

# Representative Democracy and the Implementation of Majority-preferred Alternatives

Katherine Baldiga\*

The Ohio State University

August 30, 2012

## Abstract

In this paper, we study representative democracy, one of the most popular classes of collective decision-making mechanisms, and contrast it with direct democracy. In a direct democracy, individuals have the opportunity to vote over the alternatives in every choice problem the population faces. In a representative democracy, the population commits to a candidate *ex ante* who will then make choices on its behalf. While direct democracy is normatively appealing, representative democracy is the far more common institution because of its practical advantages. The key question, then, is whether representative democracy succeeds in implementing the choices that the group would make under direct democracy. We find that, in general, it does not. We analyze the theoretical setting in which the two methods are most likely to lead to the same choices, minimizing potential sources of distortion. We model a population as a distribution of voters with strict preferences over a finite set of alternatives and a candidate as an ordering of those alternatives that serves as a binding, contingent plan of action. We focus on the case where the direct democracy choices of the population are consistent with an ordering of the alternatives. We show that even in this case, where the normative recommendation of direct democracy is clear, representative democracy may not elect the candidate with this ordering.

---

\*A special thank you to Jerry Green for many helpful conversations about this work. This author also gratefully acknowledges the comments and suggestions of Steven Brams, Lucas Coffman, Drew Fudenberg, Yuichiro Kamada, Klaus Nehring, Alvin Roth, Kenneth Shepsle, Tomasz Strzalecki, William Zwicker, seminar participants at Harvard University, and conference participants at the New York University Graduate Student Political Economy Conference, EconCon 2011, and the Judgment Aggregation and Voting Theory Workshop.

# 1 Introduction

Direct democracy is a fundamental principle of collective decision-making. When a choice problem arrives, individuals should have the opportunity to express preferences over the alternatives. A good decision-making rule then aggregates these preferences into a choice that reflects the will of the group. While many aspects of social decision-making have been debated, this individual right to direct participation has remained a normative ideal from both a theoretical and popular standpoint.

Yet, despite its normative appeal, direct democracy is relatively rare in practice. Just over half of the states in the United States allow for recalls and/or popular referendums and no forms of direct democracy exist at the federal level (NCSL.org). Direct democracy plays a similarly minor role in the governments of countries around the world, with the well-studied exception of Switzerland (Frey 1994). Most institutions instead take the form of representative democracies, under which elected officials make decisions on behalf of the electorate. From a practical perspective, representative democracies have an edge over direct democracies as they dramatically reduce transaction costs and shift the burden of decision-making to a small group of well-informed leaders. The question, then, is whether representative democracy, with its practical advantages, can successfully implement the choices that would be reached under the more normatively attractive direct democracy.

This paper tackles this question from a theoretical perspective. The first and most crucial step of our analysis is to build models of both of these forms of governance. We begin with a population of individuals with strict preferences over a finite set of alternatives. Then, we develop and apply theories of how outcomes are reached under each system: how alternatives are chosen under direct democracy and how candidates are elected under representative democracy. With these models in place, we ask whether the choices made by the population via direct democracy are preserved under representative democracy.

Our goal is to construct models of direct and representative democracy that are as similar to each other as possible, so that we may study the case that is most likely to lead to positive results. As we describe below, both models depend only on individuals' preferences over pairs of alternatives. Thus, our theories map the same preference information from the population into two sets of outcomes.

We take a majoritarian approach to modeling both direct and representative democracy. Majoritarianism requires pairwise comparisons of alternatives, with the majority preference dictating the social decision. In addition to being a widely-implemented and well-accepted political principle, majoritarianism has been shown to be the collective decision-making rule that works best over the largest class of domains (Dasgupta Maskin 2008). We use

tournament theory, a classic approach to modeling majoritarian systems (Laslier 1997). Under tournament theory, we first compare all alternatives pairwise. Then, we use this complete set of pairwise comparisons to determine the winners.

We use the tournament generated by majority rule over alternatives to analyze outcomes under direct democracy. Under this system, alternatives are compared pairwise, with the majority preference dictating the winner of each contest. We focus on populations whose majority rule tournaments have no cycles. When there are no cycles in the majority rule tournament, there exists a clear ranking of alternatives, with each alternative majority-preferred to all others below it in the ranking.

In our model of representative democracy, members of the population vote over candidates rather than alternatives. To simplify our setting, we define a candidate as an ordinal ranking of alternatives: a binding, contingent plan of action for future choice problems. We assume that when a choice problem of alternatives arrives, the social decision is made according to the ordering of the elected candidate.

The key modeling assumption is how individuals vote over candidates. We map individuals' preferences over alternatives into preferences over candidates by assuming that an individual votes for the candidate with whom she is most likely to agree about the choice from a randomly-selected pair of alternatives. While we could determine preferences over candidates in other ways, this method parallels the direct democracy case in a unique way. In this version of representative democracy, individuals vote for candidates solely based upon the decisions those candidates would make in choice problems that consist of pairs of alternatives. These are exactly the choice problems faced by individuals in the direct democracy tournament. Therefore, individuals' pairwise preferences over alternatives completely determine both their votes over alternatives and over candidates.

Once we know individuals' preferences over candidates, we can compare these candidates pairwise. In a comparison of two candidates, the winning candidate is the candidate who earns a majority of the population's votes: the candidate with whom the majority of the population is most likely to agree. This type of political action generates a new type of tournament, a tournament over the candidates.

These models provide a method for mapping the preferences of the population into two distinct tournaments: a tournament over alternatives and a tournament over candidates. Each tournament is a complete, asymmetric binary relation. That is, the tournament describes all possible pairwise comparisons within the set. This information provides the basis for determining which alternatives or which candidates would be the winners under majoritarian voting methods. In the simplest case, there will be an alternative or candidate that defeats all others in pairwise comparisons, a Condorcet winner. In these cases, it is

easy to see that any majoritarian method will elect the Condorcet winner. Indeed, for our tournaments over alternatives, because we focus exclusively on populations whose majority preferences do not cycle, there will always be a Condorcet winner of the tournament. However, the tournaments over candidates generated by these same populations may cycle. Identifying the winners of these tournaments is non-trivial.

We use the uncovered set of the tournament as our solution concept. The uncovered set has several nice theoretical properties: it is a subset of the top cycle, it is Condorcet consistent, it contains no Pareto dominated alternatives, and it contains most other popular tournament solution concepts (Miller 1980, Laslier 1997). From an applied standpoint, the uncovered set is a useful tournament solution concept because it characterizes the outcomes under many familiar majoritarian voting methods (Miller 1980, Shepsle and Weingast 1984). For example, Miller (1980) shows that the uncovered set contains the set of electable candidates under a simple model of two-party electoral competition.

With this machinery in place, we can address the question of whether representative democracy achieves the same outcomes as direct democracy. For example, suppose a population has a transitive majority preference over three alternatives:  $a_1 \succ a_2 \succ a_3$ . In this case, we ask whether the winning candidate under representative democracy is the ordering  $a_1 \succ a_2 \succ a_3$ , since this candidate would make the same choices as the population does under direct democracy.

We find that for problems with a small number of alternatives, representative democracy does succeed in electing candidates that implement the choices made under direct democracy. We show that when there are only three alternatives, if majority rule over the alternatives is consistent with an ordering, then this ordering is the sole member of the uncovered set of the tournament over candidates. Thus, our models predict that direct and representative democracy lead to the same choices for these small problems.

But, for general problems, we derive a negative result. We show that even for the most well-behaved populations, where majority preferences over the alternatives are consistent with an ordering, representative democracy may not elect the candidate with this ordering. That is, when given a chance to commit *ex ante* to a binding plan of action, the population may select a plan which contradicts the majority preferences on some choice problems. This result is surprising in light of the fact that both tournaments depend only on the population's pairwise preferences over alternatives. Our models of direct and representative democracy operate on the same preference information and arrive at different outcomes.

Importantly, the failure of representative democracy that we identify here fundamentally differs from many of the existing negative results in the fields of social choice and political economy. Our negative result arises in the case where preference aggregation is typically

least problematic: when there are no cycles in majority rule over alternatives. Thus, it is clear that the divergence of outcomes under direct and representative democracy does not stem from a tension between matching majority preferences over alternatives and producing a social ordering.

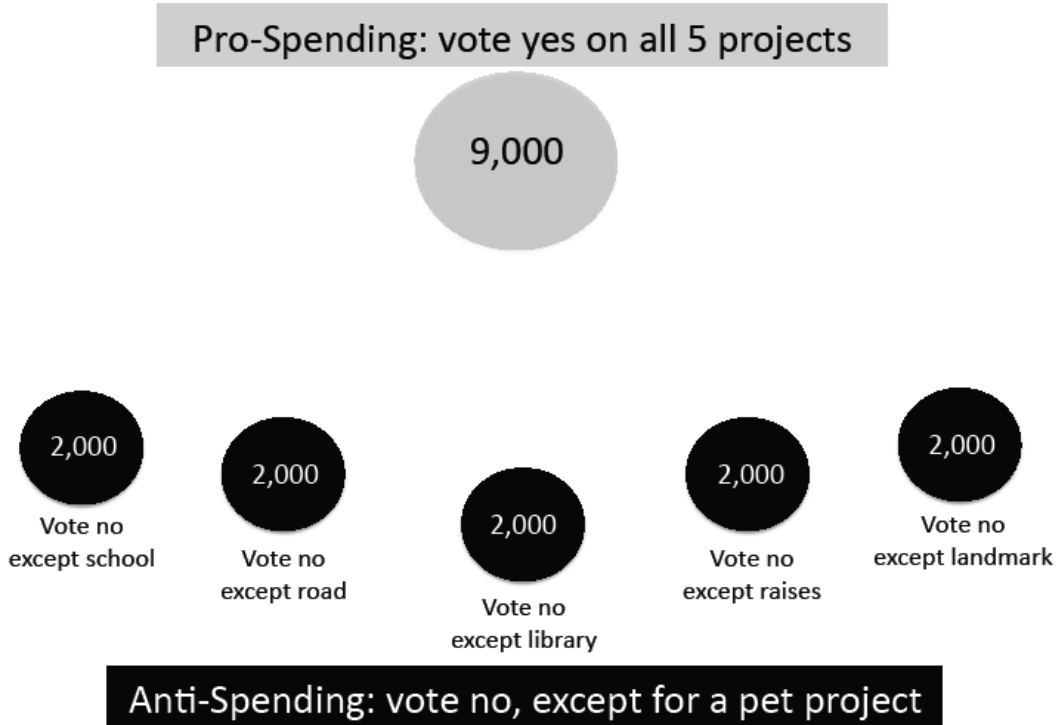
We can provide intuition for our result with a simple example. Consider a community that needs to make spending decisions. There are five possible projects that the community could fund: they could build a new school, repair a local road, improve the public library, give raises to public employees, and/or restore a historic landmark. For simplicity, let's assume that the community has the budget to fund all five projects, so the problem is not about how to allocate scarce funds: they can fund as many (or as few) projects as they would like. Individuals in this community each have a preference about which projects should be funded. In this basic example, we will describe an individual's preference as just a yes/no vote on each issue. For example, an individual might have the preference, "yes" on building the new school and repairing the local road, but "no" on the other three projects.

In Figure 1, we depict the distribution of preferences in the community. This 19,000-person community can be thought of as having two groups, or parties. There is a pro-spending group, consisting of 9,000 voters, each of whom prefers to fund all five projects (depicted in light gray). The rest of the community consists of voters that are, generally speaking, anti-spending (depicted in black). This anti-spending group is more diverse. It consists of five different smaller groups of voters, each consisting of 2,000 members. Within each of these groups, the voters mostly prefer to not spend money on these projects - however, they each have one pet project they would like to see the community fund. For example, a voter within the anti-spending group might vote yes for building the new school, but no on the four other issues.

This community could make decisions via direct democracy. For instance, when deciding whether or not to build the new school, they hold a town hall meeting. Members of the community attend the meeting and cast a vote of yes or no for whether or not to build the school. The majority preference then dictates the community's decision. A separate vote of this type is done for each spending project, giving each community member the opportunity to vote for or against each project. Under this system, it is clear to see that the community chooses to fund each of the five projects. On any particular project, the 9,000 pro-spending voters and 2,000 of the anti-spending voters vote yes.

Alternatively, the community could use representative democracy to make decisions. In this case, rather than voting on one project at a time, the community instead elects a representative, such as a mayor or a governor, who then makes a decision about each project on the community's behalf. The representative serves as a binding plan for decisions on each

Figure 1: The population of preferences in the example community



of these projects. We use the likelihood of agreement rule to determine voters' preferences over candidates, assuming that when deciding between two candidates, an individual votes for the candidate with whom she agrees on the greatest number of spending projects. We can show that this community never elects a representative who votes yes on all five projects - that is, they never elect the representative that always agrees with the majority preference.

Suppose there is a pro-spending candidate, a candidate who would vote yes on all five projects. The anti-spending voters would be able to agree upon a candidate whom they all preferred to this pro-spending candidate. For instance, they could nominate an anti-spending candidate who would vote no on all five projects. In an election between the pro-spending candidate and the anti-spending candidate, the anti-spending candidate would earn the vote of all 10,000 anti-spending voters, since each of them agree with this candidate about four of the five projects, while they agree with the pro-spending candidate about only one of the five projects. Thus, the anti-spending candidate would earn a majority of the votes when competing with the pro-spending candidate. We can use this reasoning to conclude that in order for a representative to be elected by this community, she would have to vote no on at least some of the projects. Formally, we could show that the pro-spending candidate is not a member of the uncovered set of the tournament over candidates - that is, the candidate that agrees with the majority preferences of the population on all five projects cannot be elected.

We've used a political example as a way to discuss our result, but these types of collective decision problems are prevalent in a variety of other settings. Just as communities have to decide which projects to fund, hiring committees must decide which applicants to make offers to and companies must decide which products to develop. This paper makes the point that if a group commits *ex ante* to a plan of action rather than deciding on issues one at a time, its choices may not respect the majority preference on some issues.

In the last section of this paper, we discuss restrictions on the distribution of preferences that are sufficient to guarantee that representative democracy implements the same choices as direct democracy. The sufficiency condition we propose can be interpreted as an analog to the well-known single-peakedness result proposed by Black (1948) and Arrow (1951), as it describes a class of populations for which this form of majority rule (majority rule over candidates) functions well. The domain restriction rules out populations similar to the one depicted in Figure 1. Populations like those in Figure 1 are problematic because they contain two relatively large groups of like-minded individuals - like the pro-spending and anti-spending groups in our example. Our population restriction rules out these types of bimodal populations. It states that as we move away from the ordering consistent with the majority preferences over alternatives, we cannot encounter another large cluster of similar preferences.

The rest of the paper proceeds as follows. Section 2 introduces notation and our models of direct and representative democracy. In Section 3, we explore the structure of the tournaments over candidates and prove that McGarvey's Theorem does not hold in this context. That is, given an arbitrary binary relation, it may not be possible to find a population which will generate this relation as the outcome of a tournament over candidates. Section 4 considers the relationship between direct and representative democracy, proving consistency results for the case of  $n = 3$  but negative results for more general problems. Section 5 concludes.

## 2 Notation and Model

### 2.1 Notation

We use a finite and discrete space of **alternatives**,  $X = \{a_1, a_2, \dots, a_n\}$ . A **preference**, denoted  $\pi$ , is an ordering of the alternatives, where  $\pi$  corresponds to a permutation of the integers  $\{1, \dots, n\}$ ; given the preference  $\pi = (a_{\pi(1)}a_{\pi(2)}\dots a_{\pi(n)})$ ,  $a_i$  is preferred to  $a_j$  if and only if  $\pi^{-1}(i) < \pi^{-1}(j)$ . It will be useful to write  $e$  to represent the alphabetical ordering of the alternatives,  $e = (a_1a_2\dots a_n) \in \Pi$ . The set of all  $n!$  preferences over  $X$  is  $\Pi$ .

