in the direct election (otherwise, they would not receive any votes in the primary system either and would essentially be a null candidate).

We now perform a quantitative comparison between the two systems. We distinguish the following two cases, depending on the parity of k:

Case 1: k is even, $k = 2p$

We start by presenting two straightforward observations, which greatly help us in narrowing down the possible winners in the two systems.

Observation 1. *A candidate a_i with $1 < i < m$ can have at most $2p + 1$ votes in the direct system. The candidate will have $2p + 1$ votes only if $a_i \lhd a_{i-1}$ and $a_i \lhd a_{i+1}$.*

Proof. For the voters v_j who vote for candidate a_i, the following must hold: $|j\delta - 2pi\delta| \leq p\delta \Leftrightarrow -p\delta \leq j\delta - 2pi\delta \leq p\delta \Leftrightarrow p(2i-1) \leq j \leq p(2i+1)$, and there are at most $(2i+1)p - (2i-1)p + 1 = 2p+1$ votes for a_i. $\square$

Observation 2. *In the direct system, the first candidate, a_1, located at position $2p\delta$ will have at least $2p+2$ votes for $p > 2$. In fact, they will have exactly $3p-1$ or $3p$ votes.*

Proof. Clearly, all the voters at position $\delta, 2\delta, ..., 2p\delta$ will vote for the candidate at position $2p\delta$. Moreover, if $p > 2$, voters at positions $(2p+1)\delta$ and $(2p+2)\delta$ will also vote for candidate a_1, because $d(a_1, v_{2p+1}) = (2p+1)\delta - 2p\delta < d(a_1, v_{2p+2}) = (2p+2)\delta - 2p\delta = 2\delta < (2p-2)\delta = 4p\delta - (2p+2)\delta = d(a_2, v_{2p+2}) < d(a_2, v_{2p+1})$. The same argument can be used to show that, in fact, voters $v_{2p+1}, ..., v_{3p-1}$ will surely vote for a_1, while the vote of v_{3p}, with $d(a_1, v_{3p}) = 3p\delta - 2p\delta = \delta = 4p\delta - 3p\delta = d(a_2, v_{3p})$, will go to either a_1 or a_2, depending on the tie-breaking rule. $\square$

From Observations 1 and 2, it follows that the winner in the direct system would either be the first candidate, a_1, or the last candidate, a_m.

Let's now analyse the primary system. Clearly, Observation 2 still holds for candidate a_1 for both the primary and the general election, should a_1 be the primary winner. Observation 1 also holds for candidates $a_2, ..., a_{m_1-1}$, participating in the primary, so the winner of the primary for party 1 can only be a_1 or a_{m_1}. However, candidate a_{m_1} can have at most $3p+1$ votes and it can have as few as $p-1$ votes. A more tedious case analysis is now required, depending on the number of votes a_{m_1} receives in their party's primary.

1. a_{m_1} has $3p+1$ votes. This only happens when $\rho(a_{m_1+1}) = \rho(v_{n_1})$ and $a_{m_1} \lhd a_{m_1-1}$. In this case, a_{m_1} would win the primary for party 1.

   i. If $n > 2mp + 2p$ candidate a_m would get at least $p - 1 + (2mp + 2p - 2mp + 1) + 1 = 3p + 1$ votes and would win the primary for party -1 and would also win against candidate a_1 in the direct system. In the general election between a_m and a_{m_1}, $sc_g(a_{m_1}, \{a_{m_1}, a_m\}) \leq n_1 - 2p + mp - m_1p$ and $sc_g(a_m, \{a_{m_1}, a_m\}) \geq$

$mp - m_1p - 1 + n - 2mp + 1$. $sc_g(a_m, \{a_{m_1}, a_m\}) - sc_g(a_{m_1}, \{a_{m_1}, a_m\}) \geq n - 2mp - n_1 + 2p > 0 \Leftrightarrow n > n_1 + 2p(m-1)$. In this case the winner in both systems would be a_m and there would be no difference between them. Moreover, if $a_m \lhd a_{m_1}$, for $n = n_1 + 2mp - 2p$ or $n = n_1 + 2mp - 2p - 1$ or $n = n_1 + 2mp - 2p - 2$ the winner would still be a_m. If $a_{m_1} \lhd a_m$ and $n = n_1 + 2mp - 2p$, then $sc_g(a_m, \{a_{m_1}, a_m\}) = sc_g(a_{m_1}, \{a_{m_1}, a_m\})$ and a_{m_1} would win the general election. But in this particular case $C(a_{m_1}) - C(a_m) = 2p\delta(m - m_1)$ and so the primary system would produce a winner with a higher utilitarian social cost - an example is shown in Figure 4. However, if $n \leq n_1 + 2p(m-1) - 1 = 2mp + 2m_1p - 1$, then the winner in the primary system would be a_{m_1}. But $C(a_{m_1}) = \delta(2m_1p - 1 + ... + 0 + 1 + 2 + ... + n - 2m_1p) = \delta\left(\frac{(2m_1p-1)2m_1p}{2} + \frac{(n-2m_1p)(n-2m_1p+1)}{2}\right)$ and $C(a_m) = \delta(2mp - 1 + ... + 0 + 1 + ... + n - 2mp) = \delta\left(\frac{(2mp-1)2mp}{2} + \frac{(n-2mp)(n-2mp+1)}{2}\right)$. $C(a_m) - C(a_{m_1}) = 2\delta(m - m_1)\left[2p^2(m + m_1) - p - np\right] \geq 0$, because clearly $m > m_1$ and $n \leq 2mp + 2m_1p - 1 \Leftrightarrow np \leq 2p^2(m + m_1) - p \Leftrightarrow 2p^2(m + m_1) - np - p \geq 0$, so in this case the winner in the primary system would not have a higher utilitarian social cost than the winner in the direct system.

ii. For $n = 2mp + 2p, n = 2mp + 2p - 1, n = 2mp + 2p - 2$, a_m would win the primary for party -1, but they could lose against a_1 in the direct election. a_m would only win against a_{m_1} in the general election if $m_1 = 1$, because for $m_1 > 1$ we would need $n \geq 2mp + 2m_1p - 2 > 2mp + 2p$. If $m_1 > 1$, $C(a_m) - C(a_{m_1}) = 2p\delta(m - m_1)(2mp + 2m_1p - n - 1) > 0$, because $m > m_1$ and $n \leq 2mp + 2p < 2mp + 2m_1p \Rightarrow 2mp + 2m_1p - n - 1 \geq 0$ and $C(a_1) - C(a_{m_1}) = 2p\delta(1 - m_1)(2pm_1 + 2p - 1 - n) > 0$, because $0 < 1 - m_1$ and $n > 2pm \geq 2pm_1 + 2p \Rightarrow 2pm_1 + 2p - n - 1 < 0$. So a_{m_1} has a lower utilitarian social cost than the possible winners in the direct system. For $m_1 = 1$, if $n = 2mp + 2p$, then $C(a_1) < C(a_m)$ and the primary system would produce a worse outcome only if a_1 would be the winner in the direct system and a_m would be the winner in the primary system. However, this is not possible, because for a_m to win the general election, it must hold that $a_m \lhd a_1$ and in the direct system a_m would have at least as many votes as a_1 and would also win in case of a tie, so they would also win in the direct system. If $n = 2mp + 2p - 1$, then $C(a_1) = C(a_m)$ and the two systems produce winners with the same utilitarian social cost. If $n = 2mp + 2p - 2$, $C(a_m) < C(a_1)$ and the primary system would produce a worse outcome only if a_m were the winner in the direct system and a_1 the winner in the primary system. Again, this is not possible, because a_1 would only win the general election if $a_1 \lhd a_m$, but then, because a_1 would have at least as many votes as a_m in the direct system, they would also win the direct election.

iii. If $n \leq 2mp + 2p - 3$, candidate a_m would lose against candidate a_1 in the direct system. If $n > 2mp + p + 2$, a_m would still win the primary of party -1, as it would have at least $2p + 2$ votes. In the general election $sc_g(a_{m_1}, \{a_{m_1}, a_m\}) - sc_g(a_m, \{a_{m_1}, a_m\}) \geq 2pm_1 + mp - pm_1 - 1 - (mp - pm_1 + 1 + n - 2mp) = 2mp + 2m_1p - n - 2 \geq 2mp + 2m_1p - 2mp - 2p + 3 - 2 = 2m_1p - 2p + 1 > 0$ so a_{m_1} would win the general election. $C(a_1) = \delta(2p - 1 + ... + 1 + 0 + 1 + ... + n - 2p) = \delta\left(\frac{(2p-1)2p}{2} + \frac{(n-2p)(n-2p+1)}{2}\right)$ and $C(a_1) - C(a_{m_1}) = \delta(8p^2 - 4p - 4np + 4m_1p + 4nm_1p - 8{m_1}^2p^2) = 4p\delta\left[2p(1 - m_1^2) - (1 - m_1) - n(1 - m_1)\right] = 4p\delta(1 - m_1)(2p + 2m_1p - n - 1) > 0$. because $1 - m_1 \leq 0$ and $n > 2mp \geq 2m_1p + 2p \Leftarrow 2m_1p + 2p - n - 1 < -1 < 0$, so $C(a_1) \geq C(a_{m_1})$ and the primary system would again produce a better result. Lastly, if $n \leq 2mp + p + 2$, assume the winner of the primary for party -1 is a_i, with $m_1 < i \leq m$. If a_{m_1} were to win in the general election against a_i, we have already shown that $C(a_1) > C(a_{m_1})$ (this is because it continues to hold that $n \geq 2m_1p + 2p$), so we are only interested in the case where a_i wins in the general election. $C(a_i) = \delta(2ip - 1 + ... + 1 + 0 + 1 + ... + n - 2ip) = \delta\left(\frac{(2ip-1)2ip}{2} + \frac{(n-2ip)(n-2ip+1)}{2}\right)$ and $C(a_1) - C(a_i) = \frac{\delta}{2}(4ip + 4nip - 8i^2p^2 + 8p^2 - 4p - 4np) = 2p\delta\left[2p(1 - i)(1 + i) - n(1 - i) - (1 - i)\right] = 2p\delta(1 - i)(2p + 2pi - n - 1) > 0$, because $1 - i < 0$ and for $m_1 > 1, n \geq 2ip + 2m_1p - 2 \geq 2ip + 4p - 2 > 2ip + 2p - 1 \Leftarrow 2p + 2ip - n - 1 < 0$ (the inequality $n \geq 2ip + 2m_1p - 2$ follows from the fact that for a_i to win against a_{m_1}, it must hold that $n \geq 2ip + 2m_1p - 2$). If $m_1 = 1$, for $n = 2pi + 2p - 2, m = i + 1$ and $a_i \lhd a_1$, a_i would win in the general election against a_1, but $C(a_i) - C(a_1) = 2p\delta(i - 1)$ and the primary system would produce a worse outcome - an example is shown in Figure 5.

2. a_{m_1} has $3p$ votes. This happens when $\rho(a_{m_1+1}) = \rho(v_{n_1})$ and $a_{m_1-1} \lhd a_{m_1}$ or $\rho(a_{m_1+1}) = \rho(v_{n_1+1})$ and $a_{m_1} \lhd a_{m_1-1}$.

   i. If a_1 also has $3p$ votes, i.e. $a_1 \lhd a_2$, and $a_1 \lhd a_{m_1}$ then a_1 wins the primary for party 1. If $n > 2mp + 2p$, a_m would win the primary for party -1, the general election against a_1 and they would also win in a direct election, so the two systems produce the same winner. For $n = 2mp + 2p, n = 2mp + 2p - 1, n = 2mp + 2p - 2$ the analysis is identical to the one before, for $m_1 = 1$, for which we have shown that the primary system cannot produce a worse outcome. If $n \leq 2mp + 2p - 3$, a_1 would win in the direct election. Moreover, if a_i were the primary winner of party -1, $m_1 < i \leq m$, then for $n = 2pi + 2p - 2$, $m = 2pi + 2p$ and $a_i \lhd a_1$, a_i would win in the general election against a_1 and $C(a_i) - C(a_1) = 2p\delta(i - 1)$, resulting in a worse outcome. For $n > 2pi + 2p - 2$, $C(a_1) - C(a_i) = 2p\delta(1 - i)(2p + 2pi - n - 1) > 0$, because $0 < 1 - i$ and $n \geq 2pi + 2p - 1 \Leftarrow 0 \geq 2pi + 2p - n - 1$. If $a_{m_1} \lhd a_1$, then a_{m_1} would win the primary for party 1 and the analysis is identical to the previous case.

   ii. If a_1 has less than $3p$ votes, then candidate a_{m_1} would win the primary for party 1 and the analysis is identical to 1.

3. a_{m_1} has $3p - 1$ votes. This happens when $\rho(a_{m_1+1}) = \rho(v_{n_1+1})$ and $a_{m_1-1} \lhd a_{m_1}$ or $\rho(a_{m_1+1}) = \rho(v_{n_1+2})$ and $a_{m_1} \lhd a_{m_1-1}$.

   If a_1 has $3p$ votes or a_1 has $3p - 1$ votes, i.e. $a_2 \lhd a_1$ and $a_1 \lhd a_{m_1}$, then a_1 wins the primary for party 1 and the analysis is the same as 2.i. Similarly, if a_1 has $3p - 1$ votes and $a_{m_1} \lhd a_1$ the analysis is identical to 1.

4. a_{m_1} has less than $3p - 1$ votes. Then a_1 would win the primary for party -1.

   i. If $n > 2mp + 2p + 2$, candidate a_m wins in both systems.

   ii. If $n = 2mp + 2p + 2$ or $n = 2mp + 2p + 1$ or $n = 2mp + 2p$, and $C(a_1) - C(a_i) = 2p\delta(1 - i)(2p + 2pi - n - 1) > 0$, because $1 - i < 0$ and $2p + 2pi - n - 1 \leq 2pi - 2pm - 1 <$

$0, \forall i. m_1 < i \leq m$. $C(a_m) - C(a_i) = 2p\delta(m-i)(2pm + 2pi - 1 - n) > 0$, because $m - i > 0$ and $2pm + 2pi - 1 - n \geq 2pm + 2pi - 1 - 2pm - 2p - 2 > 0, \forall i. m_1 < i \leq m$. So, the primary system does not produce a winner with a higher utilitarian social cost.

iii. If $n = 2mp + 2p - 1$, clearly $C(a_1) = C(a_m)$, so the winner in the direct system does not matter and we are only interested in the possibility of candidate a_i, $m_1 < i < m$ winning in the primary system. $C(a_1) - C(a_i) = 2p\delta(1-i)(2p + 2pi - n - 1) = 2p\delta(1-i)(2pi - 2pm) > 0$, because $1 - i < 0$ and $i < m \Rightarrow 2pi - 2pm < 0$. So, the primary system does not produce a winner with a higher utilitarian social cost.

iv. If $n = 2mp + 2p - 2$, $C(a_1) - C(a_i) = 2p\delta(1-i)(2p + 2pi - n - 1) = 2p\delta(1-i)(2pi - pm + 1) < 0, \forall i, m_1 < i < m$. For $i = m$, candidate a_m must win in both the primary and the general election against a_1. This only happens if $a_m \lhd a_1$ and $a_m \lhd a_{m-1}$, but in this case a_m would also win in the direct system. $C(a_m) - C(a_i) = 2p\delta(m-i)(2pm + 2pi - 1 - n) = 2p\delta(m-i)(2pi - 2p + 1) > 0, \forall i, m_1 < i \leq m$. Again, the primary system does not produce a winner with a higher utilitarian social cost.

v. If $n \leq 2mp + 2p - 3$, a_m would lose against candidate a_1 in the direct system and in a possible general election of the primary system. Suppose candidate a_i, $m_1 < i < m$ wins the primary for party -1 and the general election against a_1, then $n \geq 2pi + 2p - 2$. $C(a_1) - C(a_i) = 2p\delta(1-i)(2p + 2pi - n - 1) > 0$, unless $n = 2pi + 2p - 2$. However, for $n = 2pi + 2p - 2, m = i + 1$ such that a_i wins the primary for party -1 and the general election against a_1, with $a_i \lhd a_1, C(a_i) > C(a_1)$, resulting in a winner with a higher utilitarian social cost than the direct system winner. Note that in order for a_i to win the primary for party -1, candidate a_{m_1+1} must have at most as many votes as a_i in the primary. This is only influenced by the position of voter v_{n_1} and we would need to require that $\rho(v_{n_1}) \geq 2pm_1 + 3p$ as well as possibly enforce that a_i wins in case of a tie against a_{i-1} or a_{m_1+1}.

Case 2: k is odd, $k = 2p + 1$.

We formulate similar observations to those for the even k case.

Observation 3. *A candidate a_i, with $1 < i < m$ will have exactly 2p+1 votes in the direct system.*

Proof. A voter v_j will vote for a_i only if $\left|j\delta - (2p+1)i\delta\right| \le p\delta \Leftrightarrow -p\delta \le j\delta - 2pi\delta - i\delta \le p\delta \Leftrightarrow 2pi+i-p \le j\ 2pi+i+p$, and there are at most $2pi+i+p-2pi-i+p+1 = 2p+1$ votes for a_i. But because $d(a_i, v_{2pi-p+i}) = \delta(2pi + i - 2pi + p - i) = p\delta$ and $d(a_{i-1}, v_{2pi-p+i}) = \delta(2pi-p+i-2pi+2p-i+1) = (p+1)\delta > p\delta$ and similarly $d(a_i, v_{2pi+p+i}) < d(a_{i+1}, v_{2pi+p+i})$, a_i must have exactly $2p + 1$ votes. □

Observation 4. *In the direct system, candidate a_1, located at position $(2p+1)\delta$ will have exactly $3p + 1$ votes.*

Proof. Voters at positions $\delta, 2\delta, ..., (2p+1)\delta$ will clearly vote for the candidate at position $(2p+1)\delta$. Because $d(a_1, v_{3p+1}) = (3p+1)\delta - (2p+1)\delta = p\delta < (p+1)\delta = (4p+2)\delta - (3p+1)\delta = d(a_2, v_{3p+1})$, voters $v_{2p+2}, ..., v_{3p+1}$ will also vote for a_1 in the direct system. So a_1 will have precisely $3p + 1$ votes. □

From Observation 3 and 4, it follows that the winner in the direct system will either be the first candidate, a_1, or the last candidate a_m. Similarly to the previous case, Observation 3 continues to hold for candidates $a_2, ..., a_{m_1-1}$ in the primary of party 1, so the winner of the primary can only be a_1 or a_{m_1}.

The analysis for this case is very similar to the even k case and we obtain the following instances for which the primary system produces a winner with a higher utilitarian social cost than the winner in the direct system:

1. $m_1 = 1, n = (2p + 1)m - 2, a_{m-1} \lhd a_1, m$ even and a_{m-1} winning the primary for party -1. For this to happen, we require that $\rho(v_{n_1}) \ge (3p + 1)\delta$ (this ensures that candidate a_2 can get at most $2p + 1$ votes) and $\forall i, 1 < i < m - 1$, if

$sc_p^{-1}(a_i) = 2p + 1$, then $a_{m-1} \lhd a_i$. This is because in the primary for party -1, a_{m-1} will get exactly $2p + 1$ votes and they need to win in case of ties with other candidates. So a_{m-1} wins the primary for party -1 and in the general election against $a_1, sc_g(a_1, \{a_1, a_{m-1}\}) = \frac{(2p+1)m}{2} - 1 = \frac{2pm+m-2}{2} = \frac{4pm+2m-4-2pm-m+2}{2} = \frac{2n-(2pm+m-2)}{2} = n - \frac{2pm+m-2}{2} = sc_g(a_{m-1}, \{a_1, a_{m-1}\})$ and because $a_{m-1} \lhd a_1, a_{m-1}$ would win the in general election. Clearly, a_1 would win in the direct election, as it would have $3p+1$ votes, and no other candidate would have more than $2p+1$ votes. Lastly, substituting $(2p+1) = k, C(a_1) = \delta(k-1+....+1+0+1+...+mk-2-k) = \delta\left(\frac{(k-1)k}{2} + \frac{(mk-k-2)(mk-k-1)}{2}\right), C(a_{m-1}) = \delta[k(m-1)-1+...+1+0+1+...+(k-2)] = \delta\left(\frac{(km-k-1)(km-k)}{2} + \frac{(k-2)(k-1)}{2}\right)$ and $C(a_{m-1}) - C(a_1) = \delta k(m-2)$. Note that we require m to be even, because otherwise, in the general election, candidate a_{m-1} would have 1 vote less than a_1 and could not beat them.

2. $\rho(v_{n_1}) = \rho(a_{m_1+1}) \Leftrightarrow n_1 = km_1+k, k = 2p+1, n = n_1+km-k = km+km_1$, m_1 and m have the same parity, $a_{m_1} \lhd a_m$. In this case, a_m clearly wins in the direct election, as well as the primary for party -1. Candidate a_{m_1} wins the primary for party 1, because a_1 has $3p+1$ votes and a_{m_1} has $3p+2$ votes. In the general election, a_m and a_{m_1} will have the same number of votes and because $a_{m_1} \lhd a_m$, a_{m_1} would win in the primary system. Note that if m and m_1 had different parities, a_m would have one more vote than a_{m_1} in the general election and would win. The parity restriction also ensures that a_{m_1} gets the vote from $v_{\frac{m_1+m}{2}}$ due to the tie-breaking rule, because $d(a_m, v_{\frac{m_1+m}{2}}) = d(a_{m_1}, v_{\frac{m_1+m}{2}})$. $C(a_m) = \delta(mk-1+...+1+0+1+...+n-mk) = \delta\left(\frac{(mk-1)mk}{2} + \frac{(n-mk)(n-mk+1)}{2}\right), C(a_{m_1}) = \delta(m_1k-1+...+1+0+1+...+n-m_1k) = \delta\left(\frac{(m_1k-1)m_1k}{2} = \frac{(n-m_1k)(n-m_1k+1)}{2}\right)$ and $C(a_{m_1}) - C(a_m) = \delta k(m_1-m)(km_1+km-n-1) = \delta k(m-m_1)$.

3. $\rho(v_{n_1+1}) = \rho(a_{m_1+1}) \Leftrightarrow n_1 = km_1 + k - 1, n = n_1 + km - k + 1, m_1$ and m have the same parity, $a_{m_1} \lhd a_1, a_{m_1} \lhd a_m$. The condition $\rho(v_{n_1+1}) = \rho(a_{m_1+1})$ implies that candidates a_1 and a_{m_1} both receive $3p + 1$ votes in the primary for party 1, and the condition $a_{m_1} \lhd a_1$ ensures that a_{m_1} wins the primary. Moreover, we require m and m_1 to have the same parity so that in the general election they have the same number of votes and, the tie-breaking rule will make a_{m_1} win against a_m in the

general election. Similarly, as above, we obtain $C(a_{m_1}) - C(a_m) = \delta k(m - m_1)$.

In our quantitative analysis, we have distinguished two main cases depending on the parity of k. We have showed that, no matter whether k is even or odd, election instances where the direct system is better than the primary system still exist. However, such cases arise rather sparsely. We now present two examples of instances where the direct system outperforms the primary one.

For the example in Figure 4 we have the following setting: $\delta = 1, k = 8, p = 4, n_1 = 24, m_1 = 2, m = 4, n = 48$. The voters and candidates from party 1 are coloured in yellow, while the voters and candidates from party -1 are coloured in green. We also require $a_2 \triangleleft a_4$, $a_2 \triangleleft a_1$ and note that at v_{24} and a_3 are located at the same position, with $\pi(v_{24}) = 1, \pi(a_3) = -1$.

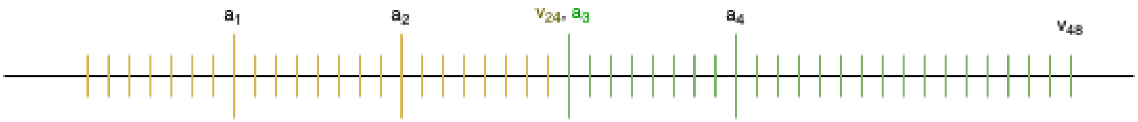


Figure 4: Example where the primary system produces a worse outcome for an even k and $m_1 > 1$

In the direct system a_4 clearly wins. a_2 wins the primary for party 1, because $sc_p^1(a_2) = 13$, while $sc_p^1(a_1) = 11$ and a_4 wins the primary for party -1. In the general election between a_2 and a_4, although $sc_g(a_2, \{a_2, a_4\}) = sc_g(a_4, \{a_2, a_4\}) = 24$, a_2 ends up winning, due to the tie-breaking rule. But $C(a_4) = 632$ and $C(a_2) = 648$.

