

CAHIER DE RECHERCHE #2110E  
DÉPARTEMENT DE SCIENCE ÉCONOMIQUE  
FACULTÉ DES SCIENCES SOCIALES  
UNIVERSITÉ D'OTTAWA

WORKING PAPER #2110E  
DEPARTMENT OF ECONOMICS  
FACULTY OF SOCIAL SCIENCES  
UNIVERSITY OF OTTAWA

## **Round-Robin Political Tournaments: Abstention, Truthful Equilibria, and Effective Power**

Roland Pongou<sup>1</sup> and Bertrand Tchantcho<sup>2</sup>

30 August 2021

---

<sup>1</sup> Department of Economics, University of Ottawa, Canada; [rpongou@uottawa.ca](mailto:rpongou@uottawa.ca); Center for African Studies, Harvard University, United States of America; [rpongou@fas.harvard.edu](mailto:rpongou@fas.harvard.edu).

<sup>2</sup> Department of Mathematics, École Normale Supérieure, University of Yaoundé I, Cameroon; Université de Cergy Pontoise, THEMA, France; [btchantcho@yahoo.fr](mailto:btchantcho@yahoo.fr).

## ***Abstract***

*A round-robin political tournament is an election format where multiple candidates contest in pairs, and votes are aggregated using a general rule to form a social ranking. We formalize this tournament as a strategic form game and provide a necessary and sufficient condition under which truthful voting is a Nash equilibrium. Building on this analysis, we study the concept of effective power, defined as a voter's ability to bring about a social ranking that maximizes his preferences. We show that the classical theories of political power do not translate into effective power in general. We then provide a full characterization of the classes of political tournaments and utility metrics for which these theories capture effective power. We offer both structural and behavioral interpretations of the findings, and derive practical implications for the design of political tournaments that are compatible with truth-telling.*

**Keywords:** Round-robin Political Tournaments; Ranked Voting; Hyper-preferences; Truthful Equilibria; Effective Power; Psychology; Political Design

**JEL classification:** D72, C72, D02, D90, P16

# 1 Introduction

We study voter behavior and effective power in round-robin political tournaments. A tournament involves a finite number of competing political leaders or candidates. Each candidate is pitted against each of the other candidates in a series of independent pairwise contests. In each contest, each voter votes in favor of one of the two candidates who are in competition or abstains. The winner of each contest is selected under a deterministic voting rule that aggregates the expressed opinions of voters. The tournament (or social) ranking is determined based on the outcomes of these local pairwise contests. Each voter derives utility from how close his preferences over candidates are to the tournament ranking.

We address two main questions. First, we analyze voters' behavior and identify conditions under which truthful voting is always a Nash equilibrium. Second, we study a voter's *effective power* in a tournament, defined as his ability to bring about a social ranking that maximizes his preferences. Building on this analysis, we further analyze conditions under which the classical theories of voting power—the influence relation, the Shapley-Shubik power index, and the Banzhaf power index—translate into effective power.

In a political tournament, the utility that each voter derives from the social ranking is objectively measured using a natural generalization of a classical measure of disagreement between preference relations. In our setting, each voter  $i$  has a preference relation  $R^i$  over candidates, which is a weak order. Each local contest involves two candidates  $a$  and  $b$  chosen from the set of candidates. Each voter supports one of the two candidates or abstains. An aggregation function based on a voting rule aggregates individual votes and yields the winner of the contest, if any, or equivalently the social ranking over the pair  $\{a, b\}$ . We use the large class of voting rules with abstention (Tchantcho (2008)), which includes both anonymous and non-anonymous rules (Freixas and Zwicker (2009), Guemmegne and Pongou (2014)), with majority rule (Dasgupta and Maskin (2008, 2020) being a particular case. If the social ranking coincides with the preference of voter  $i$  over  $\{a, b\}$ , then  $i$  is fully satisfied and hence his dissatisfaction level is zero. If on the contrary, the social ranking is the complete opposite of  $i$ 's preferences over  $\{a, b\}$ , then  $i$  is fully dissatisfied with the outcome and his dissatisfaction level is normalized to 1. There are other possibilities between these two extreme cases. The dissatisfaction level of a voter  $i$  is:  $\alpha$  if  $i$  is indifferent between  $a$  and  $b$  but the society strictly prefers  $a$  to  $b$ ;  $\beta$  if  $i$  strictly prefers  $a$  to  $b$  but  $a$  and  $b$  are socially incomparable<sup>4</sup>; and  $\lambda$  if  $i$  is indifferent between  $a$  and  $b$  but  $a$  and  $b$  are socially incomparable. The parameters  $\alpha$ ,  $\beta$  and  $\lambda$  are strictly positive and are no greater than 1, the highest level of dissatisfaction. The utility of each voter  $i$  is obtained by taking the sum of  $i$ 's utilities (or dissatisfaction levels) over all the pairs of candidates  $\{a, b\}$ .

In our environment, a voter's utility function can be viewed as representing what is popularly known in the utility theory literature as a *hyper-preference* relation or a moral meta-ranking (Sen (1974)). A hyper-preference relation is a ranking of the rankings of a set of alternatives. Sen (1974) argues that resorting to hyper-preferences is needed for a proper account of choice. The predominance of hyper-preferences in choice theory stems from the fact that one generally has to assess individuals' utilities from a social state or consensus by evaluating how close their preferences are to this state (Sen (1974)), especially as the latter is generally represented by a ranking of all possible social alternatives.

Our measure of utility belongs to a very large class of parameterized functions called  $d_{\alpha,\beta,\lambda}$ . This class generalizes the well-known measure known as the Kemeny distance.<sup>5</sup> The Kemeny distance is defined over

---

<sup>4</sup>Social incomparability between two candidates occurs when neither beats the other. For this reason, it is popularly interpreted as a social tie. But a proper social tie—a situation where candidate  $a$  beats candidate  $b$  and  $b$  beats  $a$ —is a mathematical impossibility under majority-like voting rules, be they anonymous or not. This is because of the well-known properness condition that defines these rules. According to this condition, the complement of any winning coalition is a losing coalition. This condition avoids plausible social instability that could arise from two non-overlapping winning coalitions having completely opposite opinions on how the society should be run.

<sup>5</sup>Can and Storcken (2018) note that the Kemeny distance goes back to the eighteenth century (Cramer (1750)). Kemeny (1959) provided a first characterization of this metric on weak preferences, a beautiful refinement of which can be found in Can and Storcken (2018). The Kemeny distance is a classical measure of disagreement between complete and transitive binary relations. It measures the number of ranking disagreements between two binary relations over pairs of elements of a finite set.

the set of binary relations that are complete and transitive and so do not admit incomparability, whereas our generalized functions are defined over the set of all binary relations, encompassing strict rankings, indifferences and incomparability. Our dissatisfaction functions also generalize loss functions, a natural way to measure dissatisfaction with an outcome when compared to an ideal.

It is well-known that in a tournament such as ours, voters might misrepresent their preferences. We formalize a round-robin political tournament as a strategic form game, and provide a necessary and sufficient condition on the parameters  $\alpha$ ,  $\beta$  and  $\lambda$  for truthful voting to be a Nash equilibrium (Theorem 1). We find that truthful reporting of preferences is a Nash equilibrium if and only if  $\alpha = \lambda$ . The latter condition is natural and admits several interpretations in most real-life settings, especially those that use tie-break rules to resolve situations of social incomparability. It says that an individual who is indifferent between two candidates  $a$  and  $b$  is equally satisfied whether the society strictly ranks them or is unable to rank them. Under the class of voting rules we analyze, two candidates are socially incomparable if neither defeats the other. When social incomparability occurs in real-life politics, one candidate is generally selected using a probabilistic or a non-probabilistic tie-break rule.<sup>6</sup> When such rules are used to resolve social incomparability, the condition  $\alpha = \lambda$  becomes natural because an individual who is indifferent between two candidates should be equally satisfied whether the society considers one of them superior if the vote is decisive or by applying a tie-break rule when the vote does not yield a clear winner; this is because in both situations, the two candidates are ranked, and although the rankings may be different, this does not matter to an individual who is indifferent.<sup>7</sup>

The fact that the way in which social incomparability is resolved in practice affects neither the validity nor the interpretation of the condition  $\alpha = \lambda$  means that our result is robust across different tournament formats. Some tournaments use random rules to resolve social incomparability, whereas others use the alphabetical tie-break rule, and still others use the status quo rule. Some institutions may also use a combination of these rules. A policy implication is that the adoption of a tie-break rule in the design of political tournaments is compatible with truth-telling. Since the nature of the tie-break rule does not matter, a wide range of tournament formats can be accommodated.

The aforementioned result assumes that  $d_{\alpha,\beta,\lambda}$  is a *metric*. Assuming that  $d_{\alpha,\beta,\lambda}$  is a metric may indeed be appealing if one is concerned about some form of consistency in the expression of utility. But we show that our finding continues to hold even when  $d_{\alpha,\beta,\lambda}$  is not a metric (Theorem 2). This is true because the proof of Theorem 1 does not rely on the fact that  $d_{\alpha,\beta,\lambda}$  is a metric. This result means that truthful voting is a Nash equilibrium if only if  $\alpha = \lambda$ , wherein the parameters  $(\alpha, \beta, \lambda)$  belong to an even larger set than the set implied by Theorem 1. This allows for a broader interpretation of our result. In particular, if the property known as “triangle inequality” can be viewed as a form of consistency in the expression of hyper-preferences (just like transitivity is for basic preferences), then our finding holds even when voters do not display such consistency. Indeed, there is a *behavioral* interpretation of the condition  $\alpha = \lambda$  that is consistent with a voter feeling an equal level of disappointment whenever the social ranking disagrees with his individual rankings of candidates. For such a voter,  $\alpha = \lambda = \beta$  (note that  $d_{\alpha,\beta,\lambda}$  in this case is not a metric). We view the fact that our results can be interpreted to accommodate both structural and behavioral interpretations as an appeal of our analysis.

We extend this result to a sequential setting.<sup>8</sup> In this setting, pairwise contests between candidates are

---

In statistics, it is used to define measures of rank correlation (Kendall (1938), Kruskal (1958)). In utility theory, it is used as a classical measure of hyper-preferences. The importance of the Kemeny distance in formal political theory is explained by the fact one is usually interested in how an individual’s rankings of policy alternatives is far or close to the social ranking.

<sup>6</sup>For instance, the United Nations Security Council uses a tie-break rule that consists of just keeping the status quo. In fact, in this institution, the incumbent policy is replaced with a new policy only if the latter policy defeats the former in the vote.

<sup>7</sup>To be more concrete, if candidate  $b$  defeats candidate  $a$  in the pairwise contest opposing them, then the society ranks  $b$  above  $a$ ; however, if neither candidate defeats the other and the society uses a tie-break rule to rank them, say the alphabetical tie-break rule,  $a$  will be ranked above  $b$ . A voter who is indifferent between the two candidates should be equally satisfied in both cases as the ranking does not matter given his indifference.

<sup>8</sup>We thank a referee for suggesting this setting.

conducted sequentially, and when a pair of candidates is presented to voters, they vote simultaneously over this pair. We still find that truthful voting is a Nash equilibrium if and only if  $\alpha = \lambda$  (Theorem 3). This result again proves the robustness of our main finding.<sup>9</sup> In all, the findings do not depend on whether pairwise contests are conducted sequentially or simultaneously. This again suggests a wide range of formats in the design of a political tournament that is compatible with truth-telling.

Building on the analysis above, as our second goal, we study the concept of *effective power* and investigate structural and psychological conditions under which the classical theories of political power capture this notion in a tournament. A voter’s effective power is his ability to bring about a tournament ranking that maximizes his preferences. Diffo Lambo and Moulen (2000) propose a binary relation, that we call the *effective power relation*, to formalize the concept of effective power. According to this relation, a voter  $i$  is said to be at least as *effective* as another voter  $j$  if for every preference profile, the effect of swapping preferences, so that  $i$  now has the preference relation of  $j$  and vice-versa and the preferences of other voters are unchanged, is that the post-swap social preference (or tournament ranking) is at least as close to  $j$ ’s original preferences as it was before the swap. In other words, if given the same conditions as  $j$  with respect to preferences,  $i$  will do better than  $j$  in terms of minimizing his dissatisfaction with respect to the tournament ranking.