A **population**,  $\lambda$ , is a distribution over  $\Pi$ . Let  $\Lambda$  be the set of all distributions over  $\Pi$ .

We model a **candidate** as a strict ordering of the alternatives in  $X$ . We will write candidate  $\pi$  to denote the candidate with ordering  $\pi$ .

A **choice problem**,  $A \subseteq X$ , is a non-empty subset of alternatives; the set of all choice problems is  $\mathcal{X}$ . To each preference (and candidate),  $\pi$ , we can associate a **rational choice function**,  $c_\pi : \mathcal{X} \rightarrow X$ , where for each  $A$ ,  $c_\pi(A)$  is the element in  $A$  that is preferred to all other elements in  $A$  according to  $\pi$ . Our model assumes that an elected candidate,  $\pi$ , implements choices according to  $c_\pi$ .

We model these majoritarian systems using tournament theory. A tournament is a complete, asymmetric binary relation. Our analysis considers two types of tournaments: tournaments on the alternative space and tournaments on the candidate space. We use  $\Gamma(X)$  to denote a **tournament on the space of alternatives**; we reserve the traditional  $T$  to refer to a **tournament on the space of candidates**,  $T(\Pi)$ . In both cases, a tournament depends upon the preferences of the population; therefore, we write  $\Gamma^\lambda$  or  $T^\lambda$  to denote the tournaments generated by a population  $\lambda$ .

## 2.2 Models of Direct and Representative Democracy

Under direct democracy, members of the population vote over alternatives. We use the majority tournament to model these decisions. In this tournament, the relationship between any pair of alternatives is determined by majority rule:  $a_i \Gamma^\lambda a_j$  if  $\sum_{\pi \in \Pi} (\lambda(\pi | a_i \succ a_j)) > \frac{1}{2}$ . We write  $a_i \Gamma^\lambda a_j$  if  $a_i$  beats  $a_j$  in the tournament  $\Gamma^\lambda(X)$ . In the populations we consider below, majority rule over alternatives will be consistent with an ordering. We study this case for two reasons. Firstly, when the majority rule tournament is consistent with an ordering, the normative recommendation for representative democracy is clear: it should select precisely this ordering that is consistent with the direct democracy choices. Thus, this case serves as the most straightforward test of consistency between the two methods. Secondly, we hypothesize that this case has the highest likelihood of leading to positive results, as it minimizes potential sources of distortion. Our negative result for this case suggests we would derive similar negative results in environments where there is additional tension between matching direct democracy choices and producing an ordering.

Building a model of representative democracy requires a theory of how individuals choose to vote over candidates. That is, we must map individuals' preferences over alternatives into preferences over candidates. Our goal is to construct a model of representative democracy that is as similar as possible to our model of direct democracy, as this provides the strictest test of our negative result. Because our direct democracy model depends only on preferences over pairs of alternatives, we build a model of representative democracy that uses only this

same preference information.

We assume that an individual votes for the candidate with whom she is most likely to agree about the choice from a randomly-selected pair of alternatives. To present this theory formally, we use choice-based metrics. Baldiga and Green (2011) define choice-based metrics on the space of ordinal preferences as follows. First, we define a **probability distribution**,  $\nu$ , over  $\mathcal{X}$ . Then, the distance between a pair of orderings can be described as the probability that the orderings disagree about the optimal choice from a randomly-selected subset of alternatives.

$$f(\pi, \pi'; \nu) = \nu\{A \in \mathcal{X} | c_\pi(A) \neq c_{\pi'}(A)\}$$

Our model of representative democracy will use the choice-based metric associated with the distribution over choice problems where each choice problem is a pair, and all pairs are equally likely. This particular choice-based metric is equivalent to the Kemeny distance (also known as the bubble sort distance and the Kendall distance), which defines the distance between  $\pi$  and  $\pi'$  as the number of pairs of alternatives the two orderings rank differently (Kemeny 1959 and Kemeny Snell 1962). It is perhaps the most familiar metric on the space of ordinal preferences. Following the notation of Baldiga and Green, we denote this choice-based metric by  $f(\pi, \pi'; \mu^K)$ .

We use the Kemeny distance to define individuals' preferences over candidates: voters prefer candidates whose orderings are closer to their own. Given two candidates  $\pi$  and  $\pi'$ , a voter with preference  $\pi''$  will prefer  $\pi$  if and only if  $f(\pi'', \pi; \mu^K) < f(\pi'', \pi'; \mu^K)$ . Previous work in this area has also used the Kemeny distance to determine voters' preferences over orderings. For instance, in their investigation of the strategy-proofness of social welfare functions, Bossert and Storcken (1992) assume, as we do, that individuals' preferences over orderings are determined by relative proximity under the Kemeny distance. In Figure 2, we provide an illustration of the Kemeny distances between the six orderings in  $\Pi$  for the  $n = 3$  case.

Voters' preferences over candidates generate a tournament on  $\Pi$ , which we denote  $T^{\lambda, \mu^K}$ . We say that  $\pi'$  **attracts a majority against**  $\pi$  if there exists a subset of preferences  $\Pi_1$  with  $\lambda(\Pi_1) > \frac{1}{2}$  such that  $f(\pi'', \pi'; \mu^K) < f(\pi'', \pi; \mu^K)$  for all  $\pi'' \in \Pi_1$ . We write  $\pi' T^{\lambda, \mu^K} \pi$  if  $\pi'$  attracts a majority against  $\pi$ . In the case where  $f(\pi'', \pi'; \mu^K) = f(\pi'', \pi; \mu^K)$ , we assume that the votes of those with preference  $\pi''$  are split evenly between  $\pi$  and  $\pi'$ .<sup>1</sup> It

---

<sup>1</sup>This assumption means we employ *relative* majority rule. We could alternatively define our tournaments in terms of *absolute* majority rule. Under this assumption, we would have  $\pi' T^{\lambda, \mu^K} \pi$  only if a majority of voters are strictly closer to  $\pi'$  than  $\pi$ . For absolute majority rule, if neither candidate is closer to more than half of the population, the two candidates would tie in the tournament relation.

In our framework, voters will often be indifferent between candidates. Each voter has a set of indifference

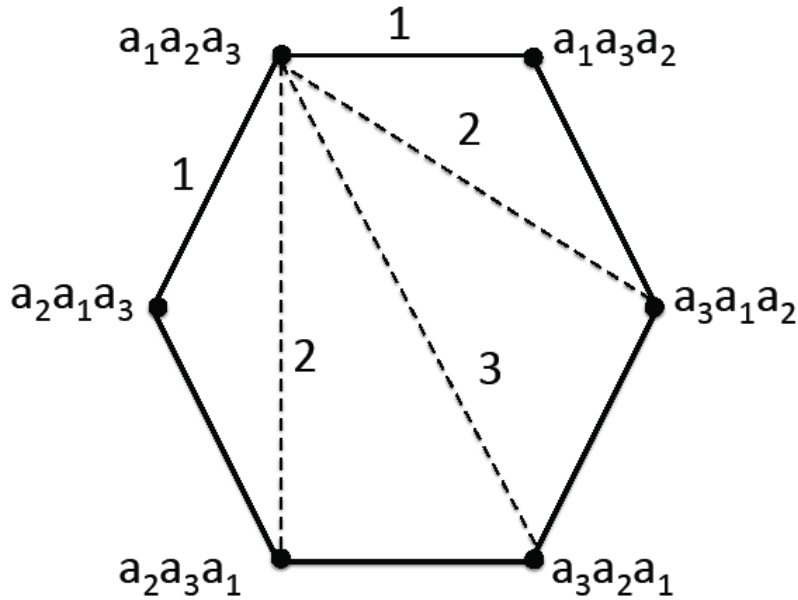


Figure 2: We use a hexagon to demonstrate the space of orderings for the  $n = 3$  case, with each vertex representing an ordering. Above, we illustrate the Kemeny distances from  $e = a_1a_2a_3$  to each of the other orderings in the space.

straightforward to check that this binary relation is complete and asymmetric.

One could imagine defining tournaments over orderings in other ways. However, the tournament over orderings induced by  $f(\pi, \pi'; \mu^K)$  parallels the tournament over alternatives induced by majority rule in a unique way. They both depend on the same preference information. The tournaments defined by  $f(\pi, \pi'; \mu^K)$  depend only on each voter's pairwise preferences over alternatives, just as the tournaments over alternatives defined by majority rule do. Tournaments over orderings induced by other choice-based metrics would take into account voters' preferences over subsets other than pairs, information that is absent from the majority rule tournament over alternatives. This way of defining tournaments over orderings serves as a “best case” analysis; it is the environment in which we would expect the most

---

curves: a voter most prefers the candidate with his own ordering, then he equally prefers all candidates with whom he disagrees about the choice from one pair of alternatives, and next he equally prefers all candidates with whom he disagrees about the choice from two pairs of candidates, etc.

Given the large amount of indifference in our population, choosing to use absolute majority rule would result in a large number of ties in our tournaments over orderings. These ties would disregard the information we have on the voters who are not indifferent. For instance, in a tie between candidate  $\pi$  and candidate  $\pi'$ , we may have that 30% of the voters are indifferent between candidate  $\pi$  and candidate  $\pi'$ , 49% of the voters prefer  $\pi$  to  $\pi'$ , and only 21% of voters prefer  $\pi'$  to  $\pi$ . Despite the large disparity in the number of voters that strictly prefer  $\pi$  to  $\pi'$  as opposed to  $\pi'$  to  $\pi$ , these two orderings would tie in the tournament relation. By using relative majority rule, we use this information on strict preference, even in the cases where large subsets of voters are indifferent between the two candidates.

consistency between direct and representative democracy. We prove negative results in this framework, illustrating that even in the case when these forms of direct and representative democracy have the same informational basis, they may not lead to the same outcomes.<sup>2</sup>

## 2.3 Tournament Solutions

We consider only the simplest direct democracy tournaments, restricting our attention to those tournaments generated by populations in which majority rule over alternatives is consistent with an ordering. In these cases, there is a clear winner of the tournament: the alternative which defeats all others in the tournament relation, the Condorcet winner. Furthermore, because there are no cycles, we can describe the tournament by an ordering where each alternative is majority-preferred to all those below it.

However, for general tournaments, including the tournaments over orderings we study below, identifying the winners is non-trivial. In the simplest tournaments, there may be a Condorcet winner, but usually, the tournament relation will cycle. Therefore, we need a tournament solution, a mapping  $S : T \rightarrow 2^\Pi \setminus \emptyset$  that will determine the best elements given an arbitrary tournament structure. We use the uncovered set as our tournament solution. Miller (1980) provides the first characterization of the uncovered set. Here, we follow the definition given by Laslier (1997), applied to our tournaments over orderings. First, define the covering relation of  $T$ . For a given  $T$ , we say  $\pi_i$  **covers**  $\pi_j$  if and only if:

- (a)  $\pi_i T \pi_j$ , and
- (b)  $\forall \pi_k \in \Pi, \pi_j T \pi_k \Rightarrow \pi_i T \pi_k$

The **uncovered set of  $T$**  is the set of maximal elements of the covering relation:  $\pi_i \in UC(T)$  iff  $\nexists \pi_j \in \Pi$  such that  $\pi_j$  covers  $\pi_i$ .

As Miller (1980) first described, the uncovered set has a number of appealing properties. The uncovered set is always a non-empty subset of the top cycle. And, unlike the top cycle, it contains only Pareto undominated orderings. It is Condorcet consistent: if a Condorcet winner exists, it will be the sole member of the uncovered set. Miller also noted the following game-theoretic interpretation of the uncovered set. Consider the following two-player zero sum game. The feasible strategies are the orderings,  $\pi_1, \dots, \pi_n \in \Pi$ . If player  $i$  and player  $j$  play  $(\pi_i, \pi_j)$ , then player  $i$  receives a payoff of 1 if  $\pi_i T \pi_j$  and a payoff of  $-1$  otherwise. Then the set of undominated strategies in this game is exactly the uncovered set of the tournament  $T(\Pi)$ .

---

<sup>2</sup>In the Appendix, we explore a model of representative democracy based upon expected utility maximization. The inclusion of cardinal information is a sharp departure from the ordinal nature of the direct democracy tournament; thus, as we explain in the Appendix, we expect even less consistency between direct and representative democracy in this expected utility framework.

The uncovered set also characterizes the outcomes under a variety of familiar voting rules. Miller (1980) and Shepsle and Weingast (1984) have shown that a number of voting procedures under both sincere and sophisticated voting implement elements of the uncovered set. We discuss three broad classes of those procedures here.

The first is the amendment procedure under sophisticated voting. The amendment procedure begins with a majority vote over a pair of alternatives. The loser of the pairwise contest is eliminated from contention; the winner of that pairwise contest advances to face another alternative in the next stage. This process continues through  $n - 1$  stages, with the surviving alternative selected as the winner. Miller proves that the amendment procedure under sophisticated voting must select a member of the uncovered set as a winner. Further, Shepsle and Weingast show that an alternative  $a_i \in X$  can be the sophisticated voting outcome of an amendment procedure agenda containing  $a_j$  in the first stage if and only if  $a_j$  does not cover  $a_i$ . In this way, the uncovered set characterizes the set of implementable outcomes under the amendment procedure with sophisticated voting.

Cooperative models of voting also lead to the selection of elements of the uncovered set. Miller shows that when voters may form coalitions and play cooperatively, communicating and making binding agreements to vote together, then any majoritarian procedure will implement an element of the uncovered set.

Finally, Miller discusses a model of strategic two-party electoral competition that chooses elements of the uncovered set. Suppose there are two political parties, A and B. Each party must nominate a candidate from the set of possible candidates  $\pi \in \Pi$ . Voters are assumed to have rational preferences over the candidates; they have a dominant strategy to vote sincerely for the party whose candidate they prefer. Under these assumptions, this model is a symmetric, 2-person, zero-sum game for the parties A and B. Miller shows that the set of undominated strategies is exactly the uncovered set of the tournament on  $\Pi$ . Thus, this theoretical construct, the uncovered set of the tournament, provides a nice characterization of likely outcomes under one stylized model of two-party political systems.<sup>3</sup>

These results motivate our use of the uncovered set. By working with this tournament solution, we avoid making specific institutional assumptions. Instead, we identify the likely winners under majoritarian voting rules more generally. The negative result we prove in this framework can be applied to any voting procedure which leads to the election of a member of the uncovered set.

---

<sup>3</sup>McKelvey (1986) proves similar results in the context of multidimensional choice spaces and quasiconcave preferences.

## 2.4 Our Approach

We are now ready to pose our question more formally. We consider a population of voters with preferences over a finite set of alternatives. We look at the choices this population would make under direct democracy by computing the tournament over alternatives induced by majority rule. Then, we determine which candidates this population would elect under representative democracy by computing the tournament over orderings induced by the Kemeny distance. We ask whether the ordering consistent with the majority preferences over alternatives is contained in the uncovered set of the tournament over orderings. This approach is summarized in Figure 3.