Similarly to the previous example, in Figure 5, we have $\delta = 1, k = 8, p = 4, n_1 = 16, m_1 = 1, m = 4, n = 30$ and the voters and candidates are coloured as before. We require $a_3 \triangleleft a_1$.

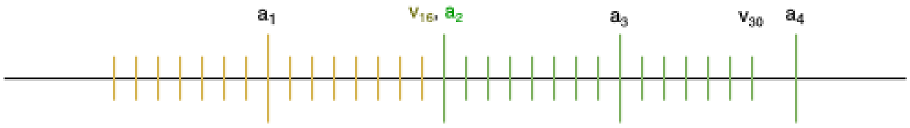


Figure 5: Example where the primary system produces a worse outcome for an even k and $m_1 = 1$

a_1 wins in the direct election, because they get at least 11 votes and no other candidate can get more than 9 votes and they are also uncontested in the primary for party 1. The primary for party -1 is won by a_3, because $sc_p^{-1}(a_3) \geq 7$, while $sc_p^{-1}(a_2) \leq 4$ and $sc_p^{-1}(a_4) \leq 4$. In the general election, $sc_g(a_1, \{a_1, a_3\} = sc_g(a_3, \{a_1, a_3\}) = 15$ and the tie-breaking rule makes a_3 the winner. However, $C(a_1) = 281$ and $C(a_3) = 297$.

4.2 Condorcet Winners

The results from this section apply to all Condorcet-consistent voting rules (e.g. Copeland, maximin), due to the median voter theorem ([Black, 1948]), which states that, if voters and candidates can be positioned in one dimension and voters' preferences are only determined by their proximity to candidates, the outcome of any Condorcet-consistent voting rule is the most preferred candidate by the median voter. It is worth noting that, in the one dimensional setting, a Condorcet winner is guaranteed to exist.

As before, we focus on the uniformly distributed voters and candidates case. Let $V = \{v_1, ..., v_n\}$ be the set of voters, with $v_1, ..., v_{n_1}$ being affiliated with party 1 and the others with party -1 and $A = \{a_1, ..., a_m\}$ be the set of candidates, with $a_1, ...a_{m_1}$ being affiliated with party 1 and the others with party -1. Candidates and voters are uniformly distributed, with $\rho(v_i) = i \times \delta$ and $\rho(a_i) = i \times k \times \delta$, for some $\delta > 0$ and $k \in \mathbb{N}$. The most preferred candidate by the median voter is declared the winner. We formulate the following conjecture and the aim of our analysis is to verify it:

Conjecture 3. *For separable election instances, if the voters and candidates are uniformly distributed, for any Condorcet-consistent voting rule, the utilitarian social cost of the winner in the primary system is lower than that of the winner in the direct system.*

Refutation of Conjecture 3. We not only refute the conjecture, but also conduct a complete analysis of the possible outcomes in the two systems using Condorcet-consitent voting rules.

If n is odd, $n = 2n'+1$, the median voter is $v_{n'+1}$. Let's observe that the candidate closer to the median voter is optimal in terms of utilitarian social cost. To see this, assume a_i is the closest candidate to $v_{n'+1}$. Without loss of generality, let $\rho(a_i) = \rho(v_{n'+p}) \Leftrightarrow ki = n'+1+p$, with $p \geq 1$ and let's consider any other candidate a_j, with $d(a_i, v_{n'+1}) \leq d(a_j, v_{n'+1})$. Moreover, we can assume that $\rho(a_j) = \rho(v_{n'+p+x})$, for $x > 0$ - otherwise, if candidate a_j were situated to the left of $v_{n'+1}$, we could simply consider the candidate located at the position corresponding to the the symmetric of $\rho(a_j)$ with respect to $\rho(v_{n'+1})$, who will have the same utilitarian social cost as a_j. $C(a_i) = \delta(n' + p - 1 + ... + 1 + 0 + 1 ... + 2n' + 1 - n' - p) = \delta\left(\frac{(n'+p-1)(n'+p)}{2} + \frac{(n'+1-p)(n'+2-p)}{2}\right)$ and $C(a_j) = \delta(n' + p + x - 1 + ... + 1 + 0 + 1 + ... + 2n' + 1 - n' - p - x) = \delta\left(\frac{(n'+p+x-1)(n'+p+x)}{2} + \frac{(n'-p-x+1)(n'-p-x+2)}{2}\right)$, with $C(a_j) - C(a_i) = \delta(x^2 + 2px - 2x) > 0$, because $2px \geq 2x$ and $x > 0$. This means that unless the primary winner produces the same winner as the direct system, the winner in the primary system will always have a greater utilitarian social cost.

For an even $n, n = 2n'$, there will be 2 median voters, $v_{n'}, v_{n'+1}$ and the winning candidate could depend on the tie-breaking rule. However, the only time when the optimal candidate in terms of utilitarian social cost does not win in the direct system is when $d(a_i, v_{n'}) = d(a_{i+1}, v_{n'+1}) + 1$ and $a_i \lhd a_{i+1}$ or $d(a_i, v_{n'}) + 1 = d(a_{i+1}, v_{n'+1})$ and $a_{i+1} \lhd a_i$. This is because in both instances, a_i and a_{i+1}, respectively, will have the same number of votes as a_{i+1} and a_i, respectively, but will win because of the tie-breaking rule. We'll only consider the case when $d(a_i, v_{n'}) = d(a_{i+1}, v_{n'+1}) + 1$ and $a_i \lhd a_{i+1}$ to show that a_i does indeed have a higher utilitarian social cost than a_{i+1}. Note that this implies that $2n' = 2ki + k$. $C(a_i) = \delta(ki - 1 + ... + 1 + 0 + 1 + ... + 2n' - ki) = \delta\left(\frac{(ki-1)ki}{2} + \frac{(2n'-ki)(2n'-ki+1)}{2}\right), C(a_{i+1}) = \delta((ki + k - 1 + ... + 1 + 0 + ... + 2n' - ki - k) = \delta\left(\frac{(ki+k-1)(ki+k)}{2} + \frac{(2n'-ki-k)(2n'-ki-k+1)}{2}\right)$ and $C(a_i) - C(a_{i+1}) = \delta(2n'k + k - k^2 - 2k^2 i) = \delta(2k^2 i + k^2 + k - k^2 - 2k^2 i) = \delta k$. The existence of such instances means that the primary system may produce a better outcome. However, these are limited to a_{i+1} and a_i, respectively, winning in the primary system. If we restrict our attention to the case considered before, when $d(a_i, v_{n'}) = d(a_{i+1}, v_{n'+1}) + 1$ and $a_i \lhd a_{i+1}$, then this automatically implies that $\pi(a_{i+1}) = -1$, as otherwise a_{i+1} could never be the winner of the primary. Moreover, we need the median voter/s of party -1 to prefer a_{i+1} more than any other candidate and we also need a_i to not win the primary of party 1. Note that, although possible, election instances of this type are not necessarily expected, because for larger n, n_1 would need to be very small, in order for a_{i+1} to be the closest to the median voter of party -1. Lastly, we don't impose $\pi(a_i) = -1$, but, there is precisely one such election instance with $\pi(a_i) = 1$, for a fixed δ shown in Figure 6, which is extremely restrictive.

To see that the primary system can produce a winner with a lower utilitarian social cost than that of the winner in the direct system, irrespective of the party affiliation of the Condorcet winner in the direct system, we begin with the example from Figure 6, where the Condorcet winner in a direct election is affiliated to party 1, and $\delta = 1, n = 10, n_1 = 4, k = 2, m_1 = 2, m = 5$ and let $a_2 \lhd a_3, a_3 \lhd a_4, a_1 \lhd a_2$.

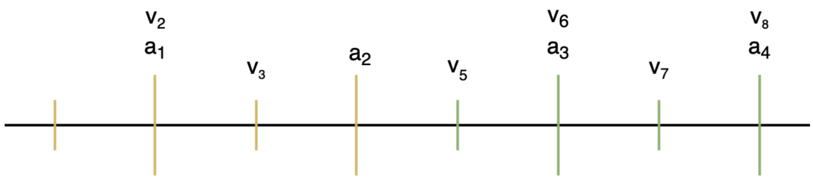


Figure 6: The primary system produces a better outcome with Condorcet winners and $\pi(a_i) = 1$

In the direct system, the median voters are v_5 and v_6, with a_2 winning because $a_2 \lhd a_3$. In the primary system, a_1 wins the primary for party 1, a_3 wins the primary for party -1 and the general election is won by a_3. $C(a_2) = 27$, $C(a_3) = 25$.

We continue with an example, shown in Figure 7, where the Condorcet winner in the direct system is affiliated to party -1 and the primary system still produces a better outcome. We have: $\delta = 1, n = 30, n_1 = 11, m_1 = 1, m = 5, i = 2, k = 6$. Let's also assume that $a_2 \lhd a_3$ and $a_3 \lhd a_4$.

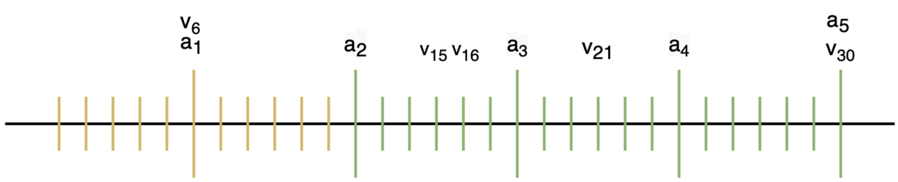


Figure 7: The primary system produces a better outcome with Condorcet winners and $\pi(a_i) = -1$

The median voters in the direct system are v_{15} and v_{16} and because $d(v_{15}, a_2) = d(v_{15}, a_3)$ and $a_2 \lhd a_3$, a_2 is the Condorcet winner and hence, the winner in the direct system. In the primary system, a_1 clearly wins the primary for party 1. For party -1, v_{21} is the median voter, with $d(v_{21}, a_3) = d(v_{21}, a_4)$. Because $a_3 \lhd a_4$, a_3 wins the primary for party -1 and they also clearly win in the general election against a_1. In this case, $C(a_2) = 237$ and $C(a_3) = 231$ and the primary produces a better outcome. □

4.2.1 Symmetric Case

We can see that in the cases where the Condorcet winner in the direct system does not correspond to the optimal candidate in terms of their utilitarian social cost, the last candidate is somewhat at a disadvantage, because there are fewer voters to their right, compared to other candidates and, more specifically, compared to the number of voters to the left of the first candidate. Therefore, it is worth focusing our attention on the case when the number of voters to the left of a_1 is the same as the number of voters to the right on a_m, if we think about the voters and candidates' positions on the line. Formally, we require $\rho(a_m) + (k-1) = \rho(v_n) \Leftrightarrow n = mk + k - 1$. It turns out that in this case, the direct system should always be chosen over the primary system.

Theorem 1. *For separable instances, if voters and candidates are uniformly distributed, the primary system will never produce a winner with a lower utilitarian social cost than that of the winner in the direct system.*

Proof. We distinguish the following two cases, depending on the parity of m:

1. m is odd, $m = 2m' + 1 \Rightarrow n = 2m'k + 2k - 1$. The median voter in the direct system is then $v_{m'k+k}$ and $\rho(v_{m'k+k}) = \delta(m'k + k) = \delta k(m' + 1) = \rho(a_{m'+1})$, so $a_{m'+1}$ would be the Condorcet winner in the direct system. Moreover, $a_{m'+1}$ is the optimal candidate in terms of utilitarian social cost, so the primary system could at most produce a winner with the same utilitarian social cost as $a_{m'+1}$ and it is never better than the direct system.

2. m is even, $m = 2m' \Rightarrow n = 2m'k + k - 1$.

i. If k is even, $k = 2p$, the median voter in the direct system is $v_{p(m+1)}$ and $\rho(v_{p(m+1)}) = \delta p(m+1) = \delta p(2m'+1) = \delta(2pm'+p) = \delta(km'+p)$. In this case, there are two candidates who could be the Condorcet winner, depending on the tie-breaking rule, namely $a_{m'}$ or $a_{m'+1}$ $d(v_{p(m+1)}, a_{m'}) = \delta(km'+p-km') = \delta p = \delta(2p-p) = \delta(k-p) = \delta(km'+k-km'-p) = d(v_{p(m+1)}, a_{m'+1})$. However, the following hold: $C(a_{m'}) = \delta(m'k-1+...+1+0+1+...+n-m'k) = \delta\left(\frac{(m'k-1)m'k}{2} + \frac{(n-m'k)(n-m'k+1)}{2}\right)$ and $C(a_{m'+1}) = \delta(m'k+k-1+...+1+0+1+...+n-m'k-k) = \delta\left(\frac{(m'k+k-1)(m'k+k)}{2} + \frac{(n-m'k-k)(n-m'k-k+1)}{2}\right)$ with $C(a_{m'+1}) - C(a_{m'}) = \delta(2m'k^2 - kn + k^2 - k) = \delta[8m'p^2 - 2p(4m'p+2p-1) + 4p^2 - 2p] = \delta(8m'p^2 - 8m'p^2 - 4p^2 + 2p + 4p^2 - 2p) = 0$, so $C(a_{m'}) = C(a_{m'+1})$ and the winner in the direct system would be an optimal candidate.

ii. If k is odd, $k = 2p+1$, there are two median voters in the direct system: $v_{m'k+p}, v_{m'k+p+1}$ and there are two possible Condorcet winners, depending on the tie-breaking rule, $a_{m'}$ and $a_{m'+1}$, with $d(v_{m'k+p}, a_{m'}) = \delta(m'k+p-m'k) = \delta p = \delta(m'k+2p+1-m'k-p-1) = d(v_{m'k+p+1}, a_{m'+1})$. Using a similar approach as before, we obtain $C(a_{m'+1}) - C(a_{m'}) = \delta(2m'k^2 - kn + k^2 - k) = \delta\{2m'(2p+1)^2 - (2p+1)[2m'(2p+1)+2p+1-1] + (2p+1)^2 - 2p - 1\} = \delta(8m'p^2 + 8pm' + 2m' - 8m'p^2 - 4pm' - 4p^2 - 4pm' - 2m' - 2p + 4p^2 + 4p + 1 - 2p - 1) = 0$, so $C(a_{m'}) = C(a_{m'+1})$ and the winner in the direct system would once again be an optimal candidate.

$\square$

Remark 1. *The only case when the primary system would produce a winner with the same utilitarian social cost as that of the winner in the direct system is when the median voter in the primaries is closest to the socially optimal candidate. This happens in very restricted cases, when the number of voters from one party is disproportionately small to the number of voters from the other party. We conclude that the primary system is strictly worse than the direct system in most of the cases, with very few instances when the winners in the two systems have the same utilitarian social cost and that it is never better than the direct* *system.*

4.3 Single Transferable Vote

Again, we only consider the uniformly distributed voters and candidates case. It is useful to note that STV may fail to select the Condorcet winner even in the direct system. Taking into account Observations 2 and 4 from Section 4.1.2, we know that few voters will give their point to the Condorcet winner in the direct system. Moreover, if the Condorcet winner wins in pairwise elections without being preferred by the tie-breaking rule, they could be the candidate with the lowest score and be one of the first eliminated candidates in both the direct and the primary systems. Similarly, even if the Condorcet winner is preferred over any other candidate, according to the tie-breaking rule, they can still be eliminated. This shows that STV heavily relies on the tie-breaking rule in one dimension and a quantitative analysis is almost impossible for a large number of candidates, for any of the two types of systems.

In our next example, we show that the Condorcet winner may be eliminated by STV, even if they are the top candidate in the tie-breaking rule. For the example shown in Figure 8, with $\delta = 1, n = 25, k = 6$. Let $a_2 \lhd a_1 \lhd a_3 \lhd a_4$. Clearly a_2 is the Condorcet winner, as they are the closest to the median voter, v_{13}. Using STV, initially, $sc(a_1) = 8, sc(a_2) = 7, sc(a_3) = 6$ and $sc(a_4) = 4$, so a_4 is eliminated. The process continues, as no candidate has at least 13 points, with the following, updated, scores: $sc(a_1) = 8, sc(a_2) = 7, sc(a_3) = 10$. As a result, a_2 is eliminated, even though they are the Condorcet winner and also the most preferred candidate, according to the tie-breaking rule.

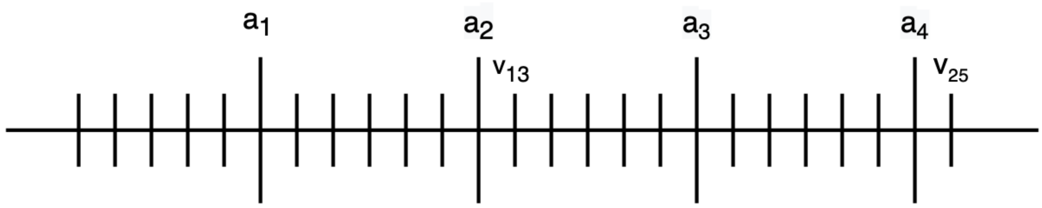


Figure 8: The Condorcet winner may be eliminated by STV

4.4 Average-case Distortion

In this section, we compare the distortion of the seven voting rules described in Section 3: plurality, anti-plurality, plurality with run-off, Borda, Harmonic Borda, Copeland and STV under the two systems. Moreover, we consider two classes of instances. Our analysis differs from that of Borodin et al. [2019], who focus on computing the average distortion over 1000 election instances for various voting rules. As a consequence, our emphasis is on comparing which of the two systems produces a better outcome. While we also look at the margins of the difference between the distortion of the mentioned voting rules under the two systems, this is not our main objective. Moreover, in their simulations, Borodin et al. [2019] generate the positions of voters and candidates at uniformly random locations, while we, apart from the case where voters and candidates are uniformly distributed, as per our theoretical analysis, also consider more general settings, by taking samples of a given size to obtain the voters and candidates' positions.

For the first class, we investigate both instances where we enforce party separability and instances where we do not, which we refer to as "random" (we note that the word "random" might be confusing, however, since it is being used by Borodin et al. [2019] with the same meaning, for consistency, we continue with the same notation) in one, three and five dimensions.

For all of the experiments related to the first class, we choose a random integer n between 200 and 1000, representing the number of voters and a random value for m, between $\lfloor n/10 \rfloor - \lfloor n/100 \rfloor * 2$ and $\lfloor n/10 \rfloor + \lfloor n/100 \rfloor * 2$, representing the number of candidates. We then set the number of voters and candidates affiliated with party 1, ranging from 30% to 70%, in increments of 5%, of the total number of voters and candidates, respectively and for each such instance, we randomly generate the tie-breaking rule. We apply the aforementioned voting rules for both the primary and the direct system and compute the utilitarian social cost of the resulting winners.

To achieve party separability, we follow the same approach as Borodin et al. [2019] by firstly sorting the voters based on the value of their last coordinate. Next, we set a value for $n_1 = p * n$, $p \in \{0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.6, 0.65, 0.7\}$, representing the number of voters from party 1. We associate the first n_1 sorted voters to party 1 and the remaining ones to party -1. This ensures that the voters are separable. Lastly, we set a value for $m_1 = p * m$, $p \in \{0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.6, 0.65, 0.7\}$, representing the number of candidates from party 1, and we randomly generate m_1 points with the last coordinate at most equal to the last coordinate of the last voter for party 1. Similarly, we randomly generate $m - m_1$ candidates associated to party -1.

For the second class, we try to simulate real-life elections, where, for each party, voters are likely to be concentrated around candidates. For this reason, we use a Gaussian distribution in one and two dimensions. We start by having the two distribution close to each other and we increase the distance between the two distributions, until the instances are almost separable. We call this "polarization", as the further the two distributions get from each other, the more likely voters will vote for a candidate from the same party, even in the general election.

For these experiments, the two parties have the same number of voters, a random integer between 100 and 500, as well as the same number of candidates, a random integer between $\lfloor n/10 \rfloor - \lfloor n/100 \rfloor * 2$ and $\lfloor n/10 \rfloor + \lfloor n/100 \rfloor * 2$ so that no party is disadvantaged when the distributions are further away from each other.

4.4.1 1 Dimension

We note that the results for anti-plurality do no offer any real insights, as voters are asked to "penalise" their least preferred candidate, and when voters and candidates can be positioned on a line, the first and the last candidate will be the only ones to receive negative scores. As a result, the winner determined by the anti-plurality voting rule is very much dependent on the tie-breaking rule. However, for completeness, we do include the results for anti-plurality.

The findings from Table 1 show the number of election instances for which one of the systems produced a winner with a lower utilitarian social cost ("PB" stands for "Primary is better" and "DB" stands for "Direct is better), or the two systems produce the same winner ("ND" stands for "No difference"), for uniformly distributed and general, separable and random instances (we remind the reader that in a random election instance party separability is not enforced) in one dimension. We use "general" to highlight that we no longer require the voters and candidates to be uniformly distributed. Rather, we randomly generate n voters and m candidates within the range $(0, 5n)$. For uniformly distributed voters and candidates, we use the most restrictive setting, where for each candidate there is a co-located voter, as most of our theoretical analysis focuses on that case.

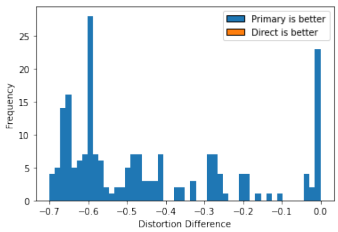
	Uniformly distributed						General					
	Separable			Random			Separable			Random		
	PB	DB	ND	PB	DB	ND	PB	DB	ND	PB	DB	ND
Plurality	184	0	22	204	0	2	154	8	62	133	62	29
Anti-plurality	192	8	6	123	79	4	114	53	57	147	73	4
Plurality Run-off	86	80	40	144	38	24	82	62	80	142	41	41
Borda	0	206	0	97	10	99	0	224	0	88	58	78
Harmonic Borda	0	203	3	24	115	68	85	24	115	132	49	43
Copeland	0	206	0	0	17	189	0	224	0	0	81	143
STV	46	105	55	123	54	29	43	108	73	132	78	15

Table 1: Comparison of the two systems in one dimension, in terms of the utilitarian social cost of the winner produced

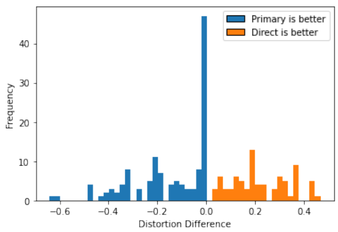
Let's begin by noticing that, although in our quantitative instance-wise comparison between the two systems, for the uniformly distributed case under plurality voting, we showed that there are several instances where the direct system is better, in practice this is unexpected. On the reverse, a similar argument holds for the Copeland voting rule, which is Condorcet-consistent: on average, the direct system is strictly better than the direct system for separable instances, while for random instances, in most of the cases the winners have the same utilitarian social cost. Also, the results for Borda and separable instances are not surprising either, in fact, this enforces the results in the literature, which claim that "Borda is close to being Condorcet-consistent" [Saari, 1985; Gehrlein, 1987; Gehrlein and Valognes, 2001; Gehrlein and Plassmann, 2014; Gehrlein et al., 2017].

An interesting trend can be observed when comparing Borda and Harmonic Borda for separable and random instances. While for separable instances, the direct system is almost always better, for random instances, Borda performs better in the primary system. Harmonic Borda, however, exhibits an interesting behaviour for random instances: the direct system appears to be better in the uniformly distributed case, while for the general case, the primary systems outperforms the direct one. Similarly, for STV, with party separability, the direct system seems to be better, while for random instances, the primary system becomes better. For Copeland, we must note that for random instances, the two systems produce a different winner very rarely, for the uniformly distributed case, with a slight increase for the general case. Lastly, plurality with run-off produces very similar results in each of the two cases, displaying the most consistency out of the voting rules we considered in one dimension.