We analyze the ordinal properties of the effective power relation in order to evaluate its internal consistency. This analysis uncovers a large class of voting rules and utility parameters for which this binary relation is complete and transitive. We show that, for  $\beta < 1$ , the effective power relation is complete and transitive if and only if the voting rule is swap-robust (or linear) (Theorem 5). The notion of swap-robustness is an extension to voting rules with abstention of a similar notion defined for the class of voting rules known as simple games (Taylor and Zwicker (1999)). The condition that  $\beta < 1$  is natural in that it captures the fact that a voter’s level of dissatisfaction is not the highest if that voter strictly ranks two candidates and the social ranking is unable to compare them. This finding implies that the effective power relation yields a complete ranking of voters in most real-life institutions. In fact, the voting rules used in most prominent institutions and legislative bodies (such as the United Nations Security Council, the United States Senate, the International Monetary Fund, the World Bank, the European Union, etc.) and in several large elections are swap-robust. The class of swap-robust voting rules also includes symmetric (or anonymous) and non-symmetric quota rules.

Finally, addressing the question of whether the classical theories of voting power translate into *effective power* leads to interesting characterizations. The classical theories we analyze are the influence relation (Isbell (1958), Taylor and Zwicker (1999)), the power index proposed by Shapley and Shubik (1954), and the power index proposed by Banzhaf (1965). These notions were initially defined for binary voting rules, where voters can only vote yes and no on a proposal, and the outcome is the adoption of the proposal if it is supported by a winning coalition or its failure otherwise. A voter  $i$  is said to be more influential than another voter  $j$  under a given voting rule if whenever  $j$  can turn a losing coalition into a winning coalition by joining it,  $i$  can do the same *ceteris paribus*. The Shapley-Shubik power index and the Banzhaf power index measure how likely a voter can transform a losing coalition into a winning coalition by joining it, although they differ in their mathematical formulation. These theories of power have since been generalized to voting rules with abstention (see, e.g., Felsenthal and Machover (1997, 1998), Tchantcho et al. (2008), Pongou et al. (2011)), and have been compared to one another in this large class of voting rules in Tchantcho et al. (2008), Pongou et al. (2011), Parker (2012), and Pongou et al. (2014).

We find that the classical theories of power do not translate into effective power in general. We provide a full characterization of the classes of voting rules and dissatisfaction functions for which they capture effective power (Theorem 4 and Theorem 6). In particular, we find that the influence relation traduces the ability of a voter to secure a social outcome that maximizes his preferences if and only if  $\beta < 1$  (Theorem 4). The

---

<sup>9</sup>Our result also goes through if the election is conducted as ranked voting where all candidates are presented simultaneously and voters submit a ballot consisting of their rankings of the candidates.

Shapley-Shubik and Banzhaf power indices measure such an ability if and only if  $\beta < 1$  and the voting rule is swap-robust (Theorem 6). It follows that more stringent conditions should be met for the latter theories to measure effective power. These characterizations reveal that classical notions of voting power translate into the ability to affect voting outcomes only under certain structural and behavioral (or psychological) conditions. They have practical implications for how to design voting rules under which a priori voting rights effectively translate into such an ability.

To our knowledge, our study is the first to analyze voter behavior and effective power in a round-robin political tournament and derive conditions under which truth-telling is a Nash equilibrium. We view our paper as contributing to two distinct literatures. Our characterization of conditions under which truthful voting is a Nash equilibrium in a political tournament can be viewed as a *possibility* result. This result stands in contrast to classical impossibility theorems such as those obtained by Arrow (1951), Gibbard (1973), and Satterthwaite (1975). A literature has developed arguing that some of the assumptions underlying the results in these pioneering works are too strong. As a result, possibility theorems have been obtained after relaxing some of these conditions (see, e.g., Maskin (2020), Bahel and Sprumont (2020), Barbie et al. (2006)). Our finding can be viewed as contributing to this latter literature. However, our setting and the questions we address are very different. Indeed, all of our findings are new. These findings also have clear implications for the design of political tournaments and ranked voting that are compatible with truth-telling, and where the distribution of power is aimed at capturing the ability to affect social outcomes.

Our paper is also related to the literature on political power. Guemmegne and Pongou (2014) analyze the structure of power in the class of (3,3) political games, but they do not study their equilibrium properties. Laruelle and Valenciano (2005) propose a model for measuring a voter’s success and decisiveness in yes-no voting games and assess different power indices on the basis of this model (see also Laruelle et al. (2006)). Closer to our work is Diffo Lambo et al. (2012) who introduce the notion of locally generated metrics which they use to study the ordinal equivalence between the effective power relation and the classical notions of voting power. They however restrict their analysis to yes-no voting games and they do not study equilibrium existence. Moreover, we find that the locally generated functions need not be a metric for our findings to hold. Freixas and Parker (2015) study the existence of Nash equilibrium in games with multiple levels of approval in the input and the output. Our paper differs in its setup and scope. To our knowledge, our paper is the first to formalize a political tournament as a strategic non-cooperative game and to study truth-telling behavior. We identify preference-based conditions under which truthful voting is a Nash equilibrium in the class of voting games with abstention. Moreover, we uncover preference-based as well as structural conditions under which the classical power notions effectively measure the ability to minimize voter dissatisfaction in a political tournament. In this sense, our analysis can be viewed as an attempt to reconcile different approaches to political power. All of our results are new.

The rest of this paper is organized as follows. Section 2 introduces the notion of a round-robin political tournament. Section 3 formalizes a round-robin political tournament as a strategic form game. Section 4 provides a necessary and sufficient condition for truth-telling to be a Nash equilibrium. Section 5 studies the effective power relation and provides necessary and sufficient conditions under which the classical theories of political power translate into effective power. Section 6 concludes. For clarity in the exposition, all the proofs are collected in the appendix.

## 2 Round-robin political tournaments: Definition

A round-robin political tournament is a series of pairwise voting contests among a finite set of candidates taking place under a fixed voting rule. It is modeled as  $(N, A, \mathcal{P}_2(A), (\succeq_i)_{i \in N}, V, D)$ , where  $N = \{1, 2, \dots, n\}$  is a finite set of voters,  $A$  is a finite set of candidates,  $\mathcal{P}_2(A)$  is the set of all the pairs of elements of  $A$

representing the pairwise contests,  $(\succeq_i)_{i \in N}$  is a profile of voters' preferences over candidates,  $V$  is a voting rule, and  $D$  is an aggregation rule that combines individual votes to produce the tournament ranking of candidates. We define each of these ingredients of the model below. These preliminary definitions provide the basis for modeling a round-robin political tournament as a strategic form game and deriving a necessary and sufficient condition under truthful voting is a Nash equilibrium.

## 2.1 Individual preferences

In a round-robin political tournament, each voter has preferences over candidates. A voter  $i$ 's preferences are represented by a binary relation  $\succeq_i$  over  $A$ , where  $\succeq_i$  is assumed to be complete, reflexive and transitive.<sup>10</sup> In a contest opposing two candidates, a voter who expresses truthful preferences votes in favor of one of the candidates or abstain. The asymmetric and the symmetric components of  $\succeq_i$  are denoted by  $\succ_i$  and  $\sim_i$ , respectively. A voter  $i$ 's preference relation will also be denoted  $R^i$ . A preference profile is a collection of individual preferences denoted  $(\succeq_i)_{i \in N}$  or  $R = (R^i)_{i \in N}$ .

## 2.2 Pairwise contests

In a round-robin political tournament, each contest opposes two distinct candidates. Let  $A = \{x_1, x_2, \dots, x_m\}$  be the set of candidates. A pair  $\{x_p, x_q\}$  represents a pairwise contest between candidates  $x_p$  and  $x_q$ . We denote by  $\mathcal{P}_2(A) = \{\{x_p, x_q\} : p, q \in \{1, 2, \dots, m\} \text{ and } p < q\}$  the set of all the contests. Remark that the elements of  $\mathcal{P}_2(A)$  can be listed as follows:

$$\mathcal{P}_2(A) : \begin{cases} \{x_1, x_2\}, \{x_1, x_3\}, \dots, \{x_1, x_m\}, \\ \{x_2, x_3\}, \dots, \{x_2, x_m\}, \\ \dots \\ \{x_{m-2}, x_{m-1}\}, \{x_{m-2}, x_m\}, \\ \{x_{m-1}, x_m\}. \end{cases}$$

It follows that there are  $(m-1) + (m-2) + \dots + 2 + 1 = \frac{m(m-1)}{2}$  pairwise contests. Noting that  $R^i$  is voter  $i$ 's preference relation over  $A$ , we denote by  $R_{\{x_p, x_q\}}^i$  the restriction of  $R^i$  to a pair of candidates  $\{x_p, x_q\}$ . If  $x_p \sim_i x_q$ , this is also denoted by  $R_{\{x_p, x_q\}}^i = (x_p x_q)$ , and if  $x_p \succ_i x_q$ , this is also denoted by  $R_{\{x_p, x_q\}}^i = x_p x_q$ . Thus,  $R_{\{x_p, x_q\}}^i \in \{(x_p x_q), x_p x_q, x_q x_p\}$ .

## 2.3 Voting rule

In a round-robin political tournament, all contests take place under a fixed voting rule. When invited to choose between two candidates  $a$  and  $b$  in a local contest, each voter votes in favor of  $a$  or in favor of  $b$ , or abstains. Let  $S_1$  be the set of voters who vote in favor of  $a$ ,  $S_2$  the set of voters who abstain, and  $S_3$  the set of voters who vote in favor of  $b$ . The vote profile is therefore  $S = (S_1, S_2, S_3)$ , where the sets  $S_1$ ,  $S_2$ , and  $S_3$  are mutually disjoint and their union is  $N$ .  $S$  is called a tripartition of  $N$  and can be viewed as a function  $S$  from  $N$  to  $\{-1, 0, 1\}$  such that  $S_1 = S^{-1}\{1\}$ ,  $S_2 = S^{-1}\{0\}$  and  $S_3 = S^{-1}\{-1\}$ . We denote by  $3^N$  the set of all tripartitions of  $N$ . For  $S, S' \in 3^N$ , we write  $S \subseteq^3 S'$  if  $S$  can be transformed into  $S'$  by shifting one or more voters to higher levels of approval, and  $S \subset^3 S'$  if  $S \subseteq^3 S'$  and  $S \neq S'$ .

The outcome of any local contest  $\{x_p, x_q\}$  is determined by a voting rule that aggregates voters' opinions over the two candidates. A voting rule is a distribution of decisive power over tripartitions.

**Definition 1.** A voting rule for  $N$  is a function  $V : 3^N \rightarrow \{0, 1\}$  such that:

- 1) For all tripartitions  $S, S' \in 3^N$  such that  $S \subseteq^3 S'$ , if  $V(S) = 1$ , then  $V(S') = 1$ ;

<sup>10</sup>Preference transitivity is assumed to be consistent with the general view in the literature that this property expresses individual rationality and consistency in decision making. However, it is not needed for the results uncovered in this paper to hold.

- 2)  $V(N, \emptyset, \emptyset) = 1$ ; and  
 3)  $V(S_3, S_2, S_1) = 0$  whenever  $V(S_1, S_2, S_3) = 1$ .

A tripartition  $S$  is said to be winning under a voting rule  $V$  if  $V(S) = 1$  and losing if  $V(S) = 0$ . A winning tripartition is said to be *minimal* if whenever a voter decreases his approval level, the resulting tripartition is losing. More formally, a winning tripartition  $S$  is minimal under a voting rule  $V$  if  $V(S') = 0$  for all tripartition  $S'$  such that  $S' \subset^3 S$ . We denote by  $W$  the set of winning tripartitions and by  $W_m$  the set of all minimal winning tripartitions under a voting rule  $V$ .