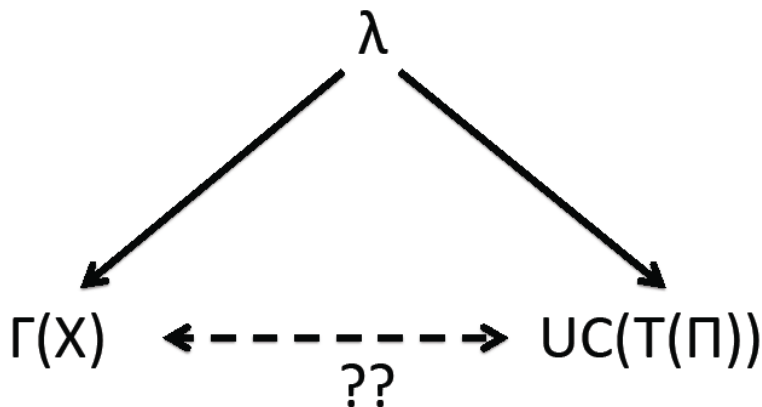


Figure 3: Our Approach

## 3 The Structure of Tournaments over Orderings

We begin by considering the structural properties of the new type of tournament we analyze: the tournament over the orderings. One of the most influential results in social choice theory is **McGarvey's Theorem**. It states that given an arbitrary tournament over a set of  $n$  alternatives  $X$ , a population of strict preferences exists which will generate this tournament as the outcome of majority rule (McGarvey 1953). Thus, when it comes to majority rule over alternatives, anything goes; the structure of the tournament space over alternatives is unconstrained.

This result does not hold for tournaments over orderings defined by the Kemeny distance. Consider an arbitrary binary relation over the orderings. We prove that it is not generally possible to find a population  $\lambda$  that will generate this binary relation as the outcome of

the tournament over orderings where voters' preferences are determined by proximity under  $f(\pi, \pi'; \mu^K)$ . That is, the mapping  $\lambda \rightarrow T^{\lambda, \mu^K}$  is not onto. This result is interesting from a mathematical standpoint and it is also computationally helpful, as restrictions on the space of tournaments over orderings reduce the number of cases to consider in the analysis that follows.

The additional structure on the space of tournaments over orderings stems from the tournament's reliance on the underlying majority relation over alternatives. We illustrate this dependence with a simple example for the  $n = 3$  case. Recall that the distance between  $\pi$  and  $\pi'$  as measured by the metric  $f(\pi, \pi'; \mu^K)$  is proportional to the number of pairwise disagreements between the two orderings. Therefore, when determining whether a voter with ordering  $\pi''$  will vote for candidate  $\pi$  or  $\pi'$ , it is enough to count the number of pairs on which  $\pi''$  disagrees with each of the two orderings. The candidate with more pairwise choices in common with  $\pi''$  will win the support of voter  $\pi''$ . As a result, we can use majority rule over pairs of alternatives to pin down the relationship between certain pairs of orderings in the tournament over orderings.

Consider a population whose majority preference over the pair  $(a_1, a_2)$  is  $a_1 \Gamma^\lambda a_2$ . This determines the relationship between pairs of candidates whose orderings disagree only about the relative ranking of  $a_1$  and  $a_2$ . For example, consider the pair of candidates  $a_1 a_2 a_3$  and  $a_2 a_1 a_3$ . Since we have  $a_1 \succ a_2$  for a majority of the population, we must have that a majority of the population is closer to  $a_1 a_2 a_3$  than  $a_2 a_1 a_3$  according to  $f(\pi, \pi'; \mu^K)$ . Similarly, we can deduce that a majority of the population is closer to  $a_3 a_1 a_2$  than  $a_3 a_2 a_1$ . If the majority preference on  $(a_1, a_2)$  were reversed, the relationship between both of these pairs of orderings would also be reversed. Therefore, we have:

$$\begin{aligned} a_1 \Gamma^\lambda a_2 &\Leftrightarrow a_1 a_2 a_3 T^{\lambda, \mu^K} a_2 a_1 a_3 \\ a_1 \Gamma^\lambda a_2 &\Leftrightarrow a_3 a_1 a_2 T^{\lambda, \mu^K} a_3 a_2 a_1 \\ a_1 a_2 a_3 T^{\lambda, \mu^K} a_2 a_1 a_3 &\Leftrightarrow a_3 a_1 a_2 T^{\lambda, \mu^K} a_3 a_2 a_1 \end{aligned}$$

This makes it clear that there is no population that would generate a tournament such that  $a_1 a_2 a_3 T^{\lambda, \mu^K} a_2 a_1 a_3$  but  $a_3 a_2 a_1 T^{\lambda, \mu^K} a_3 a_1 a_2$ .

Because of the close tie between a population's majority preferences over pairs of alternatives and the tournament relation  $T^{\lambda, \mu^K}$ , it will be helpful to classify all  $n = 3$  populations according to their majority preferences. The first class of populations has a transitive majority preference of the type  $a_1 \Gamma^\lambda a_2$ ,  $a_2 \Gamma^\lambda a_3$ , and  $a_1 \Gamma^\lambda a_3$ . There are six possible majority relations in this class, one corresponding to each of the six possible orderings. We can show

that for any one of these transitive majority relations, there are only two possible tournaments over orderings  $T^{\lambda, \mu^K}$  that can be generated. The second class of populations has a cyclic majority preference; there are only two possible cyclic majority relations:  $a_1 \Gamma^\lambda a_2$ ,  $a_2 \Gamma^\lambda a_3$ ,  $a_3 \Gamma^\lambda a_1$  and  $a_2 \Gamma^\lambda a_1$ ,  $a_1 \Gamma^\lambda a_3$ ,  $a_3 \Gamma^\lambda a_2$ . In the Appendix, we show that for either one of these cyclic majority preferences, there are only six possible tournaments over orderings  $T^{\lambda, \mu^K}$  that can be generated. Thus, we can conclude that in the  $n = 3$  case only 24 of the 32,768 ( $2^{15}$ ) possible asymmetric binary relations can be generated by a tournament over orderings induced by a population  $\lambda$  and the metric  $f(\pi, \pi'; \mu^K)$ .

Before we move to our main results, we discuss a few useful observations about the general structure of  $T^{\lambda, \mu^K}$ . These results again use the majority preferences over alternatives to draw conclusions about the pairwise relationships between certain pairs of orderings.

**Proposition 1** *Consider a population where majority preferences are consistent with  $e$ . Take any ordering  $\pi$  that has at least one pair of adjacent alternatives ordered according to  $e$ . Obtain  $\pi'$  by performing one transposition of adjacent alternatives that appeared in the natural order in  $\pi$ . Then, we have  $\pi T^{\lambda, \mu^K} \pi'$  for all such  $\pi'$ . Furthermore, we can take any of these  $\pi'$  that have at least one pair of adjacent alternatives ordered according to  $e$  and obtain  $\pi''$  by performing one transposition of alternatives that appeared in the natural order in  $\pi$  and  $\pi'$ . Then, we have  $\pi T^{\lambda, \mu^K} \pi''$ .*

*Proof:* Intuitively, when  $\pi$  and  $\pi'$  agree on all but a single pair of alternatives, a voter with the ordering  $\pi''$  will be closest to whichever of these orderings it agrees with on the pair in the question. Since  $\pi$  agrees with the majority preference on the pair in question, more than half of the population must be closer to  $\pi$ , yielding  $\pi T^{\lambda, \mu^K} \pi'$ .

We can take this logic one step further to prove the claim for cases where  $\pi$  and  $\pi''$  agree on all but two pairs of alternatives. Denote by  $\lambda_\alpha$  the subset of the population that agrees with  $\pi$  on the first pair in question (the one transposed to obtain  $\pi'$ ). Denote by  $\lambda_\beta$  the subset of the population that agrees with  $\pi$  on the second pair in question (the one transposed to obtain  $\pi''$ ). Because  $\pi$  is consistent with the majority preference on both of these pairs, we know that  $\lambda_\alpha > \frac{1}{2}$  and  $\lambda_\beta > \frac{1}{2}$ . Therefore:

$$2 \times (\lambda_\alpha \cap \lambda_\beta) + \lambda_\alpha \setminus \lambda_\beta + \lambda_\beta \setminus \lambda_\alpha > 1$$

We can rearrange this expression to show that:

$$\lambda_\alpha \cap \lambda_\beta > 1 - (\lambda_\alpha \setminus \lambda_\beta + \lambda_\beta \setminus \lambda_\alpha + \lambda_\alpha \cap \lambda_\beta)$$

The left-hand side of this equation is the fraction of the population that agrees with  $\pi$  on both of the pairs in question and will be closer to  $\pi$ . The right-hand side of this equation is

the fraction of the population that disagrees with  $\pi$  on both of the pairs in question and will be closer to  $\pi''$ . The rest of the population will be equidistant. Thus, the equation tells us that a larger fraction of the population will be closer to  $\pi$  than  $\pi''$ , yielding  $\pi T^{\lambda, \mu^K} \pi''$ . ■

The following corollary is a straightforward implication of Proposition 1 and will prove useful in the following sections.

**Corollary 1** *An ordering that is consistent with the majority preferences of a population  $\lambda$ , call this  $\pi^*$ , must beat all orderings that are fewer than two transpositions away from it; that is, we must have  $\pi^* T^{\lambda, \mu^K} \pi$  for all  $\pi$  such that  $f(\pi^*, \pi; \mu^K) \leq 2$ .*

## 4 Consistency Results for Direct and Representative Democracy

### 4.1 Defining Consistency

With this knowledge of the structure of tournaments over orderings, we turn now to computing solutions. For our populations, in which the majority preferences over alternatives are consistent with an ordering, there is a clear test of whether representative democracy implements the choices made under direct democracy. We ask whether the ordering consistent with the majority preferences over alternatives is a member of the uncovered set of the tournament over orderings.

**Definition 1** *Strong Order Consistency: If majority preferences over alternatives are consistent with an ordering, then this ordering is the sole member of the uncovered set of  $T^{\lambda, \mu^K}$ .*

Or, we might require a weaker condition, which simply requires inclusion of this ordering in the uncovered set of the tournament over alternatives:

**Definition 2** *Order Consistency: If majority preferences over alternatives are consistent with an ordering, then this ordering is a member of the uncovered set of  $T^{\lambda, \mu^K}$ .*

In the analysis that follows, we show that strong order consistency holds for the case of  $n = 3$  but fails for larger problems with  $n > 3$ . Then, we provide a counterexample which illustrates that for  $n \geq 10$ , even order consistency does not hold.<sup>4</sup>

---

<sup>4</sup>In the Appendix, we address the case in which majority preferences over alternatives cycle. There, we discuss notions of consistency for these cyclic populations and prove that for the case of  $n = 3$ , cyclic populations satisfy these reasonable notions of consistency.

## 4.2 Consistency for $n=3$

We exploit our knowledge of the structure of tournaments over orderings to prove results for the  $n = 3$  case for populations with transitive majority preferences consistent with an ordering. Here, we work through the case where  $a_1 \Gamma^\lambda a_2$ ,  $a_2 \Gamma^\lambda a_3$ , and  $a_1 \Gamma^\lambda a_3$ . We prove that a population with this underlying majority relation can generate only two possible tournaments. We show that simply knowing the majority preference over the three pairs of alternatives is enough to determine 13 of the 15 pairwise relationships between candidates and the uncovered set of  $T^{\lambda, \mu^K}$ ; the remaining two relationships, which do not impact the uncovered set, are determined jointly by the relative sizes of the majorities.

Our first step is to map majority preferences over pairs of alternatives onto  $T^{\lambda, \mu^K}$ . The majority preference over each of the three pairs  $(a_1, a_2)$ ,  $(a_1, a_3)$ , and  $(a_2, a_3)$  determines the relationship between three pairs of orderings. To see this, we work out the full set of implications of  $a_1 \Gamma^\lambda a_2$ .

Let  $a_1 \Gamma^\lambda a_2$ . Then, we know:

$$\sum_{\pi_i \in \Pi} (\lambda(\pi_i) | a_1 \succ_{\pi_i} a_2) > \sum_{\pi_i \in \Pi} (\lambda(\pi_i) | a_2 \succ_{\pi_i} a_1)$$

$$\lambda(a_1 a_2 a_3) + \lambda(a_1 a_3 a_2) + \lambda(a_3 a_1 a_2) > \lambda(a_2 a_1 a_3) + \lambda(a_2 a_3 a_1) + \lambda(a_3 a_2 a_1)$$

This inequality implies a set of  $T^{\lambda, \mu^K}$  relations. As made clear by the middle diagram in Figure 4, the orderings on the left-hand side of the inequality are closer to  $a_1 a_2 a_3$  than  $a_2 a_1 a_3$ , closer to  $a_1 a_3 a_2$  than  $a_2 a_3 a_1$ , and closer to  $a_3 a_1 a_2$  than  $a_3 a_2 a_1$ . Therefore:

$$a_1 a_2 a_3 T^{\lambda, \mu^K} a_2 a_1 a_3$$

$$a_1 a_2 a_3 T^{\lambda, \mu^K} a_2 a_1 a_3$$

$$a_3 a_1 a_2 T^{\lambda, \mu^K} a_3 a_2 a_1$$

We can do this for each of the three pairs of alternatives, pinning down nine of the relationships in  $T^{\lambda, \mu^K}$ . Figure 4 illustrates this process.

Next, we can map pairs of majority preferences into  $T^{\lambda, \mu^K}$ . For instance, since we have  $a_1 \Gamma^\lambda a_2$  and  $a_1 \Gamma^\lambda a_3$ , we know:

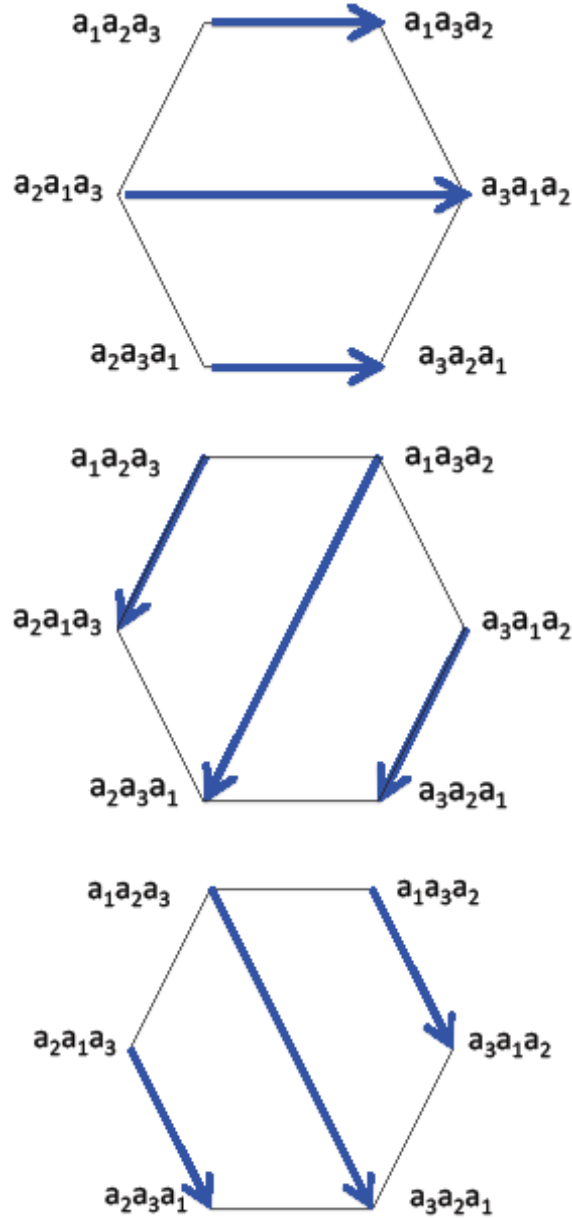


Figure 4: We use a hexagon with an ordering on each vertex to represent the  $n = 3$  space of orderings. Throughout this paper, we'll use an arrow to illustrate a  $T$  relationship, with the arrow directed toward the defeated ordering. In the top diagram, the arrows illustrate the three  $T^{\lambda, \mu^K}$  relationships determined by  $a_2\Gamma^\lambda a_3$ . In the middle diagram, the arrows illustrate the three  $T^{\lambda, \mu^K}$  relationships determined by  $a_1\Gamma^\lambda a_2$ . In the bottom diagram, the arrows illustrate the three  $T^{\lambda, \mu^K}$  relationships determined by  $a_1\Gamma^\lambda a_3$ .