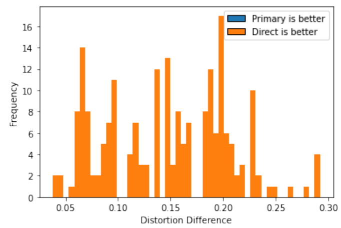
Another useful metric would be looking at the margins between the utilitarian social costs of the two winners. To obtain the following histograms, for each voting rule and for each instance, we compute the difference between its distortion under the two systems. Figures 9-12 display the difference in distortion of uniformly distributed and general, separable and random instances, respectively, in one dimension:



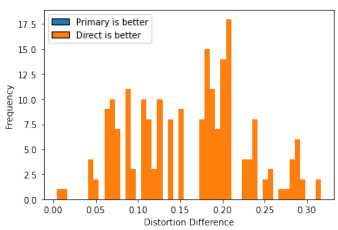
(a) Plurality



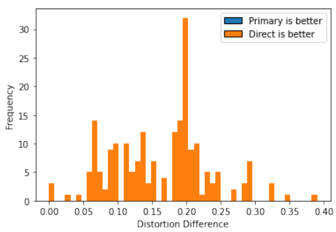
(b) Plurality Run-off



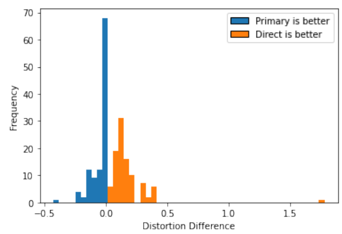
(c) Copeland



(d) Borda

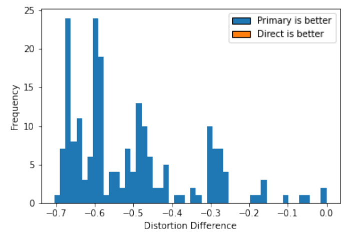


(e) Harmonic Borda

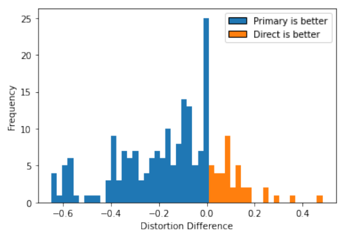


(f) STV

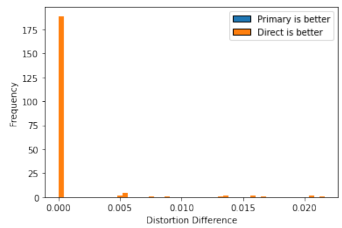
Figure 9: Difference in distortion between the primary and direct system for separable, uniformly distributed election instances



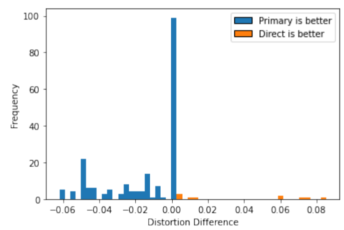
(a) Plurality



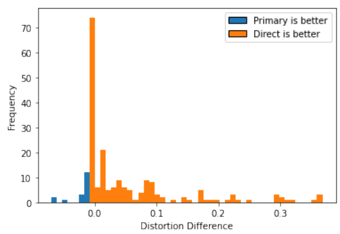
(b) Plurality Run-off



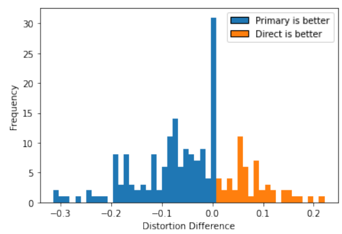
(c) Copeland



(d) Borda

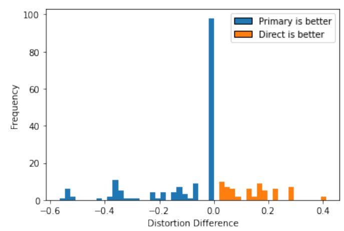


(e) Harmonic Borda

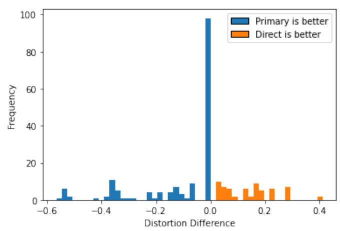


(f) STV

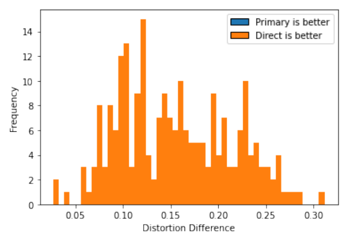
Figure 10: Difference in distortion between the primary and direct system for random, uniformly distributed election instances



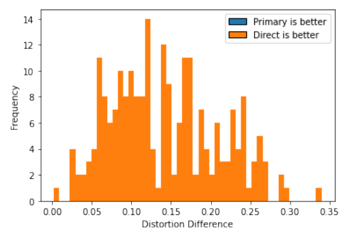
(a) Plurality



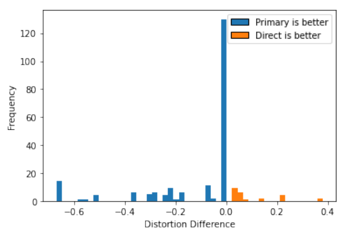
(b) Plurality Run-off



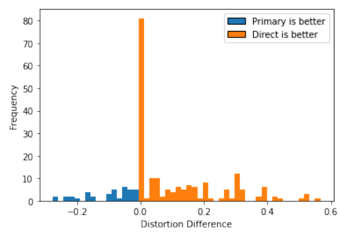
(c) Copeland



(d) Borda

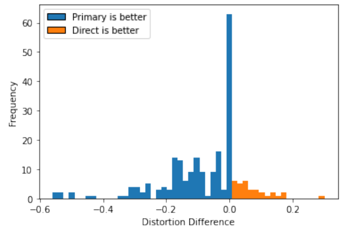


(e) Harmonic Borda

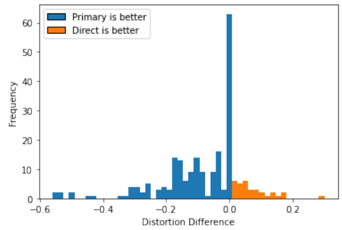


(f) STV

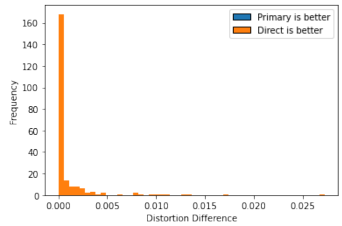
Figure 11: Difference in distortion between the primary and direct system for separable general one dimensional election instances



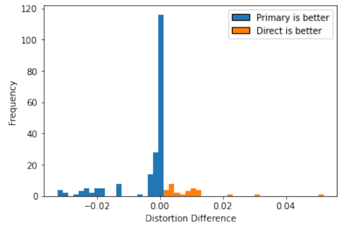
(a) Plurality



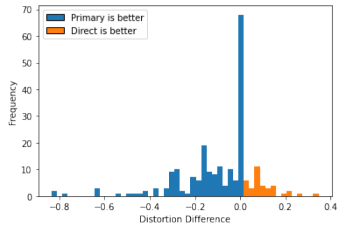
(b) Plurality Run-off



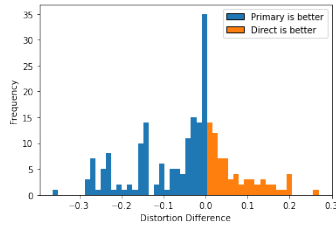
(c) Copeland



(d) Borda



(e) Harmonic Borda



(f) STV

Figure 12: Difference in distortion between the primary and direct system for general one dimensional election instances

The differences are fairly varied. We observe larger margins for plurality and plurality with run-off for all of the instances we considered, with the primary system clearly outperforming the direct one under plurality, in the uniformly distributed case. On the contrary, we notice the very small differences in random instances, for Copeland and Borda, with a substantial increase for separable instances. It is also worth remarking that the margins are very similar for the uniformly distributed and general cases, with differences only in the frequency. Hence, this can be seen as another justification for considering the, seemingly restrictive, case where the voters and candidates are uniformly distributed in one dimension.

4.4.2 3 and 5 Dimensions

We now increase the dimensions of our metric space to 3 and 5, respectively. To obtain the positions of the n voters, we randomly generate 3 and 5 samples of size n from the range $(0, 5n)$ for the 3 and 5 dimensions cases, respectively, and the position of voter i is represented by the i^{th} value from each sample. We use the same technique to position the candidates in the metric space.

Table 2 displays the comparison between the utilitarian social cost of the two winners, when voters and candidates are positioned into $\mathbb{R}^3$ and $\mathbb{R}^5$, for both types of instances:

	3 Dimension						5 Dimension					
	Separable			Random			Separable			Random		
	PB	DB	ND	PB	DB	ND	PB	DB	ND	PB	DB	ND
Plurality	106	17	102	155	41	29	111	37	77	144	32	49
Anti-plurality	130	33	62	154	65	6	151	31	43	171	44	10
Plurality Run-off	102	37	86	158	39	28	82	47	96	138	43	44
Borda	5	162	58	19	14	192	9	94	123	2	20	203
Harmonic Borda	68	71	86	141	34	50	43	86	96	69	52	104
Copeland	2	165	58	4	18	203	5	97	123	2	15	208
STV	67	99	59	135	42	48	40	83	102	90	32	103

Table 2: Comparison of the two systems in three and five dimensions, in terms of the utilitarian social cost of the winner produced

We note that the results for 3 and 5 dimensions are very similar for each voting rule, and they also resemble those for the general one dimensional case: voting under plurality, anti-plurality or plurality with run-off is more suitable for the primary system, while the direct system outperforms the primary one, if Borda or Copeland is used. Harmonic Borda continues to perform slightly better under the primary system for random instances, while for separable instances, the direct system seems to be more suitable. We also observe the same trend for STV, with quite a substantial switch from the direct system to the primary one, for separable and random instances, respectively.

We include the histograms containing the difference between the distortion of each voting rule under the two systems in Appendix A, because of their similarity to the ones we have already presented. However, we note that the margins generally get smaller, even for plurality and plurality with run-off.

4.4.3 Polarization

We now aim to simulate real-life election, where candidates are likely to be concentrated around distinguished candidates. We model this setting by using Gaussian distributions in one and two dimensions. Initially, we start by having the means of the distributions at a small distance to each other, which we then gradually increase up to the point where the two distributions are almost separable. As a consequence, we investigate the performance of the two systems in terms of the polarization of the voters, as for the case where the distributions are further apart, most voters will vote for the candidate from their own party in the general election as well. The same argument clearly does not hold when the distributions are close to each other and the voters and candidates from the two parties are mixed with each others.

For the one dimensional case, we define two normal distributions with means at 1 and $1 + dist$, for $dist \in \{1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5\}$ representing the voters from party 1 and -1 respectively. We then place our candidates uniformly between the first and last voter from each party. For the results in Table 3 (because of space constraints, we use "AP" to refer to anti-plurality, "PRO" to denote plurality with run-off and "HB" for Harmonic Borda), we compare the utilitarian social cost of the winner in the two systems, depending on the distance between the means of the two distributions.

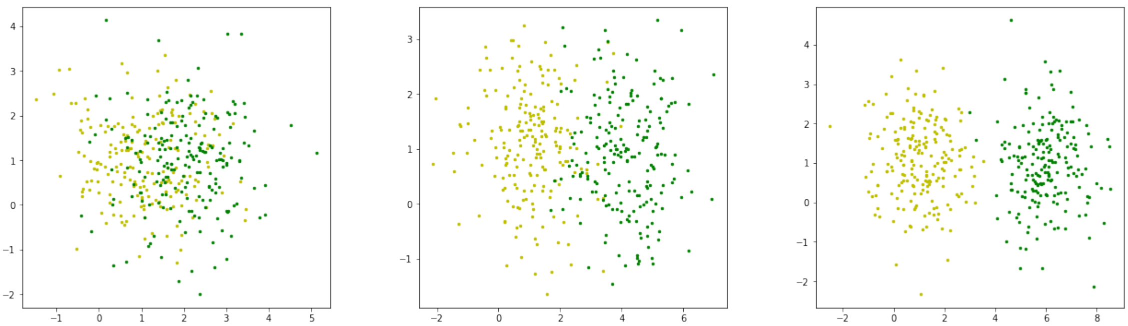
	dist = 1			dist = 1.5			dist = 2			dist = 2.5			dist = 3		
	PB	DB	ND	PB	DB	ND	PB	DB	ND	PB	DB	ND	PB	DB	ND
Plurality	13	10	2	15	5	5	10	11	4	16	2	7	17	2	6
AP	7	12	6	6	17	2	4	19	2	13	6	6	10	7	8
PRO	13	10	2	15	5	5	10	11	4	16	2	7	17	2	6
Borda	1	23	1	1	24	0	0	25	0	0	25	0	0	25	0
HB	8	11	6	10	10	5	5	19	1	12	5	8	14	2	9
Copeland	0	24	1	0	24	1	0	25	0	0	25	0	0	25	0
STV	13	9	3	6	18	1	7	13	5	11	6	8	9	11	5

	dist = 3.5			dist = 4			dist = 4.5			dist = 5		
	PB	DB	ND	PB	DB	ND	PB	DB	ND	PB	DB	ND
Plurality	15	0	10	10	3	12	10	2	13	15	4	6
AP	14	3	8	15	2	8	13	6	6	12	6	7
PRO	5	5	15	3	12	10	6	8	11	3	16	6
Borda	0	25	0	0	25	0	0	25	0	0	24	1
HB	13	3	9	10	5	10	12	1	12	15	3	7
Copeland	0	25	0	0	25	0	0	25	0	0	25	0
STV	2	13	10	2	6	17	3	4	18	1	4	20

Table 3: One dimensional polarization

We extend our previous experiment to two dimensional Gaussian distributions. For this, we set the mean vectors for party 1 and -1 to $(1, 1)$ and $(1 + dist, 1)$, for $dist \in$

$\{1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5\}$, respectively and both covariance matrices to $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. The examples from Figure 13 display two samples of size 200, for different values of $dist$. We can see as the value of $dist$ increases, the two parties go from overlapping significantly to almost no overlap. This means that for a value of 5 for $dist$, every voter is very likely to vote for a candidate from their own party in the general election as well, which is clearly not necessarily true for a value of 1.



(a) dist = 1 (b) dist = 3 (c) dist = 5

Figure 13: Two dimensional Gaussian distributions

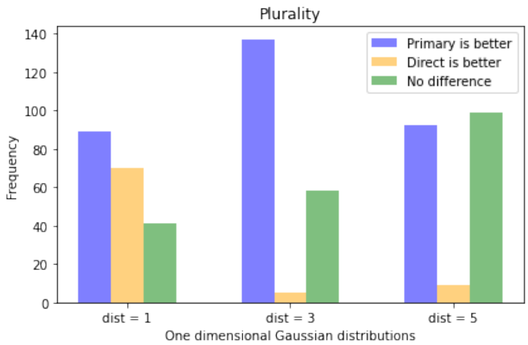
Similarly to the one dimensional case, we compare the utilitarian social cost of the two winners, under the scenarios described above, to obtain the results from Table 4:

We begin by noticing that for each voting rule, the results are very similar. Generally, the primary system appears to be better, especially for plurality, anti-plurality, plurality with run-off and Harmonic Borda. In contrast, both Borda and Copeland produce a better outcome in the direct system, just as in our previous analysis. Lastly, we observe two interesting trends. Firstly, for plurality, the number of instances where the primary system is better remains somewhat consistent, but the number of instances where the winner is the same increases with the distance between the distributions. Secondly, STV follows the same behaviour, except that the direct system stars off as better, apparently at an even higher margin. However, since this analysis is only based on 25 instances, for plurality and STV, we perform 200 additional experiments, in both one and two dimensions, for each value of $dist \in \{1, 3, 5\}$ the results of which we display in Figures 14 and 15:

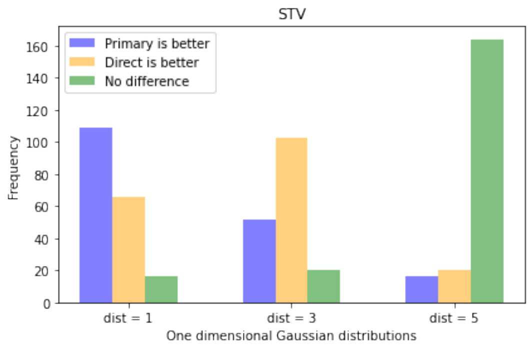
	dist = 1			dist = 1.5			dist = 2			dist = 2.5			dist = 3		
	PB	DB	ND	PB	DB	ND	PB	DB	ND	PB	DB	ND	PB	DB	ND
Plurality	13	5	7	16	3	6	13	4	8	17	3	5	16	2	7
AP	17	4	4	19	2	4	19	1	5	13	8	4	11	8	8
PRO	14	3	8	12	9	4	11	7	7	12	7	6	8	5	12
Borda	0	23	2	1	24	0	0	25	0	0	25	0	0	25	0
HB	15	3	7	11	8	6	12	4	9	10	6	9	11	4	10
Copeland	0	23	2	0	24	1	0	25	0	0	25	0	0	25	0
STV	10	12	3	11	10	4	6	18	1	6	18	1	7	11	7

	dist = 3.5			dist = 4			dist = 4.5			dist = 5		
	PB	DB	ND	PB	DB	ND	PB	DB	ND	PB	DB	ND
Plurality	12	1	12	10	1	14	12	0	13	10	1	14
AP	15	2	8	14	3	8	16	2	7	14	6	5
PRO	10	8	7	7	4	14	11	3	11	8	7	10
Borda	0	25	0	0	25	0	0	25	0	1	24	0
HB	14	3	8	12	2	11	12	4	9	7	4	14
Copeland	0	25	0	0	25	0	0	25	0	1	24	0
STV	4	9	12	8	3	14	3	7	15	5	3	17

Table 4: Two dimensional polarization

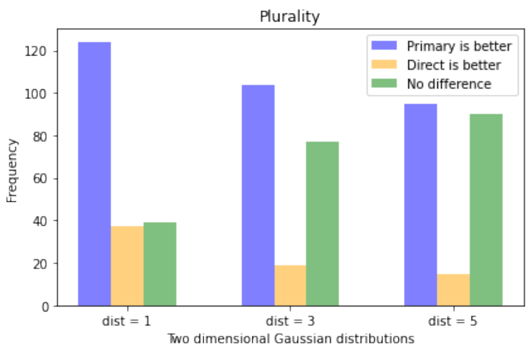


(a) Plurality

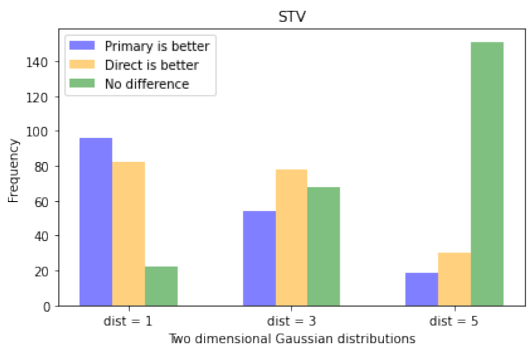


(b) STV

Figure 14: Further experiments for one dimensional Gaussian distributions



(a) Plurality



(b) STV

Figure 15: Further experiments for two dimensional Gaussian distributions

The additional experiments support our previous statements about the trends of plurality and STV. For plurality, the primary system is generally better than the direct one, especially for larger values of the distance between the means of the two distributions. As the distance increases, the two systems produce the same winner more often and the direct systems performs more poorly.

The behaviour of STV is even more interesting. The primary systems is more advantageous for a value of 1 for *dist*. Nonetheless, when the value of *dist* is increased to 3, the direct system appears to be better and finally, when the two distribution are far apart from each other, the two systems produce the same winner in most of the cases, with the direct system still holding a small lead against the primary one, in terms of the number of instances where the winner has a lower utilitarian social cost than that of the winner in the direct system.

4.4.4 Overall Results

We combine the results of all the simulations in Table 5. We note that plurality and STV contain more instances, which resulted from the additional experiments for polarization. Again, we conclude that plurality is best suited to the primary system and Copeland to the direct system. Notably, although Borda and Harmonic Borda might seem similar, they produce very different results, with Harmonic Borda, just like STV, performing approximately the same in both systems and Borda performing much better in the direct system, as expected.

	Primary is better	Direct is better	No difference
Plurality	2072	412	927
Anti-plurality	1414	500	297
Plurality Run-off	1115	512	584
Borda	224	1230	757
Harmonic Borda	765	731	715
Copeland	14	1268	929
STV	1136	1155	1120

Table 5: Overall results

5 Strategic Candidacies

In this section, we perform the first analysis of strategic candidacy games for the primary system. We continue to use the spatial model defined in Section 3, but unlike in the previous section, candidates now also have their own preferences over the set of candidates, A, which are determined by their proximity to other candidates. As a result, the candidates have self-supporting preferences, i.e. for all candidates $c \in A, c \succ_c a, \forall a \in A$.

We will consider three types of strategic candidacy games, which we call *lazy* (introduced by Obraztsova et al. [2015]), *eager* and *keen* (described by Lang et al. [2019]). In all three games, each candidate (or player) has two available actions or strategies: 1, meaning that the candidate wants to run in the election, or 0, for which the candidate prefers to abstain. Formally, the strategy of player $c \in A$ is denoted by $\mathbf{s}_c$, with $\mathbf{s}_c \in \{0, 1\}$ and a strategy profile is then a vector $\mathbf{s} = (\mathbf{s}_c)_{c \in A}$. The set of candidates participating in the election, according to a strategy profile $\mathbf{s}$ is denoted by $A(\mathbf{s}) = \{c \in A | \mathbf{s}_c = 1\}$. For each strategy profile, there will be a winner, denoted by $w(\mathbf{s}) \in A(\mathbf{s})$. We also use $w_p^1(\mathbf{s})$ and $w_p^{-1}(\mathbf{s})$, to denote the winners in the primaries of parties 1 and -1, respectively.

We remind the reader of the notation $sc_p^{\pi(a)}(a)$, used to define the score of candidate a in the primary of their party, which we extend to $sc_p^{\pi(a)}(a, \mathbf{s})$ to define the score of candidate a in the primary of their party, for the strategy profile $\mathbf{s}$. We continue to use $sc_g(a, \{a, b\})$ for the score of candidate a in the general election against candidate b.

Lastly, as it common practice in the prior work on strategic candidacy games, for simplicity, we will only be considering the plurality voting rule, for which we investigate the pure strategy Nash Equilibria and their properties.

5.1 Lazy Strategic Candidacy Games

The concept of a lazy strategic candidacy games(LSCG) was first introduced by Obraztsova et al. [2015] for the direct system. The motivation of considering such games was that it is often the case that a candidate would prefer to not take part in an election if they could not have any influence on the outcome. This could be because of the expenses associated with the electoral campaign (e.g. travel costs, televised advertisements or political consulting fees) or because an unsuccessful participation in an election might harm the candidate's reputation. It is, therefore, of interest to also investigate the properties of such games for the primary system.

We now adapt their model to the primary system. Formally, an LSCG, Γ^L, is a tuple $\Gamma^L(V, A, M, d, \rho, \pi, \triangleleft)$ (remember that (M, d) is the metric space containing the voters and candidates). In Γ^L, a candidate $c \in A$ prefers the strategy profile $\mathbf{s}$ to $\mathbf{t}$ if (i) $w(\mathbf{s}) \succ_c w(\mathbf{t})$ or (ii) $w(\mathbf{s}) = w(\mathbf{t})$ and $\mathbf{s}_c = 0$ and $\mathbf{t}_c = 1$. Hence, the term "lazy", because if candidates are not able to influence the outcome of an election to one they prefer, they would rather not participate in the election and incur the associated consequences.

5.1.1 Nash Equilibria

For newly introduced types of strategic candidacy games, the related literature is concerned with firstly analysing whether they admit a PNE, and subsequently describing properties of such equilibrium strategy profiles. We aim to do precisely that. The following results are related to the one dimensional metric space $(\mathbb{R}, d)$, where d is the standard Euclidean distance. For this setting, we know that a Condorcet winner always exists. Moreover, due to the tie-breaking rule, the Condorcet winner must be unique, since candidates cannot be tied. As we will show, one dimensional LSCGs reveal very interesting properties in the primary system.

We now formally define the concept of pure strategy Nash equilibrium (PNE) of LSCGs. A strategy profile $\mathbf{s}$ is a PNE of an LSCG Γ^L if no player prefers to deviate from $\mathbf{s}$, i.e.

for every candidate $c \in A$, there is no strategy profile $\mathbf{t}$, with $\mathbf{t}_a = \mathbf{s}_a, \forall a \in A \setminus \{c\}$ and $\mathbf{t}_c = 1 - \mathbf{s}_c$, such that c prefers $\mathbf{t}$ to $\mathbf{s}$.

Somewhat intuitively, we show that lazy candidates affiliated to the losing party have no reason to participate in the election in a PNE. This is because they have no influence over who the candidate from the opposing party is in the general election. As a result, a candidate from the losing party would prefer to run only if they would be able to change the outcome in both, their party's primary and the general election to one they prefer. The impossibility arises in the latter case, because of structural properties of one dimensional LSCGs.