In Definition 1, item (1) is a monotonicity condition which says that if a tripartition is winning, any tripartition that results from any voter increasing his approval level is winning. Item (2) means that if all voters vote for a candidate, that candidate will win against his opponent. And item (3) means that if a tripartition is winning, its complementary is losing. This latter condition rules out the possibility of two winning disjoint tripartitions supporting different candidates and creating unnecessary social instability. Items (1) and (3) can be found in Rubinstein (1980).

A voting rule as defined above is also called a voting rule with abstention. It generalizes the class of voting rules known as *simple games*, where voters can only vote yes and no. It is therefore a very large class of voting rules generalizing majority rule (Dasgupta and Maskin (2008, 2020)) and other majority-like rules (Pongou et al. (2008)). The class of voting rules with abstention includes both *anonymous* rules (Freixas and Zwicker (2009)) and *non-anonymous* rules (Guemmegne and Pongou (2014)). Rules in this class have been widely studied in the literature. They are a subclass of ternary voting games (Felsenthal and Machover (1997)), (3,2) simple games (Freixas and Zwicker (2003, 2009)), and (3,3) simple games (Guemmegne and Pongou (2014)).

### 2.3.1 Tournament ranking

Let  $N$  be a set of voters,  $A = \{x_1, x_2, \dots, x_m\}$  a set of candidates, and  $\mathcal{P}_2(A) = \{\{x_p, x_q\} : p, q \in \{1, 2, \dots, m\}\}$  the set of pairwise contests over  $A$ . How do we obtain the tournament (or social) ranking under a voting rule  $W$ ? In each pairwise contest, each voter votes in favor of one of the competing candidates or abstains. A strategy for a voter  $i$  is a sequence  $z_i = (z_i^{pq})_{1 \leq p < q \leq m}$  where  $z_i^{pq} \in \{-1, 0, 1\}$  is the vote cast by  $i$  in the contest opposing  $x_p$  and  $x_q$ ;  $z_i^{pq} = -1$  means that  $i$  has voted for  $x_q$  against  $x_p$ ;  $z_i^{pq} = 1$  means that  $i$  has voted for  $x_p$  against  $x_q$ ; and  $z_i^{pq} = 0$  means that  $i$  has abstained between  $x_p$  and  $x_q$ . A vote profile is a sequence  $z = (z_i)_{i \in N}$ .

Let  $z = (z_i)_{i \in N}$  be a vote profile. For all  $p, q$  with  $p < q$ , we denote by  $N^1(z^{pq}) = \{i \in N : z_i^{pq} = 1\}$  the set of voters who have voted for  $x_p$  against  $x_q$ , by  $N^{-1}(z^{pq}) = \{i \in N : z_i^{pq} = -1\}$  the set of voters who have voted for  $x_q$  against  $x_p$ , and by  $N^0(z^{pq}) = \{i \in N : z_i^{pq} = 0\}$  the set of voters who have abstained between  $x_p$  and  $x_q$ . In general, for any pair  $x, y$  of candidates and any preference profile  $R$ , we denote by  $N(x, y, R) = \{i \in N : R_{\{x, y\}}^i = xy\}$  the set of voters who strictly prefer  $x$  to  $y$  at the profile  $R$  and by  $N^0(x, y, R) = \{i \in N : R_{\{x, y\}}^i = (xy)\}$  the set of voters who are indifferent between  $x$  and  $y$  at  $R$ .

The tournament ranking under the voting rule  $W$  induced by a vote profile  $z$ , which we denote by  $D(z)$ , is given by:

$$x_p x_q \in D(z) \text{ if and only if } (N^1(z^{pq}), N^0(z^{pq}), N^{-1}(z^{pq})) \in W;$$

$$x_p * x_q \in D(z) \text{ if and only if } (N^1(z^{pq}), N^0(z^{pq}), N^{-1}(z^{pq})) \notin W \text{ and } (N^1(z^{qp}), N^0(z^{qp}), N^{-1}(z^{qp})) \notin W$$

where  $x_p x_q \in D(z)$  means that  $x_p$  beats  $x_q$  in the pairwise contest opposing them and so is ranked above  $x_q$  in the tournament ranking  $D(z)$ ; and  $x_p * x_q \in D(z)$  means that  $x_p$  and  $x_q$  are socially incomparable, which

means that neither of the two candidates beats the other in the tournament. Remark that under our voting rules, the social ranking cannot be indifferent between two candidates.<sup>11</sup> In the sequel,  $x_p x_q \in D(z)$  will also be denoted by  $x_p \succ_{D(z)} x_q$  or  $D_{\{x_p, x_q\}}(z) = x_p x_q$ .

The following example illustrates a round-robin political tournament.

**Example.** Consider a political tournament where  $N = \{1, 2, 3\}$ ,  $A = \{x_1, x_2, x_3\}$ , and the voting rule is simple majority rule. Under this rule, a tripartition  $(S_1, S_2, S_3)$  is winning if and only if  $|S_1| > |S_3|$ . Voters' preferences are defined as follows:  $R = (x_3 x_1 x_2; x_2(x_3 x_1); x_1(x_2 x_3))$

It can be checked that the tournament ranking over the three pairs  $\{x_1, x_2\}$ ,  $\{x_1, x_3\}$  and  $\{x_2, x_3\}$  is given by :  $D(R) = \{x_1 x_2, x_1 * x_3, x_2 * x_3\}$

It follows that  $x_1$  is ranked above  $x_2$  in the tournament, but  $x_1$  and  $x_3$  as well as  $x_2$  and  $x_3$  are socially incomparable.

### 3 A round-robin political tournament as a strategic form game

In this section, we model a round-robin political tournament  $(N, A, \mathcal{P}_2(A), (\succeq_i)_{i \in N}, V, D)$  as a strategic form game. A strategic form game is a tuple consisting of a set of players, a set of strategies for each player, and a utility (or payoff) function that maps each vector of strategies to a real number for each player.

#### 3.1 Utility function

In a round-robin political tournament, the utility that a voter derives from a strategy profile  $z$  is measured by how much that voter is satisfied with the tournament (or social) ranking  $D(z)$  induced by  $z$ . A voter  $i$ 's dissatisfaction with the social ranking  $D(z)$  induced by  $z$  is measured by how *distant* his preferences  $R^i$  are to  $D(z)$ . The distance is measured using a class of locally generated metrics defined on the set of binary relations. This class of metrics, which generalizes the well-known Kemeny distance, was introduced in Diffo Lambo et al. (2012).<sup>12</sup> We recall below the definition of a locally generated metric in Diffo Lambo et al. (2012).

Let  $\mathcal{R}_1$  and  $\mathcal{R}_2$  be two binary relations and  $\{a, b\} \in \mathcal{P}_2(A)$ . The distance between  $\mathcal{R}_1$  and  $\mathcal{R}_2$  over  $\{a, b\}$  is measured by a function  $d_{\alpha, \beta, \lambda}^{\{a, b\}}(\mathcal{R}_1, \mathcal{R}_2)$ ,  $\alpha, \beta, \lambda \in [0, 1]$ , defined in the table below, where  $a \succ b$  means that  $a$  is better than  $b$ ,  $a \simeq b$  means that  $a$  and  $b$  are indifferent, and  $a * b$  means that  $a$  and  $b$  are incomparable.

$\mathcal{R}_1$	$a \succ b$	$a \succ b$	$a \succ b$	$a \simeq b$	$a \simeq b$	$a \simeq b$	$a * b$
$\mathcal{R}_2$	$a \succ b$	$b \succ a$	$a * b$	$a \simeq b$	$a \succ b$	$a * b$	$a * b$
$d_{\alpha, \beta, \lambda}^{\{a, b\}}(\mathcal{R}_1, \mathcal{R}_2)$	0	1	$\beta$	0	$\alpha$	$\lambda$	0

Following the property of symmetry defining a metric, the table above is completed by letting  $d_{\alpha, \beta, \lambda}^{\{a, b\}}(\mathcal{R}_1, \mathcal{R}_2) = d_{\alpha, \beta, \lambda}^{\{a, b\}}(\mathcal{R}_2, \mathcal{R}_1)$ . The distance between  $\mathcal{R}_1$  and  $\mathcal{R}_2$ , denoted  $d_{\alpha, \beta, \lambda}(\mathcal{R}_1, \mathcal{R}_2)$ , is obtained by summing up  $d_{\alpha, \beta, \lambda}^{\{a, b\}}(\mathcal{R}_1, \mathcal{R}_2)$  over all the pairs  $\{a, b\} \in \mathcal{P}_2(A)$  :  $d_{\alpha, \beta, \lambda}(\mathcal{R}_1, \mathcal{R}_2) = \sum_{\{a, b\} \in \mathcal{P}_2(A)} d_{\alpha, \beta, \lambda}^{\{a, b\}}(\mathcal{R}_1, \mathcal{R}_2)$ .

Diffo Lambo et al. (2012) show that the mapping  $d_{\alpha, \beta, \lambda}$  is a *metric* if and only if the following property  $\Delta_{\alpha, \beta, \lambda}$  is satisfied:

$$\Delta_{\alpha, \beta, \lambda} : \begin{cases} \alpha + \lambda \geq \beta \geq \frac{1}{2} \\ \beta + \lambda \geq \alpha \geq \frac{1}{2}. \end{cases}$$

<sup>11</sup>In fact, social indifference between two candidates  $x_p$  and  $x_q$  would mean that  $x_p$  beats  $x_q$  and  $x_q$  beats  $x_p$  in the pairwise contest opposing them. This is impossible because of the third property of a voting rule in Definition 1. In fact,  $x_p$  beats  $x_q$  under a voting rule  $W$  if  $(N^1(z^{pq}), N^0(z^{pq}), N^{-1}(z^{pq})) \in W$  and  $x_q$  beats  $x_p$  under the same rule  $W$  if  $(N^1(z^{qp}), N^0(z^{qp}), N^{-1}(z^{qp})) \in W$ ; but it is impossible to have  $(N^1(z^{pq}), N^0(z^{pq}), N^{-1}(z^{pq})) \in W$  and  $(N^1(z^{qp}), N^0(z^{qp}), N^{-1}(z^{qp})) \in W$  simultaneously because of item (3) of Definition 1.

<sup>12</sup>This class generalizes the well-known Kemeny distance to an environment where alternatives can be equivalent or incomparable. The Kemeny distance is defined on the set of complete, reflexive, asymmetric, and transitive binary relations. The metrics we use are defined on a much larger set of binary relations.

To define a voter's utility function in a political tournament, we apply the *restriction* of these metrics to the subset of ordered pairs of binary relations  $(\mathcal{R}_1, \mathcal{R}_2)$  where  $\mathcal{R}_1$  is complete and transitive and  $\mathcal{R}_2$  does not admit ties. If  $\mathcal{R}_1$  represents an individual's preference relation and  $\mathcal{R}_2$  the social ranking,  $d_{\alpha, \beta, \lambda}^{\{a, b\}}(\mathcal{R}_1, \mathcal{R}_2)$  is interpreted as the level of dissatisfaction of that individual with respect to the social preference over  $\{a, b\}$ . Overall dissatisfaction with respect to  $\mathcal{R}_2$  is obtained by summing all of the local measures of dissatisfaction  $d_{\alpha, \beta, \lambda}^{\{a, b\}}(\mathcal{R}_1, \mathcal{R}_2)$  over all the pairs  $\{a, b\} \in \mathcal{P}_2(A)$ . A particular appeal of the class of metrics  $d_{\alpha, \beta, \lambda}$  therefore stems from the fact that it is possible to measure a voter's dissatisfaction with a tournament ranking  $D(z)$  even if  $D(z)$  is not complete or transitive.