$$\begin{aligned}
& \sum_{\pi_i \in \Pi} (\lambda(\pi_i) | a_1 \succ_{\pi_i} a_2) + \sum_{\pi_i \in \Pi} (\lambda(\pi_i) | a_1 \succ_{\pi_i} a_3) > \\
& \sum_{\pi_i \in \Pi} (\lambda(\pi_i) | a_2 \succ_{\pi_i} a_1) + \sum_{\pi_i \in \Pi} (\lambda(\pi_i) | a_3 \succ_{\pi_i} a_1) \\
2\lambda(a_1 a_2 a_3) + 2\lambda(a_1 a_3 a_2) + \lambda(a_3 a_1 a_2) + \lambda(a_2 a_1 a_3) & > \lambda(a_2 a_1 a_3) + 2\lambda(a_2 a_3 a_1) + 2\lambda(a_3 a_2 a_1) + \lambda(a_3 a_1 a_2) \\
2\lambda(a_1 a_2 a_3) + 2\lambda(a_1 a_3 a_2) & > 2\lambda(a_2 a_3 a_1) + 2\lambda(a_3 a_2 a_1) \\
\lambda(a_1 a_2 a_3) + \lambda(a_1 a_3 a_2) & > \lambda(a_2 a_3 a_1) + \lambda(a_3 a_2 a_1)
\end{aligned}$$

The orderings on the left-hand side of the inequality are closer to  $a_1 a_2 a_3$  than  $a_2 a_3 a_1$  and closer to  $a_1 a_3 a_2$  than  $a_3 a_2 a_1$ . Since the other orderings,  $a_2 a_1 a_3$  and  $a_3 a_1 a_2$  are equidistant from each pair of orderings, this is enough to determine the relationship between each of these pairs.

$$\begin{aligned}
& a_1 a_2 a_3 T^{\lambda, \mu^K} a_2 a_3 a_1 \\
& a_1 a_3 a_2 T^{\lambda, \mu^K} a_3 a_2 a_1
\end{aligned}$$

We can draw a similar set of conclusions working from the majority preference of  $a_1 \Gamma^\lambda a_3$  and  $a_2 \Gamma^\lambda a_3$ , deducing:<sup>5</sup>

$$\begin{aligned}
& a_2 a_3 a_1 T^{\lambda, \mu^K} a_3 a_2 a_1 \\
& a_1 a_2 a_3 T^{\lambda, \mu^K} a_3 a_1 a_2
\end{aligned}$$

Figure 5 summarizes these relationships. Regardless of how the final two links are resolved, we see that  $a_1 a_2 a_3$  is the Condorcet winner of the tournament of orderings, beating every other ordering directly. Therefore, we have  $UC(T^{\lambda, \mu^K}) = \{a_1 a_2 a_3\}$ .

Relating this result to our definitions of consistency above, we can state the following proposition:

**Proposition 2** *For  $n = 3$ ,  $UC(T^{\lambda, \mu^K})$  satisfies strong order consistency.*

---

<sup>5</sup>Note that together the majority preferences over  $a_1 \Gamma^\lambda a_2$  and  $a_2 \Gamma^\lambda a_3$  only tells us that  $\lambda(a_1 a_2 a_3) > \lambda(a_3 a_2 a_1)$ , which alone is not enough to determine the  $T^{\lambda, \mu^K}$  relationship between any pair of orderings.

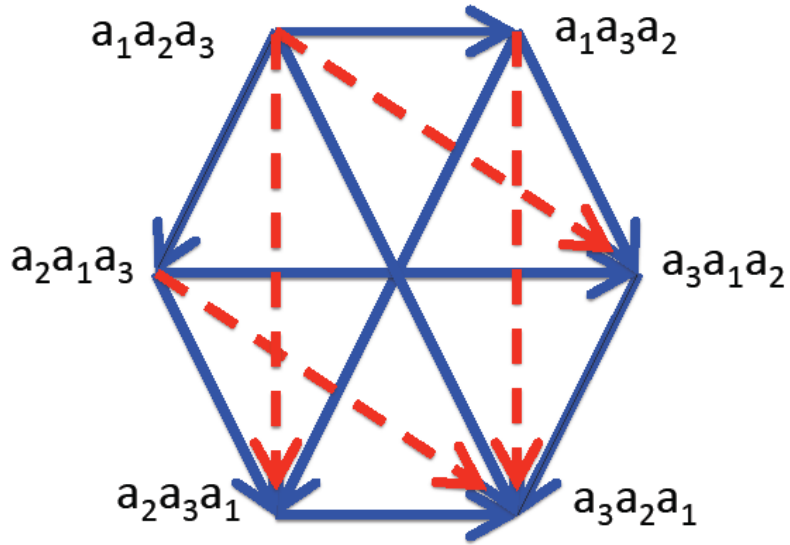


Figure 5: The solid arrows are the relationships determined by looking at the majority preferences over each single pair of alternatives. The four dashed arrows illustrate the additional  $T^{\lambda, \mu^K}$  relationships that are determined by pairs of majority preferences over alternatives.

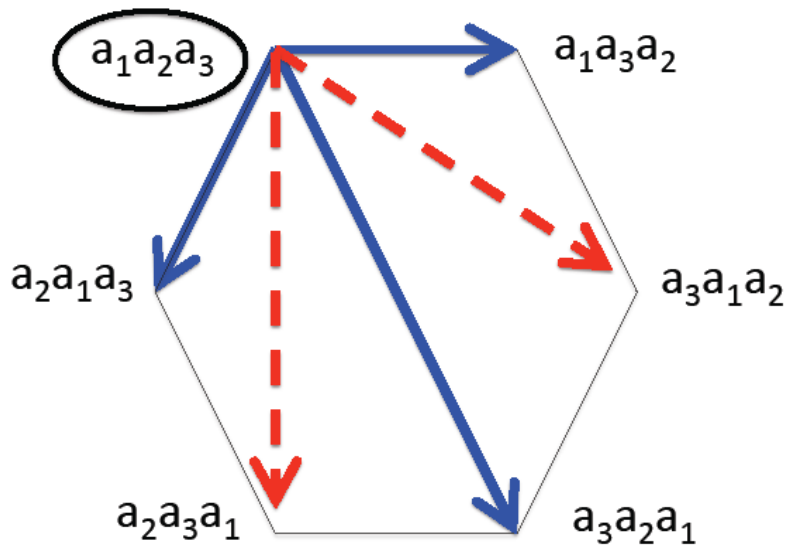


Figure 6: The diagram above illustrates the uncovered set of the transitive population for  $n = 3$ . Because  $a_1a_2a_3$  beats all other orderings, it is the Condorcet winner of the tournament and the sole member of the uncovered set.

### 4.3 Inconsistency for General Problems

First, we present an example that shows that for  $n = 4$ , strong order consistency fails. This proves that for  $n > 3$  there exist populations whose majority preferences are consistent with an ordering but whose uncovered sets of  $T^{\lambda, \mu^K}$  are multi-valued. Consider the following population:

$\pi$	$\lambda(\pi)$
$a_1a_2a_3a_4$	.399
$a_2a_4a_1a_3$	.2
$a_1a_4a_3a_2$	.2
$a_3a_4a_1a_2$	.201

It is clear that the majority preferences of this population are consistent with  $e = a_1a_2a_3a_4$ . But, there is no Condorcet winner of the tournament over candidates. We can show that there exists an ordering, denoted  $\hat{\pi}$  such that  $\hat{\pi}T^{\lambda, \mu^K}e$ . Consider  $\hat{\pi} = a_4a_1a_3a_2$ . We have  $f(\hat{\pi}, a_1a_4a_3a_2) < f(e, a_1a_4a_3a_2)$  and  $f(\hat{\pi}, a_3a_4a_1a_2) < f(e, a_3a_4a_1a_2)$ , and we have  $f(\hat{\pi}, a_2a_4a_1a_3) = f(e, a_2a_4a_1a_3)$ . So,

$$\sum_{\pi \in \Pi} [\lambda(\pi) | f(\hat{\pi}, \pi) < f(e, \pi)] > \sum_{\pi \in \Pi} [\lambda(\pi) | f(e, \pi) < f(\hat{\pi}, \pi)]$$

As a result,  $\hat{\pi}T^{\lambda, \mu^K}e$ . And, in fact, we can show that  $\hat{\pi} \in UC(T^{\lambda, \mu^K})$ , with  $UC(T^{\lambda, \mu^K}) = \{e, a_1a_2a_4a_3, a_1a_4a_2a_3, \hat{\pi}\}$ . We can extend this  $n = 4$  example to a problem with an arbitrary number of alternatives by preserving the structure above for the first four alternatives and simply appending additional alternatives in their natural order to the right end of each of the four orderings above. This leads to the following proposition.

**Proposition 3** *For  $n \geq 4$ ,  $UC(T^{\lambda, \mu^K})$  fails strong order consistency.*

This is the strongest result that we can achieve for  $n = 4$ , as it is straightforward to demonstrate that for populations whose majority preferences are consistent with an ordering, this ordering must be a member of its uncovered set. Assume majority preferences are consistent with  $e$ . We can show that there is no ordering that can cover  $e$ . The key step is to recognize that we can apply Proposition 1 to prove that any ordering fewer than five transpositions from  $e$  cannot cover  $e$ :  $eT^{\lambda, \mu^K}\pi$  for any  $\pi$  within two transpositions, so they cannot cover  $e$ , and for those three or four transpositions away, even if they beat  $e$ , they will be defeated by at least one ordering one or two transpositions from  $e$  (which  $e$  beats). So, the only orderings that could potentially cover  $e$  are five or six transpositions away from  $e$ :

$\{a_4a_3a_1a_2, a_4a_2a_3a_1, a_3a_4a_2a_1, a_4a_3a_2a_1\}$ . In the Appendix, we rule these out one at a time, proving that order consistency must hold for  $n = 4$  populations, as  $e \in UC(T^{\lambda, \mu^K})$  for all populations whose majority preferences are consistent with  $e$ .

For  $n \geq 10$ , we are able to prove a stronger result. We identify populations in which the ordering consistent with the majority preferences is covered. For problems with fewer alternatives, like the one above, we can find a  $\hat{\pi}$  such that  $\hat{\pi}T^{\lambda, \mu^K}e$ , but this  $\hat{\pi}$  does not beat everything that  $e$  beats. That is, we can apply Proposition 1 to find a chain of the type  $eT^{\lambda, \mu^K}\pi T^{\lambda, \mu^K}\hat{\pi}$  where  $\pi$  is just two transpositions away from both  $e$  and  $\hat{\pi}$ . In order to prevent these types of chains,  $\hat{\pi}$  must be at least five transpositions from  $e$ . We can do this most simply for problems with at least 10 alternatives. However, as we discuss after the proof, it may be possible to find other types of populations that violate order consistency when there are fewer than 10 alternatives.

**Theorem 1** *For  $n \geq 10$ ,  $UC(T^{\lambda, \mu^K})$  fails order consistency.*

Proof: We prove this through a general counterexample. First, we construct the population. The majority preferences of our population  $\lambda$  will be consistent with  $e$  on all pairs of  $n$  alternatives. Select  $j$  pairs of alternatives, where  $5 \leq j \leq \frac{n}{2}$ . Each pair should consist of two adjacent elements in the natural ordering, and all pairs should be mutually exclusive. For example, it would be permissible to select  $\{a_1, a_2\}$  and  $\{a_3, a_4\}$  as two of the pairs, but one could not select  $\{a_1, a_2\}$  and  $\{a_2, a_3\}$ , or  $\{a_1, a_3\}$  and  $\{a_4, a_6\}$ . It will be helpful to have notation for the  $j$  pairs; let them be denoted  $p_1, p_2, \dots, p_j$ . Note that since the pairs consist of adjacent and mutually exclusive alternatives, it is always possible to find an ordering  $\pi$  that agrees with the majority preference on any particular subset of the pairs  $\{p_1, p_2, \dots, p_j\}$  exactly.

We will associate with each of the  $j$  pairs a particular ordering,  $\pi_{p_i}$ , where  $\pi_{p_i}$  agrees with the majority preference on pair  $p_i$ , disagrees with the majority preferences on the other  $j - 1$  pairs in the set  $\{p_1, p_2, \dots, p_j\}$ , and agrees with the majority preference on all other pairs of alternatives.

Allocate the population as follows. Let  $\lambda(e) = \frac{1}{2} - \varepsilon$ , where  $\varepsilon < \frac{1}{2j}$ . Divide the rest of the population evenly among the orderings  $\{\pi_{p_1}, \pi_{p_2}, \dots, \pi_{p_j}\}$ , creating  $j$  equal masses of size  $\frac{\frac{1}{2} + \varepsilon}{j}$ .

It is straightforward to check that this population produces majority rule over alternatives that agrees with  $e$ . For each pair not included in  $\{p_1, p_2, \dots, p_j\}$ , the population unanimously prefers  $a_i$  to  $a_{i+k}$ , where  $i \in \{1, \dots, n - 1\}$  and  $k \in \{1, \dots, n - i\}$ . For each of the pairs in  $\{p_1, p_2, \dots, p_j\}$ , we have that  $(\frac{1}{2} - \varepsilon) + (\frac{\frac{1}{2} + \varepsilon}{j}) > \frac{1}{2}$  agrees with  $e$ .

Now we will show that for this population  $e \notin UC(T^{\lambda, \mu^K})$ . We do so by identifying an ordering which covers  $e$ . Consider the ordering which disagrees with the majority preferences on all of the pairs  $\{p_1, p_2, \dots, p_j\}$  and agrees with the majority preference on all other pairs. Denote this ordering  $\hat{\pi}$ . Since we are working with  $f(\pi, \pi'; \mu^K)$ , the distance between any two orderings is, up to a scale factor, the number of pairs over which the two orderings disagree. For simplicity, we'll scale our distances below to the number of pairwise disagreements.

First we'll show that  $\hat{\pi}Te$ . We have  $f(\pi_{p_i}, \hat{\pi}; \mu^K) = 1$  and  $f(\pi_{p_i}, e; \mu^K) = j - 1 \forall i \in \{1, \dots, j\}$ . Therefore, we have  $\frac{1}{2} + \varepsilon$  of the population that is closer to  $\hat{\pi}$  than  $e$ , so  $\hat{\pi}Te$ .

In order to prove that  $\hat{\pi}$  covers  $e$ , we must show that there cannot exist a  $\pi'$  such that  $eT\pi'$  but  $\pi'T\hat{\pi}$ . Suppose there did exist such a  $\pi'$ .

We have that  $eT\pi'$ . This implies that we have  $f(e, \pi_{p_i}; \mu^K) \leq f(\pi', \pi_{p_i}; \mu^K)$  for at least some  $i \in \{1, \dots, j\}$ . Because we know  $f(e, \pi_{p_i}; \mu^K) = j - 1$ , this implies that  $f(\pi', \pi_{p_i}; \mu^K) \geq j - 1$  for at least some  $i \in \{1, \dots, j\}$ .

We also know that  $\pi'T\hat{\pi}$ . Then we must have that  $f(\pi', \pi_{p_k}; \mu^K) \leq f(\hat{\pi}, \pi_{p_k}; \mu^K)$  for at least some  $k \in \{1, \dots, j\}$ . We know that  $f(\hat{\pi}, \pi_{p_k}; \mu^K) = 1$  for all  $k \in \{1, \dots, j\}$ , which implies that  $f(\pi', \pi_{p_k}) \leq 1$  for at least some  $k \in \{1, \dots, j\}$ .