Proposition 1. *Let $\Gamma^L = \Gamma^L(V, A, \mathbb{R}, d, \rho, \pi, \triangleleft)$. Then Γ^L has no PNE with candidates running from both parties.*

Proof. Suppose, for a contradiction, that there exists a PNE $\mathbf{s}$ of Γ^L, such that $\exists a, b \in A(\mathbf{s})$, with $\pi(a) \neq \pi(b)$ and $w_p^1(\mathbf{s}) = a, w_p^{-1}(\mathbf{s}) = b$. Without loss of generality, assume that the winner of the general election is a: $w(\mathbf{s}) = a$. Then if b were not to run, because $\mathbf{s}$ is a PNE, there must exist $c \in A(\mathbf{s})$, with $\pi(c) = \pi(b)$ and $a \succ_b c$ that would win both the primary of party -1 and the general election. Firstly, if b were the only candidate from party -1, them not running would mean a winning the general election, but then they would prefer not to run and deviate, which is not possible, because $\mathbf{s}$ is a PNE. Secondly, c must be the winner of the general election, otherwise b would still prefer not to participate and it must also hold that $a \succ_b c$, otherwise, if $c \succ_b a$, then b would want to deviate.

Let's also assume, without loss of generality, that b is positioned to the right of a on the line. We distinguish the following cases:

I. c is positioned to the left of a. If c were not to run, because $\mathbf{s}$ is a PNE, there must exist $d \in A(\mathbf{s})$ with $\pi(d) = \pi(c)$ and $a \succ_c d$ that would win both the primary of party -1 and the general election. To see this, if c not running would result in b winning the primary, then the general election winner would be a, and c would want to deviate.

- d cannot be positioned between c and a, because then $d \succ_c a$.

- Neither can d be positioned between a and b. We know that c beats a in the general election, so $sc_g(c, \{a, c\}) \geq sc_g(a, \{a, c\})$ and if $sc_g(c, \{a, c\}) = sc_g(a, \{a, c\})$, then $c \lhd a$. If d were positioned between a and b and $sc_g(c, \{a, c\}) > sc_g(a, \{a, c\})$, then in the general election against a, $sc_g(d, \{a, d\}) < sc_g(a, \{a, d\})$, because a would get at least $sc_g(c, \{a, c\})$, votes, while d would get at most $sc_g(a, \{a, c\})$ votes and a would be the winner. If $sc_g(c, \{a, c\}) = sc_g(a, \{a, c\})$, then $c \lhd a$. But in the general election between d and a, a would again have at least $sc_g(c, \{a, c\})$ votes, while d would have at most $sc_g(a, \{a, c\})$, so in order for d to beat a, we would need $sc_g(d, \{a, d\}) = sc_g(a, \{a, d\})$ and $d \lhd a$. We know that when b, c and d run, b has the most votes and that if b were not to run, c would be the primary winner. However, in that case all the votes for b would go to d, so when all of b, c and d run, b and c must have the same number of votes in the primary (otherwise, when b does not run, d has more votes than c and wins the primary) and $b \lhd c$ and d must have 0 votes. This means that when b does not run, in order for c to win the primary against d, they must also have the same number of votes (as all votes for b go to d) and $c \lhd d$. Lastly, we know that if c does not run, d wins the primary against b - this only happens when d and b have the same number of votes and $d \lhd b$. However, we have reached a contradiction because it cannot hold at the same time that $b \lhd c \lhd d$ and $d \lhd b$.

- Lastly, d cannot be positioned to the left of b, because d would get no extra votes when c does not run and could not beat b in the primary.

This implies that d must be positioned to the left of c. If d were not to run, then b could not win the primary, because a would win the general election and d would rather deviate. Also, c could not win the primary, because it would then beat a in the general election and $c \succ_d a$, which means that there must exist $e \in A(\mathbf{s})$ with $\pi(e) = \pi(d)$ that would win both the primary of party -1 and the general election against a and $a \succ_d e$. Using similar arguments as above, e would need to be positioned to the left of d. Continuing in this manner, because the set of possible candidates is finite, we obtain that $\exists b, c, d, e, f_1, f_2, ..., f_\nu \in A(\mathbf{s})$, with $\pi(b) = \pi(c) = \pi(d) = \pi(e) = \pi(f_1) = ... = \pi(f_\nu) = -1$ positioned, from left to right, in the order: $f_\nu, ..., f_1, e, d, c, a, b$. But in this case, if f_ν

were not to run, $f_{\nu-1}$ would win the primary and then the general election against a and because $f_{\nu-1} \succ_{f_\nu} a$, f_ν would want to deviate - contradiction.

II. c is positioned to the right of a. Because $a \succ_b c$, c must be positioned to the right of b. Again, if c were not to run, then there must exist $d \in A(\mathbf{s})$, with $\pi(d) = \pi(c)$ and $a \succ_c d$ that would win both the primary of party -1 and the general election. Clearly, d cannot be positioned to the left of b, because d would get no additional votes when c does not run and could not win the primary. Similarly, d cannot be positioned between b and c, because when b does not participate in the election, d would get b's votes, rather than c, so c could not be the winner. The only exception is when $sc_p^{-1}(b, \mathbf{s}) = sc_p^{-1}(c, \mathbf{s})$ and $b \lhd c$ and $sc_p^{-1}(d, \mathbf{s}) = 0$. But in this case, when b does not run, all their votes go to d and in order for c to win the primary, it must hold that $sc_p^{-1}(d, \mathbf{s'}) = sc_p^{-1}(c, \mathbf{s'})$ and $c \lhd d$, where $\mathbf{s'}_x = \mathbf{s}_x, \forall x \in A \setminus \{b\}$ and $\mathbf{s'}_b = 0$. Also, when c does not run, all their votes must go to d and in order for d to win the primary, it must hold that $sc_p^{-1}(d, \mathbf{s''}) = sc_p^{-1}(b, \mathbf{s''})$ and $d \lhd b$, where $\mathbf{s''}_x = \mathbf{s}_x, \forall x \in A \setminus \{b\}$ and $\mathbf{s''}_c = 0$. However, we have once again reached a contradiction, because we have: $b \lhd c \lhd d$ and $d \lhd b$. Therefore, d must be positioned to the right of c.

We can now follow a similar approach to obtain candidates $b, c, d, e_1, ..., e_\nu \in A(\mathbf{s})$ with $\pi(b) = \pi(c) = \pi(d) = \pi(e_1) = ... = \pi(e_\nu) = -1$ positioned, from left to right, in the order: $a, b, c, d, e_1, ..., e_\nu$. But in this case, if e_ν were not to run, $e_{\nu-1}$ would win the primary and then the general election against a and because $e_{\nu-1} \succ_{e_\nu} a$, e_ν would want to deviate - contradiction. This now completes our proof.

□

Again, because of the same structural properties in one dimension, the scenario becomes even more restrictive. It turns out that not even the candidates from the winning party are motivated to participate in the election in a PNE.

Proposition 2. *Let $\Gamma^L = \Gamma^L(V, A, \mathbb{R}, d, \rho, \pi, \lhd)$. Γ^L has no PNE with more than one candidate running.*

Proof. We already know that Γ^L has no PNE with agents running from both parties from Proposition 1, so let's suppose, for a contradiction, that there exists a PNE $\mathbf{s}$, such that $\exists a, b \in A(\mathbf{s})$, with $\pi(a) = \pi(b)$. Clearly, the winner of the primary becomes the winner of the general election. Without loss of generality, assume that the winner of the general election is candidate a: $w(\mathbf{s}) = a$. Because $\mathbf{s}$ is a PNE, if b were not to run, a could not win the primary (because b would prefer not to run), so there must exist $c \in A(\mathbf{s}), \pi(c) = \pi(b)$ and $a \succ_b c$ that would win the primary. We can now follow an identical approach as in the proof of Proposition 1, to show that there would need to exist candidates $c, d, e, f_1, ..., f_\nu \in A(\mathbf{s})$, positioned in the order $f_\nu, ..., f_1, e, d, c$ from left to right or right to left, respectively, depending on whether candidate b is positioned to the left or right of candidate a, respectively. However, f_ν would prefer to deviate, contradiction.

□

At this point, because of the results from Propositions 1 and 2, the notion of a Condorcet winner becomes relevant, as the only case where an LSCG admits a PNE is when the Condorcet winner in one of parties' primaries would also beat all the candidates from the opposing party in a general election. Note that this does not imply that such a candidate would also be the Condorcet winner in the direct system, because they could lose against a candidate from their own party if the voters from the opposing party were also allowed to vote. As a consequence, the candidate who is the Condorcet winner in the primary of their party is only guaranteed to win against any candidate from their party, in a two-candidate primary election, where only the voters affiliated to the same party are allowed to vote.

Observation 5. *Let* $\Gamma^L = \Gamma^L(V, A, \mathbb{R}, d, \rho, \pi, \lhd)$. Γ^L *can have at most one PNE.*

Proof. Let $\mathbf{s}$ be a PNE of Γ^L. It follows from Propositions 1 and 2 that $\exists a^* \in A(\mathbf{s})$ such that $\mathbf{s}_{a^*} = 1$ and $\forall a \in A \setminus \{a^*\}, \mathbf{s}_a = 0$. In order for $\mathbf{s}$ to be a PNE, a^* must win in the primary of their party against any $a \in A$, with $\pi(a) = \pi(a^*)$ and we can assume, without loss of generality that $\pi(a^*) = 1$. This means that a^* is the Condorcet winner in the primary of party 1. Moreover, since $\mathbf{s}$ is a PNE, a^* must also win against any $b \in A$. with $\pi(b) = -1$, which implies that a^* also wins against the Condorcet winner from the primary of party -1. Hence a strategy profile in which the only candidate running is the Condorcet winner of party -1 cannot be a PNE of Γ^L. However, since there can only be a unique Condorcet winner in the primary of party 1, $\mathbf{s}$ must be unique, and the proof is complete.

$\square$

It now becomes clear that not all LSCGs are guaranteed to admit a PNE. Figure 16 contains an example of a LSCG that admits no PNE. The candidates and voters from parties 1 and -1 are coloured in yellow and green, respectively. We focus our attention only on the strategy profiles with one candidate running, due to Proposition 2 and we prove that none of those are PNE. Clearly, a_2 is the Condorcet winner of party 1 and a_5 is the Condorcet winner of party -1.

- For the strategy profile $\mathbf{s}$ where a_2 is the only candidate running, $A(\mathbf{s}) = \{a_2\}$, if a_3 were to run, they would be uncontested in the primary of party -1 and would win against a_2 in the general election, as $sc_g(a_3, \{a_2, a_3\}) = 6$ and $sc_g(a_2, \{a_2, a_3\}) = 4$. Therefore, $\mathbf{s}$ is not a PNE.
- The strategy profiles $\mathbf{s}$, **s'** and **s"** with $A(\mathbf{s}) = \{a_1\}, A(\mathbf{s'}) = \{a_4\}$ and $A(\mathbf{s''}) = \{a_3\}$ are not PNE either, because for $\mathbf{s}$ and **s'**, a_2 would prefer to participate in the election, while for **s"**, a_5 would prefer to participate, as they would become the overall winners, being the Condorcet winner in the primary of their party.
- The strategy profile $\mathbf{s}$ where candidate a_5 is the only one running, $A(\mathbf{s}) = \{a_5\}$ is not a PNE, because if a_2 were to run, they would compete in the general election, with $sc_g(a_5, \{a_2, a_5\}) = 4$ and $sc_g(a_2, \{a_2, a_5\}) = 6$, so a_2 would win.

Hence, there is no PNE.

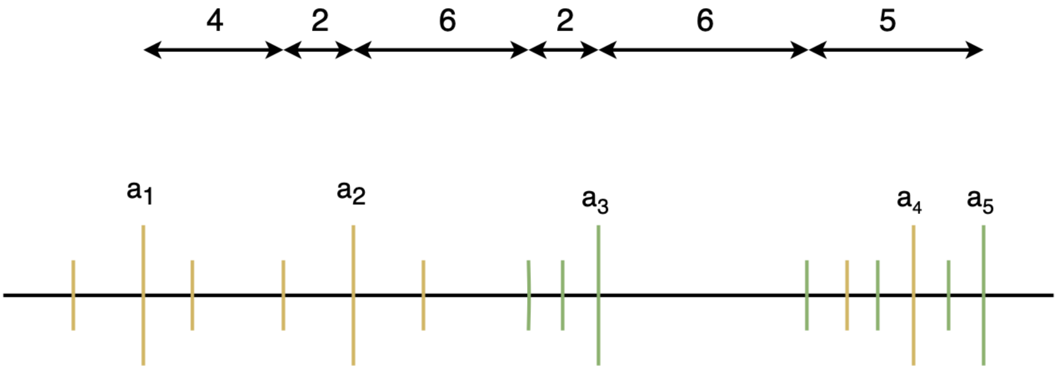


Figure 16: Example of a lazy strategic candidacy game with no PNE

5.1.2 Best-response Dynamics

The existence of a PNE for a game is generally not sufficient, from a practical point of view, as, depending on the initial state, it may be possible that the players (or candidates) never reach such a state dynamically. Based on this, we consider myopic dynamics of LSCG, where states are represented by strategy profiles and a transition from a state to another is valid only if there is precisely one player that beneficially deviates from their strategy.

5.1.2.1 J- and W-dynamics

Firstly, we consider a more restrictive type of myopic dynamics, called J-dynamics and W-dynamics, which were introduced by Obraztsova et al. [2015] for LSCGs in the direct system. For J- dynamics and W-dynamics, at each state one candidate can join or withdraw from the election, respectively, should they prefer to. Moreover, once a candidate has joined or withdrawn, they are no longer allowed to withdraw or join at a later state. Intuitively, for J-dynamics, we start from the state with no candidates running and for W-dynamics we start from the state with all candidates running. Although seemingly restrictive, the dynamics do capture an attribute of real-life election; specifically, the fact that candidates are generally not permitted to re-enter the election after they had withdrawn. Nonetheless, it is easy to see that, in this setting, convergence of such dynamics is guaranteed in at most m steps, with m being the number of candidates.

In their analysis, Obraztsova et al. [2015] investigate whether a certain state (e.g. representing a PNE of the game) can be reached by J- and W-dynamics. There are positive results (e.g. J-dynamics can reach the desired state), as well as negative ones (e.g. there exists an LSCG for which no W-dynamics terminates in a state satisfying a certain condition). Moreover, the authors investigate whether the state with all candidates running can be reached by such dynamics. Thus, we are interested in verifying whether their results can be extended to one dimensional LSCGs in the primary system.

For a LSCG $\Gamma^L = \Gamma^L(V, A, \mathbb{R}, d, \rho, \pi, \triangleleft)$, suppose $a^*, \pi(a^*) = 1$, is the Condorcet winner in the primary of party 1 and that they would win against any candidate from party -1 in a two-candidate general election. It follows from Observation 5, that the strategy profile $\mathbf{s}$ with $A(\mathbf{s}) = \{a^*\}$ is a PNE of Γ^L. We only focus on this scenario, as otherwise the game would have no PNE.

In Proposition 3, we outline two results that extend to the primary system and notably, with Proposition 4, we present a proof that there must exist a W-dynamics that terminates in a state with candidate a^* running, in contrast to the results for the direct system. One might think that, because the game only admits one PNE, all W-dynamics will converge to that state, however, this is not generally true, as a W-dynamics might terminate in a state where a candidate who has previously withdrawn would prefer to re-enter the election, but they are not allowed to.

Proposition 3. *For the LSCG Γ^L with PNE $\mathbf{s}$, $A(\mathbf{s}) = \{a^*\}$, there is a J-dynamics that leads to the state $\mathbf{s}$, but not all J-dynamics terminate in a state $\mathbf{s'}$ with $a^* \in A(\mathbf{s'})$.*

Proof. Trivially, $\emptyset \to \{a^*\}$ is a terminating J-dynamics that results in the state $\mathbf{s}$. However, if we consider the example is Figure 17, with $A = \{a, b, c, d, e\}, \pi(a) = \pi(b) = \pi(c) = 1, \pi(d) = \pi(e) = -1$ and $c \triangleleft a$, let's observe that b is the Condorcet winner of party 1 and that b would win against d or e in a general election. $\emptyset \to \{e\} \to \{d, e\} \to \{c, d, e\} \to \{a, c, d, e\}$ is a terminating J-dynamics. In the strategy profile $\mathbf{s'}$ with $A(\mathbf{s'}) = \{a, c, d, e\}$, $w_p^1(\mathbf{s'}) = a$ and $w_p^{-1}(\mathbf{s'}) = d$, candidates a and d compete in the general election where $sc_g(a, \{a, d\}) = 5$ and $sc_g(d, \{a, d\}) = 3$, so a wins the general election against d. If b were to join, c would win the primary for party 1, because c would be tied for the highest number of votes (2) with a, but preferred by the tie-breaking rule, as well as the general election against d, with 5 to 3 votes. However, because $a \succ_b c$, as $d(b, a) < d(b, c)$, b does not want to join the election and the J-dynamics terminates at **s'**. $\square$

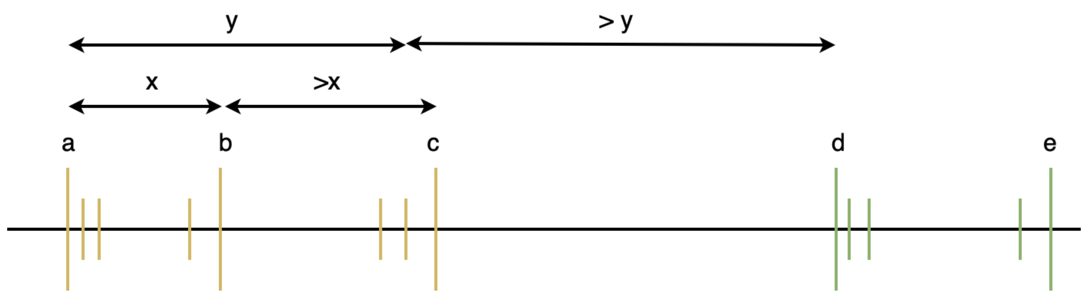


Figure 17: J-dynamics for a lazy strategic candidacy game

Proposition 4. *For the LSCG Γ^L with PNE* $\boldsymbol{s}$*,* $A(\boldsymbol{s}) = \{a^*\}$*, there must exist a W-dynamics that terminates in a state* $\boldsymbol{s}$*' with* $a^* \in A(\boldsymbol{s}')$*.*

Proof. Suppose, for a contradiction, that all W-dynamics terminate in a state **s'** with $a^* \notin A(\mathbf{s}')$. Then for each W-dynamics, there must exist a state **s"**, where a^* is the only candidate who would prefer to withdraw.

- If $\pi(w(\mathbf{s}")) = 1$, let's note that a^* cannot be the only candidate from party 1 in $A(\mathbf{s}")$, because they would then be the winner of the general election and would not want to deviate from their strategy. So there must exist $b \in A(\mathbf{s}"), \pi(b) = 1$. Moreover, there must also exist $c \in A(\mathbf{s}"), \pi(c) = 1$, since a^* is the Condorcet winner of party 1 and would win the primary of party 1 against b, as well as the general election.

  We now show that there cannot be any candidate from party -1 running in **s"**. Suppose, for a contradiction, that there exists $x \in A(\mathbf{s}")$, with $\pi(x) = -1$. Then x cannot be the only candidate running from party -1, because they would then prefer to withdraw, so there must exist $y \in A(\mathbf{s}"), \pi(y) = -1$. We can assume, without loss of generality, that x is the winner in the primary of party -1. If x and y were the only candidates running from party -1, then y would prefer to withdraw, so there must exist $z \in A(\mathbf{s}"), \pi(z) = -1$. Because there are at least three candidates running from party -1, at least two of them must be positioned to the left or to the right of $w(\mathbf{s"})$. We assume they are positioned to the left of $w(\mathbf{s"})$. Now, if we denote by a_{-1} the leftmost candidate running from party -1, they would prefer to withdraw, because by their withdrawal, the winner in the general election would still be $w(\mathbf{s"})$, or the first candidate from party -1 to the right of a_{-1} would win both the primary and the general election, contradiction.

Let's now observe that a^* must be the leftmost or rightmost candidate running from party 1 and that $w(\mathbf{s"})$ must be the rightmost of leftmost candidate running from party 1, respectively. This means that there is at least one more candidate from party 1, a_1 positioned between a^* and $w(\mathbf{s"})$. However, in this setting, a^* cannot be the Condorcet winner of party 1. To see this, $w(\mathbf{s"})$ cannot have more votes than a^* in the primary of party 1 in $w(\mathbf{s"})$, because a_1 would then beat a^* in a two-candidate primary. So, $w(\mathbf{s"})$ and a^* must be tied in the primary of party 1 in $\mathbf{s"}$ and $w(\mathbf{s"}) \lhd a^*$ and a_1 cannot get any votes in the primary, as they would again beat a^* in a two-candidate primary. However, in a two-candidate primary, a^* would then lose to $w(\mathbf{s"})$, contradiction.

- If $\pi(w(\mathbf{s"})) = -1$, using a similar argument to the previous case, it follows that there must be at least three candidates running from party 1 in $\mathbf{s"}$: a^*, b, c, so at least two candidates are positioned to the left or to the right of $w(\mathbf{s"})$. We assume they are positioned to the left of $w(\mathbf{s"})$. Then, the leftmost candidate, who cannot be a^*, would prefer to withdraw, contradiction.

□

We are also interested in determining whether the two types of dynamics can terminate in a state with all candidates running. Clearly, for W-dynamics, this is not true, since that state is not a PNE in our LSCG. However, there exists a J-dynamics that reaches a state with all candidates running. More specifically, the example in Figure 18 contains one such J-dynamics: $\emptyset \to \{e\} \to \{d, e\} \to \{c, d, e\} \to \{b, c, d, e\} \to \{a, b, c, d, e\}$.

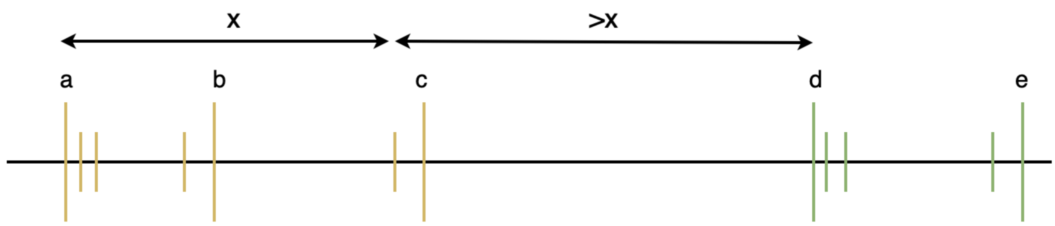


Figure 18: J-dynamics terminating in a state with all candidates running for a lazy strategic candidacy game

5.1.2.2 Equilibrium Dynamics

We extend the previous setting of myopic dynamics by allowing our dynamics to start from any initial state, and candidates may now change their status, from participating in the election to abstaining and reversely, any number of times, should they prefer to. We term this "equilibrium dynamics".

A similar concept has been studied by Polukarov et al. [2015] for regular strategic candidacy games in the direct system. However, the main focus of the paper was on dynamic candidacies with refusing voters, i.e. voters that block their most preferred candidate, if they had previously withdrawn from the election, with a given probability.

Even though such dynamics are more unusual in political scenarios, we once again note that, although we frequently use the term "election" in this project, we are not constrained to political elections. Rather, our results apply to general decision-making processes. As a consequence, as mentioned by Polukarov et al. [2015], this type of dynamics can be encountered in sale campaigns of online shops, as well as online photography, art or literature competitions, where users are not restricted to adding or removing items a certain amount of times.

Again, we focus on the case when a LSCG $\Gamma^L = \Gamma^L(V, A, \mathbb{R}, d, \rho, \pi, \triangleleft)$ admits a PNE $\mathbf{s}$ with $A(\mathbf{s}) = \{a^*\}$. We can assume, without loss of generality, that $\pi(a^*) = 1$, so a^* is the Condorcet winner in the primary of party 1 and they would win against any candidate from party -1 in a two-candidate general election. We show that the equilibrium state will be reached with probability 1 by all equilibrium dynamics. Trivially, once the equilibrium state has been reached, no agents want to deviate and the dynamics terminate. Our results from Theorem 2 are in line with those for the direct system (Polukarov et al. [2015] proved that the probability of convergence of such dynamics to an equilibrium state, for games with refusing voters, is also 1).

Theorem 2. *For the LSCG Γ^L, with a PNE $\boldsymbol{s}, A(\boldsymbol{s}) = \{a^*\}$, with probability 1, any equilibrium dynamics will reach the state $\boldsymbol{s}$.*

Proof. We define a Markov chain, with the set of states $S = \{s_1, \ldots, s_{2^m}\}$, where $m = |A|$ and each state represents a strategy profile of Γ^L, and the initial state $s_{\text{init}} \in S$. We use $w(s_i)$ to denote the winner of the election in state $s_i \in S$. Given a state $s' \in S$, if the set of candidates who would prefer to deviate is denoted by D, with $|D| = x > 0$, then there will be precisely x transitions, each having a probability of $\frac{1}{x}$, to the states where candidates in D individually change their strategy. In the case that s' corresponds to the strategy profile $\mathbf{s}$, we add a self-loop on that state with probability 1. Formally, the transition probability matrix is defined as follows:

$$P(s_i, s_j) = \begin{cases} \dfrac{1}{|A_i|} & \exists a \in A_i, s_{j_a} = 1 - s_{ia} \text{ and } \forall x \in A \setminus \{a\}, s_{j_x} = s_{ix} \\ 1 & i = j \text{ and } A_i = \emptyset \\ 0 & \text{otherwise} \end{cases}$$

We note that the Markov property does indeed hold: the probability of transitioning from a state to another only depends on the current state and is independent of any previously visited states.