Given a metric  $d_{\alpha, \beta, \lambda}$ , we measure the utility that a voter  $i$  derives from a strategy profile  $z = (z_1, \dots, z_n)$  by the function  $U_i^d(z) = \sum_{p < q} U_i^{pq}(z)$ , where  $U_i^{pq}(z) = 1 - d_{\alpha, \beta, \lambda}^{\{x_p, x_q\}}(R^i, D(z))$ . In fact, given that  $d_{\alpha, \beta, \lambda}^{\{x_p, x_q\}}(R^i, D(z))$  is the level of dissatisfaction (or disutility) that  $i$  derives from the tournament ranking of candidates  $x_p$  and  $x_q$ ,  $U_i^{pq}(z) = 1 - d_{\alpha, \beta, \lambda}^{\{x_p, x_q\}}(R^i, D(z))$  can be viewed as the utility that he derives from that ranking.<sup>13</sup> It logically follows, by summing up utilities over all the pairs of candidates, that  $U_i^d(z) = \sum_{p < q} U_i^{pq}(z)$  is the utility that  $i$  derives from the tournament ranking  $D(z)$ .<sup>14</sup>

### 3.2 Strategic form game

We now model a round-robin political tournament as a strategic form game. Assume that in a tournament  $(N, A, \mathcal{P}_2(A), (\succeq_i)_{i \in N}, V, D)$  (or  $(N, A, \mathcal{P}_2(A), R = (R^i)_{i \in N}, W, D)$ ), each voter  $i$ 's utility is measured by a utility function  $U_i^d$  as defined in the preceding section. We associate to this tournament the strategic form game  $\Gamma_R = (N, (Z_i), (U_i^d)) = (N, Z, U)$  where:

- for all  $i \in N$ ,  $Z_i = \{z_i = (z_i^{pq})_{1 \leq p < q \leq m}\}$ , with  $z_i^{pq} \in \{-1, 0, 1\}$ , is the strategy set of  $i$ ; and
- for all  $i \in N$ ,  $U_i^d(z) = \sum_{p < q} U_i^{pq}(z)$ , with  $z = (z_1, \dots, z_n)$  and  $U_i^{pq}(z) = 1 - d_{\alpha, \beta, \lambda}^{\{x_p, x_q\}}(R^i, D(z))$ , is the payoff function of  $i$ .

## 4 When is truthful voting a Nash equilibrium?

In this section, we provide a necessary and sufficient condition for truthful voting to be a Nash equilibrium of a political tournament game  $\Gamma_R = (N, (Z_i), (U_i^d))$ . Define by  $s^*(R)$  the profile of truthful votes as follows:

$$\text{For all voter } i, s_i^*(R) = (s_i^{*pq}(R))_{1 \leq p < q \leq m} \text{ with } s_i^{*pq}(R) = \begin{cases} 1 & \text{if } R_{\{x_p, x_q\}}^i = x_p x_q \\ 0 & \text{if } R_{\{x_p, x_q\}}^i = (x_p x_q) \\ -1 & \text{if } R_{\{x_p, x_q\}}^i = x_q x_p. \end{cases}$$

We provide below a full characterization of the parameters  $\alpha, \beta, \lambda$  for which truthful voting is a Nash equilibrium.

**Theorem 1.** *Let  $d_{\alpha, \beta, \lambda}$  be a metric with  $\alpha, \beta, \lambda \in ]0, 1]$ . Given  $N$  and  $A$  with  $|N| \geq 3$  and  $|A| \geq 2$ , a necessary and sufficient condition under which for all voting rule  $W$ , all preference profile  $R$ , the truthful vote profile  $s^*(R)$  is a Nash equilibrium for the associated strategic form game  $\Gamma_R = (N, Z, U)$  is that  $\alpha = \lambda$ .*

<sup>13</sup>Note that all our findings go through if we replace  $1 - d_{\alpha, \beta, \lambda}^{\{x_p, x_q\}}(R^i, D(z))$  by  $-d_{\alpha, \beta, \lambda}^{\{x_p, x_q\}}(R^i, D(z))$  or  $\frac{1}{1 + d_{\alpha, \beta, \lambda}^{\{x_p, x_q\}}(R^i, D(z))}$  or any decreasing function of  $d_{\alpha, \beta, \lambda}^{\{x_p, x_q\}}(R^i, D(z))$ . The idea here is simply to express the notion that a smaller value of  $d_{\alpha, \beta, \lambda}^{\{x_p, x_q\}}(R^i, D(z))$  increases voter  $i$ 's satisfaction with the social ranking of  $x_p$  and  $x_q$ .

<sup>14</sup>We remark that this definition of the utility function assumes that each voter cares equally about each comparison of candidates. Our results are still valid if each comparison receives a different weight. That is, our results are maintained if the utility function of a voter  $i$  is defined as  $U_i^d(z) = \sum_{p < q} \theta_i^{pq} U_i^{pq}(z)$ , where  $\theta_i^{pq} > 0$  measures the importance that  $i$  places on the comparison of the candidates  $x_p$  and  $x_q$ .

For clarity, the proofs of this Theorem 1 and of all the subsequent results are provided in the appendix.

Theorem 1 says that truthful voting is a Nash equilibrium for all voting rule and all preference profile if, and only if, whenever a voter is indifferent between two candidates  $a$  and  $b$ , he is equally satisfied whether the society strictly ranks these candidates or is unable to compare them ( $\alpha = \lambda$ ). Note that although we provide a proof that holds for any number of candidates, an alternative proof could be obtained by just focusing on two candidates and then use the fact that utilities over all pairs of candidates are separable to generalize the argument to any number of candidates.

Regarding the interpretation of the condition  $\alpha = \lambda$ , we note that it is a natural condition in a large majority of real-life settings, especially in those settings that use a tie-break rule to select a winner when two candidates are socially incomparable. Indeed, in most institutions, a situation in which neither candidate defeats the other in a pairwise contest is treated as a tie. When there is a tie, one candidate is generally selected using a probabilistic or a non-probabilistic tie-break rule. When such rules are used to resolve social incomparability, the condition  $\alpha = \lambda$  becomes natural because an individual who is indifferent between two candidates should be equally satisfied whether the society considers one of them superior following the vote or by applying a tie-break rule when the vote does not yield a clear winner; this is because in both situations, the two candidates are ranked, and although the rankings may be different, this does not matter to an individual who is indifferent. To be more concrete, if candidate  $b$  defeats candidate  $a$  in the pairwise contest opposing them, then the society ranks  $b$  above  $a$ ; however, if neither candidate wins, and the society uses a tie-break rule such as the alphabetical tie-break rule,  $a$  will be ranked above  $b$ . Note that a voter who is indifferent between the two candidates is equally satisfied in both cases as the ranking does not matter given his indifference. As argued in the Introduction, the fact that the way in which social incomparability is resolved in practice affects neither the validity nor the interpretation of the condition  $\alpha = \lambda$  means that our result is robust across different tournament formats.<sup>15</sup>

Other interpretations of the condition  $\alpha = \lambda$  are possible. For example, there is a behavioral interpretation of this condition that is consistent with a voter feeling an equal level of disappointment whenever the social ranking disagrees with his individual rankings of candidates. For such a voter, we have  $\alpha = \lambda = \beta$ . Remark that  $d_{\alpha,\beta,\lambda}$  in this case is not a metric as  $\Delta_{\alpha,\beta,\lambda}$  is violated; but Theorem 2 states that  $d_{\alpha,\beta,\lambda}$  need not be a metric for our result to hold. We view the fact that our results can accommodate different interpretations as an appeal of our analysis.

The example below illustrates Theorem 1 and provides the intuition underlying its proof.

**Example.** Consider a political tournament where  $N = \{1, 2, 3\}$ ,  $A = \{x_1, x_2\}$ , and the voting rule is simple majority rule, where a tripartition  $(S_1, S_2, S_3)$  is winning if and only if  $|S_1| > |S_3|$ . Denote this rule by  $W$ . Let  $R$  be a preference profile defined as follows:  $R^1 = x_1x_2$ ,  $R^2 = x_2x_1$  and  $R^3 = (x_1x_2)$ . The truthful vote profile is  $S^*(R) = (1, -1, 0)$  and the corresponding social ranking is  $D(S^*(R)) = \{x_1 * x_2\}$ . The utilities of voters from this social ranking at the preference profile  $R$  are :

$$u_1(S^*(R)) = u_2(S^*(R)) = 1 - \beta \text{ and } u_3(S^*(R)) = 1 - \lambda.$$

- If voter 1 deviates from 1 to 0, then  $D((0, -1, 0)) = \{x_2x_1\}$  and his utility is  $u_1((0, -1, 0)(R)) = 0$ . If he deviates from 1 to -1, then  $D((-1, -1, 0)) = \{x_2x_1\}$  and his utility is  $u_1((-1, -1, 0)(R)) = 0$ . Since  $\beta \in ]0, 1]$ , voter 1 has no incentive to deviate from his true preferences.
- For the same reason as voter 1, voter 2 has no incentive to deviate from his true preferences.

---

<sup>15</sup>In results not shown here but available upon request, we prove that truthful voting is always a Nash equilibrium if we explicitly incorporate into the tournament aggregation rule a probabilistic or an alphabetical tie-break rule to resolve situations of social incomparability. This result, which imposes no conditions on the parameters  $\alpha$ ,  $\lambda$ , and  $\beta$ , confirms that our interpretation of the condition  $\alpha = \lambda$  is natural if a voting mechanism uses a probabilistic or an alphabetical tie-break rule to resolve social incomparability in our setting.

- If voter 3 deviates from 0 to 1, then  $D((1, -1, 1)) = \{x_1x_2\}$  and his utility is  $u_3((1, -1, 1)(R)) = 1 - \alpha$ . If he deviates from 0 to -1, then  $D((1, -1, -1)) = \{x_2x_1\}$  and his utility is  $u_3((1, -1, -1)(R)) = 1 - \alpha$ . If  $\lambda > \alpha$ , then  $u_3(S^*(R)) = 1 - \lambda < 1 - \alpha = u_3(S_{-3}^*(R), 1)$ . Hence in this case, voter 3 can deviate from his true preferences to increase his utility. But if  $\lambda = \alpha$  (the condition stated in Theorem 1), he has no such incentive, implying that truthful voting is a Nash equilibrium.

Remark that there are other Nash equilibria in the game defined in the preceding example when  $\lambda = \alpha$ ; such an equilibrium is the vote profile  $(1, 0, -1)$ . It is also worth mentioning that truth-telling is not necessarily socially optimal.

#### 4.1 An extension: When $d_{\alpha,\beta,\lambda}$ is not a metric

Theorem 1 assumes that  $d_{\alpha,\beta,\lambda}$  is a *metric*. This is an appealing assumption if one believes that the properties of a metric embed some form of consistency in the expression of utility or hyper-preferences. But we show below that our finding continues to hold even when  $d_{\alpha,\beta,\lambda}$  is not a metric (Theorem 2). In fact, one can argue that since in the expression  $d_{\alpha,\beta,\lambda}(\mathcal{R}_1, \mathcal{R}_2)$ ,  $\mathcal{R}_1$  (individual preferences) and  $\mathcal{R}_2$  (social ranking) do not have the same properties in general in the context of a political tournament<sup>16</sup>,  $d_{\alpha,\beta,\lambda}$  need not be a metric. Consider the following approach to measuring an individual's dissatisfaction with the social ranking  $D(R)$  of a pair of candidates  $\{a, b\}$  at a preference profile  $R = (R^i)_{i \in N}$ :

Individual preference : $R^i$	$a \succ b$	$a \succ b$	$a \succ b$	$a \preceq b$	$a \preceq b$
Social preference : $D(R)$	$a \succ b$	$b \succ a$	$a * b$	$a \succ b$	$a * b$
Pairwise disagreement $d_{\alpha,\beta,\lambda}^{\{a,b\}}(R^i, D(R))$	0	1	$\beta$	$\alpha$	$\lambda$

The dissatisfaction of a voter  $i$  with preferences  $R^i$ , with respect to the social ranking  $D(R)$  is still given by  $d_{\alpha,\beta,\lambda}(R^i, D(R)) = \sum_{\{a,b\} \in P_2(A)} d_{\alpha,\beta,\lambda}^{\{a,b\}}(R^i, D(R))$ . We have the following result.