Finally, we know that  $f(\pi_{p_i}, \pi_{p_k}; \mu^K) \leq 2$  for any  $i, k \in \{1, \dots, j\}$ .

This creates the following violation of the triangle inequality:  $f(\pi', \pi_{p_k}; \mu^K) \leq 1$ ,  $f(\pi_{p_i}, \pi_{p_k}; \mu^K) \leq 2$ , and  $f(\pi', \pi_{p_i}; \mu^K) \geq j - 1$ , where  $j \geq 5$ . This is a contradiction.

Therefore, there can exist no  $\pi'$  such that  $eT\pi'$  but  $\pi'T\hat{\pi}$ . As a result, we can conclude that  $\hat{\pi}$  covers  $e$ . ■

This proof describes a method for constructing populations for which direct and representative democracy yield different choices. It shows that when there are 10 or more alternatives, it is always possible to construct a population such that the majority preferences are consistent with an ordering but the uncovered set of  $T^{\lambda, \mu^K}$  does not contain this ordering. These populations have a rather intuitive interpretation. Let's think about a population constructed by the method above for the case of  $n = 10$ . First, we note the distinction between the five "contested" issues ( $\{a_1, a_2\}, \{a_3, a_4\}, \{a_5, a_6\}, \{a_7, a_8\}, \{a_9, a_{10}\}$ ) and the other 40 pairwise choices which are decided unanimously. The largest mass of voters, just under half of them, have the preference  $e$ . Let's call these our "mainstream" voters. The remaining voters are divided evenly among five smaller minority preferences. Each minority agrees with the mainstream preference on just one of the contested issues; on the other hand, each minority group agrees with every other minority group on three of the five contested issues. In this way, the minority preferences are all more similar to one another than to the mainstream voters.

Let's think about the choices this population would make under direct and representative democracy. When voting directly over the alternatives, the population implements choices consistent with  $e$ . Most of the choices are unanimous; and, for the five contested pairs, the mainstream voters and one of the minority groups form a majority. Though the minority groups have similar preferences, when voting issue-by-issue, they never vote all together on a contested issue. As a result, the mainstream voters are able to implement their preferred choices. We can contrast this with the dynamic under representative democracy. In this setting, candidate  $e$  cannot be elected. Though  $e$  attracts the mainstream voters, there exist candidates which all five minority groups prefer to  $e$ . Consider candidate  $\hat{\pi}$ , which agrees with the unanimous choices of the population but disagrees with the majority preference on the five contested pairs. All five minority preferences prefer  $\hat{\pi}$  to  $e$ ; together, they consist of a majority of the population and can succeed in electing this compromise candidate. The choices of the selected candidates in representative democracy are more closely aligned with the minority preferences than the choices under direct democracy. The ability of the minority groups to compromise under representative democracy produces choices that are different than those implemented under direct democracy but that are preferred by more than half the population.

While our model is quite different, our results echo the findings of Besley and Coate (2008), who study the question of whether citizens' initiatives, which allow citizens to cast votes directly over issues, improve upon the outcomes reached under electoral competition among representatives. Their model consists of a two party political system, where the population makes decisions in a two-dimensional policy space. Using this framework, they show that the elected candidates may implement policies that are odds with the majority preferences of the population. They attribute these errors to the bundling of issues that is inherent in the election of a representative. As in our model, when issues are decided upon concurrently, via the choice of a representative, decisions may diverge from those made when citizens are able to vote directly over issues, one at a time. Ahn and Oliveros (2012) prove a similar result in the case where individuals' preferences over issues are not separable. Our paper shows that even in the case where there are no complementarities or substitutabilities across issues, the bundling of choices may be distortionary.

The mechanics behind this result also reveal a connection to the political science literature on vote trading, or log rolling (see Tullock 1970, Riker and Brams 1973). A vote trade may occur when two voters (or parties) are on opposite sides of two issues, with each being a pivotal voter on exactly one the two. In these circumstances, both voters may stand to gain by agreeing to vote against their true preference on the issue in which they are pivotal, in exchange for the other doing the same. Through these types of vote trades,

voters may succeed in implementing their preferred choices on the issues they value most, when they would otherwise have been out-voted by the majority. A sizable literature exists documenting the potential for this type of vote trading to alter outcomes under majority rule. As Schwartz (1977) observes, in cases in which majority rule does not cycle and thus would lead to a unique, stable outcome, the outcome under vote trading can be different. While the specifics of the two models are rather different, the underlying intuition is similar. In the model presented here, representative democracy, like vote trading, provides a way for small parties to impact choices through joint action. By each sacrificing their preferred choice on one issue, the minority parties are able to agree upon a compromise candidate which is an improved outcome for all of them.

While our result is similar in spirit to the vote trading literature, we point out that it is rather different from most other negative results about majority rule because it does not stem from a problem of cyclical collective preferences. It would be less surprising to find inconsistencies between direct and representative democracy when there was no ordering consistent with the direct democracy choices. But in our example, there is no tension between producing an ordering and matching the majority preferences of the population. Additionally, the divergence noted here is distinct from the type of conflict that has been identified between majoritarian methods and scoring rules in the work of Saari (see Saari 1995). For the populations in the counterexample above, the ordering generated by applying a scoring rule to the preferences over alternatives is exactly the ordering that is consistent with majority rule over alternatives. Thus, the outcome under representative democracy in this setting is at odds with the predictions of most previously-studied voting rules, majoritarian or not, which all agree that  $e$  is the “correct” ordering for the population. This makes the result more surprising, and perhaps more troublesome. We have proven the result for a relatively well-behaved population; we expect that for populations with cyclical majority preferences or for which scoring rules and majoritarian methods predict different outcomes, the choices under direct and representative democracy will diverge to a greater extent. Better understanding the tournaments over orderings associated with more general classes of populations is a topic for future study.

#### 4.4 Population Restrictions

A natural question to ask in this context is whether we can impose restrictions on the distribution of preferences that would guarantee order consistency. This domain restriction approach has been adopted by many social choice theorists in attempts to rule out other paradoxical outcomes; perhaps most classic is the single-peakedness restriction pioneered

independently by Black (1948) and Arrow (1951). Their goal was to describe a class of populations for which majority rule over alternatives would not cycle. The domain restriction they proposed requires populations to be unimodal in the sense that all members of the population, for any particular triple of alternatives, must be able to agree on an alternative that is not worst. Assuming the number of voters is odd, this condition is sufficient for transitive majority rule.

Clearly, this restriction will not be enough to assure order consistency, as the class of populations we consider in our counterexample above are indeed single-peaked in terms of preferences over alternatives. However, we can use a similar idea, that of restricting the number of modes in the distribution, in order to derive a sufficient condition for order consistency in our framework. The class of populations with transitive majority rule consistent with the ordering  $e$  can be thought of as having a “peak” or cluster of weight around  $e$ . Our sufficiency condition says that as we move away from  $e$ , we must not encounter another cluster of similar orderings. In order to state this condition more formally, it will be useful to introduce some new terminology. When referring to a population with transitive majority rule consistent with  $e$ , we will call any pairwise disagreement with  $e$  an **error**. For example, we will say that an ordering  $\pi$  that is  $m$  transpositions from  $e$  contains  $m$  errors. We can state our sufficiency condition in terms of these errors.

**Proposition 4** *Consider the class of populations with transitive majority rule consistent with  $e$ . Then, order consistency holds if for any set of  $m$  errors,  $m \geq 5$ , we have*

$$\sum \lambda(\pi | \pi \text{ contains at least } \frac{1}{2} \text{ of these } m \text{ errors}) < \frac{1}{2}$$

**Proof.** Suppose  $e \notin UC(T^{\lambda, \mu^K})$ . We will show there must exist a set of  $m$  errors,  $m \geq 5$ ,

such that  $\sum \lambda(\pi | \pi \text{ contains at least } \frac{1}{2} \text{ of these } m \text{ errors}) < \frac{1}{2}$ . Since  $e \notin UC(T^{\lambda, \mu^K})$ , we

know there exists  $\hat{\pi}$  such that  $\hat{\pi}$  covers  $e$ . Let  $\hat{\pi}$  contain  $m$  errors; we know  $m \geq 5$  in order for  $\hat{\pi}$  to cover  $e$ . Since  $\hat{\pi}$  and  $e$  agree on all pairwise choices outside of the  $m$  errors, we know that  $f(\pi, \hat{\pi}; \mu^K)$  and  $f(\pi, e; \mu^K)$  are determined only by how many of the  $m$  errors  $\pi$  contains. Those  $\pi$  that have less than  $\frac{1}{2}$  of the  $m$  errors have  $f(\pi, e; \mu^K) < f(\pi, \hat{\pi}; \mu^K)$ . So, suppose the set of orderings that had at least  $\frac{1}{2}$  of these  $m$  errors in common with  $\hat{\pi}$  had mass less than  $\frac{1}{2}$ . Then, we would have  $\sum \lambda(\pi | f(\pi, e; \mu^K) < f(\pi, \hat{\pi}; \mu^K)) > \frac{1}{2}$ . This would imply  $e \in T^{\lambda, \mu^K} \hat{\pi}$ , contradicting  $\hat{\pi}$  covers  $e$ . ■

This sufficiency condition has a straightforward intuition. If we encounter a population that contains a mass of orderings that are both (a) relatively distant from  $e$ , and (b) relatively

close to one another, then we may have the type of counterexample presented above. This type of cluster of similar orderings far from  $e$  may be able to agree upon a compromise candidate which covers  $e$ , but only if together they constitute a majority. The condition rules out this possibility by assuring that no such cluster of mass greater than  $\frac{1}{2}$  exists.

This is a sufficient but not necessary condition for order consistency. A gray area exists between the class of populations described in our counterexample above and the class of populations described by this sufficiency condition. For populations that fall into neither of these classes, perhaps the best answer to whether  $e$  is a member of  $UC(T^{\lambda, \mu^K})$  is it depends. For some populations that fail the condition above, the distribution of mass on orderings far from  $e$  may be too dispersed to agree upon an ordering like  $\hat{\pi}$  which could beat everything that  $e$  beats. One might ask whether we could improve the sufficiency condition by restricting this set of distant orderings to fall within a certain radius of one another. For an example that illustrates why this strategy fails, please see the Appendix.

## 4.5 Discussion

We summarize our results for populations with order consistent majority preferences in the table below.

	$n = 3$	$n = 4$	$5 \leq n \leq 9$	$n \geq 10$
strong order consistency	✓	x	x	x
order consistency	✓	✓	?	x

Order consistency results for  $5 \leq n \leq 9$  remain an open question. Our strategy here focuses on generating populations by manipulating preferences over mutually exclusive, non-overlapping pairs. Preferences over one pair of alternatives are independent of preferences over a mutually exclusive non-overlapping pair. As a result, we are able to combine arbitrary preferences over these pairs into an ordering. This makes generating a population with a particular set of majority preferences much simpler. This approach has been fruitful in constructing populations for which order consistency fails, though it may require more alternatives than might be necessary with other approaches. An important question is what is smallest number of alternatives necessary to have order consistency break down.

We have chosen to use the uncovered set as the solution concept for our tournaments. How heavily does our result depend on this choice? One of the attractive features of the uncovered set is that many other popular tournament solution concepts are subsets of the uncovered set (Laslier 1997). Therefore, it is possible to extend our negative result to many other solution concepts. This includes the basic refinements of the uncovered set, the iterated uncovered

set and the minimal covering set. It also includes the Banks solution and the Bipartisan set. Another well-studied method for identifying tournament winners is ranking the members of the tournament based upon their victories and losses within the tournament. The most popular of these ranking methods include the Copeland solution, the Markov solution, and the Slater solution, each of which is also a refinement of the uncovered set.

However, if we expanded our solution concept to the top cycle, which contains the uncovered set, we could prove a positive result. For any population with majority preferences consistent with an ordering, this ordering is a member of the top cycle of the tournament over orderings. We can prove this using Proposition 1. By performing one transposition at a time, each time transposing a pair of alternatives that were ordered according to the majority preference, we can construct a chain,  $eT^{\lambda,\mu^K}\pi_iT^{\lambda,\mu^K}\pi_jT^{\lambda,\mu^K}\dots T^{\lambda,\mu^K}\pi$ , from  $e$  to any other ordering  $\pi$ . Thus, we must have  $e \in TC(T^{\lambda,\mu^K})$ . But for populations like the one in our counterexample, the top cycle is large; it will contain many other orderings, including our  $\hat{\pi}$ . Thus, while order consistency would hold, strong order consistency would fail.

Our models operate on populations of strict preferences, however, the negative result that we have identified can be derived in a more basic setting as well. Consider our example from the introduction. We had a population of individuals who each had a yes/no preference on five different spending projects: building a school, repairing a road, improving the library, giving raises to public employees, and restoring a historic landmark. We can apply our same rule for mapping individuals' preferences over alternatives into preferences over candidates. When deciding between two candidates, we assume that an individual votes for the candidate with whom she agrees on the greatest number of issues. In our example, we showed that while the majority preference was to vote yes on each of these five issues, under representative democracy, the population never elected the candidate who voted yes on all five issues.

This example shows that our result does not depend on our assumption that voters' preferences take the form of orderings. What is important for our result is the existence of multiple issues. Under a direct democracy, voters express preferences over issues one at a time. Therefore, voters who may be very similar with respect to their decisions over the entire set of issues - they would agree most of the time - can end up voting differently on any one particular issue. But in a representative democracy, voters are able to express preferences over the entire set of issues by voting for a candidate. In this way, representative democracy can serve to unite groups of voters with similar, but not identical preferences.

## 5 Conclusion

In this paper, we introduce a tournament over orderings as a natural model of majoritarian forms of representative democracy. Instead of looking at pairwise comparisons of alternatives, as traditional tournaments do, our tournaments are built from pairwise comparisons of candidates, modeled as orderings of alternatives. We assume that when comparing two candidates, an individual votes for the candidate with whom she is most likely to agree about the choice from a randomly-selected pair of alternatives.

We find that when applied to the same problem, direct democracy, as modeled by tournaments over alternatives, and representative democracy, as modeled by tournaments over candidates, often predict different outcomes. In fact, even when the choices of the population under direct democracy are consistent with an ordering, the candidate with this ordering may not be selected under representative democracy. We conclude that the choice of institution plays a significant role in determining outcomes. While this is a familiar theme in much of the existing social choice and political economy literature, our result is novel because of the context in which we prove it. We consider a majoritarian method and find that even in settings when a Condorcet winner exists, a candidate may be elected who will not implement this alternative in certain choice problems.

Many open questions remain. One avenue to pursue would be considering different methods for modeling representative democracy. Because we are interested in comparing our results to the outcome of tournaments defined by majority preferences over pairs of alternatives, we focus on a model of representative democracy that depends only on the population's preferences over pairs of alternatives. However, it would be interesting to consider tournaments generated by other choice-based metrics, corresponding to different probability distributions over choice problems. We could ask how the likelihood of a choice problem relates to the probability of representative democracy implementing the outcome recommended by direct democracy. That is, are representative and direct democracy more likely to agree on choice problems which are very likely to be faced?

We could also explore ways of adapting this framework to more applied settings. In this work, we study a simple, abstract model of representative democracy. In most institutional settings, representative democracy has more structure than what we've imposed here: there may be party systems that shape which candidates are available, restrictions on how many issues or alternatives can be decided upon at once, or priority given to status quo alternatives or candidates. And, the process is dynamic. Considering these modifications would enrich the model, leading to sharper predictions and perhaps positive results for some forms of representative democracy.