Since $\mathbf{s}$ is a PNE of Γ^L, there must exist a state $s_* \in S$, corresponding to the strategy profile $\mathbf{s}$, that is absorbing, i.e. $P(s_*, s_*) = 1$. Moreover, because $\mathbf{s}$ is the only PNE of Γ^L, s_* must be the only absorbing state.

We can now show that the state s_* is reachable from any other state, and we do so by induction. Specifically, we aim to show that from a state with k candidates running, s_* can always be reached, $\forall k, 1 \leq k \leq m$:

- Base case: $k = 1$. Clearly, if a_* is the only candidate running, we are done. Otherwise, if a different candidate $a, a \neq a_*$ is the only candidate running, a_* would prefer to join, irrespective of the party affiliation of candidate a, as they would become the winner of the general election and a would then prefer to withdraw: $\{a\} \to \{a_*, a\} \to \{a_*\}$.

- Inductive step: We assume that from any state with k candidates running, s_* can always be reached, and we show that s_* is also reachable from any state with $k+1$ candidates running. Let s_x be a state with $k+1$ candidates running.

  1. $w(s_x) = a_*$. If there exists a candidate a_{-1} running from party -1 in s_x, then they would prefer to withdraw, as a_* would win against any candidate from party -1. From that state, by the induction hypothesis, we could then reach s_*. If there is no candidate running from party -1, then the leftmost or the rightmost candidate from party 1 would prefer to withdraw, depending on whether a_* is the rightmost or the leftmost candidate and by the induction hypothesis, we could then reach s_*.

  2. $w(s_x) \neq a_*$. If a_* is running in s_x, then there must be at least three candidates running from party 1, including a_*, as otherwise a_* would win the primary of party 1 and the general election. If $\pi(w(s_x)) = -1$, then at least two candidates from party 1 are positioned to the left or to the right of $w(s_x)$ and the leftmost or rightmost candidate, respectively, who cannot be a_*, would prefer to withdraw.

     If $\pi(w(s_x)) = 1$ and there is only one candidate running from party -1, that candidate would prefer to withdraw. If there are two candidates running from party -1, the candidate losing in the primary of party -1 would prefer to withdraw. Lastly, if there are at least three candidates running from party -1, at least two must be positioned to the left, or to the right of $w(s_x)$ and the leftmost or the rightmost candidate, respectively, would prefer to withdraw.

  We have shown that we can reach a state with only k candidates running and by the induction hypothesis, we could then reach s_*.

By induction, we conclude that we can reach s_* from any state with k candidates running, $\forall k, 1 \leq k \leq m$. Thus, our Markov chain is absorbing, and the probability of being in s_* after ν steps tends to 1, as ν tends to infinity [Kemeny et al., 1960].

$\square$

Remark 2. *We note that the results from Theorem 2 do not imply that all equilibrium dynamics will reach the state* $\boldsymbol{s}$*. This is also justified by the fact that the PCTL formula* $P_{\geq 1}[\Diamond \boldsymbol{s}]$ *and the CTL formula* $\forall \Diamond \boldsymbol{s}$ *are not equivalent (i.e. starting from the initial state, if all infinite paths, or in our case, sequences of states, with probability greater than 0 eventually reach the state* $\mathbf{s}$*, it does not mean that all possible paths from the initial state, including those with a probability of* 0*, will also reach that state). If we again consider the example from Figure 17,* $(\{a, b, c\} \rightarrow \{a, c\} \rightarrow \{a\} \rightarrow \{a, b\})^\omega$ *is an equilibrium dynamics that never reaches the state corresponding to the strategy profile* $\{b\}$*, the only PNE of the game, but such a path has probability* 0 *in our Markov chain.*

5.2 Eager Strategic Candidacy Games

As it could be seen, lazy strategic candidacy games do not extend very well for the primary system, at least in one dimension, since they can have at most one PNE, nonetheless, with only one candidate running. For this reason, we define a new type of games, which we term "eager strategic candidacy games" (ESCG), where in addition to being lazy, a candidate would prefer to run in the election, even if they could not become the general winner, but would win in the primary of their party. Consider, for example, the 2020 presidential election in the US, with the main contenders being Biden and Trump. It is very likely that Trump would still have preferred to participate in the election, even if he knew that he was going to lose to Biden in the general election, at least for the impact and influence he could have gained over the Republican party.

An ESCG, Γ^E, is then a tuple $\Gamma^E(V, A, M, d, \rho, \pi, \triangleleft)$. Formally, in Γ^E a candidate $c \in A$ prefers the strategy profile $\mathbf{s}$ to $\mathbf{t}$ if (i) $w(\mathbf{s}) \succ_c w(\mathbf{t})$ or (ii) $w(\mathbf{s}) = w(\mathbf{t})$, $w_p^{\pi(c)}(\mathbf{s}) = c$ and $w_p^{\pi(c)}(\mathbf{t}) \neq c$ or (iii) $w(\mathbf{s}) = w(\mathbf{t})$,$w_p^{\pi(c)}(\mathbf{s}) \neq c$, $w_p^{\pi(c)}(\mathbf{t}) \neq c$ and $\mathbf{s}_c = 0$ and $\mathbf{t}_c = 1$.

5.2.1 Nash Equilibria

We follow a similar approach to that for LSCGs, in that we begin by investigating the properties of PNE of ESCGs. The following results are related to the one dimensional metric space $(\mathbb{R}, d)$, where d is the standard Euclidean distance. For this setting, we know that a Condorcet winner always exists. In addition, just as for LSCGs, the Condorcet winner is unique.

Firstly, we formally define the PNE of ESCGs. A strategy profile $\mathbf{s}$ is a PNE of an ESCG Γ^E if no player prefers to deviate from $\mathbf{s}$, i.e. for every candidate $c \in A$, there is no strategy profile $\mathbf{t}$, with $\mathbf{t}_a = \mathbf{s}_a, \forall a \in A \setminus \{c\}$ and $\mathbf{t}_a = 1 - \mathbf{s}_a$ such that c prefers $\mathbf{t}$ to $\mathbf{s}$.

Although we have relaxed the notion of lazy candidates, so that candidates representing the losing party also get some incentive to participate in the election, it turns out that in a PNE, the candidates from the losing party are still not present in high numbers. Rather, there can only be one candidate running from the losing party in a PNE.

Proposition 5. *Let $\Gamma^E = \Gamma^E(V, A, \mathbb{R}, d, \rho, \pi, \triangleleft)$. Then Γ^E has no PNE with more than one candidate running from the losing party.*

Proof. Let $\mathbf{s}$ be a PNE with $w(\mathbf{s}) = a$ and suppose, for a contradiction, that $\exists b, c \in A(\mathbf{s}), \pi(b) = \pi(c), \pi(b) \neq \pi(a)$, with $w_p^{\pi(b)}(\mathbf{s}) = b$. Then, if c were not to run, for $\mathbf{s}$ to be a PNE, there must exist $d \in A(\mathbf{s})$, with $\pi(d) = \pi(c)$ and $a \succ_c d$ that would win the primary of their party and the general election. At this point, we follow the same arguments as in the proof of Proposition 1 to obtain a contradiction. □

Unsurprisingly, we come across the notion of Condorcet winners again, as the candidate running from the losing party can only be the Condorcet winner from that party's primary. The reasoning is also quite obvious, as they are the only candidate who are guaranteed beat any other candidate in the primary of their party, and thus, no other candidate from the same party has any reason to participate in the election.

Observation 6. *Let $\Gamma^E = \Gamma^E(V, A, \mathbb{R}, d, \rho, \pi, \triangleleft)$. In any PNE of Γ^E, the only running* *candidate from the losing party is the Condorcet winner from the associated primary.*

Proof. Let $\mathbf{s}$ be a PNE of Γ^E. It follows from Proposition 5 that $\exists a^* \in A(\mathbf{s})$ such that $\pi(a^*) \neq \pi(w(\mathbf{s})), s_{a^*} = 1$ and $\forall a \in A \setminus \{a^*\}$, with $\pi(a) = \pi(a^*)$, $s_a = 0$. In order for $\mathbf{s}$ to be a PNE, a^* must win against any other candidate in a two-candidate primary of their party, which implies that a^* is the Condorcet winner in their party's primary. □

Note that in Observation 6, we do not impose any restrictions on the number of candidates running from the winning party. For the example displayed in Figure 19, with $A = \{a_{-1}, a_1, a_2, ..., a_n\}, \pi(a_1) = ... = \pi(a_n) = 1, \pi(a_{-1}) = -1$, we show that there exists a PNE with all the candidates from party 1 running in the election. To see this, consider the strategy profile $\mathbf{s}$, with $\mathbf{s}_i = 1, \forall i \in A$. In the primary of party 1, a_i gets k voters for $2 \leq i \leq n$, while a_1 gets $k+1$ votes, so a_1 wins the primary and competes against a_{-1}, who is uncontested in their party, in the general election.

- Clearly a_1 wins the general election, as a_{-1} receives no votes. So a_1 has no reason to deviate and neither does a_{-1}, because they win the primary of their party.
- If a_i were not to run, with $3 \leq i \leq n$, all the voters who previously voted for a_i would vote for a_{i-1}, so a_{i-1} would have $2k$ votes in the primary of party 1 and would be the winner of the primary. In the general election, a_{i-1} would receive $(n-1)k$ votes, while a_{-1} would receive $(n-1)k+1$ votes and the winner would be a_{-1}. However, $a_1 \succ_{a_i} a_{-1}, \forall i$, with $3 \leq i \leq n$, so a_i would not want to deviate.
- If a_2 were not to run, a_3 would receive $2k$ votes in the primary and become the winner of the primary, but we have already shown that a_{-1} would beat a_3 in the general election and because $a_1 \succ_{a_2} a_{-1}, a_2$ would not want to deviate either.

Hence $\mathbf{s}$ is a PNE.

The number of participating candidates in PNE of ESCGs stands in a drastic contrast with our results for LSCGs and it proves that the presence of even a single candidate from the opposing party can have an important impact on whether the candidates from the winning

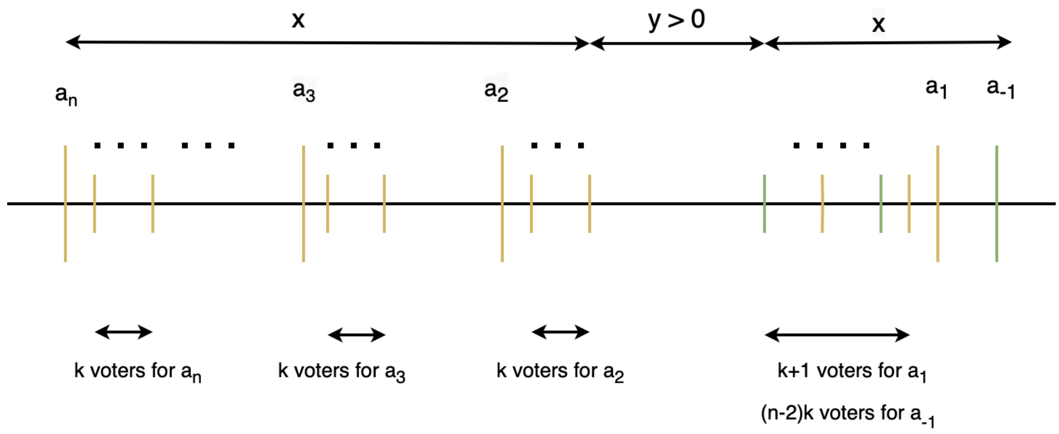


Figure 19: Example of a PNE in eager strategic candidacy games

party decide to participate in the election or not. This is because if a candidate from the winning party changes their strategy and a new candidate wins in the primary of their party, that candidate has to compete in a general election, and they could potentially lose. So it does not matter that the winner in the primary was more preferred by the deviating candidate if the overall winner is not. As a result, candidates might have to run just to ensure that they prevent someone from winning in the primary of their party and then losing the general election.

Another difference from LSCGs, for which there were instances that did not admit a PNE, is that for any ESCGs we are able to describe a characterization of a PNE in which there are only two candidates running, who, unsurprisingly, are the Condorcet winners in the primaries of each party.

Proposition 6. *Let* $\Gamma^E = \Gamma^E(V, A, \mathbb{R}, d, \rho, \pi, \triangleleft)$. *The strategy profile* $\boldsymbol{s}$, $\boldsymbol{s}_{a_1{}^*} = \boldsymbol{s}_{a^*_{-1}} = 1$ *and* $\boldsymbol{s}_a = 0, \forall a \in A \setminus \{a_1{}^*, a^*_{-1}\}$ *is a PNE of* Γ^E, *where* $a_1{}^*$ *and* a^*_{-1} *are the Condorcet winners in the primary of party* 1 *and* -1, *respectively.*

Proof. Neither of $a_1{}^*, a^*_{-1}$ wants to withdraw, since one of them wins the general election and the other is the primary winner of their party. Moreover, no candidate wants to join the election either, since they would lose in the primary of their party against the respective Condorcet winner. Thus, $\mathbf{s}$ is a PNE. $\square$

We now investigate the possible equilibrium outcomes of ESCGs in which there exists a candidate who is Pareto-dominated - the dominating candidate is more preferred than the dominated candidate by all voters. In other words, let candidate a Pareto-dominate candidate b (i.e. $a \succ_v b$ for all $v \in V$). It might be natural to assume that candidate b can never be the winner of an election in an equilibrium state of an ESCG, but as we demonstrate in Proposition 7 and in the example from Figure 20, the tie-breaking rule plays a crucial role in establishing whether such equilibrium states exist.

Proposition 7. *Let $\Gamma^E = \Gamma^E(V, A, \mathbb{R}, d, \rho, \pi, \lhd)$ and suppose candidate a Pareto-dominates candidate b. If $a \lhd b$ and $\pi(a) = \pi(b)$, then b cannot be the election winner in a PNE of Γ^E.*

Proof. Suppose, for a contradiction, that there exists a PNE $\mathbf{s}$ with $w(\mathbf{s}) = b$. Without loss of generality, let $\pi(b) = 1$. If $t = sc_p^1(b, \mathbf{s}) > 0$, then $a \notin A(\mathbf{s})$, because if a were to run, b would get no votes in the primary and hence could not be the winner of the election. Let's observe that by considering the modified strategy profile $\mathbf{s'}$ with $\mathbf{s'}_c = \mathbf{s}_c, \forall c \in A \setminus \{a\}$ and $\mathbf{s'}_a = 1$, it follows that: $sc_p^1(a, \mathbf{s'}) \geq sc_p^1(b, \mathbf{s}) \geq sc_p^1(c, \mathbf{s}) \geq sc_p^1(c, \mathbf{s'}), \forall c \in A \setminus \{a, b\}, \pi(c) = 1$. If $\exists c \in A \setminus \{a, b\}, \pi(c) = 1$ such that $sc_p^1(c, \mathbf{s}) = sc_p^1(b, \mathbf{s}) \Rightarrow b \lhd c$, because b is the winner of the election. Lastly, because $a \lhd b$ it follows that $a = w_p^1(\mathbf{s'})$. Moreover, a would also win the general election in $\mathbf{s'}$, since they would receive at least as many votes as b did in $\mathbf{s}$ and is also preferred by the tie-breaking rule, in case of a tie (a tie would only occur if b were also tied in the general election in $\mathbf{s}$ against the candidate from party -1, a_{-1}, so $b \lhd a_{-1}$). As a result, a would prefer to run in the election. Hence, a would prefer to deviate in $\mathbf{s}$, which means that $\mathbf{s}$ is not a PNE of Γ^E, contradiction. $\square$

We show why the condition $a \lhd b$ is necessary by considering the example from Figure 20, where $A = \{a, b, c, d, a_{-1}\}$, with $\pi(a) = \pi(b) = \pi(c) = \pi(d) = 1$ and $\pi(a_{-1}) = -1$ and $V = \{v_1, v_2, v_3, v_4, v_5, v_{-1}, v_{-2}, v_{-3}\}$, with $\pi(v_1) = \pi(v_2) = \pi(v_3) = \pi(v_4) = \pi(v_5) = 1$ and $\pi(v_{-1}) = \pi(v_{-2}) = \pi(v_{-3}) = -1$. Moreover, we assume that $b \lhd c \lhd a$.

Consider the strategy profile $\mathbf{s}$: $\mathbf{s}_a = 0, \mathbf{s}_b = \mathbf{s}_c = \mathbf{s}_d = \mathbf{s}_{a_{-1}} = 1$. Clearly, a_{-1} is the only candidate of party -1, so they win the primary and participate in the general election,

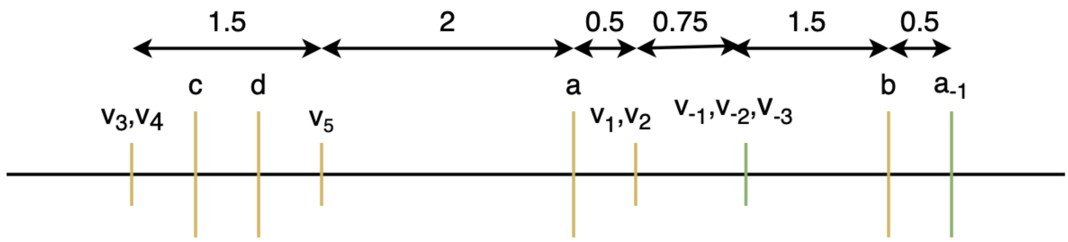


Figure 20: Example of a PNE where candidate b is the winner

against the primary winner from party 1. For this strategy profile, the primary winner of party 1 is b, because $sc_p^1(b, \mathbf{s}) = sc_p^1(c, \mathbf{s}) = 2, sc_p^1(d, \mathbf{s}) = 1$ and because $b \lhd c$, b becomes the winner of the primary. So the general election is between b and a_{-1}, with b being the winner, as $sc_g(b, \{b, a_{-1}\}) = 8$ and $sc_g(a_{-1}, \{b, a_{-1}\}) = 0$. It remains to show that $\mathbf{s}$ is a PNE.

- If candidate a were to run, then in the primary for party 1, $sc_p^1(a, \mathbf{s'}) = sc_p^1(c, \mathbf{s'}) = 2$ and $sc_p^1(d, \mathbf{s'}) = 1$, where $\mathbf{s'}_a = 1$ and $\mathbf{s'}_x = \mathbf{s}_x, \forall x \in A \setminus \{a\}$. Because $c \lhd a$, c becomes the winner of the primary. In the general election against a_{-1}, $sc_g(c, \{c, a_{-1}\}) = 3$ (votes from v_3, v_4, v_5) and $sc_g(a_{-1}, \{c, a_{-1}\}) = 5$, so a_{-1} wins the general election. But clearly $b \succ_a a_{-1}$, so a would not want to deviate.

- b does not want to deviate, because they are the winner of the election and candidates have self-supporting preferences.

- If candidate c were not to run, then for the primary of party 1, $sc_p^1(d, \mathbf{s'}) = 3$ and $sc_p^1(b, \mathbf{s'}) = 2$ and d would compete in the general election against a_{-1}, where $\mathbf{s'}_c = 0$ and $\mathbf{s'}_x = \mathbf{s}_x, \forall x \in A \setminus \{c\}$. In the general election, $sc_g(d, \{d, a_{-1}\}) = 3$ and $sc_g(a_{-1}, \{d, a_{-1}\}) = 5$, so a_{-1} would be the winner. But $b \succ_c a_{-1}$, so c would not want to deviate.

- If d were not to run, in the primary of party 1, $sc_p^1(c, \mathbf{s'}) = 3$ and $sc_p^1(b, \mathbf{s'}) = 2$ so c would be the primary winner, where $\mathbf{s'}_d = 0$ and $\mathbf{s'}_x = \mathbf{s}_x, \forall x \in A \setminus \{d\}$. In the general election, $sc_g(c, \{c, a_{-1}\}) = 3$ and $sc_g(a_{-1}, \{c, a_{-1}\}) = 5$ and a_{-1} would be the winner. Because $b \succ_d a_{-1}$, d would not want to deviate either.

- Lastly, if candidate a_{-1} were not to run, there would be no contestant in the general election from party -1, so b would win the general election. But, a_{-1} would prefer to at least win the primary, so they would not want to deviate.

Therefore $\mathbf{s}$ is indeed a PNE.

5.2.2 Complexity of PNE

As our analysis has highlighted, even for one dimensional ESCGs, the number of participating candidates in an equilibrium state can be as low as two, but also arbitrarily large. As a consequence, discovering all PNE might prove to be challenging, fact that motivates investigating the computational complexity of finding PNE in ESCGs.

For this purpose, we define the following decision problems under plurality voting:

- EAGERNE: An instance is an ESCG $\Gamma^E = \Gamma^E(V, A, M, d, \rho, \pi, \lhd)$ and a strategy profile $\mathbf{s}$. The answer is *true* if $\mathbf{s}$ is a PNE of Γ^E and *false* otherwise.
- EAGER∃NE: An instance is an ESCG $\Gamma^E = \Gamma^E(V, A, M, d, \rho, \pi, \lhd)$. The answer is *true* if Γ^E has a PNE and *false* otherwise.
- EAGER∃WNE: An instance is an ESCG $\Gamma^E = \Gamma^E(V, A, M, d, \rho, \pi, \lhd)$ and a candidate $c \in A$. The answer is *true* is there exists a PNE $\mathbf{s}$ of Γ^E, with $w(\mathbf{s}) = c$ and *false* otherwise.

Before we delve into the analysis, it is worth mentioning that analogous decision problems are considered by Obraztsova et al. [2015] for lazy strategic candidacy games in the direct system (note that their model does not use a metric space and their results would not hold for one dimensional games). Essentially, if we considered higher metric spaces for our LSCGs, all their results would also apply to our setting, since we could consider an instance with no candidates affiliated to one of the parties. Relevant to us, however, is their LAZY∃WNE problem, shown to be NP-complete, for which an instance is a LSCG $\Gamma = \Gamma(A, V, P^V, P^C, \lhd)$ and a candidate c, where P^V and P^C represent the preference profiles of the voters and candidates, respectively, and the answer is *true* if there exists a PNE $\mathbf{s}$ of Γ with $w(\mathbf{s}) = c$ and *false* otherwise.

Coming back to our analysis, it is clear that EAGERNE is in P, since for a given strategy profile $\mathbf{s}$, for each candidate c, we need to check whether c would prefer the strategy profile $\mathbf{s'}$ to $\mathbf{s}$, where $\mathbf{s'}_c = 1 - \mathbf{s}_c$ and $\mathbf{s'}_i = \mathbf{s}_i, \forall i \in A \setminus \{c\}$. Hence, we only need to check m strategy profiles, where m is the number of candidates.

Moreover, if we are dealing with one dimensional games, we know that Condorcet winners always exist for the primaries of both parties and it follows from Proposition 6 that EAGER∃NE is also in P. Generally, however, this is no longer true, if we consider metric spaces of larger dimensions.

Theorem 3. *For eager strategic candidacy games, the decision problems* EAGER∃NE *and* EAGER∃WNE *are NP-complete.*

Proof. It follows immediately, from the fact that EAGERNE is in P, that EAGER∃NE and EAGER∃WNE are in NP. To show NP-hardness, we base our proof on the findings of Obraztsova et al. [2015]. More specifically, we describe a reduction from LAZY∃WNE described above.

Let's first notice that a game with $|A| = m$ and arbitrary voters and candidates' preferences can be encoded by the hypercube of dimension m. We explicitly choose the LSCG constructed in their proof, $\Gamma(A, V, P^V, P^C, \triangleleft)$ and construct an eager strategic candidacy game $\Gamma^E = \Gamma^E(V', A', \mathbb{R}^{6r+4}, d, \rho', \pi, \triangleleft)$, where $|A| = m = 6r + 3$, with $A' = A \cup \{a_{-1}\}$, $V' = V \cup \{v_{-1}\}$ such that $\pi(a) = 1, \forall a \in A$, $\pi(a_{-1}) = -1$ and $\pi(v) = 1, \forall v \in V$, $\pi(v_{-1}) = -1$. Moreover, we keep the same tie-breaking rule, for which we additionally require $a \triangleleft a_{-1}, \forall a \in A$. Lastly, our game can be encoded in the hypercube of dimension $m+1$, so we can set $\rho'(a) = (\rho(a), 0), \forall a \in A; \rho'(a_{-1}) = (\underbrace{0, 0, ..., 0}_{\text{m zeroes}}, max_{v \in V, a \in A}(d(v, a)+1)$ and similarly, $\rho'(v) = (\rho(v), 0), \forall v \in V; \rho'(v_{-1}) = (\rho(w_1), 0)$, where $\rho : V \cup A \to \mathbb{R}^m$ is a function that positions voters and candidates of Γ into $\mathbb{R}^m$, in order to obtain the preference profiles P^V and P^C of Γ.