**Theorem 2.** : *Let  $\alpha, \beta, \lambda \in ]0; 1]$ . Given  $N$  and  $A$  with  $|N| \geq 3$  and  $|A| \geq 2$ , a necessary and sufficient condition under which for all voting rule  $W$ , all preference profile  $R$ , the truthful vote profile  $s^*(R)$  is a Nash equilibrium for the associated strategic form game  $\Gamma_R = (N, Z, U)$  is that  $\alpha = \lambda$ .*

The proof of this result is similar to that of Theorem 1; this is because the proof of Theorem 1 never really uses the fact that  $d_{\alpha,\beta,\lambda}$  is a metric.

Theorem 2 implies that truthful voting is a Nash equilibrium even when the parameters  $(\alpha, \beta, \lambda = \alpha)$  belong to a larger set than the set assumed in Theorem 1. This allows for a broader interpretation of our result, especially when psychological or behavioral considerations are taken into consideration as argued earlier.

#### 4.2 An extension to a sequential setting

So far, we have assumed that pairwise contests take place simultaneously as in ranked voting. In this section, we assume that the game is played sequentially in the following sense: pairwise contests are conducted sequentially, and when a pair of candidates is presented to voters, they vote simultaneously over that pair. The tournament therefore consists of  $m(m - 1)/2$  sequential pairwise contests. We do not impose any particular order (or agenda) on the contests. The theorem below offers a necessary and sufficient condition for truthful voting to be a Nash equilibrium.

**Theorem 3.** : *Let  $\alpha, \beta, \lambda \in ]0; 1]$ . Given  $N$  and  $A$  with  $|N| \geq 3$  and  $|A| \geq 2$ , assume that pairwise contests are conducted sequentially as described above. A necessary and sufficient condition under which for all voting rule*

<sup>16</sup>For instance, individual preferences are complete whereas social preferences are not complete in general.

$W$ , all preference profile  $R$ , the truthful vote profile  $s^*(R)$  is a Nash equilibrium for the associated strategic form game  $\Gamma_R = (N, Z, U)$  is that  $\alpha = \lambda$ .

The proof is similar to that of Theorem 1. This is true because the proof of Theorem 1 does not depend on the format of the tournament. Note also that although the pairwise contests are sequential, this does not result in a game in extensive form because we do not assume that voters vote sequentially. It follows that the Nash equilibrium is still a relevant equilibrium notion in this setting.

Theorems 1, 2 and 3 show that our findings are robust across a wide range of tournament formats.

## 5 Effective power and the classical theories of political power

In this section, we study the concept of *effective power*, defined as a voter’s ability to bring about a social ranking that maximizes his preferences. Then, building on this study, we answer the question of when the classical theories of political power translate into effective power.

Our notion of effective power is conceptually close to that of “outcome power” (Dowding (1991)), which is the ability of an actor to bring about or help bring about an outcome. It is also consistent with the definition of power given by the 17th century Political Philosopher Thomas Hobbes. Hobbes defines power as an individual’s “present means, to obtain some future apparent good” (Hobbes (1651)). In a voting context, an individual’s effective power can be seen as his ability to secure a social outcome that is as close as possible to his preferences, given the institutional constraints. In our treatment of this problem, voters have voting rights that they derive from a voting rule  $W$ . How a voter’s voting rights translate into his ability to obtain desirable outcomes is formalized in the next section and illustrated in a few examples.

We then build on this analysis to uncover conditions under which the classical notions of political power—the influence relation, the Shapley-Shubik power index, and the Banzhaf power index—translate into effective power. This exercise is useful because the latter notions of power are generally viewed as measures of *a priori* voting power, but we know very little about whether they reflect the actual ability of voters to secure personal advantages in a competitive political context. Our analysis can therefore be seen as an attempt to reconcile different views of political power. In doing so, we provide new “behavioral” characterizations of classical power notions.

### 5.1 The effective power relation

In a political tournament, any voter  $i$  with individual preference relation  $R^i$  uses his “effective power” to secure a collective preference  $D(R)$  that is as close as possible to  $R^i$ . This notion of effective power is formalized in Difo Lambo and Moulen (2000) and Moulen et al. (2012) as follows. Let  $i$  and  $j$  be two voters involved in a tournament. To any preference profile  $R = (R^k)_{k \in N}$ , we assign another profile  $R_{ij} = (R_{ij}^k)_{k \in N}$  defined by:

$$R_{ij}^k = \begin{cases} R^k & \text{if } k \notin \{i, j\} \\ R^j & \text{if } k = i \\ R^i & \text{if } k = j. \end{cases}$$

Put in words, the profile  $R_{ij}$  is obtained from voters  $i$  and  $j$  swapping their preferences in  $R$ . Let  $d_{\alpha, \beta, \lambda}$  be a dissatisfaction function defined over the set of binary relations. Then,  $d_{\alpha, \beta, \lambda}(R^j, D(R))$  measures the dissatisfaction of  $j$  at the preference profile  $R$  and  $d_{\alpha, \beta, \lambda}(R^j, D(R_{ij}))$  measures the dissatisfaction of  $i$  when the preference profile is  $R_{ij}$ . When  $d_{\alpha, \beta, \lambda}(R^j, D(R_{ij})) \leq d_{\alpha, \beta, \lambda}(R^j, D(R))$ , the difference  $d_{\alpha, \beta, \lambda}(R^j, D(R)) - d_{\alpha, \beta, \lambda}(R^j, D(R_{ij}))$  quantifies the extent to which “ $i$  does better than  $j$ ” at the profile  $R$ . Therefore, if  $d_{\alpha, \beta, \lambda}(R^j, D(R_{ij})) \leq d_{\alpha, \beta, \lambda}(R^j, D(R))$  holds for every profile  $R$ , this can be interpreted as “voter  $i$  being at least as effective at influencing the tournament ranking (or social preference) as  $j$ ”. The following definition formalizes the notion of effective power.

**Definition 2.** Let  $N$  be a set of voters,  $W$  a voting rule,  $A$  a set of candidates,  $d_{\alpha,\beta,\lambda}$  a dissatisfaction function defined over the set of binary relations defined over  $A$ , and  $i$  and  $j$  two voters.

1)  $i$  is said to be at least as effective as  $j$  under the rule  $W$  for the function  $d_{\alpha,\beta,\lambda}$ , denoted  $i \geq_{P,d_{\alpha,\beta,\lambda}} j$ , if for every preference profile  $R$ ,  $d_{\alpha,\beta,\lambda}(D(R_{ij}), R^j) \leq d_{\alpha,\beta,\lambda}(D(R), R^j)$ .

2)  $i$  is strictly more effective than  $j$  under the rule  $W$  for the function  $d_{\alpha,\beta,\lambda}$ , denoted  $i >_{P,d_{\alpha,\beta,\lambda}} j$ , if  $i \geq_{P,d_{\alpha,\beta,\lambda}} j$  and  $\text{not}(j \geq_{P,d_{\alpha,\beta,\lambda}} i)$ .

3)  $i$  is as effective as  $j$  under the rule  $W$  for the function  $d_{\alpha,\beta,\lambda}$ , denoted by  $i \sim_{P,d_{\alpha,\beta,\lambda}} j$ , if  $i \geq_{P,d_{\alpha,\beta,\lambda}} j$  and  $j \geq_{P,d_{\alpha,\beta,\lambda}} i$ .

For ease of notation, we will sometimes denote the effective power relation  $\geq_{P,d_{\alpha,\beta,\lambda}}$  simply by  $\geq_{P,\alpha,\beta,\lambda}$ .

## 5.2 Illustrating the effective power relation

In this section, we illustrate the effective power relation using examples from real-life legislative bodies and political institutions.

### 5.2.1 The United Nations Security Council

The United Nations Security Council (UNSC) has 5 permanent members and 10 non-permanent or rotating members. It enforces a resolution if at least nine members support it and no permanent member opposes it. This voting rule can be formalized as follows. Let  $P = \{i_k : k = 1, \dots, 5\}$  be the set of permanent members and  $\bar{P} = \{j_k : k = 1, \dots, 10\}$  the set of non-permanent members. A tripartition  $(S_1, S_2, S_3)$  is winning if and only if:  $|S_1| \geq 9$  and  $P \cap S_3 = \emptyset$ . Denote this voting rule by  $W^u$ .

We would like to rank the UNSC's members using the effective power relation. Let  $A$  be a set of policy alternatives, and  $d_{\alpha,\beta,\lambda}$  a dissatisfaction function with  $\alpha \leq \lambda$ . Let  $a, b \in A$  be two distinct policy alternatives. We claim that:

1. for all  $u \in \{1, \dots, 5\}$  and  $v \in \{1, \dots, 10\}$ ,  $i_u >_{P,d_{\alpha,\beta,\lambda}} j_v$ ;
2. for all  $u, v \in \{1, \dots, 5\}$ ,  $i_u \sim_{P,d_{\alpha,\beta,\lambda}} i_v$ ;
3. for all  $u, v \in \{1, \dots, 10\}$ ,  $j_u \sim_{P,d_{\alpha,\beta,\lambda}} j_v$ .

To show that for all  $u \in \{1, \dots, 5\}$  and  $v \in \{1, \dots, 10\}$ ,  $i_u >_{P,d_{\alpha,\beta,\lambda}} j_v$ , we have to show that:

- (i) for all preference profile  $R = (R^i)_{i \in P \cup \bar{P}}$ ,  $d_{\alpha,\beta,\lambda}(R^{j_u}, D(R_{i_u j_v})) \leq d_{\alpha,\beta,\lambda}(R^{j_u}, D(R))$ ; and
- (ii) there exists at least one preference profile  $R = (R^i)_{i \in P \cup \bar{P}}$  for which  $d_{\alpha,\beta,\lambda}(R^{j_u}, D(R_{i_u j_v})) < d_{\alpha,\beta,\lambda}(R^{j_u}, D(R))$ .

To show (i), let  $R$  be a preference profile and let us assume for simplicity that there are only two alternatives  $a$  and  $b$ .

- If  $d_{\alpha,\beta,\lambda}(R^{j_u}, D(R)) = 1$ , then there is nothing to prove.
- Assume that  $d_{\alpha,\beta,\lambda}(R^{j_u}, D(R)) = \alpha$ .

In this case,  $a \simeq_{R^{j_u}} b$  and  $a \succ_{D(R)} b$ . Then either  $i_u \in N(a, b, R)$  or  $i_u \in N^0(a, b, R)$ . Regardless, under the voting rule  $W^u$ , it is clear that  $|N(a, b, R_{i_u j_v})| \geq 9$  and  $N(b, a, R_{i_u j_v}) \cap P = \emptyset$ ; this implies that  $a \succ_{D(R_{i_u j_v})} b$ , and therefore we have  $d_{\alpha,\beta,\lambda}(R^{j_u}, D(R_{i_u j_v})) = \alpha = d_{\alpha,\beta,\lambda}(R^{j_u}, D(R))$ .

- Assume that  $d_{\alpha,\beta,\lambda}(R^{j_u}, D(R)) = \beta$ .

In this case,  $a \succ_{R^{j_u}} b$  and  $a *_{D(R)} b$ . Then, either  $a \succ_{D(R_{i_u j_v})} b$  or  $a *_{D(R_{i_u j_v})} b$ , and this implies that  $d_{\alpha,\beta,\lambda}(R^{j_u}, D(R_{i_u j_v})) \in \{0, \beta\}$ . Once more, we have  $d_{\alpha,\beta,\lambda}(R^{j_u}, D(R_{i_u j_v})) \leq d_{\alpha,\beta,\lambda}(R^{j_u}, D(R))$ .

- Assume that  $d_{\alpha,\beta,\lambda}(R^{j_u}, D(R)) = \lambda$ .