## 6 Appendix

### 6.1 Expected Utility Tournaments

In the analysis above, we consider a model of representative democracy generated by the choice-based metric,  $f(\pi, \pi'; \mu^K)$ . Alternatively, we could use an expected utility framework to model representative democracy. To create a cardinal model, we must enrich our basic model in a few key ways. We introduce utility functions over the alternatives,  $U : X \rightarrow R$ ; a utility function,  $u$ , assigns a real number to each alternative in  $X$ . Then, we redefine a population,  $\lambda_u$ , as a distribution over utility functions rather than a distribution over orderings. We can easily generate a model of direct democracy,  $\Gamma^{\lambda_u}(X)$ , via majority preference in this setting:  $a_i \Gamma^{\lambda_u} a_j$  if  $\sum_{u \in U} [\lambda_u(u|u(a_i) > u(a_j))] > \frac{1}{2}$ .<sup>6</sup>

To define preferences over the candidates, we build up from our basic utility function over alternatives. Consider a member of the population with utility function,  $u$ . She evaluates a candidate  $\pi$  according to an expected utility function,  $\omega_{u,\nu}$ :

$$\omega_{u,\nu}(\pi) = \sum_{A \in \mathcal{X}} \nu(A) \times u(c_\pi(A))$$

She evaluates candidate  $\pi$  by considering the choices  $\pi$  would make from each possible choice problem. She takes a weighted average of the utility of each choice weighted by the likelihood of that choice problem arriving, which is given by  $\nu$ . From here, it is straightforward to define a rule for voting over candidates. We say that  $\pi'$  **attracts a majority against**  $\pi$  if there exists a subset of utility functions,  $U_1$ , with  $\lambda_u(U_1) > \frac{1}{2}$  such that  $\omega_{u,\nu}(\pi) < \omega_{u,\nu}(\pi')$  for all  $u \in U_1$ . In the case where  $\omega_{u,\nu}(\pi) = \omega_{u,\nu}(\pi')$ , we assume that the votes of those with utility function  $u$  are split evenly between  $\pi$  and  $\pi'$ . It is clear that this rule also generates a tournament over the orderings.

A natural question to ask is what is the relationship between the ordinal, metric-based model and the expected utility model here. Consider a tournament over orderings generated by a population  $\lambda$  and a choice-based metric,  $f(\pi, \pi'; \nu)$ . Is it possible to find a population of utility functions,  $\lambda_u$ , that generates this same tournament under  $\omega_{u,\nu}$ ? In general, the answer to this question is no. And in fact, for general  $n$ , there is no expected utility tournament that corresponds exactly to the tournament over orderings induced by  $f(\pi, \pi'; \mu^K)$ . If we restrict our attention to the  $n = 3$  case and those  $\nu$  that only put positive probability on choice problems of order two, we can find such a mapping. More importantly, despite the fact that the tournaments generated by expected utility and by  $f(\pi, \pi'; \mu^K)$  may not be identical, the

---

<sup>6</sup>For simplicity, assume  $u$  assigns unique values to each  $a_i \in X$ . This avoids the need to define tie-breaking rules.

solutions of these tournaments, and the results we proved above, may often correspond.

**Proposition 5** *Consider a tournament over orderings generated by a population  $\lambda$  and  $f(\pi, \pi'; \mu^K)$ . For  $n > 3$ , it is not generally possible to find a population of utility functions,  $\lambda_u$ , that generates this same tournament under  $\omega_{u,\nu}$ .*

Proof: We can show this with a quick counterexample. Let  $n = 4$ . Consider a member of the population whose preference is  $e$ . We will show that we cannot find a utility function  $u$  such that his preferences over orderings according to  $\omega_{u,\nu}(\pi)$  are the same as his preferences over orderings induced by  $f(e, \pi; \mu^K)$ . We know that for this agent  $\omega_{u,\nu}(\pi)$  must be the same for all  $\pi$  such that  $f(e, \pi; \mu^K) = \frac{1}{9}$  (these are all the  $\pi$  that are exactly one transposition from  $e$ ). Thus, we must have that  $u(a_i) - u(a_{i+1}) = u(a_{i+1}) - u(a_{i+2})$  for  $i \in \{1, 2\}$ . For convenience, let us label this distance  $u(a_i) - u(a_{i+1}) = \alpha$ . Now consider the following pair of orderings:  $a_3a_1a_2a_4$  and  $a_2a_1a_4a_3$ . For both of these  $\pi$ , we have  $f(e, \pi; \mu^K) = \frac{2}{9}$ . But, in terms of expected utility according to  $\omega_{u,\nu}(\pi)$ , when choices are made according to  $a_3a_1a_2a_4$ , an agent with preference  $e$  has a utility loss of  $2\alpha$  (for choosing  $a_3$  rather than  $a_1$  from  $\{a_1a_3\}$ ) plus a utility loss of  $\alpha$  (for choosing  $a_3$  rather than  $a_2$  from  $\{a_2a_3\}$ ). When choices are made according to  $a_2a_1a_4a_3$ , an agent with preference  $e$  has a utility loss of  $\alpha$  (for choosing  $a_2$  rather than  $a_1$  from  $\{a_1a_2\}$ ) plus a utility loss of  $\alpha$  (for choosing  $a_4$  rather than  $a_3$  from  $\{a_3a_4\}$ ). So,  $\omega_{u,\nu}(a_2a_1a_4a_3) > \omega_{u,\nu}(a_3a_1a_2a_4)$  even though  $f(e, \pi; \mu^K) = \frac{2}{9}$  for both of these  $\pi$ . Therefore, we cannot find a utility function that induces the same preferences over orderings as  $f(e, \pi; \mu^K)$  does. ■

**Proposition 6** *Suppose  $n = 3$ . Consider a tournament over orderings generated by a population  $\lambda$  and a choice-based metric,  $f(\pi, \pi'; \nu)$ , with  $\nu$  such that  $\nu(A) \succ 0$  only if  $|A| = 2$ . Then it is possible to find a population of utility functions,  $\lambda_u$ , that generates this same tournament under  $\omega_{u,\nu}$ .*

Proof: Suppose that there are three alternatives,  $\{a_1, a_2, a_3\}$ , and the likelihood of the four choice problems are given by  $\nu : \nu(a_1a_2) = p, \nu(a_2a_3) = q, \nu(a_1a_3) = 1 - p - q$ , where  $p, q \in [0, 1]$ . Without loss of generality, consider a member of the population with ordering  $\pi = a_1a_2a_3$ . Then we have:

$$\begin{aligned}
f(\pi, \pi; \nu) &= 0 \\
f(\pi, a_1 a_3 a_2; \nu) &= q \\
f(\pi, a_2 a_1 a_3; \nu) &= p \\
f(\pi, a_2 a_3 a_1; \nu) &= 1 - q \\
f(\pi, a_3 a_1 a_2; \nu) &= 1 - p \\
f(\pi, a_3 a_2 a_1; \nu) &= 1
\end{aligned}$$

We will show that we can find a utility function,  $u$ , to assign to this member of the population such that his preferences over the orderings induced by  $\omega_{u,\nu}$  will be the same as his preferences induced by the metric  $f(\pi, \pi'; \nu)$ . Without loss of generality, let  $u(a_1) = x$ ,  $u(a_2) = y$ ,  $u(a_3) = 0$ , with  $x > y > 0$ . Then, the expected payoff from each ordering is:

$$\begin{aligned}
\omega_{u,\nu}(\pi) &= [1 - q]x + qy \\
\omega_{u,\nu}(a_1 a_3 a_2) &= [1 - q]x \\
\omega_{u,\nu}(a_2 a_1 a_3) &= [p + q]y + [1 - p - q]x \\
\omega_{u,\nu}(a_2 a_3 a_1) &= [p + q]y \\
\omega_{u,\nu}(a_3 a_1 a_2) &= px \\
\omega_{u,\nu}(a_3 a_2 a_1) &= py
\end{aligned}$$

Straightforward calculations reveal that  $f(\pi, \pi'; \nu)$  and  $\omega_{u,\nu}$  induce the same preference over orderings for 11 of the 15 possible pairwise comparisons. That is, for 11 of the 15 pairs of orderings, we have  $f(\pi, \pi'; \nu) < f(\pi, \pi''; \nu)$  if and only if  $\omega_{u,\nu}(\pi') > \omega_{u,\nu}(\pi'')$ . There are four remaining relationships to be determined:  $(a_1 a_3 a_2, a_2 a_1 a_3)$ ,  $(a_1 a_3 a_2, a_2 a_3 a_1)$ ,  $(a_2 a_1 a_3, a_3 a_1 a_2)$ , and  $(a_2 a_3 a_1, a_3 a_1 a_2)$ .

The four relationships in question are governed by the following inequalities, shown simplified here.

$$a_1 a_3 a_2 \succ a_2 a_1 a_3 \Leftrightarrow px > (p + q)y \quad (1)$$

$$a_1 a_3 a_2 \succ a_2 a_3 a_1 \Leftrightarrow (1 - q)x > (p + q)y \quad (2)$$

$$a_2a_1a_3 \succ a_3a_1a_2 \Leftrightarrow (p+q)y + (1-p-q)x > px \quad (3)$$

$$a_2a_3a_1 \succ a_3a_1a_2 \Leftrightarrow (p+q)y > px \quad (4)$$

It is clear that for all values of  $p, q, x,$  and  $y,$  4 implies 3 and that 1 implies not 4. There are five possible cases to consider. For each case, we list the relevant ranges of  $p$  and  $q.$  Then, we list the corresponding preference over orderings induced by  $\nu$  for this range of probabilities.

- (1)  $1-p > p > q$  : 1, 2, and 3 hold and 4 does not
- (2)  $p > 1-p, p > q, (1-q) > p$  : 1 and 2 hold and 3 and 4 do not
- (3)  $p > 1-p, p > q, (1-q) < p$  : 1 holds and 2, 3, and 4 do not
- (4)  $1-q > q > p$  : 1 does not hold and 2, 3, and 4 do hold
- (5)  $q > p, q > 1-q$  : 1 and 2 do not hold and 3 and 4 do hold

For each case, we must prove that we can find a  $u$  (values of  $x, y$ ) such that we deduce the same preferences over orderings when looking at  $\omega_{u,\nu}.$

Case 1: Need to show  $\exists x, y$  such that 1, 2, and 3 hold and 4 does not. In this case, since  $p < (1-q),$  1 implies 2. Therefore, we only need to check:

$$\begin{aligned} (p+q)y &< px < (p+q)y + (1-p-q)x \\ 0 &< px - (p+q)y < (1-p-q)x \\ 0 &< p - (p+q)\frac{y}{x} < (1-p-q) \\ -p &< -(p+q)\frac{y}{x} < (1-2p-q) \\ \frac{p}{p+q} &< \frac{y}{x} < \frac{2p+q-1}{p+q} \end{aligned}$$

Since the left-hand side is weakly less than 1, we just need to check that  $p < 2p+q-1.$  This holds as long as  $p < 1-q,$  which is true in this case.  $\square$

Case 2: Need to show  $\exists x, y$  such that 1 and 2 hold and 3 and 4 do not. In this case, since  $p < (1-q),$  1 implies 2. And, not 3 implies 1. Therefore, we only need to check:

$$\begin{aligned}
px &> (p+q)y + (1-p-q)x \\
p &> (p+q)\frac{y}{x} + (1-p-q)x \\
2p+q-1 &> (p+q)\frac{y}{x} \\
\frac{y}{x} &< \frac{2p+q-1}{p+q}
\end{aligned}$$

So, we just need to have  $0 < 2p+q-1$ . This holds as long as  $p > 1-p-q$ , which is true in this case since  $p > 1-p$ .  $\square$

Case 3: Need to show  $\exists x, y$  such that 1 holds and 2, 3, and 4 do not. In this case, not 3 implies 1. Therefore, we only need to check:

$$\begin{aligned}
(1-q)x &< (p+q)y < px - (1-p-q)x \\
(1-q) &< (p+q)\frac{y}{x} < 2p+q-1 \\
\frac{(1-q)}{p+q} &< \frac{y}{x} < \frac{2p+q-1}{p+q}
\end{aligned}$$

Since  $p > 1-q$ , the left-hand side is less than 1. We just need to check that  $(1-q) < 2p+q-1$ . This holds as long as  $p > 1-q$ , which is true in this case.  $\square$

Case 4: Need to show  $\exists x, y$  such that 1 does not hold and 2, 3, and 4 do hold. In this case, 4 implies 3 and not 1. Therefore, we only need to check:

$$\begin{aligned}
px &< (p+q)y < (1-q)x \\
p &< (p+q)\frac{y}{x} < 1-q \\
\frac{p}{p+q} &< \frac{y}{x} < \frac{1-q}{p+q}
\end{aligned}$$

Since the left-hand side is weakly less than 1, we just need  $(1-q) > p$ , which is true in this case.  $\square$

Case 5: Need to show  $\exists x, y$  such that 1 and 2 do not hold and 3 and 4 do hold. In this case, 4 implies 3 and not 1. And, since  $(1-q) > p$ , we have not 2 implies 4. Therefore, we only need to check:

$$\begin{aligned} (p+q)y &> (1-q)x \\ \frac{y}{x} &> \frac{1-q}{p+q} \end{aligned}$$

We just need  $(1-q) < p+q$ , which is true in this case since  $q > 1-q$ .  $\square$

By exhausting these five cases, we've completed our proof. We've shown that for any  $n = 3$  tournament over orderings generated by a population  $\lambda$  and a choice-based metric,  $f(\pi, \pi'; \nu)$ , with  $\nu$  such that  $\nu(A) \succ 0$  only if  $|A| = 2$ , it is possible to find a population of utility functions,  $\lambda_u$ , that generates this same tournament under  $\omega_{u,\nu}$ .  $\blacksquare$

Despite the differences between the choice-based metric models and the expected utility models, the counterexample presented in Section 4 does work for a family of expected utility generated tournaments. Restrict all members of the population to the same scale of cardinalization; that is, each utility function assigns the alternative in the  $i$ th spot the same real number  $\forall i \in 1, \dots, n$ . Then if  $u(a_k) - u(a_{k-1}) = u(a_{k-1}) - u(a_{k-2})$ , the counterexample is still valid. While the structure of the tournament over candidates may be slightly different, we would still have that  $e$  is covered by our  $\hat{\pi}$ . Furthermore, holding fixed any particular utility function, common to all members of the population, it is likely possible to construct counterexamples that would allow us to prove an expected utility analog to 4.3. This is a topic for future work.

## 6.2 Cyclic Populations

### 6.2.1 Defining Consistency

Here, we consider a few reasonable notions of consistency between direct and representative democracy for cyclic populations. In the next subsection, we'll use these concepts to evaluate cyclic populations for  $n = 3$ .

Our first notion of consistency makes an effort to adapt the idea of order consistency to the case of cyclic populations. While majority preferences over alternatives are not generally consistent with an ordering, the Kemeny rule provides a way to identify an ordering that is "closest" to the population's preferences (Kemeny 1959, Kemeny and Snell 1962). The Kemeny ordering is the ordering which minimizes the expected distance to a randomly-drawn member of the population, where distance is given by  $f(\pi, \pi'; \mu^K)$ :

$$\pi^*(\lambda, \mu^K) = \arg \min_{\pi} \sum_{\pi' \in \Pi} f(\pi, \pi'; \mu^K) \lambda(\pi')$$

In the case where the population's majority preferences are consistent with an ordering, this ordering is the Kemeny ordering. This suggests a reasonable pair of consistency conditions for our tournament comparison:

**Definition 3** *Strong Kemeny Order Consistency:* For any population,  $\pi^*(\lambda, \mu^K) = UC(T^{\lambda, \mu^K})$ .