In the considered instance, Γ, $w_1 \in A$ is a distinguished candidate, because if Γ admits a PNE $\mathbf{s}$, then $w(\mathbf{s}) = w_1$.

We now show that if we have started with a *true* instance of LAZY∃WNE, then Γ^E has a PNE $\mathbf{s'}$ with $w(\mathbf{s'}) = w_1$, and if we have started with a *false* instance, Γ^E has no PNE. This is enough to prove that both problems are NP-hard.

If Γ has a PNE $\mathbf{s}$ with $w(\mathbf{s}) = w_1$, we show that $\mathbf{s'}$, with $\mathbf{s'}_a = \mathbf{s}_a, \forall a \in A$ and $\mathbf{s'}_{a_{-1}} = 1$ is a PNE for Γ^E with $w(\mathbf{s'}) = w_1$. Let's observe that the winner in the primary of party 1 is precisely w_1, the winner in $\mathbf{s}$. Candidate a_{-1} is uncontested in the primary of party -1, so they also advance to the general election. The general election is clearly won by w_1, as a_{-1} receives no votes, so $w(\mathbf{s'}) = w_1$. To see that $\mathbf{s'}$ is a PNE we note that:

- w_1 does not want to withdraw, as they win the general election.
- a_{-1} does not want to withdraw, as they win the primary of party -1.
- Any $a \in A \setminus \{w_1\}$ with $\mathbf{s'}_a = 0$ does not want to deviate, because of the fact that $\mathbf{s}$ is a PNE of Γ. Suppose, for a contradiction, that a would prefer to run in $\mathbf{s'}$. Then the winner of the primary of party 1 cannot be w_1. Let's assume that $a^*, \pi(a^*) = 1$, is the winner of the primary. It then follows that $a^* \succ_a w_1$, contradiction, as a would prefer to run in $\mathbf{s}$, so $\mathbf{s}$ could not be a PNE of Γ.
- Any $a \in A \setminus \{w_1\}$ with $\mathbf{s'}_a = 1$ does not want to deviate either. Suppose, for a contradiction, that a would prefer not to run in $\mathbf{s'}$. This means that the winner of the primary of party 1 could still be w_1, or the winner of the primary is $a^*, \pi(a^*) = 1$, and $a_* \succ_a w_1$. In both cases, however, it follows that a would prefer to withdraw in $\mathbf{s}$ as well, contradiction.

So $\mathbf{s'}$ is a PNE of Γ^E and $w(\mathbf{s'}) = w_1$.

Conversely, if Γ^E has a PNE $\mathbf{s}$, we prove that $\mathbf{s'}$, with $\mathbf{s'}_a = \mathbf{s}_a, \forall a \in A$ is a PNE of Γ and $w(\mathbf{s'}) = w_1$.

Clearly $\mathbf{s}_{a_{-1}} = 1$, because a_{-1} is uncontested in the primary of party -1 and they prefer winning the primary to not participating in the election. Morever, since Γ does not have a Condorcet winner [Obraztsova et al., 2015], then there is no Condorcet winner for the primary of party 1 either, so there are no PNE with only two candidates running. Lastly, we observe that the winner in the primary of party 1 corresponds to the winner of Γ and because in Γ all PNE have w_1 as the winner, it must hold that in any PNE of Γ^E, w_1 wins the primary of party 1. This directly implies that in any PNE of Γ^E, w_1 is also the winner, since a_{-1} receives no votes in a general election against w_1. So $w(\mathbf{s}) = w_1$ and $w(\mathbf{s'}) = w_1$.

Lastly, we show that **s'** is a PNE of Γ.

- w_1 does not want to withdraw, because they are the winner of the election.
- Any $a \in A \setminus \{w_1\}$ with $\mathbf{s'}_a = 0$ does not want to deviate, because of the fact that **s** is a PNE of Γ^E. Suppose, for a contradiction, that a would prefer to run in **s'**. Then the winner of the election cannot be w_1. Let's assume that $a^*, \pi(a^*) = 1$, is the winner of election. It then follows that a^* is also the winner in the primary of party 1 and the winner of the general election in **s** and $a^* \succ_a w_1$, contradiction, as a would prefer to run in **s**, so **s** could not be a PNE of Γ^E.
- Any $a \in A \setminus \{w_1\}$ with $\mathbf{s'}_a = 1$ does not want to deviate either. Suppose, for a contradiction, that a would prefer not to run in **s'**. This means that the winner of the election could still be w_1, or the winner of the election is $a^*, \pi(a^*) = 1$, and $a_* \succ_a w_1$. In both cases, however, it follows that a would prefer to withdraw in **s** as well, contradiction.

So **s'** is a PNE of Γ with $w(\mathbf{s'}) = w_1$. □

5.3 Keen Strategic Candidacy Games

Keen strategic candidacy games (KSCG) are yet another type of strategic candidacy games considered in the literature for the direct system [Lang et al., 2019]. This type of games models the fact that candidates may gain utility from participating in the election, even if they do not end up winning the election or the primary of their party. To support this statement, we cite from Bol et al. [2016]: "parties need constant visibility and are likely to endorse a candidate for an election even if she has no chance of winning". Indeed, this is a very strong argument, as it is of high importance for parties to constantly have their platform exposed, so that they receive support from the general public. In this section, we extend the model for direct elections to primary systems, and begin by investigating whether analogous results to those obtained by Lang et al. [2019] also hold in the primary system.

Formally, a KSCG, Γ^K, is a tuple $\Gamma^K(V, A, M, d, \rho, \pi, \epsilon, \triangleleft)$. We denote the participation bias by ϵ, with $\epsilon > 0$. For this, we define the utility of a candidate c for a strategy profile $\mathbf{s}$: $U_c(\mathbf{s}) = d(\rho(c), \rho(w(\mathbf{s}))) - b_c(\mathbf{s})$, where $b_c(\mathbf{s}) = 0$, if $\mathbf{s}_c = 0$ and $b_c(\mathbf{s}) = \epsilon$, if $\mathbf{s}_c = 1$. A candidate c then prefers the strategy profile $\mathbf{s}$ to $\mathbf{t}$ if $U_c(\mathbf{s}) < U_c(\mathbf{t})$. To help with readability, rather than using $d(\rho(a), \rho(b))$ to denote the distance between the positions of candidates a and b in the metric space, we will use $d(a, b)$.

5.3.1 Nash Equilibria

Similar to our work on the other two types of strategic candidacy games, we start by investigating the properties of PNE of KSCGs. The following results are related to the one dimensional metric space $(\mathbb{R}, d)$, where d is the standard Euclidean distance.

A strategy profile $\mathbf{s}$ is a PNE of a KSCG Γ^K if no player prefers to deviate from $\mathbf{s}$, i.e. for every candidate $c \in A$, there is no strategy profile $\mathbf{t}$, with $\mathbf{t}_a = \mathbf{s}_a, \forall a \in A \setminus \{c\}$ and $\mathbf{t}_a = 1 - \mathbf{s}_a$ such that c prefers $\mathbf{t}$ to $\mathbf{s}$.

We have seen that for the other two types of games we considered, candidates were only motivated to participate in an election if they could have some sort of influence over the outcome. In other words, we could think that there was a negative bias associated to the participation in the election. On the other hand, since in KSCGs candidate gain additional utility for taking part in the election, it is expected the properties of their PNE differ from what we have encountered so far for LSCGs and ESCGs. As a result, in every PNE of a KSCG, each party must be represented by at least one candidate. Moreover, in contrast with the results for ESCGs, there must be at least two candidates running from the losing party.

Proposition 8. *Let* $\Gamma^K = \Gamma^K(V, A, \mathbb{R}, d, \rho, \pi, \epsilon, \triangleleft)$. *Every PNE of* Γ^K *must have candidates running from both parties. Moreover, if there are at least two candidates associated with the losing party, at least two of them must run in any PNE of* Γ^K.

Proof. Suppose, for a contradiction, that there exists a PNE $\mathbf{s}$ of Γ^K, such that $\exists a \in A(\mathbf{s})$ and $\forall b \in A$, with $\pi(b) \neq \pi(a), \mathbf{s}_b = 0$. Moreover, we can assume, without loss of generality that $w(\mathbf{s}) = a$. Consider a candidate $b \in A$, with $\pi(b) \neq \pi(a)$. If we consider the strategy profile $\mathbf{s'}$, with $\mathbf{s'}_c = \mathbf{s}_c, \forall c \in A \setminus \{b\}$ and $\mathbf{s'}_b = 1$, then candidate a would still win the primary of their party and would compete against b in the general election. However, $U_b(\mathbf{s'}) = d(b, w(\mathbf{s'})) - \epsilon \leq d(b, a) - \epsilon < -d(b, a) = d(b, w(\mathbf{s})) = U_b(\mathbf{s})$ and b would want to deviate, contradiction.

For the second part, suppose again, for a contradiction, that there exists a PNE $\mathbf{s}$ of Γ^K, with $w(\mathbf{s}) = a$ and that $\exists b \in A(\mathbf{s})$ with $\pi(b) \neq \pi(a)$ and $\forall c \in A \setminus \{b\}, \pi(c) = \pi(b), \mathbf{s}_c = 0$. We know that $|\{c | c \in A \setminus \{b\}, \pi(c) = \pi(b)\}| > 0$, so there must exist $c \in A \setminus \{b\}, \pi(c) = \pi(b)$. By considering the strategy profile $\mathbf{s'} : \mathbf{s'}_x = \mathbf{s}_x, \forall x \in A \setminus \{c\}, \mathbf{s'}_c = 1$, because $U_c(\mathbf{s'}) = d(c, w(\mathbf{s'})) - \epsilon \leq d(c, a) - \epsilon < d(c, a) = d(c, w(\mathbf{s})) = U_c(\mathbf{s})$, candidate c would want to join the election, contradiction. □

For LSCGs, we characterised the condition that needs to be fulfilled so that a game admits a PNE, while for ESCGs, we have shown that each game must have at least one PNE, which we managed to identify precisely. It turns out, that if we focus on KSCGs where the size of the candidates' set is three, even for a small value of the participation bias (i.e. the value of the participation bias cannot influence a candidate's preference over two different outcomes), there must exist at least one PNE. Note that we are unable to fully characterise such a PNE: it could be that the strategy profile with all candidates running or some strategy profile with two candidates, from different parties, running.

Observation 7. *Let* $\Gamma^K = \Gamma^K(V, A, \mathbb{R}, d, \rho, \pi, \epsilon, \triangleleft)$. *If* $\epsilon < min_{c,a,b \in A, a \neq b}|d(c,a) - d(c,b)|$ *and* $|A| = 3$, Γ^K *must have at least one PNE.*

Proof. Suppose, for a contradiction, that Γ^K has no PNE. We can assume, without loss of generality that $A = \{a, b, c\}, \pi(a) = \pi(b) = 1, \pi(c) = -1$. Then, because the strategy profile $\mathbf{s}$, with all candidates running, is not a PNE, one candidate must want to deviate. Let's notice that it must hold that $\pi(w(\mathbf{s})) = -1$, because otherwise, $w(\mathbf{s})$ would clearly not want to deviate and neither would the other candidates, because their withdrawal would not change the general winner, so they would prefer to participate in the election. Let's assume that candidate a wins against candidate b in the primary for party -1. Because $\mathbf{s}$ is not a PNE, a would want to withdraw, which implies that in the strategy profile **s'**, with $A(\mathbf{s'}) = \{b, c\}$, i.e. only candidates b and c running, b would beat c in the general election and $U_a(\mathbf{s'}) < U_a(\mathbf{s})$. However, **s'** would then be a PNE of Γ^K, because a would not want to join the election and neither would b or c want to withdraw, contradiction. □

However, if we increase the size of the candidates' set by only 1, the existence of PNE of a KSCG is no longer guaranteed for games satisfying certain conditions (even though such conditions are very specific, they are of theoretical importance). Next, we present an analysis of the existence of PNE in such KSCGs and describe the conditions than need to be fulfilled so that a KSCG admits no PNE.

Analysis of the existence of PNE in KSCGs with $\epsilon < min_{c,a,b \in A, a \neq b}|d(c,a) - d(c,b)|$ and $|A| = 4$:

Let $\Gamma^K = \Gamma^K(V, A, \mathbb{R}, d, \rho, \pi, \epsilon, \triangleleft)$, with $\epsilon < min_{c,a,b \in A, a \neq b}|d(c,a) - d(c,b)|$ and $|A| = 4$. Suppose that Γ^K has no PNE and $A = \{a, b, c, d\}$. We distinguish two cases:

1. One party is represented by only one candidate: $\pi(a) = \pi(b) = \pi(c) = 1, \pi(d) = -1$. The strategy profile $\mathbf{s}$, $\mathbf{s}_x = 1, \forall x \in A$ is not a PNE, so one candidate from party 1 must prefer to withdraw.

   i. If $w(\mathbf{s}) = d$, let's assume that candidate a would prefer to withdraw. If a were the winner of the primary of party 1, in the strategy profile $\mathbf{s'}$ with $A(\mathbf{s'}) = \{b, c, d\}$, the winner of the primary must also win in the general election against d and $d(a, d) > d(a, w(\mathbf{s'}))$. However, $\mathbf{s'}$ is a PNE, because a would not prefer to join and neither would any of b, c, d want to withdraw: one of b, c wins the general election and the other one cannot change the election winner, so they would rather participate, due to the participation bias, contradiction. If a were not the winner of the primary of party 1, let's assume b were. Then, in the strategy profile $\mathbf{s'}$ with $A(\mathbf{s'}) = \{b, c, d\}$, candidate c must win the primary and the general election against d. But again, $\mathbf{s'}$ must be a PNE, for the same arguments as above, contradiction.

ii. If $w(\mathbf{s}) \neq d$, let's suppose that $w(\mathbf{s}) = a$. Clearly a would not want to withdraw, so let's suppose that b would prefer to withdraw. Then, in the strategy profile $\mathbf{s'}$, with $A(\mathbf{s'}) = \{a, c, d\}$, c must win the primary for party 1 and the winner of the general election must be d as otherwise, with c winning the general election, as before, $\mathbf{s'}$ would be a PNE. This implies that $d(b, a) > d(b, d)$. Because $\mathbf{s'}$ is also not a PNE, and since b does not want to join and a, d do not want to withdraw, c must prefer to withdraw and because a wins against d in a general election, $d(c, d) > d(c, a)$. Thinking about how candidates a, b, c can be ordered on a line from left to right, we dismiss the orders: b, a, c and c, a, b, because by b's withdrawal, c would receive no extra votes and would not beat a in the primary.

- For the ordering a, b, c, because $d(c, d) > d(c, a)$, d can only be positioned to the left of a, in which case $d(b, d) = d(b, a) + d(a, d)$ and $d(b, a) > d(b, d)$ cannot be true, or to the right of c and we have $d(b, a) > d(b, d) = d(b, c) + d(c, d) > d(b, c) + d(c, a) = d(b, c) + d(c, b) + d(b, a)$, contradiction. The ordering c, b, a is symmetric to this ordering.
- For the ordering a, c, b, because $d(b, a) > d(b, d)$, d cannot be positioned to the left of a. Moreover, because $d(c, d) > d(c, a)$, d must must be positioned to the right of c. However, because a beats d in the general election, c must have at least as many votes as d in a general election, because they would get at least as many votes as a would in a general election against d. This means that there

exist instances in which there is no PNE, namely, those when a is tied with d in the general election, so $a \lhd d$, c is tied with d in the general election, so $d \lhd c$ and by b's withdrawal, c receives enough extra votes to win in the primary against a, without the help of the tie-breaking rule. The ordering b, c, a is also symmetric to this ordering.

2. Each party is represented by two candidates: $\pi(a) = \pi(b) = 1, \pi(c) = \pi(d) = -1$. Let's assume that in strategy profile $\mathbf{s}$, $\mathbf{s}_x = 1, \forall x \in A, w(\mathbf{s}) = d$ and that candidate a wins the primary of party 1. If the following conditions hold, the game has no PNE:

   – In the strategy profile $\mathbf{s'}$, $\mathbf{s'}_a = 0, \mathbf{s'}_x = 1, \forall x \in A \setminus \{a\}$, b wins against d in the general election and $d(a, b) < d(a, d)$.

   – In the strategy profile $\mathbf{s''}$, $\mathbf{s''}_a = \mathbf{s''}_d = 0, \mathbf{s''}_b = \mathbf{s''}_c = 1$, c wins against b in the general election and $d(d, c) < d(d, b)$.

   To see this, from Proposition 8, any strategy profile with only one or two candidates running, cannot be a PNE. Moreover, $\mathbf{s}$, $\mathbf{s}_x = 1, \forall x \in A$ is not a PNE, because a would rather withdraw. Similarly, $\mathbf{s'}$, $\mathbf{s'}_a = 0, \mathbf{s'}_x = 1, \forall x \in A \setminus \{a\}$ is not a PNE, because d would rather withdraw. Lastly, the strategy profiles with candidates $\{a, b, c\}$, $\{a, b, d\}$ and $\{a, c, d\}$, respectively, running are not PNEs, because d, c and b, respectively, would rather join the election.

Remark 3. *Even if we enforce party separability between candidates and voters, we can still find KSCGs with four candidates that admit no PNE. Consider the example from Figure 21, where $A = \{a, b, c, d\}, \pi(a) = \pi(b) = \pi(c) = 1, \pi(d) = -1$ and suppose $a \lhd d \lhd c$ and $d(b, d) < d(b, a)$. For simplicity, we refer to $\{x, y\}$ as a strategy profile with only x and y running and we consider all possible strategy profiles:*

- From Proposition 8, we know that any strategy profile with one candidate running, or with two candidates running from the same party is not a PNE. This means that we can rule out the following strategy profiles: $\{a\}$, $\{b\}$, $\{c\}$, $\{d\}$, $\{a, c\}$, $\{a, b\}$, $\{c, b\}$.

- For $\{a, b, c, d\}$, a wins the primary for party 1, as $sc_p^1(a) = 3$, while $sc_p^1(b) = sc_p^1(c) = 2$ a. In the general election, $sc_g(a, \{a, d\}) = sc_g(d, \{a, d\}) = 5$, but since $a \triangleleft d$, a wins the general election. However, b would prefer to withdraw, because c would then win the primary for party 1 and d would beat c in the general election. So $\{a, b, c, d\}$ is not a PNE.

- $\{a, b, d\}$ is not a PNE, because a wins in the general election against d and c would prefer to join because of the participation bias.

- $\{c, b, d\}$ is not a PNE, because d wins in the general election against c and a would prefer to join, since they would become the winner.

- $\{a, c, b\}$ is not a PNE, because there is no candidate running from party -1, so d would prefer to join.

- $\{a, c, d\}$ is not a PNE, because d wins against c in the general election and c would rather withdraw, as that would make a the winner and $d(c, a) < d(c, d)$.

- $\{b, d\}$ is not a PNE, because a would prefer to join, as they would win the general election.

- $\{a, d\}$ is not a PNE, because a wins against d in the general election and b would prefer to join, due to the participation bias.

- $\{c, d\}$ is not a PNE, because d wins in the general election and a would prefer to join, due to the participation bias.

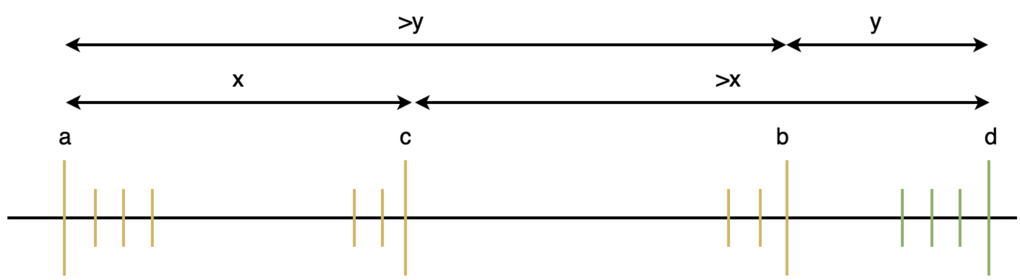


Figure 21: Example of a keen strategic candidacy game with 4 candidates, party separability and no PNE

Remark 4. *Further, even if there are two candidates representing each party, party separability is not enough to enforce the existence of a PNE. Considering the example from Figure 22,* $A = \{a, b, c, d\}, \pi(a) = \pi(b) = 1, \pi(d) = \pi(c) = -1$ *and suppose* $c \lhd b \lhd d \lhd a$*. Again, we refer to* $\{x, y\}$ *as a strategy profile with only* x *and* y *running and consider all strategy profiles:*

- From Proposition 8, we can rule out the following strategy profiles: $\{a\}$, $\{b\}$, $\{c\}$, $\{d\}$, $\{a, b\}$, $\{c, d\}$, $\{a, c\}$, $\{a, d\}$, $\{b, c\}$, $\{b, d\}$.
- $\{a, b, c, d\}$ is not a PNE, because a loses against d in the general election, due to the tie-breaking rule and they would prefer to withdraw, since b would win against d in the general election.
- $\{a, b, d\}$ is not a PNE, because d wins in the general election against a and c would prefer to join because of the participation bias.
- $\{c, b, d\}$ is not a PNE, because b wins in the general election against d and d would prefer to withdraw, as c would win against b in the general election.
- $\{a, c, b\}$ is not a PNE, because d would prefer to join, as they would win the general election.
- $\{a, c, d\}$ is not a PNE, because d wins against a in the general election and b would prefer to join because of the participation bias.

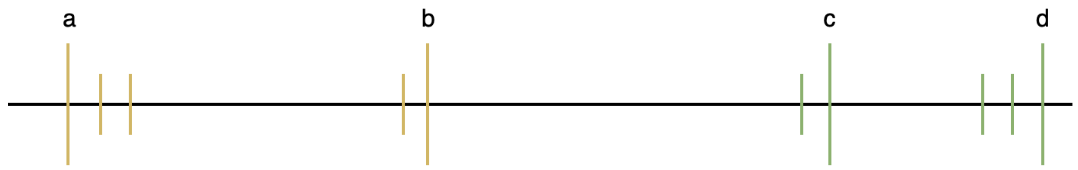


Figure 22: Further example of a keen strategic candidacy game with 4 candidates, party separability and no PNE

Lastly, we present a result which also holds for KSCGs in the direct system and provides us with a complete characterisation of PNE of KSCGs for large values of the participation bias (i.e. the value of the participation bias overrides a candidate's preference over two different outcomes). In this case, the most important thing for all candidates is participating in the election, as they gain the most utility from doing so.

Observation 8. *Let* $\Gamma^K = \Gamma^K(V, A, \mathbb{R}, d, \rho, \pi, \epsilon, \lhd)$, $\epsilon > max_{c,a,b \in A, a \neq b}|d(c,a) - d(c,b)|$. *Trivially, the strategy profile* $\boldsymbol{s}$*, with* $\boldsymbol{s}_a = 1, \forall a \in A$ *is the unique PNE of* Γ^K.

5.3.2 Complexity of PNE

Our theoretical analysis has highlighted that it is not trivial to decide whether a KSCG with a small value for the participation bias admits a PNE, even if the size of the candidates' set is small. Our results from the average-case analysis on the number of equilibria of KSCGs with four candidates, presented in the next section, also show that such a game may have up to three PNE. We could only imagine that as the size of the candidates' set increases, investigating the existence of PNE becomes even harder.

To this end, we define the following decision problems, analogous to the decision problems introduced for ESCGs:

- KEENNE: An instance is a KSCG $\Gamma^K = \Gamma^K(V, A, \mathbb{R}, d, \rho, \pi, \epsilon, \lhd)$, for which $\epsilon < min_{c,a,b \in A, a \neq b}|d(c,a) - d(c,b)|$ and a strategy profile $\mathbf{s}$. The answer is *true* if $\mathbf{s}$ is a PNE of Γ^K and *false* otherwise.
- KEEN∃NE: An instance is a KSCG $\Gamma^K = \Gamma^K(V, A, \mathbb{R}, d, \rho, \pi, \epsilon, \lhd)$, for which $\epsilon < min_{c,a,b \in A, a \neq b}|d(c,a) - d(c,b)|$. The answer is *true* if Γ^K has a PNE and *false* otherwise.
- KEEN∃wNE: An instance is KSCG $\Gamma^K = \Gamma^K(V, A, \mathbb{R}, d, \rho, \pi, \epsilon, \lhd)$, for which $\epsilon < min_{c,a,b \in A, a \neq b}|d(c,a) - d(c,b)|$ and a candidate $c \in A$. The answer is *true* is there exists a PNE $\mathbf{s}$ of Γ^K with $w(\mathbf{s}) = c$ and *false* otherwise.