In this case,  $a \simeq_{R^{j_u}} b$  and  $a *_{D(R)} b$ . Thanks to the definition of the voting rule  $W^u$ , we have  $a \succ_{D(R_{i_u j_v})} b$ , or  $b \succ_{D(R_{i_u j_v})} a$ , or  $a *_{D(R_{i_u j_v})} b$ . Remark that we will have  $a \succ_{D(R_{i_u j_v})} b$ , for example, if  $N(a, b, R)$  contains 9 voters without containing player  $i_u$  who belongs to  $N(b, a, R)$ . In this last case,  $d_{\alpha,\beta,\lambda}(R^{j_u}, D(R_{i_u j_v})) =$

$\alpha \leq \lambda = d_{\alpha,\beta,\lambda}(R^{j_u}, D(R))$ . In the other case where  $a *_{D(R_{i_u j_v})} b$ , we have  $d_{\alpha,\beta,\lambda}(R^{j_u}, D(R_{i_u j_v})) = \lambda$  and the inequality holds.

In all of these cases, we have  $d_{\alpha,\beta,\lambda}(R^{j_u}, D(R_{i_u j_v})) \leq d_{\alpha,\beta,\lambda}(R^{j_u}, D(R))$ , which shows that  $i_u \geq_{P,d_{\alpha,\beta,\lambda}} j_v$ .

To show (ii), let us again assume for simplicity that there are only two alternatives  $a$  and  $b$ . Let  $R$  be the preference profile in which  $R^k = ab$  for all  $k \in P \cup \{j_1, \dots, j_4\}$  and  $R^k = ba$  for  $k \in \{j_5, \dots, j_{10}\}$ .

For this preference profile we have:  $D(R) = ab$ ,  $d_{\alpha,\beta,\lambda}(R^{i_u}, D(R)) = 0$ ,  $D(R_{i_u j_v}) = a * b$  and thus  $d_{\alpha,\beta,\lambda}(R^{i_u}, D(R_{i_u j_v})) = \beta > 0$ , which proves that  $\text{not}(j_v \succ_{P,d_{\alpha,\beta,\lambda}} i_u)$ . We conclude that  $i_u \succ_{P,d_{\alpha,\beta,\lambda}} j_v$ .

Showing (2) and (3) is obvious because same-type voters are interchangeable.

### 5.2.2 The United States Senate

Decisions in the two houses of the United States Congress are governed by a variety of voting rules, depending on the nature of the issue at hand. We will focus on the rule used to pass an ordinary bill or an amendment in the US Senate. The US Senate has 100 senators and the US vice-president. The US vice-president leads the Senate and casts tie-break votes in certain circumstances. Passing a bill requires a simple (or strict) majority of those who vote ‘yes’; when the number of ‘yes’ and ‘no’ votes is equal, the vice president casts a tie-breaking vote.<sup>17</sup> For example, if 81 senators abstain or are absent, 10 senators vote in favor of a bill, and 9 senators vote against it, the bill passes regardless of whether or not the vice-president is opposed to it.

This voting rule can be modeled as follows:

$N = \{i_k : k = 1, \dots, 100\} \cup \{v\}$  where  $i_k$  is senator  $k$  and  $v$  is the vice-president. A tripartition  $(S_1, S_2, S_3)$  is winning if and only if:  $|S_1| \geq |S_3 \setminus \{v\}|$ . One can show that senators are equally effective and that the vice-president is more effective than any senator: for all  $k, k' \in \{1, \dots, 100\}$ ,  $i_k \sim_{P,d_{\alpha,\beta,\lambda}} i_{k'}$  and  $v \succ_{P,d_{\alpha,\beta,\lambda}} i_k$ .

### 5.2.3 Anonymous voting rules

The two legislative institutions described above use voting rules that are not anonymous in that certain members do not hold interchangeable positions in the distribution of constitutional power, which results in some members being more effective than others. Many institutions, however, use anonymous voting rules (Freixas and Zwicker (2009)) to select policies. Based on the effective power relation, one can show that the members of such institutions are equally effective. This is the case even if preferences are heterogeneous. In fact, our definition of the effective power relation adjusts for preferences when comparing the ability of voters to pull the social ranking as close as possible to their individual preferences.

Note that adjusting for preferences in the general evaluation of power is important because it avoids situations in which a null player  $i$  may be seen as detaining more effective power than another null player  $j$  simply because  $i$ 's preferences happen to be closer to those of more powerful players than  $j$ 's preferences are. For instance, in a dictatorship, if  $i$  has more power than  $j$  simply because  $i$  has similar preferences as the dictator, then  $i$  will lose his power in case the dictator changes his opinion on a matter. It follows that a theory of power that evaluates a voter's power based on the proximity of his preferences to those of other voters is necessarily prone to externalities. By adjusting for preferences, one can see that in a dictatorship, all null players have the same effective power and the dictator has absolute power. In this sense, our measure of effective power is externality-free. In the same way, one can show that under any anonymous voting rule  $W$ , voters are equally effective: for all voters  $i$  and  $j$ ,  $i \sim_{P,d_{\alpha,\beta,\lambda}} j$ .

## 5.3 Effective power and the influence relation

In this section, we begin the analysis of the extent to which the classical theories of voting power measure a voter's ability to bring about a social outcome that maximizes his preferences. We uncover conditions under

<sup>17</sup>Note that this voting rule has a quorum requirement, but rules for quorums are somewhat complex, so we ignore them.

which these theories translate into effective power. We first focus on the influence relation. This relation is an ordinal measure of influence in a setting where output is produced by combining several inputs. It was introduced by Isbell (1958) for simple voting rules. Also known as the desirability relation, its properties have been extensively studied (see Taylor and Zwicker (1999) for an authoritative monograph). Tchantcho et al (2008) generalize the influence relation to voting rules with abstention.<sup>18</sup> The study of the influence relation for this larger class of rules is still in its infancy. We recall the generalization in Tchantcho et al. (2008) below.

**Definition 3.** *Let  $N$  be a set of voters,  $V$  a voting rule, and  $p$  and  $r$  two voters.  $p$  is said to be at least as influential as  $r$ , denoted  $p \geq_T r$ , if for all tripartition  $(S_1, S_2, S_3) \in 3^N$ , the following statement holds:*

- $V(S_1 \cup p, S_2 \setminus p, S_3) \geq V(S_1 \cup r, S_2 \setminus r, S_3)$  if  $p, r \in S_2$ ; and
- $V(S_1, S_2 \cup p, S_3 \setminus p) \geq V(S_1, S_2 \cup r, S_3 \setminus r)$  if  $p, r \in S_3$ ; and
- $V(S_1 \cup p, S_2, S_3 \setminus p) \geq V(S_1 \cup r, S_2, S_3 \setminus r)$  if  $p, r \in S_3$ .

The interpretation of the influence relation is as follows. A voter  $p$  is said to be more influential than another voter  $r$  under a voting rule  $V$  if whenever  $r$  can turn a losing tripartition into a winning tripartition by increasing his level of support,  $p$  can do the same *ceteris paribus*.

If  $N$  is a set of voters,  $A$  a set of candidates, and  $V$  (or  $W$ ) a voting rule with abstention, we will call  $(N, A, V)$  (or  $(N, A, W)$ ) a society.<sup>19</sup>

We now derive a necessary and sufficient condition for the influence relation to be equivalent to the effective power relation.

**Theorem 4.** *Let  $\alpha, \beta, \lambda \in ]0, 1]$  such that  $\alpha = \lambda$ . Then,  $\geq_{P_{\alpha, \beta, \lambda}}$  and  $\geq_T$  are equivalent for all society  $(N, A, V)$  if and only if  $\beta < 1$ .*

This theorem says that, under truth-telling ( $\alpha = \lambda$ ), the influence relation translates into effective power in all societies if and only if  $\beta$ —a voter’s level of dissatisfaction when he strictly ranks two candidates and the society is unable to rank them—is not the highest. Truth-telling is assumed to accommodate an implicit assumption made in the definition of the influence relation; according to this assumption, voters’ actions are symmetric (although they may not have the same effect on the outcome) since two voters are compared by swapping them within tripartitions (see Definition 3). Similarly, the truth-telling assumption guarantees that when two voters swap their preferences, they also swap their actions, with these actions reflecting the preferences they hold.<sup>20</sup>

To understand Theorem 4, recall that the highest level of dissatisfaction is reached when a voter’s ranking and the society’s ranking are completely opposite (that is, when a voter strictly prefers  $a$  to  $b$  and the tournament ranking puts  $b$  above  $a$ ). It is therefore natural to think that a voter’s dissatisfaction with the social ranking is not the highest if this voter strictly prefers  $a$  to  $b$  and the two options appear to be socially incomparable. In fact, to say that one is unable to compare  $a$  and  $b$  is not the strict opposite of  $a$  is better than  $b$ . In this sense, the condition  $\beta < 1$  is natural. This condition is also natural if social incomparability is resolved using a tie-break rule. In fact, if when the society is unable to rank two candidates strictly ranked by a voter, a probabilistic tie-break rule, for example, is used to rank them, the social ranking obtained from

<sup>18</sup>The influence relation has been studied in a variety of contexts. In a recent study, Pongou et al. (2017) derive a general form of this relation as the ranking rule of a ladder tournament in organizations.

<sup>19</sup>As noted earlier, we use  $V$  and  $W$  interchangeably. In fact, for any tripartition  $(S_1, S_2, S_3)$ ,  $V(S_1, S_2, S_3) = 1$  (that is,  $(S_1, S_2, S_3)$  is winning) is equivalent to  $(S_1, S_2, S_3) \in W$  (which also means that  $V(S_1, S_2, S_3) = 0$  is equivalent to  $(S_1, S_2, S_3) \notin W$ ). We use  $V$  or  $W$  when convenient for expositional purposes.

<sup>20</sup>Remark that the truth-telling assumption ( $\alpha = \lambda$ ) is not needed in the definition (Definition 2) or the operationalization of the effective power relation. Indeed, the analysis of its ordinal properties in Theorem 5 below is valid regardless of whether  $\alpha = \lambda$  or not.

this tie-break rule will coincide with the voter’s preferences with positive probability, resulting in a level of dissatisfaction  $\beta$  that is smaller than 1 with positive probability.

It follows from this discussion that the influence relation translates into effective power under plausible behavioral (or psychological) conditions, given that the condition  $\beta < 1$  is natural. However, one should also acknowledge the fact that for certain individuals, the highest level of disappointment may result from any disagreement no matter how small the disagreement may be perceived. If voters display such psychological conditions, the influence relation will not generally translate into effective power.

A final remark is that when  $\beta = 1$ , our result implies that there exists at least one society  $(N, A, V)$  for which the influence relation does not translate into effective power. This is obtained by contraposition of the “only if” part; an example of a society in which the influence relation does not translate into effective power is given in the proof for  $\beta = 1$ . Our result does not say that when  $\beta = 1$ , there is no society for which the influence relation captures effective power.<sup>21</sup>

Remark that the effective power relation compares two voters  $i$  and  $j$  in a society  $(N, A, V)$  only if it can compare them in all tournaments  $(N, A, \mathcal{P}_2(A), R, V, D)$  associated with this society. For this reason, one may be tempted to believe that this relation is not fully operational. However, Theorem 4 shows that, for a large class of behavioral parameters, it coincides with the influence relation, the latter being an established theory of power that is considered to be very useful (Isbell (1958), Taylor and Zwicker (1999), Tchantcho et al. (2008)). Another appeal of the effective power relation is that it can be defined more locally, by restricting it to any tournament  $(N, A, \mathcal{P}_2(A), R, V, D)$ . Moreover, we show below that it is locally more operational than the influence relation. More precisely, we show that in any political tournament, whenever the influence relation compares two voters, the effective power relation compares them as well; however, there exists at least one political tournament in which two voters are ranked by the effective power relation whereas they are incomparable by the influence relation.

**Proposition 1.** *Let  $\alpha, \beta, \lambda \in ]0, 1]$  and  $(N, A, \mathcal{P}_2(A), R, V, D)$  be a political tournament. Then,  $\geq_T$  is included in  $\geq_{P_{\alpha, \beta, \lambda}}$  ( $\geq_T \subseteq \geq_{P_{\alpha, \beta, \lambda}}$ ), and this inclusion may be strict.*

This proposition shows that the effective power relation is more likely to compare individuals involved in a tournament than the influence relation.