**Definition 4** *Kemeny Order Consistency:* For any population,  $\pi^*(\lambda, \mu^K) \in UC(T^{\lambda, \mu^K})$ .

Requiring the membership of the Kemeny ordering in the uncovered set of the tournament over alternatives is one intuitive operationalization of consistency in this context. It asks that the ordering that best fits the population's majority preferences be selected as a winner of the tournament over orderings. Given that we define our tournaments over orderings by  $f(\pi, \pi'; \mu^K)$ , it seems plausible to expect that the Kemeny ordering be a member of  $UC(T^{\lambda, \mu^K})$ .

We know, however, that the Kemeny method is just one way of producing a representative social ordering from a population of individual rational preferences. As Arrow's Theorem and the rich literature that followed has shown, there is no clear answer as to what ordering best represents the majority preferences of the population (Arrow 1951). Therefore, when developing a notion of testable consistency for our framework, it may be appealing to employ a more flexible concept. Instead of looking for any particular ordering in the uncovered set, we could ask that the orderings in  $UC(T^{\lambda, \mu^K})$  simply rank highly the winners from the tournament over alternatives. To formalize this concept, we define weak consistency:

**Definition 5** *Weak Consistency:* Suppose  $a_i \in UC(\Gamma(X))$  and  $a_j \notin UC(\Gamma(X))$ . Then, for any  $\pi \in UC(T^{\lambda, \mu^K})$ , we do not have  $a_j$  precedes  $a_i$ .

This axiom requires that members of  $UC(T^{\lambda, \mu^K})$  rank alternatives in  $UC(\Gamma^\lambda(X))$  above elements that are not in  $UC(\Gamma^\lambda(X))$ . In the case where there is a Condorcet winner among alternatives,  $a_i^* \in UC(\Gamma^\lambda(X))$ , weak consistency requires that  $a_i^*$  be ranked first in any ordering in  $UC(T^{\lambda, \mu^K})$ . Note that in contrast to the notions of consistency proposed above, weak consistency does not make any statements about the ranking of alternatives within  $UC(\Gamma^\lambda(X))$ . In this sense, weak consistency is more in keeping with the rationale of the uncovered set. Because any alternative within the uncovered set could be a winner under direct democracy, we have no grounds for requiring a specific ranking of these elements in the orderings of  $UC(T^{\lambda, \mu^K})$ .

### 6.2.2 Results

For populations with cyclic majority preferences,  $UC(\Gamma^\lambda(X)) = \{a_1, a_2, a_3\}$ . It is immediately clear that weak consistency will be vacuous in this case, as every alternative is a member of the uncovered set. But we can evaluate our other notions of consistency by looking at the structure of these tournaments. For illustration, we study the case of the cyclic majority preference  $a_1\Gamma^\lambda a_2$ ,  $a_2\Gamma^\lambda a_3$ , and  $a_3\Gamma^\lambda a_1$  and prove that this majority preference can generate six possible tournaments with identical consistency properties.

We proceed in the same fashion as above. First, we map majority preferences over pairs onto  $T^{\lambda, \mu^K}$ . Again, the majority preference over each of the three pairs  $(a_1, a_2)$ ,  $(a_1, a_3)$ , and  $(a_2, a_3)$  determines the relationship between three pairs of orderings. As shown in Figure 7, the only wrinkle is the reversal of the relationship between the pairs of orderings that depend on the majority preference of  $(a_1, a_3)$ .

Because we no longer have that  $a_1\Gamma^\lambda a_3$ , we cannot employ the technique of combining information about pairs of majority preferences over pairs. Instead, we have to rely on information about the strength of majorities in order to flesh out the rest of the tournament relation. We know that for a given cyclic majority preference, we can rank the strength of the majority for each pair; so, we can identify a strongest link (the pair for which the size of the majority is largest) and a weakest link (the pair for which the size of the majority is smallest).

Suppose in our example that the weakest link is  $a_3\Gamma^\lambda a_1$ . The fact that  $a_3\Gamma^\lambda a_1$  is supported by a smaller majority than  $a_1\Gamma^\lambda a_2$  yields the following inequality:

$$\begin{aligned} & [\lambda(a_3a_1a_2) + \lambda(a_3a_2a_1) + \lambda(a_2a_3a_1)] - [\lambda(a_1a_2a_3) + \lambda(a_1a_3a_2) + \lambda(a_2a_1a_3)] < \\ & [\lambda(a_1a_2a_3) + \lambda(a_1a_3a_2) + \lambda(a_3a_1a_2)] - [\lambda(a_3a_2a_1) + \lambda(a_2a_3a_1) + \lambda(a_2a_1a_3)] \end{aligned}$$

$$2[\lambda(a_3a_2a_1) + \lambda(a_2a_3a_1)] < 2[\lambda(a_1a_2a_3) + \lambda(a_1a_3a_2)]$$

$$[\lambda(a_3a_2a_1) + \lambda(a_2a_3a_1)] < [\lambda(a_1a_2a_3) + \lambda(a_1a_3a_2)]$$

We can determine two pairs of  $T^{\lambda, \mu^K}$  relationships from this inequality:  $a_1a_2a_3T^{\lambda, \mu^K} a_2a_3a_1$  and  $a_1a_3a_2T^{\lambda, \mu^K} a_3a_2a_1$ . Similarly, the fact that  $a_3\Gamma^\lambda a_1$  is supported by a smaller majority than  $a_2\Gamma^\lambda a_3$  tells us that  $[\lambda(a_3a_2a_1) + \lambda(a_3a_1a_2)] < [\lambda(a_1a_2a_3) + \lambda(a_2a_1a_3)]$ , which in turn determines  $a_1a_2a_3T^{\lambda, \mu^K} a_3a_1a_2$  and  $a_2a_1a_3T^{\lambda, \mu^K} a_3a_2a_1$ . See Figure 8 for a diagram of these relationships.

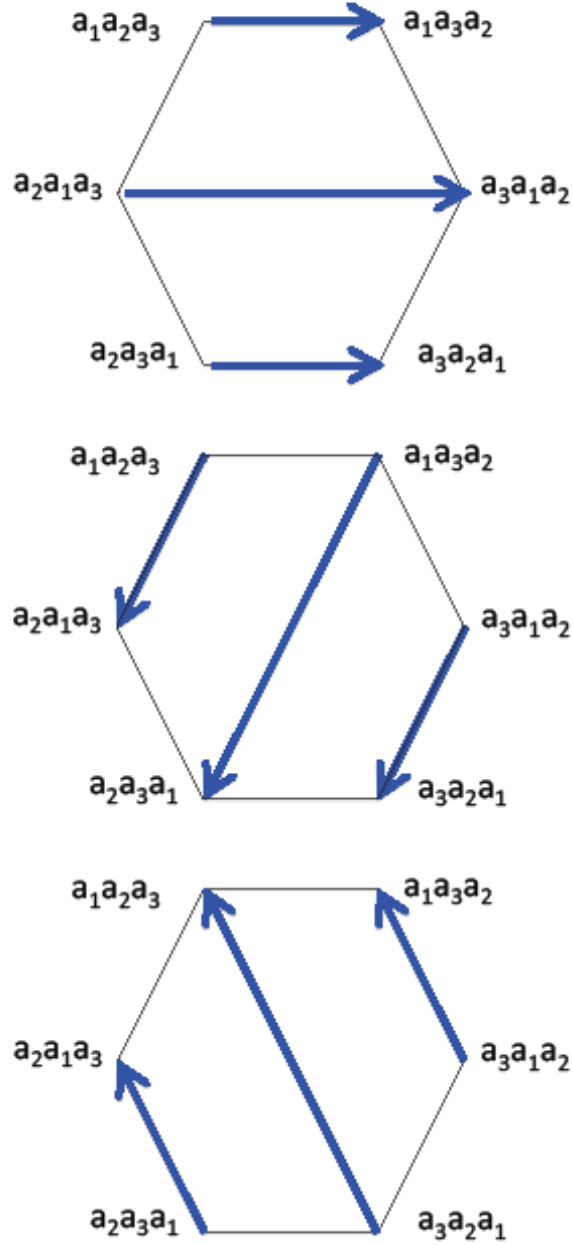


Figure 7: In the top diagram, the three arrows illustrate  $T^{\lambda, \mu^K}$  relationships determined by  $a_2\Gamma^\lambda a_3$ . In the middle diagram, the three arrows illustrate  $T^{\lambda, \mu^K}$  relationships determined by  $a_1\Gamma^\lambda a_2$ . In the bottom diagram, the three arrows illustrate  $T^{\lambda, \mu^K}$  relationships determined by  $a_3\Gamma^\lambda a_1$ .

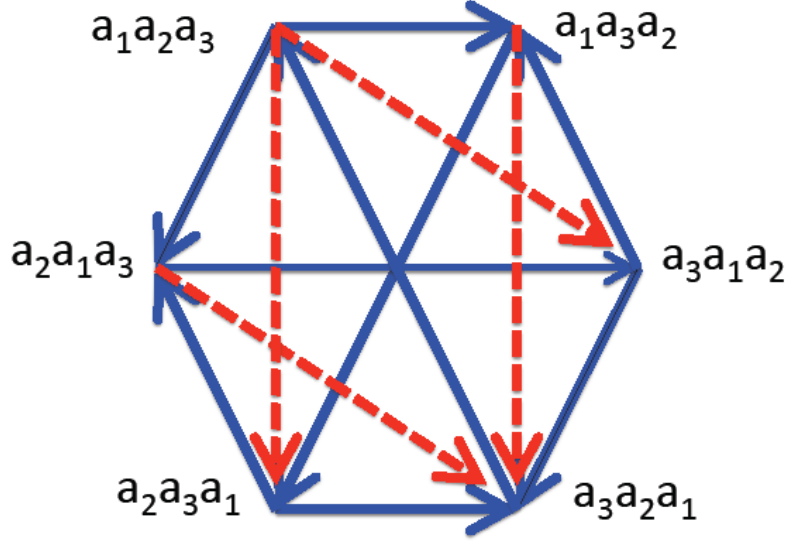


Figure 8: The dashed arrows are determined by the fact that  $a_3\Gamma^\lambda a_1$  is supported by a smaller majority than  $a_1\Gamma^\lambda a_2$  and that  $a_3\Gamma^\lambda a_1$  is supported by a smaller majority than  $a_2\Gamma^\lambda a_3$ .

We are left with the same two relationships to determine as we had in the transitive case; here, however, the uncovered set of the tournament over orderings will depend on the resolution of these links, as  $a_1a_2a_3$  is no longer a Condorcet winner. Again we can turn to the relative strength of the majorities in favor of  $(a_1, a_2)$  and  $(a_2, a_3)$ . Suppose that the strength of the majority for preference for  $a_1 \succ a_2$  is greater than the strength of the majority preference for  $a_2 \succ a_3$ . Then we deduce that  $[\lambda(a_1a_3a_2) + \lambda(a_3a_1a_2)] > [\lambda(a_2a_1a_3) + \lambda(a_2a_3a_1)]$ , and we must have  $a_1a_3a_2T^{\lambda, \mu^K} a_2a_1a_3$  and  $a_3a_1a_2T^{\lambda, \mu^K} a_2a_3a_1$ . If we instead have that the strength of the majority preference for  $a_2 \succ a_3$  is greater than the strength of the majority preference for  $a_1 \succ a_2$ , then both of these  $T^{\lambda, \mu^K}$  relationships would be reversed.

Therefore, for a given cyclic majority relation, we can have six possible tournaments. The differences stem from what pairs are supported by the weakest and strongest majorities: there are three possible weakest links, and each weakest link can be paired with two possible strongest links.

We can check these uncovered sets for Kemeny order consistency and see that it is always satisfied. The Kemeny winner of each tournament is the ordering which breaks the cycle at its weakest majority: in the top row, when the smallest majority is  $a_2 \succ a_3$ ,  $\pi^*(\lambda, \mu^K) = \{a_3a_1a_2\}$ , in the middle row, when the smallest majority is  $a_3 \succ a_1$ ,  $\pi^*(\lambda, \mu^K) = \{a_1a_2a_3\}$ , and in the last row, when the smallest majority is  $a_1 \succ a_2$ ,  $\pi^*(\lambda, \mu^K) = \{a_2a_3a_1\}$ . Thus, in every case we have  $\pi^*(\lambda, \mu^K) \in UC(T^{\lambda, \mu^K})$  and we see that Kemeny order consistency is

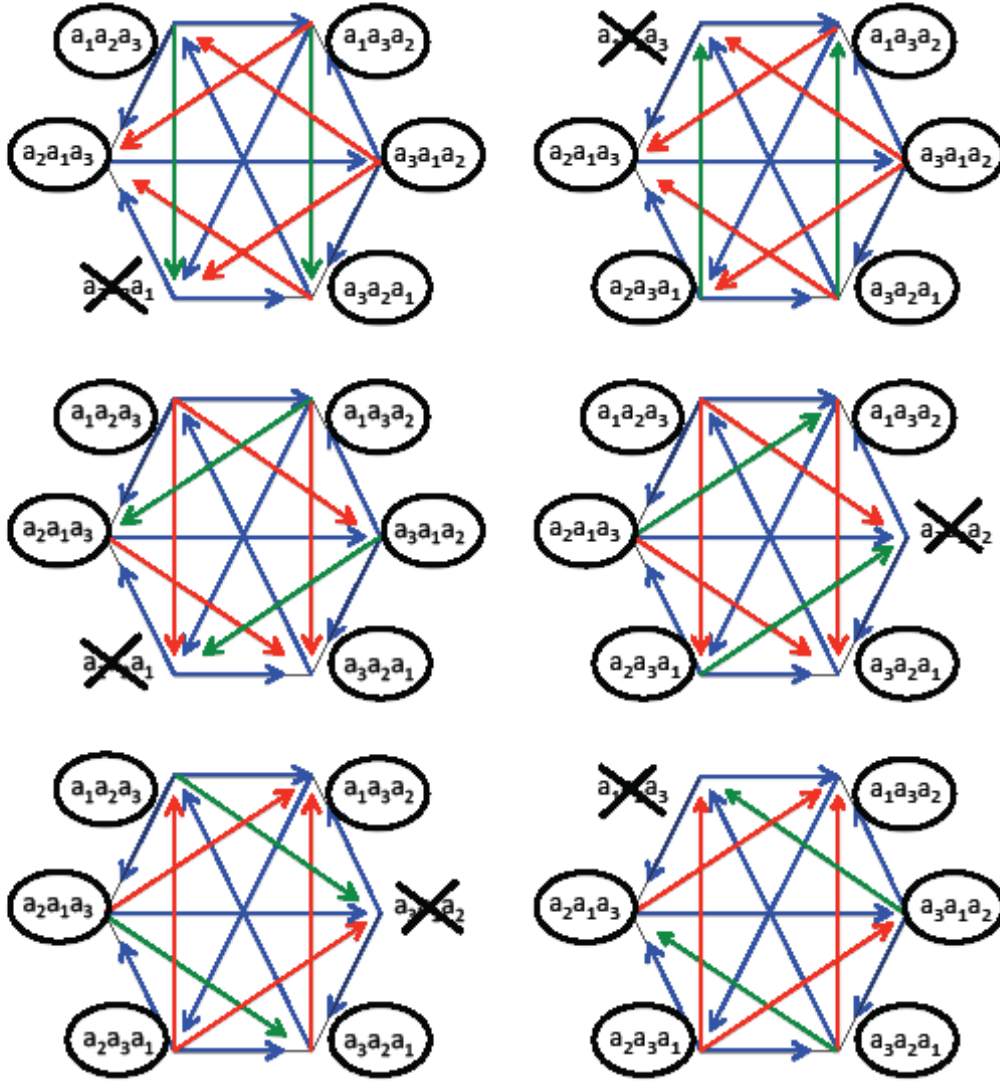


Figure 9: The above diagram illustrates the uncovered sets for the six possible tournaments in the cyclic case. A circle around the ordering indicates that it is a member of the uncovered set; an X through the ordering indicates that it is covered. The top row features the two tournaments generated when  $a_2\Gamma^\lambda a_3$  is supported by the smallest majority: on the left is the tournament generated when  $a_1\Gamma^\lambda a_2$  is strongest, one the right is the tournament generated when  $a_3\Gamma^\lambda a_1$  is strongest. In the second row,  $a_3\Gamma^\lambda a_1$  is supported by the smallest majority ( $a_1\Gamma^\lambda a_2$  strongest on the left,  $a_2\Gamma^\lambda a_3$  on the right). In the third row,  $a_1\Gamma^\lambda a_2$  is supported by the smallest majority ( $a_2\Gamma^\lambda a_3$  strongest on the left,  $a_3\Gamma^\lambda a_1$  on the right). While some of these tournaments have the same uncovered sets, no two of these tournaments are identical.

always satisfied for any  $n = 3$  population. But, strong Kemeny order consistency fails for cyclic tournaments, as the Kemeny ordering is not the sole member of the uncovered set for the populations above. In all of the cyclic cases,  $|UC(T^{\lambda, \mu^K})| = 5$ . Interestingly, note that the ordering covered in any particular case above agrees with the majority preference on two of the three pairs. The three orderings that are always in the uncovered set:  $a_1 a_3 a_2$ ,  $a_2 a_1 a_3$ , and  $a_3 a_2 a_1$  only agree with the majority preference on one of the three pairs. Thus, while the closest ordering to the majority preferences as defined by the Kemeny method is always a member of the uncovered set, agreement with the majority preferences is not in general a good predictor of membership in  $UC(T^{\lambda, \mu^K})$ .