We first observe that KEENNE is in P for precisely the same reasons as EAGERNE. The results from Theorem 4 are related to metric spaces of larger dimensions, rather than one dimensional games. It is also important to note that questions related to computational complexity for KSCGs in the direct system have not been considered by Lang et al. [2019], making our results even more significant.

Theorem 4. *Let* $\Gamma^K = \Gamma^K(V, A, \mathbb{R}, d, \rho, \pi, \epsilon, \triangleleft)$*, with* $\epsilon < min_{c,a,b \in A, a \neq b}|d(c,a) - d(c,b)|$*. The problems* KEEN∃NE *and* KEEN∃wNE *are NP-complete.*

Proof. Before we present our proof, we note that, at a first glance, the proof might seem very similar to that of Obraztsova et al. [2015] for LSCGs. However, while we adopt the same strategy for our proof and similar results hold, the reasoning behind is completely different (as we have mentioned, LSCGs and KSCGs are quite opposite concepts, in the sense that the former have a negative participation bias and the latter have a positive one).

We begin by observing that a direct consequence of KEENNE being in P is that both KEEN∃NE and KEEN∃wNE are in NP.

To show NP-hardness for the two problems, we describe a reduction from a restricted version of the exact cover by three sets, which we call RXC3, that was shown to be NP-complete by Gonzalez [1985]. Let us formally introduce the RXC3 decision problem:

- RXC3: An instance is a set of elements $U = \{u_1, u_2, ..., u_{3r}\}$, a family $\mathcal{Z} = \{Z_1, Z_2, ..., Z_{3r}\}$, with $Z_l = \{u_{i_l}, u_{j_l}, u_{k_l}\}$, of 3-element subsets of U, such that each element in U appears in precisely three subsets from $\mathcal{Z}$. The answer is *true* if there exists a sub-family $\hat{\mathcal{Z}} \subset \mathcal{Z}$ that is a partition of U, i.e. $\cup_{Z \in \hat{\mathcal{Z}}} Z = U$ and $Z_i \cap Z_j = \emptyset, \forall Z_i, Z_j \in \hat{\mathcal{Z}}, i \neq j$, and *false* otherwise.

Given an instance $(U, \mathcal{Z})$ of RXC3, with $|U| = 3r$, let $q = 30r^2$ and we define a keen strategic candidacy game, $\Gamma^K = \Gamma^K(V, A, \mathbb{R}^{6r+4}, d, \rho, \pi, \epsilon, \triangleleft)$ with the set of candidates $A = U \cup \mathcal{Z} \cup \{w_0, w_1, w_2, a_{-1}\}$ and $n = 3rq + 3q + 12r + 1$ voters and $\epsilon < min_{c,a,b \in A, a \neq b}|d(c,a) - d(c,b)|$. We set $n_1 = n - 1$, such that $\pi(a_{-1}) = -1$ and $\pi(a) = 1, \forall a \in A \setminus \{a_{-1}\}$ and describe the preference profiles of the voters and candidates in Tables 6 and 7, respectively. The voters from Blocks 1-4 in Table 6 are affiliated to party 1 and the sole voter from

Block 5 is affiliated to party -1. In these tables, we write U_i to denote the order $u_i \succ u_{i+1} \succ \dots \succ u_{3r} \succ u_1 \succ \dots \succ u_{i-1}$ over U and $U \setminus U'$, with $U' \subset U$ to denote the order U without the candidates from U'. Moreover, we write $\mathcal{Z}$, $\mathcal{U}$ and $\mathcal{A}$ to denote arbitrary orders over $\mathcal{Z}$, U and A, respectively and $\mathcal{Z}_{-i}$ to denote an arbitrary order over $\mathcal{Z} \setminus \{Z_i\}$. $\mathcal{Z} \setminus \{a_{-1}\}$ is used to denote the arbitrary order $\mathcal{Z}$ without candidate a_{-1}. Lastly, we define the following tie-breaking rule: $w_1 \lhd \mathcal{U} \lhd \mathcal{Z} \lhd w_2 \lhd w_0 \lhd a_{-1}$. This game can be encoded in the hypercube of dimension m, where $m = |A| = 6r + 4$, so there must exist a function $\rho : V \cup A \to \mathbb{R}^m$ that positions voters and candidates into $\mathbb{R}^m$ to obtain the preference profiles from Tables 6 and 7.

Block 1 (3r voters)				Block 2 (6r voters)					Block 3 (9r voters)				
Z_1	Z_2	...	Z_{3r}	Z_1	Z_1	...	Z_{3r}	Z_{3r}	...	Z_l	Z_l	Z_l	...
w_1	w_1	...	w_1	w_2	w_2	...	w_2	w_2	...	u_{i_l}	u_{j_l}	u_{k_l}	...
w_2	w_2	...	w_2	U_1	U_1	...	U_{3r}	U_{3r}	...	w_2	w_2	w_2	...
U_1	U_2	...	U_{3r}	$\mathcal{Z}_{-1}$	$\mathcal{Z}_{-1}$	...	$\mathcal{Z}_{-3r}$	$\mathcal{Z}_{-3r}$	...	$U \setminus \{u_{i_l}\}$	$U \setminus \{u_{j_l}\}$	$U \setminus \{u_{k_l}\}$	...
$\mathcal{Z}_{-1}$	$\mathcal{Z}_{-2}$	...	$\mathcal{Z}_{-3r}$	w_0	w_0	...	w_0	w_0	...	$\mathcal{Z}_{-l}$	$\mathcal{Z}_{-l}$	$\mathcal{Z}_{-l}$	...
w_0	w_0	...	w_0	w_1	w_1	...	w_1	w_1	...	w_0	w_0	w_0	...
a_{-1}	a_{-1}	...	a_{-1}	a_{-1}	a_{-1}	...	a_{-1}	a_{-1}	...	w_1	w_1	w_1	...
									...	a_{-1}	a_{-1}	a_{-1}	...

Block 4 ($3r(q-1) + 3q - 3r$ voters)							Block 5 (1 voter)
u_1	u_2	...	u_{3r}	w_2	w_1	w_0	w_1
$U_1 \setminus \{u_1\}$	$U_2 \setminus \{u_2\}$	...	$U_{3r} \setminus \{u_{3r}\}$	U_1	U_1	U_1	w_2
$\mathcal{Z}$	$\mathcal{Z}$	...	$\mathcal{Z}$	$\mathcal{Z}$	$\mathcal{Z}$	$\mathcal{Z}$	U_1
w_2	w_2	...	w_2	w_0	w_2	w_2	$\mathcal{Z}$
w_0	w_0	...	w_0	w_1	w_0	w_1	w_0
w_1	w_1	...	w_1	a_{-1}	a_{-1}	a_{-1}	a_{-1}
a_{-1}	a_{-1}	...	a_{-1}				
				$\underbrace{}_{(q-2r) \text{ copies}}$	$\underbrace{}_{(q-r) \text{ copies}}$	$\underbrace{}_{q \text{ copies}}$	
$\underbrace{}_{(q-1) \text{ copies}}$	$\underbrace{}_{(q-1) \text{ copies}}$		$\underbrace{}_{(q-1) \text{ copies}}$				

Table 6: Voters' preferences in the proof of Theorem 4

We will now show that if we have started with a *true* instance of RXC3, then Γ^K has a PNE $\mathbf{s}$ with $w(\mathbf{s}) = w_1$, and if we have started with a *false* instance, Γ^K admits no PNE. This is enough to prove that both problems are NP-hard.

Suppose there exists a sub-family $\hat{\mathcal{Z}} \subset \mathcal{Z}$ that is a partition of U, then the strategy profile $\mathbf{s}$, with $A(\mathbf{s}) = U \cup \{w_0, w_1, w_2, a_{-1}\} \cup (\mathcal{Z} \setminus \hat{\mathcal{Z}})$ is a PNE of Γ^K with $w(\mathbf{s}) = w_1$. Firstly, let's analyse the candidates' scores in the primary of party 1:

$Z_l, l = 1, ..., 3r$	$u_i, i = 1, ..., 3r$	w_1	w_2	w_0	a_{-1}
Z_l	u_i	w_1	w_2	w_0	a_{-1}
u_{i_l}	w_1	w_2	w_1	w_1	$\mathcal{A} \setminus \{a_{-1}\}$
u_{j_l}	w_0	w_0	w_0	$\mathcal{U}$	
u_{k_l}	$U_i \setminus \{u_i\}$	$\mathcal{U}$	$\mathcal{U}$	$\mathcal{Z}$	
w_1	$\mathcal{Z}$	$\mathcal{Z}$	$\mathcal{Z}$	w_2	
$U_l \setminus \{u_{i_l}, u_{j_l}, u_{k_l}\}$	w_2	a_{-1}	a_{-1}	a_{-1}	
$\mathcal{Z}_{-l}$	a_{-1}				
w_2					
w_0					
a_{-1}					

Table 7: Candidates' preferences in the proof of Theorem 4

- w_1 receives r votes from voters in Block 1 and $q - r$ votes from voters in Block 4, with a total of q votes.
- w_2 receives $2r$ votes from voters in Block 1 and $q - 2r$ votes from voters in Block 4, with a total of q votes.
- w_0 receives all their q votes from voters in Block 4.
- Candidate $Z_i \in \mathcal{Z} \setminus \hat{\mathcal{Z}}$ receives 1 vote from voters in Block 1, 2 votes from voters in Block 2 and 3 votes from voters in Block 3, with a total of 6 votes.
- Candidate $u_i \in U$ receives 1 vote from voters in Block 3 and $q - 1$ votes from voters in Block 4, with a total of q votes.

So w_1 wins the primary of party 1 due to the tie-breaking rule and they also win against a_{-1} in the general election. It follows that $w(\mathbf{s}) = w_1$.

We now argue that $\mathbf{s}$ is a PNE of Γ^K:

- w_1 is the winner of the general election, so they do not want to withdraw.
- If w_2 were to withdraw, u_1 would receive all their votes and would become the winner of the general election. Because $w_1 \succ_{w_2} u_1$, w_2 would rather run.
- If w_0 were to withdraw, u_1 would receive all their votes and would become the winner of the general election. Because $w_1 \succ_{w_0} u_1$, w_0 would rather run.

- If $u_i \in U$ were to withdraw, u_{i+1} would receive all their votes ($u_{i+1} = u_1$ for $i = 3r$) and would become the winner of the general election. Because $w_1 \succ_{u_i} u_{i+1}$, u_i would rather run.

- If $Z_l \in \mathcal{Z} \setminus \hat{\mathcal{Z}}$ were to withdraw, w_2 would receive two more votes and become the winner of the general election. Because $w_1 \succ_{Z_l} w_2$, Z_i would rather run.

- If $Z_l \in \hat{\mathcal{Z}}$ were to run, w_1, w_2 and w_0 would receive $q-1$, $q-2$ and q votes, respectively. Moreover, $u_{i_l}, u_{j_l}, u_{k_l}$ would receive $q-1$ votes, and some $u_x \in U \setminus \{u_{i_l}, u_{j_l}, u_{k_l}\}$ would receive q votes and win the primary over w_0, due to the tie-breaking rule and become the winner of the general election. Because $w_1 \succ_{Z_l} u_x$, Z_l would rather not run.

- a_{-1} would lose against any candidate from party 1, so they prefer to run due to the participation bias.

This shows that $\mathbf{s}$ is indeed a PNE of Γ^K.

Conversely, suppose Γ^K has a PNE $\mathbf{s}$. Then $a_{-1} \in A(\mathbf{s})$, as they would lose against any candidate from party 1 in the general election and they prefer to participate due to the participation bias. We begin by showing that $U \subset A(\mathbf{s})$. Firstly, suppose, for a contradiction, that $U \cap A(\mathbf{s}) = \emptyset$. If u_1 were to run, they would receive $3r(q-1) > \frac{n_1}{2}$ in the primary of party 1 and they would win the primary and the general election against a_{-1}. So u_1 would prefer to join the election, contradiction, as $\mathbf{s}$ is a PNE of Γ^K.

Suppose now, for a contradiction, that $U \setminus A(\mathbf{s}) \neq \emptyset$.

- If $U \setminus A(\mathbf{s}) = \{u_i\}$, u_{i+1} (where $u_{i+1} = u_1$ for $i = 3r$) would receive at least $2q-2$ votes in the primary of party 1. Because any other candidate in $\mathcal{Z} \cup \{w_0, w_1, w_2\}$ could only receive at most $q + 18r$ votes in the primary and $2q - 2 > q + 18r$, the winner of the primary must be u_{i+1} or some other u_j. Moreover, let's observe that all candidates in $\mathcal{Z} \cup \{w_1, w_2\}$ must run in $\mathbf{s}$, as their participation cannot influence the winner of the election, so they prefer to run due to the participation bias. If $w_0 \in A(\mathbf{s})$, the winner of the general election is u_{i+1} and u_i would then prefer to join, as that would make w_0 the winner and $w_0 \succ_{u_i} u_{i+1}$, contradiction. If $w_0 \notin A(\mathbf{s})$, then the winner is either u_1, if $u_i \neq u_1$ or u_2, if $u_i = u_1$. If $u_i = u_1$, u_2 wins the general election and u_1 would prefer to join the election due to the participation bias. If $u_i \neq u_1$, then u_1 wins the general election and u_i would prefer to join the election, due to the participation bias, contradiction.

- If $|U \setminus A(\mathbf{s})| \geq 2$ then suppose the winner of the general election is u_i.
  - If u_i is not tied with any other candidate in the primary and the votes of all candidates not running go to u_i, then any of the candidates not running would prefer to join the election, due to the participation bias, contradiction. Otherwise, there must exist a candidate $u_j \notin A(\mathbf{s})$ whose votes to go $u_{j+x} \in A(\mathbf{s})$. However, u_j would prefer to join the election, due to the participation bias, as the winner would still be u_i, contradiction.
  - If u_i is tied with at least one more candidate u_j, then $u_i \lhd u_j$. However, the candidate whose votes go to u_j when they do not run, would prefer to join the election, due to the participation bias, contradiction.

We conclude that $U \subset A(\mathbf{s})$.

We next prove that $\{w_0, w_1, w_2\} \subset A(\mathbf{s})$.

- If $w_0 \notin A(\mathbf{s})$. Then u_1 would receive at least $2q - 1$ votes in the primary, and would win the primary and general election. Also, all candidates in $\mathcal{Z} \cup \{w_1, w_2\}$ would prefer to run, due to the participation bias, since none of them could influence the result of the election. However, w_0 would prefer to join the election, as they would win, contradiction. So $w_0 \in A(\mathbf{s})$.
- If $w_1 \notin A(\mathbf{s})$. Then u_1 would receive at least $2q - r - 1$ votes in the primary, and would win the primary and general election. Also, all candidates in $\mathcal{Z} \cup \{w_0, w_2\}$ would prefer to run, due to the participation bias, since none of them could influence the result of the election. However, w_1 would prefer to join the election, as that would make w_0 the winner of the general election and $w_0 \succ_{w_1} u_1$, contradiction. So $w_1 \in A(\mathbf{s})$.

- If $w_2 \notin A(\mathbf{s})$. Then u_1 would receive at least $2q - 2r - 1$ votes in the primary, and would win the primary and general election. Also, all candidates in $\mathcal{Z} \cup \{w_0, w_1\}$ would prefer to run, due to the participation bias, since none of them could influence the result of the election. However, w_1 would prefer to join the election, as that would make w_0 the winner of the general election and $w_0 \succ_{w_2} u_1$, contradiction. So $w_2 \in A(\mathbf{s})$.

Let's show that $w(\mathbf{s}) = w_1$.

- $w(\mathbf{s}) = w_0$, if and only if $A(\mathbf{s}) = A$. However, in this case some Z_l would prefer to withdraw, as $u_{i_l}, u_{j_l}, u_{k_l}$ would receive q votes, same as w_0, and one of them would win the primary and the general election, depending on the tie-breaking rule. Let's suppose u_{i_l} would win the general election .Because $u_{i_l} \succ_{Z_l} w_0$, Z_l would prefer to withdraw, contradiction.
- If $w(\mathbf{s}) = w_2$, any u_i would prefer to withdraw, as u_{i+1} would become the winner of the general election and $u_{i+1} \succ_{u_i} w_2$, contradiction.
- If $w(\mathbf{s}) = u_i$, then candidate u_{i+1} would prefer to withdraw, as that would make u_{i+2} the winner and $u_{i+2} \succ_{u_{i+1}} u_i$, contradiction.
- Clearly candidates $Z_l \in \mathcal{Z}$ cannot be the winners, since they can receive at most 6 votes and candidate w_0 receives at least q votes, with $q > 6$.

We conclude that $w(\mathbf{s}) = w_1$. Let's now observe that there must be exactly $2r$ candidates running from $\mathcal{Z}$. If $|A(\mathbf{s}) \cap \mathcal{Z}| = 2r + x, x > 0$, in the primary of their party, w_1 would receive $q - r + r - x = q - x$ votes and w_0 would receive q votes, so w_1 could not win against w_0, contradiction. Similarly, if $|A(\mathbf{s}) \cap \mathcal{Z}| = 2r - x, x > 0$, in the primary of their party, w_1 would receive $q - r + r + x = q + x$ votes and w_2 would receive $q - 2r + 2(r + x) = q + 2x$ votes, so w_1 could not win against w_2, contradiction.

If we denote by $\hat{\mathcal{Z}} = \mathcal{Z} \backslash A(\mathbf{s})$, we have just shown that $|\hat{\mathcal{Z}}| = r$. We are now in a position to show that $\hat{\mathcal{Z}}$ forms a partition of U. Suppose, for a contradiction, that $\hat{\mathcal{Z}}$ does not form a partition of U. Then there must exist an element $u_i \in U$ that appears in $Z_j, Z_k \in \hat{\mathcal{Z}}, i \neq j$.

However, u_i would then receive at least $q+1$ votes in the primary of party 1 and would win against w_1, who only receives q votes, contradiction with $w(\mathbf{s}) = w_1$.

□

5.3.3 Number of Equilibria

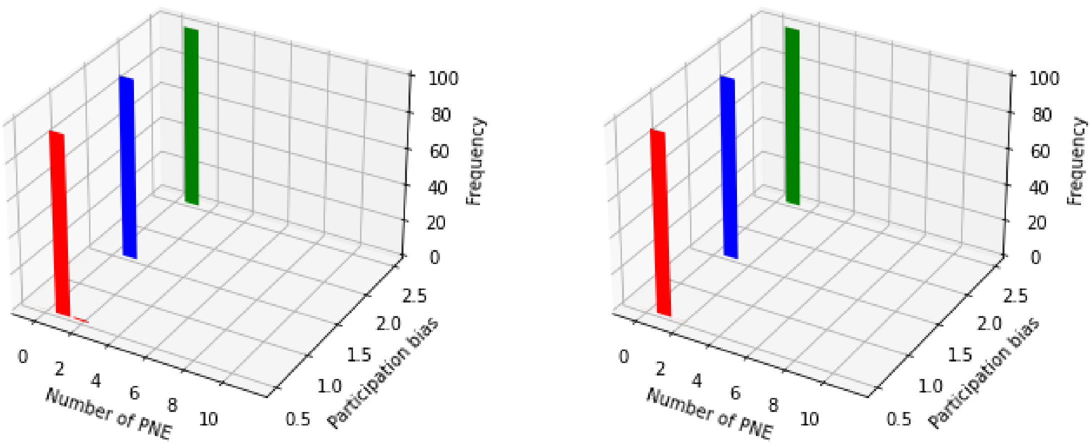
Given our theoretical analysis concerning the existence of PNE under plurality voting, for KSCGs with four candidates in one dimension, we also conduct an average-case analysis on the number of equilibria, motivated by the fact that the cases for which a KSCG admits no PNE seem very restrictive. Thus, we are interested to discover whether such instances may arise in practice. We consider both one dimensional and four dimensional KSCGs games, as for the latter we could have arbitrary voters and candidates' preferences.

In our experiments, for each candidate, we associate Borda costs to the set of candidates, depending on their proximity to each other and we only consider the plurality and Copeland voting rules. Since in KSCGs, the smaller the cost of a strategy profile, the more preferred it is by a candidate, we only need the costs in $\{0, 1, 2, 3\}$ for the candidates (as a result, a candidate will have a cost of 0 for themselves). This enables us to use values of $\{0.5, 1.5, 2.5\}$ for the participation bias, as, for example, each value between 0 and 1 for the participation bias induces the same game, as it has the same effect on the overall cost of a strategy profile.

Moreover, we randomly choose the number of candidates affiliated with party 1 for each experiment, although we do impose that each party is represented by at least one candidate. The methodology is similar to that in Section 4.4, in that we fix the positions of our candidates, and we then generate between 5 and 101 voters around the candidates. The results for 10000 instances are displayed in Table 8, with Figures 23 and 24 serving as a visual aid.

	bias=0.5				
	0 PNE	1 PNE	2 PNE	3 PNE	> 3 PNE
Plurality 1D	7	9903	89	1	0
Plurality 4D	75	9782	140	3	0
Copeland 1D	4	9990	6	0	0
Copeland 4D	7	9978	15	0	0
	bias=1.5				
	0 PNE	1 PNE	2 PNE	3 PNE	> 3 PNE
Plurality 1D	0	9998	2	0	0
Plurality 4D	16	9969	15	0	0
Copeland 1D	0	10000	6	0	0
Copeland 4D	0	10000	0	0	0
	bias=2.5				
	0 PNE	1 PNE	2 PNE	3 PNE	> 3 PNE
Plurality 1D	0	10000	0	0	0
Plurality 4D	0	10000	0	0	0
Copeland 1D	0	10000	0	0	0
Copeland 4D	0	10000	0	0	0

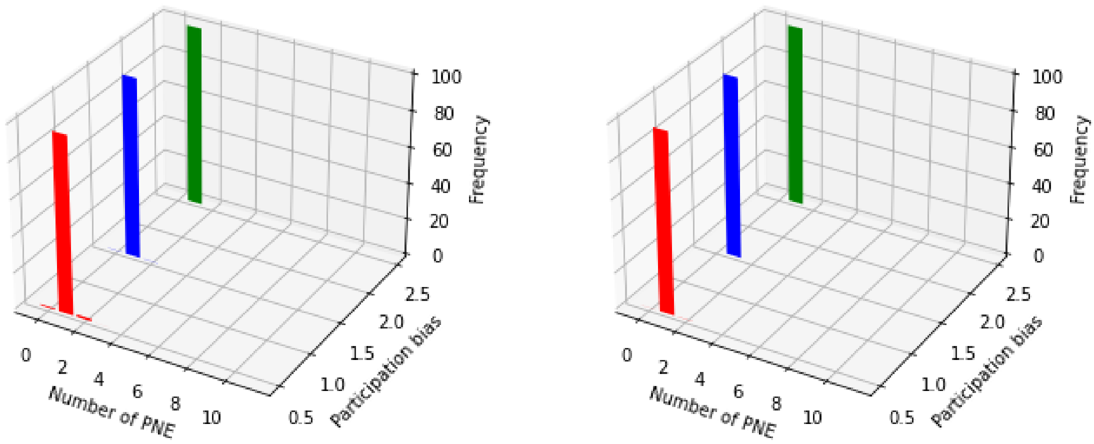
Table 8: Number of PNE in KSCGs



(a) Plurality

(b) Copeland

Figure 23: Number of PNE in one dimensional KSCGs



(a) Plurality

(b) Copeland

Figure 24: Number of PNE in four dimensional KSCGs

The results are quite intuitive: as we increase the values for the participation bias, more instances admit only one PNE (unsurprising, in those cases the unique PNE is the strategy profile with all candidates running). Also, we note that the differences between arbitrary preferences for the four dimensional case and the one dimensional case are very small, proving that we can get important insights even for one dimensional KSCGs.

Lastly, to answer the question we raised at the beginning of this section, instances that admit no PNE under plurality voting may appear in practice, although their incidence is extremely small. However, so can instances that admit more than one PNE, and they do so at a larger scale, especially for plurality. Under the Copeland voting rule, as expected, the overwhelming majority of instances, irrespective of the value of the participation bias, admit only one PNE.