## 5.4 Effective power, the Shapley-Shubik power index, and the Banzhaf power index

In this section, we study conditions under which the Shapley-Shubik and Banzhaf power indices measure voters’ ability to secure desirable social outcomes. These indices were initially defined for simple games and generalized to voting games with abstention by Felsenthal and Machover (1997). Given that the Shapley-Shubik and Banzhaf power indices are real-valued functions, each induces a complete and transitive binary relation on the set of voters. This provides a motivation for studying the conditions under which the effective power relation is complete and transitive first. We show below that this is the case when the voting rule is swap-robust and utility parameters satisfy a mild condition.

We recall the definition of a swap-robust voting rule with abstention. Let  $N$  be a set of voters,  $V$  a voting rule with abstention,  $i, j \in N$  two voters,  $S = (S_1, S_2, S_3)$  a tripartition, and  $p(i, S)$  the position of  $i$  in  $S$ .<sup>22</sup> We denote by  $\Pi_{ij}S$  the tripartition resulting from  $S$  by swapping the positions of  $i$  and  $j$  while all the other

<sup>21</sup>By analogy, the Gibbard-Satterthwaite impossibility theorem which states that all non-dictatorial rules (e.g., majority rule) are manipulable does not mean that non-dictatorial rules are manipulable in all voting games. It simply means that non-dictatorial rules are manipulable in some games.

<sup>22</sup>The position of  $i$  in  $S$ ,  $p(i, S)$ , simply indicates his level of approval (yes, no, or abstention) in  $S$ .

individuals keep their respective positions in  $S$ . We have the following definition of a swap-robust voting rule with abstention:

**Definition 4.** *Let  $N$  be a set of voters, and  $V$  a voting rule with abstention. We say that  $V$  is swap-robust if for any winning tripartitions  $S = (S_1, S_2, S_3)$  and  $T = (T_1, T_2, T_3)$  (that is,  $V(S) = V(T) = 1$ ), any voters  $i, j \in N$  such that  $(p(i, S), p(j, S)) = (p(j, T), p(i, T))$ ,  $\Pi_{ij}S$  or  $\Pi_{ij}T$  is a winning tripartition.*

In words, a voting rule  $V$  is swap-robust if for any winning tripartitions  $S$  and  $T$  and any two voters  $i$  and  $j$  such that their respective positions in  $S$  are swapped in  $T$ , at least one of the tripartitions  $\Pi_{ij}S$  or  $\Pi_{ij}T$  resulting from swapping the positions of  $i$  and  $j$  in both  $S$  and  $T$  is winning. This definition is a generalization of the same notion defined by Taylor and Zwicker (1999) for simple voting rules under which voters can only vote yes and no.

Tchantcho et al. (2008) show that the influence relation is a complete preorder for all swap-robust voting rules.

**Lemma 1.** *Let  $N$  be a set of voters, and  $V$  a voting rule with abstention. The influence relation  $\geq_T$  is complete and transitive if and only if  $V$  is swap-robust. When  $\geq_T$  is a complete preorder for  $(N, V)$ ,  $V$  is also said to be linear.*

We obtain the following result which provides a necessary and sufficient condition for the effective power relation to be complete and transitive.

**Theorem 5.** *Let  $\alpha, \beta, \lambda \in ]0, 1]$  such that  $\beta < 1$ , and  $(N, A, V)$  be a society. Then,  $\geq_{P_{\alpha, \beta, \lambda}}$  is complete and transitive if and only if  $V$  is swap-robust.*

The class of swap-robust voting rules is large. In real-life settings, these rules are also highly prevalent. Most of the rules used in prominent organizations such as the United Nations Security Council, the United States Senate, the European Union Council, the International Monetary Fund, the World Bank, and the African Union are swap-robust. Anonymous and non-anonymous quota rules are also swap-robust. Moreover, as argued in the preceding section, the behavioral condition  $\beta < 1$  is natural. This implies that the effective power relation is complete and transitive under a very large set of rules. However, it is still very useful even in settings where it is not complete. In fact, if a political tournament is such that the effective power relation is not complete, that simply means that there exist at least two voters such that neither is better than the other in terms of influencing the social ranking. When this happens, we will see that the Shapley-Shubik and Banzhaf power indices do not measure effective power either, even though they are complete. We recall the definition of each of these power notions for voting rules with abstention as given in Felsenthal and Machover (1998).

**Definition 5.** *A ternary roll-call on  $N$  is an ordered pair  $R = (sR, dR)$  such that  $sR$  is a permutation on  $N$  and  $dR$  is a tripartition of  $N$ . We denote by  $\mathcal{J}_N$  the set of all ternary roll-calls over  $N$ .*

*Two ternary roll-calls,  $R_1$  and  $R_2$ , agree up to voter  $a$  if  $sR_1 = sR_2$  and  $dR_1(x) = dR_2(x)$  for all  $x \in N$  whenever  $sR_1(x) \leq sR_1(a)$ .*

*If  $(N, V)$  is such that  $V$  is a voting rule with abstention and  $R$  is a ternary roll-call on  $N$ , we say that voter  $a$  is the pivot of  $R$  under  $V$ , denoted  $a = \text{Piv}(V, R)$ , if  $a$  is the first voter in the ordering  $sR$  satisfying the condition:  $V(dS) = V(dR)$  for every ternary roll-call  $S$  that agrees with  $R$  up to  $a$ .*

*For any  $(N, V)$  such that  $V$  is a voting rule with abstention and any voter  $a \in N$ , the Shapley-Shubik index of  $a$ , denoted  $\phi_a$ , is defined by*

$$\phi_a(V) = \frac{|R \in \mathcal{J}_N : a = \text{Piv}(V, R)|}{3^n n!}.$$

This definition is interpreted as follows. It is implicitly assumed that voters are called to vote in favor of, or against, a motion or abstain under a voting rule  $V$ . In a ternary roll-call  $R = (sR, dR)$ ,  $sR$  indicates the order in which voters are called to vote, and  $dR$  is a tripartition. The vote of a voter  $x$ , denoted by  $dR(x)$ , takes the value 1 (in favor of the motion), 0 (abstention), or  $-1$  (against the motion). Thus, to say that two ternary roll-calls,  $R_1$  and  $R_2$ , agree up to voter  $a$  means that the voters are called in the same order in both roll-calls, and the votes of all voters up to and including voter  $a$  are identical in both roll-calls.

In a roll-call  $R$ , a voter is said to be pivotal, denoted by  $Piv(V, R)$ , if he is the first voter in  $R$  whose vote makes a difference to the outcome  $V(dR)$ , and the outcome remains unaffected regardless of how the voters that vote after the pivotal voter vote. Finally, the Shapley–Shubik index for  $a$  is the number of binary roll-calls  $R$  such that  $a = Piv(V, R)$  divided by the cardinality of the set of roll-calls. In other words, the Shapley–Shubik value of a voter measures the frequency at which this voter is pivotal.<sup>23</sup>

We also recall the definition of the Banzhaf power index for voting games with abstention. We first introduce the following notation. For a voter  $a \in N$ , we denote by  $\chi_a$  the tripartition of  $N$  defined for all  $x \in N$  by  $\chi_a(x) = 1$  if  $x = a$  and  $\chi_a(x) = 0$ .

**Definition 6.** Let  $(N, V)$  be such that  $V$  is a voting rule with abstention and let  $T$  be a tripartition of  $N$ . A voter  $a$  is positively critical for  $T$  in  $V$  if  $T(a) \geq 0$ ,  $V(T) = 1$  and  $V(T - \chi_a) = -1$ .

The Banzhaf score of  $a$  under  $V$ , denoted by  $\eta_a(V)$ , is the number of tripartitions for which  $a$  is positively critical in  $V$ . The (relative) Banzhaf index for  $a$ , denoted  $\beta_a(N)$ , is given by

$$\beta_a(V) = \frac{\eta_a(V)}{\sum_{x \in N} \eta_x(V)}.$$

This definition is interpreted as follows. Under a voting rule  $V$ , a voter is positively critical (or is a swing voter) in a winning tripartition  $T$  if the tripartition resulting from this voter reducing his level of approval for a motion is losing. The Banzhaf score of a voter is simply the number of tripartitions in which this voter is critical, and his relative Banzhaf value is obtained by dividing his Banzhaf score by the sum of the Banzhaf scores of all voters. In other words, the relative Banzhaf score of a voter is the fraction of all swing votes this voter can cast.<sup>24</sup> The Banzhaf index is also called the Banzhaf–Coleman power index.

Denoting by  $\geq_S$  and  $\geq_B$  the preorderings induced by the Shapley–Shubik and the Banzhaf power indices, respectively, Tchantcho et al. (2008) prove that these preorderings coincide in the subclass of weakly equitable linear voting rules with abstention. Parker (2012) generalizes this result to the whole class of linear voting rules with abstention. Pongou et al. (2014) obtain this result as a special case of a more general result in the class of multi-choice rules. This result is summarized below:

**Lemma 2.**  $\geq_T, \geq_S$  and  $\geq_B$  coincide in all linear voting rules with abstention  $V$ .

We obtain the following result:

**Theorem 6.** Let  $\alpha, \beta, \lambda \in ]0, 1]$  such that  $\alpha = \lambda$ , and let  $(N, A, W)$  be a society.

- 1- If  $\beta < 1$  and  $W$  is swap-robust, then  $\geq_{P_{\alpha, \beta, \lambda}}, \geq_S$  and  $\geq_B$  are equivalent.
- 2- If  $\beta = 1$  or  $W$  is not swap-robust, then  $\geq_{P_{\alpha, \beta, \lambda}}, \geq_S$  and  $\geq_B$  may not be equivalent. In particular, if  $\beta < 1$  and  $V$  is not swap-robust, then  $\geq_{P_{\alpha, \beta, \lambda}}, \geq_S$  and  $\geq_B$  are not equivalent.

<sup>23</sup>The classical Shapley–Shubik power index (Shapley and Shubik (1954)) computes the power of each voter under the assumption that the grand coalition forms, which is equivalent to assuming a voting profile in which all voters vote yes. Similarly, the generalization to voting rules with abstention presented in Definition 5 above computes the power of each voter in the configuration where all voters have expressed the highest level of approval. Relaxing this assumption, Pongou and Tondji (2018) generalize the Shapley value to a more complex environment which allows to compute this value for any possible action configuration. See also Aguiar et al. (2018) and Aguiar et al. (2020) for other generalizations and for a test of the axioms of the Shapley value with a limited dataset.

<sup>24</sup>The absolute Banzhaf value of a voter is obtained by dividing his Banzhaf score by the number of tripartitions when this voter is excluded. All the results that follow hold regardless of whether we use the relative or the absolute Banzhaf power index.

In Theorem 6, truth-telling ( $\alpha = \lambda$ ) is assumed for similar reasons as in Theorem 4. In fact, the definitions of the Shapley-Shubik and Banzhaf power indices assume that voters take symmetric actions, although these actions may not carry the same weight on the voting outcome. The truth-telling assumption is made to accommodate this assumption, as it ensures that when voters swap their preferences, they also swap their actions, with the latter reflecting the preferences they hold.

Theorem 6 implies that the Shapley-Shubik power index and the Banzhaf power index measure effective power under more restrictive conditions than the influence relation. One of its implications is that a greater value of these indices for a voter does not necessarily mean a greater ability to influence the outcome of the vote. However, according to Theorem 6, this is the case for an important class of political tournaments and utility parameters. An immediate corollary of this result is that a necessary and sufficient condition for  $\geq_{P_{\alpha,\beta,\lambda}}$ ,  $\geq_S$  and  $\geq_B$  to coincide for all linear society  $(N, A, W)$  (or society where  $W$  is swap-robust) is that  $\beta < 1$ .