We can summarize the results for the  $n = 3$  case with the following proposition:

**Proposition 7** *For  $n = 3$ ,  $UC(T^{\lambda, \mu^K})$  satisfies strong order consistency, Kemeny order consistency, and weak consistency.*

Therefore, for problems with only three alternatives, our results indicate that direct and representative democracies yield similar policy outcomes. In the case of transitive majority preferences in the population, our model predicts that a candidate selected via representative democracy will implement the same policy choice that would have been agreed upon under direct democracy for any pairwise choice problem. In the case of a cyclic majority preference, the policy predictions under direct democracy are less clear - the uncovered set predicts any of the possible policies could potentially be a winner. Correspondingly, the tournament over orderings identifies a larger set of electable candidates. Importantly, this set always includes the ordering closest to the majority preferences, the Kemeny winner. However, this set does not include all three orderings that are consistent with the maximal number of direct democracy choices. One of the orderings that makes two of the three choices made under direct democracy is omitted from the uncovered set, while all three of the orderings that make only one of the two choices made under direct democracy are included. Thus, even for the  $n = 3$  case, we see that the policies under direct and representative democracy may diverge for cyclic populations.

### 6.3 Additional Results for $n=4$

In this section, we prove Proposition 3 stated in Section 4.3. For  $n = 4$ ,  $UC(T^{\lambda, \mu^K})$  satisfies order consistency.

Proof: Assume majority preferences are consistent with  $e$ . We show that there is no ordering that can cover  $e$ . The key step is to recognize that we can apply Proposition 1 to rule out any ordering fewer than 5 transpositions from  $e$ :  $eT^{\lambda, \mu^K}\pi$  for any  $\pi$  within two

transpositions, so they cannot cover  $e$ , and for those three or four transpositions away, even if they beat  $e$ , they will be defeated by at least one ordering one or two transpositions from  $e$  (which  $e$  beats). So, the only orderings that could potentially cover  $e$  are five or six transpositions away from  $e$ :  $\{a_4a_3a_1a_2, a_4a_2a_3a_1, a_3a_4a_2a_1, a_4a_3a_2a_1\}$ . We rule these out one at a time:

- We cannot have  $a_4a_3a_1a_2$  covers  $e$ , since  $eT^{\lambda, \mu^K} a_2a_3a_1a_4T^{\lambda, \mu^K} a_4a_3a_1a_2$  for any population with majority preferences consistent with  $e$ . Proposition 1 proves  $eT^{\lambda, \mu^K} a_2a_3a_1a_4$ . We cannot have  $a_2a_3a_1a_4T^{\lambda, \mu^K} a_4a_3a_1a_2$  since all of the orderings closer or equidistant to  $a_4a_3a_1a_2$  than  $a_2a_3a_1a_4$  have  $a_4$  precedes  $a_2$ . Thus, if more than half the population were closer to or equidistant to  $a_4a_3a_1a_2$ , we would not have  $a_2 \succ a_4$  in the majority preference.

- We cannot have  $a_4a_2a_3a_1$  covers  $e$ , since we must have  $eT^{\lambda, \mu^K} a_4a_2a_3a_1$ . All of the orderings closer or equidistant to  $a_4a_3a_1a_2$  than  $e$  have  $a_4$  precedes  $a_1$ . Thus, if more than half the population were closer to or equidistant to  $a_4a_2a_3a_1$  than to  $e$ , we would not have  $a_1 \succ a_4$  in the majority preference.

- We cannot have  $a_3a_4a_2a_1$  covers  $e$ , since  $eT^{\lambda, \mu^K} a_1a_4a_2a_3T^{\lambda, \mu^K} a_3a_4a_2a_1$  for any population with majority preferences consistent with  $e$ . Proposition 1 proves  $eT^{\lambda, \mu^K} a_1a_4a_2a_3$ . We cannot have  $a_3a_4a_2a_1T^{\lambda, \mu^K} a_1a_4a_2a_3$  since all of the orderings closer or equidistant to  $a_1a_4a_2a_3$  than  $a_3a_4a_2a_1$  have  $a_3$  precedes  $a_1$ . Thus, if more than half the population were closer to or equidistant to  $a_3a_4a_2a_1$ , we would not have  $a_1 \succ a_3$  in the majority preference.

- We cannot have  $a_4a_3a_2a_1$  covers  $e$ , since  $eT^{\lambda, \mu^K} a_1a_3a_2a_4T^{\lambda, \mu^K} a_4a_3a_2a_1$  for any population with majority preferences consistent with  $e$ . Proposition 1 proves  $eT^{\lambda, \mu^K} a_1a_3a_2a_4$ . We cannot have  $a_4a_3a_2a_1T^{\lambda, \mu^K} a_1a_3a_2a_4$  since all of the orderings closer or equidistant to  $a_4a_3a_2a_1$  than  $a_1a_3a_2a_4$  have  $a_4$  precedes  $a_1$ . Thus, if more than half the population were closer to or equidistant to  $a_4a_3a_2a_1$ , we would not have  $a_1 \succ a_4$  in the majority preference. ■

## 6.4 Population Restrictions

**Example 1** *Why Tightening the Sufficiency Condition by Restricting the Radius of the Outlier Orderings Does Not Work*

Consider the following population, a slight variant from the example presented in Section 4:

$\pi$	$\lambda(\pi)$
$e$	$\frac{1}{2} - 2\varepsilon$
$a_1a_2a_4a_3a_6a_5a_8a_7a_{10}a_9$	$\frac{1}{5}(\frac{1}{2} - \varepsilon)$
$a_2a_1a_3a_4a_6a_5a_8a_7a_{10}a_9$	$\frac{1}{5}(\frac{1}{2} - \varepsilon)$
$a_2a_1a_4a_3a_5a_6a_8a_7a_{10}a_9$	$\frac{1}{5}(\frac{1}{2} - \varepsilon)$
$a_2a_1a_4a_3a_6a_5a_7a_8a_{10}a_9$	$\frac{1}{5}(\frac{1}{2} - \varepsilon)$
$a_2a_1a_4a_3a_6a_5a_8a_7a_9a_{10}$	$\frac{1}{5}(\frac{1}{2} - \varepsilon)$
$a_{10}a_9a_8a_7a_6a_5a_4a_3a_2a_1$	$3\varepsilon$

Using the strategy from the proof above, we can show that  $\hat{\pi} = a_2a_1a_4a_3a_6a_5a_8a_7a_{10}a_9$  covers  $e$ , the ordering consistent with majority preferences. We need to show that (a)  $\hat{\pi}T^{\lambda,\mu^K}e$  and (b)  $\forall \pi' \in \Pi, eT^{\lambda,\mu^K}\pi' \Rightarrow \hat{\pi}T^{\lambda,\mu^K}\pi'$ . First we will show that  $\hat{\pi}T^{\lambda,\mu^K}e$ . For the five orderings in population with weight  $\frac{1}{5}(\frac{1}{2} - \varepsilon)$ , we have  $f(\pi, \hat{\pi}) = 1$  and  $f(\pi, e) = 4$ . And, we know  $a_{10}a_9a_8a_7a_6a_5a_4a_3a_2a_1$  is closer to  $\hat{\pi}$  than  $e$ , since it is maximally distant from  $e$ . Thus,  $\frac{1}{2} + 2\varepsilon$  of the population is closer to  $\hat{\pi}$  than  $e$ , so  $\hat{\pi}T^{\lambda,\mu^K}e$ . Now we need to show there cannot exist  $\pi'$  such that  $eT^{\lambda,\mu^K}\pi'$  but  $\pi'T^{\lambda,\mu^K}\hat{\pi}$ . Suppose there did exist such a  $\pi'$ . Then,  $eT^{\lambda,\mu^K}\pi'$  implies that for at least one of the orderings  $\pi$  in population other than  $e$ ,  $f(e, \pi) \leq f(\pi', \pi)$ . We know there cannot exist a  $\pi'$ ,  $\pi' \neq e$ , such that  $f(e, a_{10}a_9a_8a_7a_6a_5a_4a_3a_2a_1) \leq f(\pi', a_{10}a_9a_8a_7a_6a_5a_4a_3a_2a_1)$ . So, it must be that this is true for one of the remaining five orderings. Since for any of these orderings  $f(e, \pi) = 4$ , we must have  $f(\pi', \pi) \geq 4$  for at least one of those five orderings  $\pi$ . And, the fact that  $\pi'T^{\lambda,\mu^K}\hat{\pi}$  implies that we have at least one of the following two cases:

1. For at least one of the orderings with weight  $\frac{1}{5}(\frac{1}{2} - \varepsilon)$ , we have  $f(\pi', \pi) \leq f(\hat{\pi}, \pi)$ .
2. For both  $e$  and  $a_{10}a_9a_8a_7a_6a_5a_4a_3a_2a_1$ , we have  $f(\pi', \pi) < f(\hat{\pi}, \pi)$ .

For case 1, we know  $f(\hat{\pi}, \pi) = 1$ , so this would imply,  $f(\pi', \pi) \leq 1$  for one of the orderings with weight  $\frac{1}{5}(\frac{1}{2} - \varepsilon)$ . This leads to the same violation of the triangle inequality that we reached above, since for any two orderings with weight  $\frac{1}{5}(\frac{1}{2} - \varepsilon)$ , we have  $f(\pi_i, \pi_j) \leq 2$ . For case 2,  $f(\pi', e) < f(\hat{\pi}, e)$  implies  $f(\pi', e) < 5$ . And,  $f(\pi', a_{10}a_9a_8a_7a_6a_5a_4a_3a_2a_1) < f(\hat{\pi}, a_{10}a_9a_8a_7a_6a_5a_4a_3a_2a_1)$  implies  $f(\pi', a_{10}a_9a_8a_7a_6a_5a_4a_3a_2a_1) < 40$ . But,  $f(\pi', e) < 5$  and  $f(\pi', a_{10}a_9a_8a_7a_6a_5a_4a_3a_2a_1) < 40$  cannot both hold, since the first implies  $\pi'$  has fewer than five errors and the second implies it has more than 5 errors. This leads to a contradiction. Thus, order consistency fails for this population.

This example illustrates the difficulty we encounter if we attempt to tighten the sufficiency condition for ordering consistency by imposing a radius around the orderings with common

errors. Taking the basic counterexample from above, where the minority orderings all lie relatively close to another, we can move some weight to  $a_{10}a_9a_8a_7a_6a_5a_4a_3a_2a_1$  and still arrive at  $\hat{\pi}$  covers  $e$ . Thus, it is not always true that we need the minority orderings to be relatively close to one another in order to have order consistency fail.

## References

- [1] Ahn, D.S. and S. Oliveros. 2012. Combinatorial Voting. *Econometrica* 80(1), 89-141.
- [2] Arrow, K.J. 1951. *Social Choice and Individual Values*. Wiley, New York.
- [3] Baldiga, K.A. and J.R. Green. 2011. Assent-Maximizing Social Choice. *Social Choice and Welfare*
- [4] Besley, T. and S. Coate. 2008. Issue Unbundling via Citizens' Initiatives. *Quarterly Journal of Political Science* 3(4), pp. 379-397.
- [5] Black, D. 1948. On the Rational of Group Decision Making. *Journal of Political Economy*, 56.
- [6] Bossert, W. and T. Storcken. 1992. Strategy-proofness of social welfare functions: The use of the Kemeny distance between preference orderings. *Social Choice and Welfare*, 9, 345-360.
- [7] Bowser, Jennie D. Initiative, Referendum, and Recall. National Conference of State Legislatures. Jennie Drage Bowser. August 2012. <<http://www.ncsl.org/legislatures-elections/elections/initiative-referendum-and-recall-overview.aspx>>
- [8] Cox, G.W. 1987. The Uncovered Set and the Core. *American Journal of Political Science* 31(2), 408-422.
- [9] Dasgupta, P. and E. Maskin. (2008). On the Robustness of Majority Rule. *Journal of the European Economic Association* 6, 5, 949-973.
- [10] Frey, B. 1994. Direct Democracy: Politico-Economic Lessons from the Swiss Experience. *American Economic Review*, 84(2), 338-342.
- [11] Kemeny, J.G. 1959. Mathematics Without Numbers. *Daedalus* 88. 577-591.
- [12] Kemeny, J.G. and J.L. Snell. 1962. *Mathematical Models in the Social Sciences*, chapter 2, Cambridge: MIT Press.

- [13] Koehler, D.H. 1975. Vote Trading and the Voting Paradox: A Proof of Logical Equivalence. *The American Political Science Review* 69(3), 954-960.
- [14] Laslier, J.F. 1997. *Tournaments Solutions and Majority Voting*. Berlin, New York: Springer-Verlag.
- [15] McGarvey, D.C. 1953. A Theorem on the Construction of Voting Paradoxes. *Econometrica* 21, 608-610.
- [16] McKelvey, R.D. 1986. Covering, Dominance, and Institution-Free Properties of Social Choice. *American Journal of Political Science* 30(2), 283-314.
- [17] Miller, N.R. 1980. A New Solution Set for Tournaments and Majority Voting: Further Graph-Theoretical Approaches to the Theory of Voting. *American Journal of Political Science* 24(1), 68-96.
- [18] Riker, W.H. and S.J. Brams. 1973. The Paradox of Vote Trading. *The American Political Science Review* 67(4), 1235-1247.
- [19] Saari, D.G. 1995. *Basic Geometry of Voting*. Berlin: Springer-Verlag.
- [20] Schwartz, T. 1977. Collective Choice, Separation of Issues and Vote Trading. *The American Political Science Review* 71(3), 999-1010.
- [21] Shepsle, K.A. and B.R. Weingast. 1984. Uncovered Sets and Sophisticated Voting Outcomes with Implications for Agenda Institutions. *American Journal of Political Science* 28(1), 49-74.
- [22] Tullock, G. 1970. A Simple Algebraic Logrolling Model. *American Economic Review* 60, 419-426.
- [23] Young, H.P. 1988. Condorcet's Theory of Voting. *American Political Science Review*. 82(2). 1231-1244.
- [24] Young, H.P and A. Levenglick. 1978. A Consistent Extension of Condorcet's Election Principle, *SIAM Journal on Applied Mathematics*, 35(2), 285-300.