We remark that our results are quite in line with those of Lang et al. [2019] for KSCGs in the direct system (only the results for five candidates are presented, but the authors claim that they are similar to those for four candidates). While in their work plurality voting yields more PNE for smaller values of the participation bias, Copeland drastically reduces their number. For larger values of the participation bias, the results are almost identical.

6 Conclusions and Future Work

In this project, we have focused on two directions: extending the results of Borodin et al. [2019] on how different voting rules, commonly used in real-world scenarios, perform under the primary and direct systems, and initiating the first analysis of strategic candidacy games under the primary system.

With respect to the first direction, from a theoretical perspective, we have mainly focused on the case where voters and candidates are uniformly distributed. We showed that for plurality voting, there may be instances where the winner in the direct system has a lower utilitarian social cost than that of the winner in the primary system, with the converse holding for Condorcet-consistent voting rules. Finally, we have conducted our own average-case analysis on the distortion of different voting rules, for which we have also considered higher dimensions for the metric space, as well as the case where voters and concentrated around distinguished candidates. We have modelled the latter with Gaussian distributions and investigated whether there is a shift in the quality of the two systems as the means of the distributions are further away, a phenomenon we termed "polarization". Notably, plurality and STV have displayed interesting trends, which were supported by further simulations.

Regarding the second direction, we initiated the analysis of strategic candidacy games under the primary systems. We extended and adapted the models for LSCGs and KSCGs to investigate the properties of their PNE in one dimension. Given the theoretical results for LSCGs, for which there may exist at most one PNE, we have also considered best-response dynamics and their convergence to the unique equilibrium state. We concluded that, for equilibrium-dynamics, the equilibrium state will eventually be reached with probability 1 (once again, we note that this does not imply that the equilibrium state will be reached by any such dynamics). Moreover, we have introduced a novel model of strategic candidacy games, for which we have also proven computational complexity results, regarding the existence of PNE, for larger dimensions of the metric space. We have shown similar results for KSCGs and we also performed an average-case analysis on the number of equilibria for four candidates and a small value of the participation bias. Our results were close to those in the related literature for the direct system, with most games admitting only one PNE (in most of those instances, the strategy profile with all candidates running was the unique PNE).

There are several possible directions for further work. Firstly, for the comparison between the two systems, one could theoretically consider a probabilistic model, where each party is represented by a distribution (e.g. Gaussian distribution) and then random samples are drawn to obtain the candidates and voters' positions. We would be interested in answering the following question: "is it the case that primaries are better with high probability?". Secondly, other particular cases (e.g. where the ratio between the number of voters and candidates from each party is fixed and the candidates and voters are uniformly distributed) could be considered in both one and higher dimensions. Finally, it would be relevant to verify whether our results generalise to settings with more than two parties.

Regarding strategic candidacies under the primary system, one could investigate whether our results about the properties of PNE for each category of games generalise for higher dimensional games. Similarly, other voting rules, apart from plurality, could be considered. Lastly, another direction would be considering strategic voting, or perhaps even combining the two concepts (similar to what Brill and Conitzer [2015] do for the direct system).

6.1 Reflection

Working on this dissertation has been an incredibly enjoyable experience. Computational social choice and, more specifically, voting, is a very broad field, thus with rich possibilities for research. Having initially conducted a comprehensive background review on the topic of voting in primary systems, I started with less difficult questions and, as I was making progress on the topic, I managed to eventually consider more demanding questions.

The subjects covered in the Computational Game Theory were the most relevant to this project. However, I also believe that, focusing on writing structured proofs in the Probability and Computing, Probabilistic Model Checking and Automata, Logic and Games modules has also helped me present the results in a scholarly manner. Dealing with the COVID-19 pandemic and having to complete almost the entirety of my MSc degree remotely was not easy and, at times, very challenging. Nonetheless, I am content with what I have achieved in this thesis, not only limited to the presented results, but also to the in-depth understanding I have acquired on the topic.

References

Anshelevich, E., Bhardwaj, O., Elkind, E., Postl, J., and Skowron, P. (2018). Approximating optimal social choice under metric preferences. *Artificial Intelligence*, 264:27–51.

Anshelevich, E., Filos-Ratsikas, A., Shah, N., and Voudouris, A. A. (2021). Distortion in social choice problems: The first 15 years and beyond. *arXiv preprint arXiv:2103.00911*.

Anshelevich, E. and Postl, J. (2017). Randomized social choice functions under metric preferences. *Journal of Artificial Intelligence Research*, 58:797–827.

Arrow, K., Yale College (New Haven, C. D. o. E. C. F. f. R. i. E., commission for research in economics, C., Collection, K. W. H. C., University, Y., for Research in Economics, Y. U. C. F., and for Research in Economics at Yale University, C. F. (1963). *Social Choice and Individual Values*. Number nr. 12 in Cowles Commission Monographs. Wiley.

Black, D. (1948). On the rationale of group decision-making. *Journal of political economy*, 56(1):23–34.

Bol, D., Blais, A., Laslier, J.-F., and Macé, A. (2016). Electoral system and number of candidates: Candidate entry under plurality and majority runoff. In *Voting Experiments*, pages 303–321. Springer.

Borodin, A., Lev, O., Shah, N., and Strangway, T. (2019). Primarily about primaries. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 1804–1811.

Boutilier, C., Caragiannis, I., Haber, S., Lu, T., Procaccia, A. D., and Sheffet, O. (2015). Optimal social choice functions: A utilitarian view. *Artificial Intelligence*, 227:190–213.

Brill, M. and Conitzer, V. (2015). Strategic voting and strategic candidacy. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 29.

Caragiannis, I., Nath, S., Procaccia, A. D., and Shah, N. (2017). Subset selection via implicit utilitarian voting. *Journal of Artificial Intelligence Research*, 58:123–152.

Caragiannis, I. and Procaccia, A. D. (2011). Voting almost maximizes social welfare despite limited communication. *Artificial Intelligence*, 175(9-10):1655–1671.

Conitzer, V. and Sandholm, T. (2012). Common voting rules as maximum likelihood estimators. *arXiv preprint arXiv:1207.1368*.

de Caritat de Condorcet, J. A. N. (1785). *Essai sur l'Application de l'Analyse*.

Dennisen, S. L. and Müller, J. P. (2015). Agent-based voting architecture for traffic applications. In *German Conference on Multiagent System Technologies*, pages 200–217. Springer.

Dutta, B., Jackson, M. O., and Le Breton, M. (2001). Strategic candidacy and voting procedures. *Econometrica*, 69(4):1013–1037.

Dwork, C., Kumar, R., Naor, M., and Sivakumar, D. (2001). Rank aggregation methods for the web. In *Proceedings of the 10th international conference on World Wide Web*, pages 613–622.

Elkind, E., Faliszewski, P., Laslier, J.-F., Skowron, P., Slinko, A., and Talmon, N. (2017). What do multiwinner voting rules do? an experiment over the two-dimensional euclidean domain. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31.

Elkind, E., Faliszewski, P., and Slinko, A. (2009). On distance rationalizability of some voting rules. In *Proceedings of the 12th Conference on Theoretical Aspects of Rationality and Knowledge*, pages 108–117.

Elkind, E., Faliszewski, P., and Slinko, A. M. (2010). On the role of distances in defining voting rules. In *Proceedings of the 9th Conference on Autonomous Agents and MultiAgent Systems*, volume 10, pages 375–382.

Enelow, J. and Hinich, M. J. (1990). *The theory of predictive mappings*, volume 7. Cambridge University Press Cambridge.

Enelow, J. M. and Hinich, M. J. (1984). *The spatial theory of voting: An introduction.* CUP Archive.

Feldman, M., Fiat, A., and Golomb, I. (2016). On voting and facility location. In *Proceedings of the 2016 ACM Conference on Economics and Computation*, pages 269–286.

Gehrlein, W. V. (1987). The impact of social homogeneity on the condorcet efficiency of weighted scoring rules. *Social Science Research*, 16(4):361–369.

Gehrlein, W. V., Lepelley, D., et al. (2017). *Elections, voting rules and paradoxical outcomes*. Springer.

Gehrlein, W. V. and Plassmann, F. (2014). A comparison of theoretical and empirical evaluations of the borda compromise. *Social Choice and Welfare*, 43(3):747–772.

Gehrlein, W. V. and Valognes, F. (2001). Condorcet efficiency: A preference for indifference. *Social Choice and Welfare*, 18(1):193–205.

Ghosh, S., Mundhe, M., Hernandez, K., and Sen, S. (1999). Voting for movies: the anatomy of a recommender system. In *Proceedings of the third annual Conference on Autonomous Agents*, pages 434–435.

Gibbard, A. (1973). Manipulation of voting schemes: a general result. *Econometrica: journal of the Econometric Society*, pages 587–601.

Gkatzelis, V., Halpern, D., and Shah, N. (2020). Resolving the optimal metric distortion conjecture. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1427–1438. IEEE.

Goel, A., Krishnaswamy, A. K., and Munagala, K. (2017). Metric distortion of social choice rules: Lower bounds and fairness properties. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, pages 287–304.

Gonzalez, T. F. (1985). Clustering to minimize the maximum intercluster distance. *Theoretical computer science*, 38:293–306.

Kemeny, J. G., Snell, J. L., et al. (1960). *Finite markov chains*, volume 356. van Nostrand Princeton, NJ.

Kempe, D. (2020a). An analysis framework for metric voting based on lp duality. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 2079–2086.

Kempe, D. (2020b). Communication, distortion, and randomness in metric voting. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 2087–2094.

Lang, J., Markakis, V., Maudet, N., Obraztsova, S., Polukarov, M., and Rabinovich, Z. (2019). Stategic Candidacy with Keen Candidates. Presented at the Games, Agents and Incentives Workshop.

Lang, J., Maudet, N., and Polukarov, M. (2013). New results on equilibria in strategic candidacy. In *International Symposium on Algorithmic Game Theory*, pages 13–25. Springer.

May, K. O. (1952). A set of independent necessary and sufficient conditions for simple majority decision. *Econometrica: Journal of the Econometric Society*, pages 680–684.

Meskanen, T. and Nurmi, H. (2008). Closeness counts in social choice. In *Power, freedom, and voting*, pages 289–306. Springer.

Munagala, K. and Wang, K. (2019). Improved metric distortion for deterministic social choice rules. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, pages 245–262.

Obraztsova, S., Elkind, E., Polukarov, M., and Rabinovich, Z. (2015). Strategic candidacy games with lazy candidates. In *Twenty-Fourth International Joint Conference on Artificial Intelligence*.

Ordeshook, P. C. and McKelvey, R. (1990). A decade of experimental research on spatial models of elections and committees. *Advances in the Spatial Theory of Voting*, page 99.

Plott, C. R. (1967). A notion of equilibrium and its possibility under majority rule. *The American Economic Review*, pages 787–806.

Polukarov, M., Obraztsova, S., Rabinovich, Z., Kruglyi, A., and Jennings, N. R. (2015). Convergence to equilibria in strategic candidacy. In *Twenty-Fourth International Joint Conference on Artificial Intelligence*.

Procaccia, A. D. and Rosenschein, J. S. (2006). The distortion of cardinal preferences in voting. In *International Workshop on Cooperative Information Agents*, pages 317–331. Springer.

Saari, D. G. (1985). The optimal ranking method is the borda count. Technical report, Discussion paper.

Sabato, I., Obraztsova, S., Rabinovich, Z., and Rosenschein, J. S. (2017). Real candidacy games: a new model for strategic candidacy. In *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems*, pages 867–875.

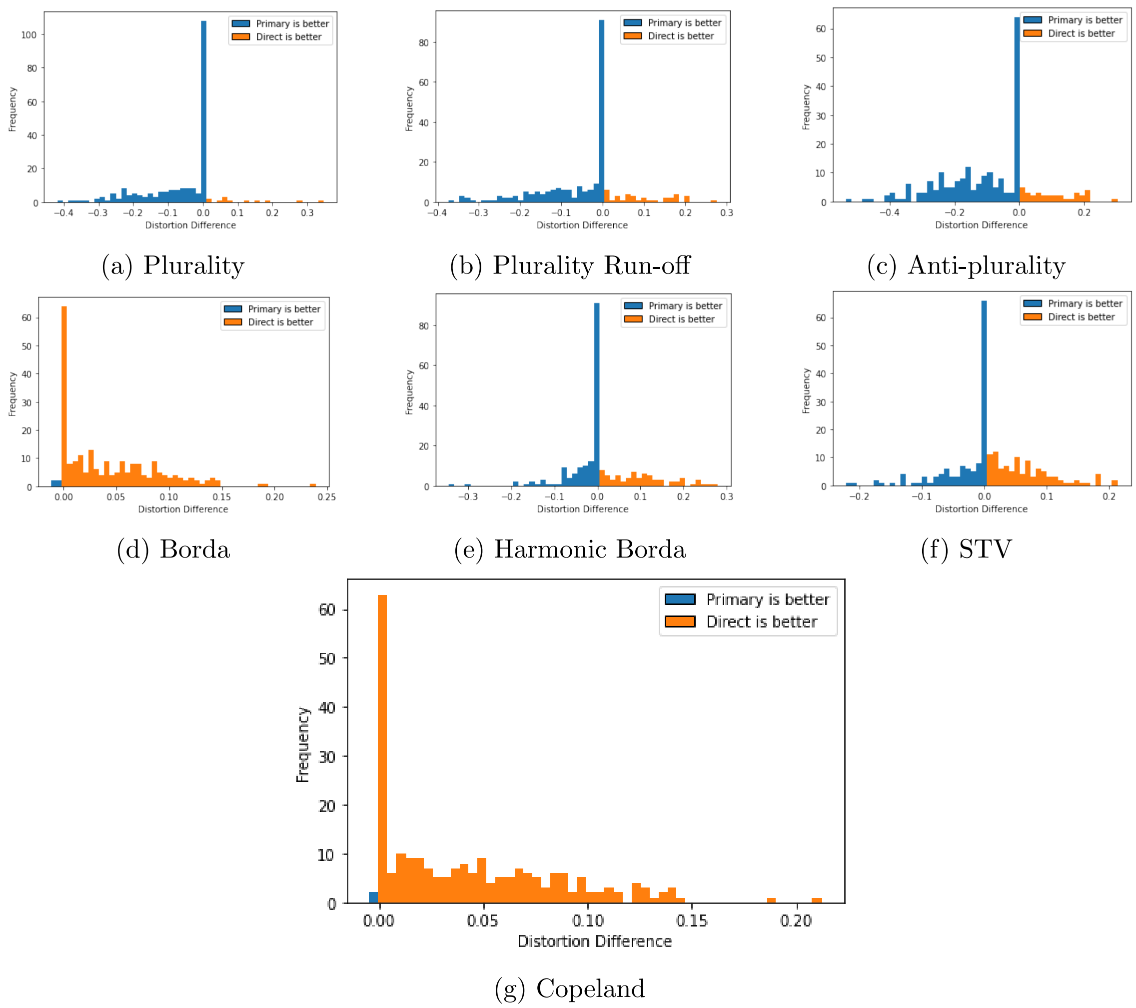
Satterthwaite, M. A. (1975). Strategy-proofness and arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of economic theory*, 10(2):187–217.

Schofield, N. (2007). *The spatial model of politics*. Routledge.

Skowron, P. K. and Elkind, E. (2017). Social choice under metric preferences: Scoring rules and STV. In *Thirty-First AAAI Conference on Artificial Intelligence*.

Young, H. P. (1975). Social choice scoring functions. *SIAM Journal on Applied Mathematics*, 28(4):824–838.

Appendix A Difference in Distortion

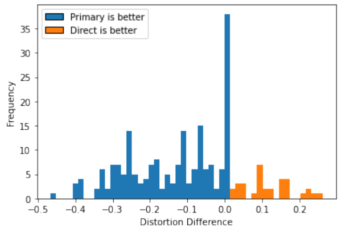


(a) Plurality (b) Plurality Run-off (c) Anti-plurality

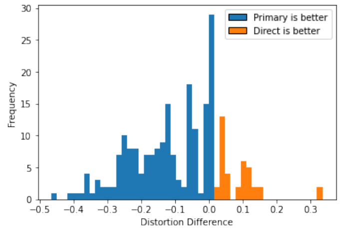
(d) Borda (e) Harmonic Borda (f) STV

(g) Copeland

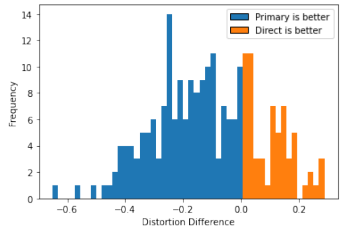
Figure 25: Difference in distortion between the primary and direct system for separable three dimensional election instances



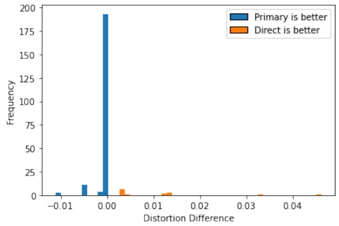
(a) Plurality



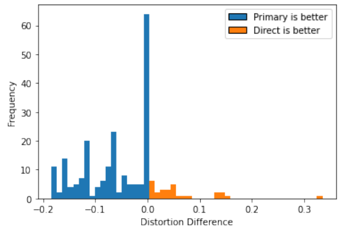
(b) Plurality Run-off



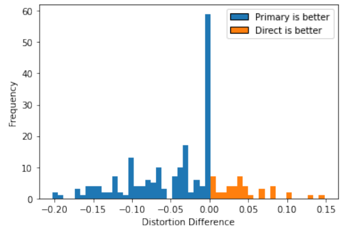
(c) Anti-plurality



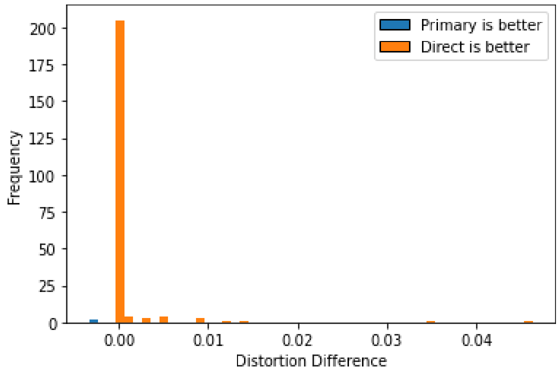
(d) Borda



(e) Harmonic Borda

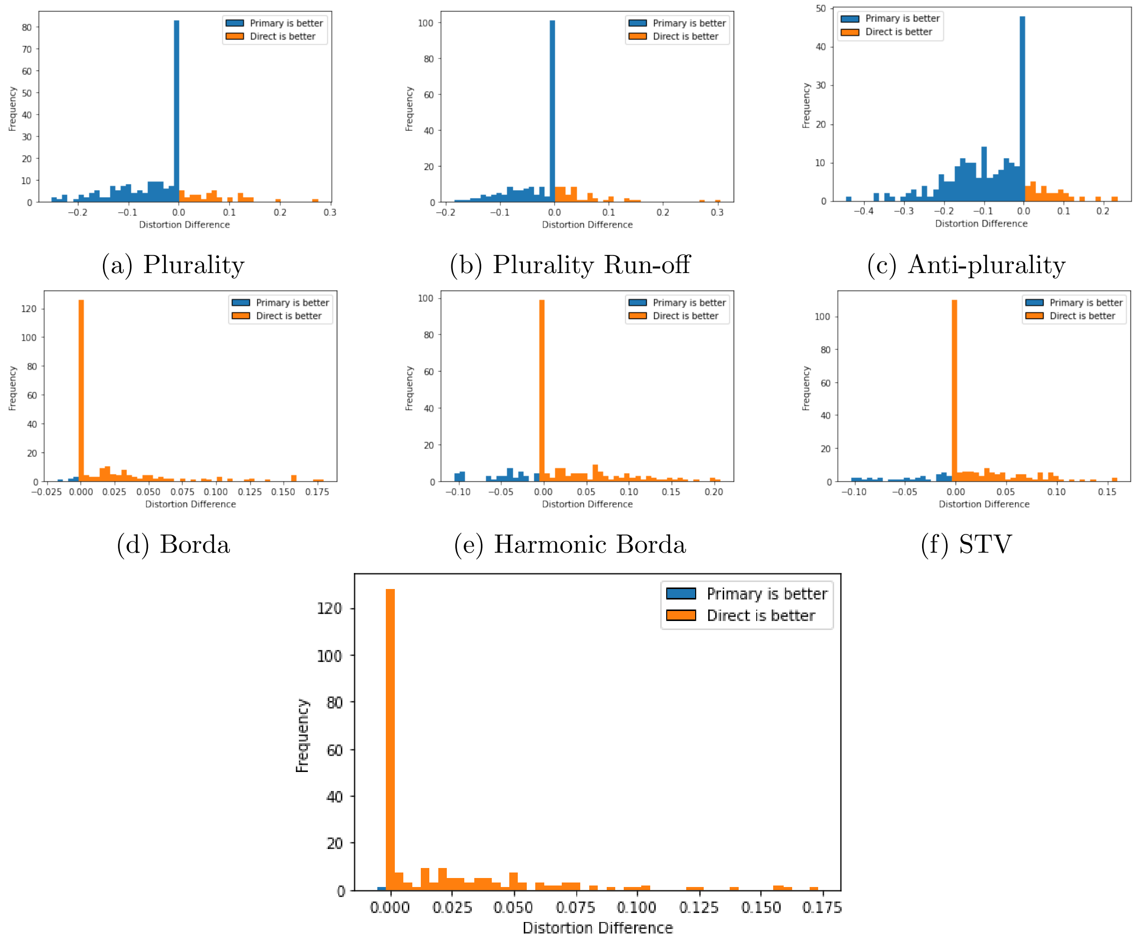


(f) STV



(g) Copeland

Figure 26: Difference in distortion between the primary and direct system for three dimensional election instances

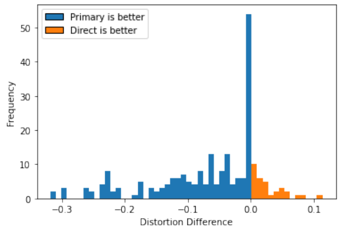


(a) Plurality (b) Plurality Run-off (c) Anti-plurality

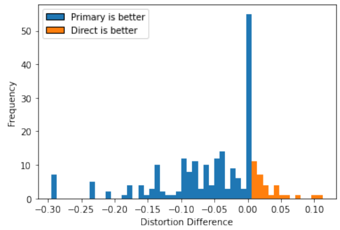
(d) Borda (e) Harmonic Borda (f) STV

(g) Copeland

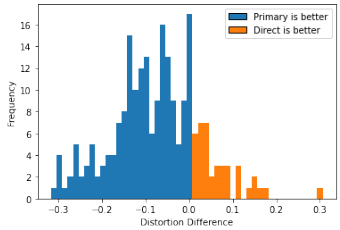
Figure 27: Difference in distortion between the primary and direct system for separable five dimensional election instances



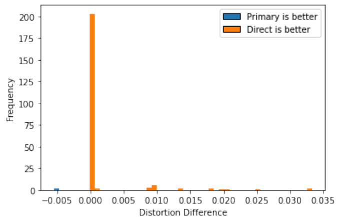
(a) Plurality



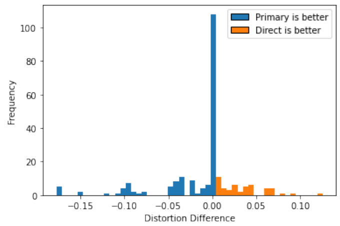
(b) Plurality Run-off



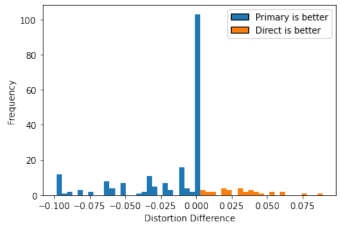
(c) Anti-plurality



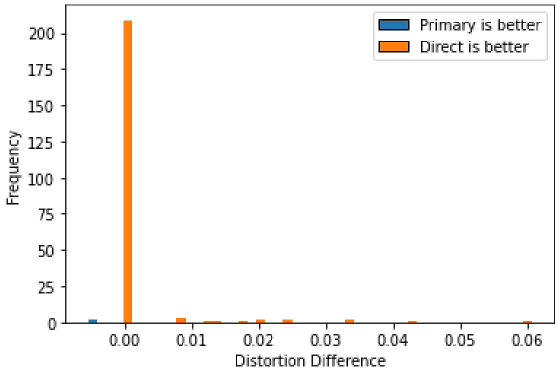
(d) Borda



(e) Harmonic Borda



(f) STV



(g) Copeland

Figure 28: Difference in distortion between the primary and direct system for five dimensional election instances