**Corollary 1.** *Let  $\alpha, \beta, \lambda \in ]0, 1]$  such that  $\alpha = \lambda$ . Then  $\geq_{P_{\alpha,\beta,\lambda}}$ ,  $\geq_S$  and  $\geq_B$  are equivalent for all linear society  $(N, A, W)$  if and only if  $\beta < 1$ .*

We illustrate these findings in the following example.

**Example.** *Let  $N = \{1, 2, 3\}$ ,  $A = \{x_1, x_2\}$  and  $W$  the voting rule defined for every tripartition  $(S_1, S_2, S_3)$  by  $(S_1, S_2, S_3) \in W$  if and only if  $1 \in S_1$  or  $1 \in S_2$  and  $|S_1| > |S_3|$ .*

1. *Let us determine the ranking induced by the Shapley-Shubik power index  $\phi(W)$  in  $(N, W)$ .*

*Voter 1 is pivotal in any ternary roll-call  $R$  on  $N$  unless he votes first and abstains or he votes second and abstains and one of the other voters votes yes or abstains. Thus we have:  $\phi_1(W) = \frac{162 - (2 \times 9 + 2 \times 2 \times 3)}{162} = \frac{132}{162}$ .*

*Moreover, given that 2 and 3 are symmetric, it follows from the fact that all the voters' power indices sum up to 1 that  $\phi_2(W) = \phi_3(W) = \frac{15}{162}$ . Hence the ranking induced by the Shapley-Shubik power index is  $1 >_S 2 \simeq_S 3$ .*

2. *Let us determine the ranking induced by the effective power relation  $\geq_{P_{\alpha,\beta,\lambda}}$ . Let  $R$  be a preference profile on  $A$ .*

- *If  $R^1 = R^2$ , then  $R_{12} = R$  and so,  $d_{\alpha,\beta,\lambda}(R^2, D(R_{12})) = d_{\alpha,\beta,\lambda}(R^2, D(R))$ .*
- *If  $R^2 = x_1x_2$  or  $R^2 = x_2x_1$ , then  $D(R_{12}) = R_{12}^1 = R^2$  and so,  $d_{\alpha,\beta,\lambda}(R^2, D(R_{12})) = 0 \leq d_{\alpha,\beta,\lambda}(R^2, D(R))$ .*
- *Assume  $R^2 = (x_1x_2)$  and  $R^1 = x_1x_2$ . Then  $D(R) = x_1x_2$  and  $d_{\alpha,\beta,\lambda}(R^2, D(R)) = \alpha$ . On the other hand, we have  $R_{12}^1 = R^2 = (x_1x_2)$  and  $R_{12}^2 = R^1 = x_1x_2$ ; so the preference of voter 3 is needed to determine the social decision  $D(R_{12})$ . If  $R^3 = x_1x_2$  or  $R^3 = (x_1x_2)$ , then  $D(R_{12}) = x_1x_2$ , and it follows that  $d_{\alpha,\beta,\lambda}(R^2, D(R_{12})) = \alpha$ ; so  $d_{\alpha,\beta,\lambda}(R^2, D(R_{12})) = d_{\alpha,\beta,\lambda}(R^2, D(R))$ . But if  $R^3 = x_2x_1$ , then  $D(R_{12}) = x_1 * x_2$  and it follows that  $d_{\alpha,\beta,\lambda}(R^2, D(R_{12})) = \lambda$ ; this implies that  $d_{\alpha,\beta,\lambda}(R^2, D(R_{12})) = d_{\alpha,\beta,\lambda}(R^2, D(R))$  since  $\alpha = \lambda$ .*
- *The case where  $R^2 = (x_1x_2)$  and  $R^1 = x_2x_1$  is similar to the previous case.*

*Hence for any preference profile  $R$ , we have  $d_{\alpha,\beta,\lambda}(R^2, D(R_{12})) \leq d_{\alpha,\beta,\lambda}(R^2, D(R))$ , that is,*

*$1 \geq_{P_{\alpha,\beta,\lambda}} 2$ . Moreover for  $R^1 = x_1x_2$  and  $R^2 = x_2x_1$ , and assuming any  $R^3$ , we have  $d_{\alpha,\beta,\lambda}(R^1, D(R_{12})) = 1 > 0 = d_{\alpha,\beta,\lambda}(R^1, D(R))$ ; it follows that not( $2 \geq_{P_{\alpha,\beta,\lambda}} 1$ ), and we conclude that  $1 >_{P_{\alpha,\beta,\lambda}} 2$ .*

*Also, remark that for any preference profile  $R$ ,  $d_{\alpha,\beta,\lambda}(R^2, D(R_{23})) = d_{\alpha,\beta,\lambda}(R^2, D(R))$ , which is equivalent to  $2 \simeq_{P_{\alpha,\beta,\lambda}} 3$ . Therefore, the social ranking induced by the effective power relation is  $1 >_{P_{\alpha,\beta,\lambda}} 2 \simeq_{P_{\alpha,\beta,\lambda}} 3$ , which coincides with the ranking induced by the Shapley-Shubik power index.*

Similarly, we can show that this ranking coincides with the ranking induced by the Banzhaf power index.

Note that in the example above, the voting rule  $W$  is swap-robust. We did not need to assume that  $\beta < 1$  for the ranking induced by the Shapley-Shubik and Banzhaf power indices to coincide with the effective power relation. Below, we show an example where the Shapley-Shubik and Banzhaf power indices do not translate into effective power when  $\beta = 1$ .

**Example.** Let  $(N, A, W)$  be a society where  $N = \{1, 2, 3, 4, 5\}$ ,  $A = \{x_1, x_2\}$  and  $W$  is such that  $W_m = \{(1, 25, 34), (14, 2, 35), (12, 4, 35)\}$ .

1. It can be checked that  $1 >_T 2 >_T 5 >_T 4 >_T 3$ . Thanks to Lemma 1, this implies that  $W$  is swap-robust, and thanks to Lemma 2, we can deduce the rankings of the Shapley-Shubik and Banzhaf preorderings as follows:  $1 >_S 2 >_S 5 >_S 4 >_S 3$  and  $1 >_B 2 >_B 5 >_B 4 >_B 3$ .
2. Assume  $\beta = 1$ . We can show that  $4 \geq_{P_{\alpha, \beta, \gamma}} 5$ ; that is, for all political tournament  $(N, A, \mathcal{P}_2(A), R = (R^i)_{i \in N}, W, D)$  in the society  $(N, A, W)$ ,  $4 \geq_{P_{\alpha, \beta, \lambda}} 5$ . In particular, consider the preference profile  $R = (xy, (xy), yx, yx, (xy))$ . We have  $D(R) = xy$  (given truthful voting ( $\lambda = \alpha$ ) and the fact that  $(1, 25, 34)$  is a winning tripartition) and  $D(R_{45}) = x * y$ . It follows that  $d(R^5, D(R)) = \alpha$ ,  $d(R^5, D(R_{45})) = \lambda = \alpha = d(R^5, D(R))$ , which implies that  $4 \sim_{P_{\alpha, \beta, \gamma}} 5$  at the profile  $R$ . However,  $5 >_S 4$  and  $5 >_B 4$ , which implies that  $\geq_{P_{\alpha, \beta, \lambda}}$ ,  $\geq_S$  and  $\geq_B$  do not coincide in general. It follows that the Shapley-Shubik and Banzhaf power indices do not translate into effective power in general when  $\beta = 1$ , even if a society is using a linear voting rule to select policies. This analysis reveals that the extent to which classical notions of voting power measure the ability to affect a society's preference depends on psychological parameters.

## 6 Conclusion

We study voter behavior and effective power in a round-robin political tournament. In this tournament, candidates contest in pairs. In each contest, each voter votes in favor of one of the candidates or abstains. Votes are aggregated using a general monotonic rule to determine the tournament ranking. Each voter's utility is measured by how close his preferences are to the tournament ranking.

We uncover a necessary and sufficient condition on utility parameters for truthful voting to be a Nash equilibrium. This condition has both structural and behavioral interpretations, and it has implications for the design of political tournaments in which truth-telling is a desirable goal. In particular, it points to the importance of incorporating tie-break rules into tournaments to resolve social incomparability.

We also study a voter's effective power in a tournament, defined as his ability to bring about a social ranking that maximizes his preferences. We uncover a necessary and sufficient condition under which the effective power relation is transitive and complete. We also show that the classical theories of political power—the influence relation, the Shapley-Shubik power index, and the Banzhaf power index—do not translate into effective power in general. We provide a full characterization of the classes of political tournaments and psychological parameters for which these theories measure effective power. These characterizations have practical implications for how to design rules under which *a priori* voting rights effectively translate into the ability to affect voting outcomes for a natural category of voters.

# Appendix

## Proof of Theorem 1

1) Assume that  $\alpha = \lambda$ . Let  $R$  be a preference profile. In order to show that  $s^*(R)$  is a Nash equilibrium of the associated game  $\Gamma_R = (N, Z, U)$ , let  $i \in N$  and  $t_i(R) \in \{-1, 0, 1\}^{\frac{m(m-1)}{2}}$ , we will prove that  $U_i(s^*(R)) \geq U_i(t_i(R), s_{-i}^*(R))$ . Since for all profile  $Q$ ,  $U_i(s(Q)) = \sum_{\{x_p, x_q\} \in \mathcal{P}_2(A)} U_i^{pq}(s(Q))$ , it suffices to show that:  $\forall \{x_p, x_q\} \in \mathcal{P}_2(A)$ ,  $U_i^{pq}(s^*(R)) \geq U_i^{pq}(t_i(R), s_{-i}^*(R))$ .

Let  $\{x_p, x_q\} \in \mathcal{P}_2(A)$ . Without loss of generality, assume  $p < q$ . Three cases are possible for  $D(s^*(R))$ .

**First case:**  $x_p x_q \in D(s^*(R))$ .

$$\text{In this case, } U_i^{pq}(s^*(R)) = \begin{cases} 1 & \text{if } R_{\{x_p, x_q\}}^i = x_p x_q \\ 0 & \text{if } R_{\{x_p, x_q\}}^i = x_q x_p \\ 1 - \alpha & \text{if } R_{\{x_p, x_q\}}^i = (x_p x_q). \end{cases}$$

If  $R_{\{x_p, x_q\}}^i = x_p x_q$ , then  $U_i^{pq}(s^*(R)) = 1 \geq U_i^{pq}(t_i(R), s_{-i}^*(R))$ .

If  $R_{\{x_p, x_q\}}^i = x_q x_p$ , then  $s_i^{*pq}(R) = -1$  and  $U_i^{pq}(s^*(R)) = 0$ . Now, if  $t_i^{pq}(R) \in \{0, 1\}$ , then

$$N^1(s^{*pq}(R)) \subseteq N^1(t_i^{pq}(R), s_{-i}^{*pq}(R)),$$

and by monotonicity of  $W$ , we have

$$(N^1(t_i^{pq}(R), s_{-i}^{*pq}(R)), N^0(t_i^{pq}(R), s_{-i}^{*pq}(R)), N^{-1}(t_i^{pq}(R), s_{-i}^{*pq}(R))) \in W \text{ since } (N^1(s^{*pq}(R)), N^0(s^{*pq}(R), N^{-1}(s^{*pq}(R)))) \in W;$$

thus,

$$U_i^{pq}(t_i(R), s_{-i}^*(R)) = 0 = U_i^{pq}(s^*(R)).$$

If  $R_{\{x_p, x_q\}}^i = (x_q x_p)$ , then  $s_i^{*pq}(R) = 0$  and  $U_i^{pq}(s^*(R)) = 1 - \alpha$ .

For  $t_i^{pq}(R) = 1$ ,  $D_{\{x_p, x_q\}}(t_i(R), s_{-i}^*(R)) = x_p x_q$  and  $U_i^{pq}(t_i(R), s_{-i}^*(R)) = 1 - \alpha = U_i^{pq}(s^*(R))$ .

For  $t_i^{pq}(R) = -1$ , by monotonicity, we have:

$$D_{\{x_p, x_q\}}(t_i(R), s_{-i}^*(R)) = x_p x_q \text{ or } D_{\{x_p, x_q\}}(t_i(R), s_{-i}^*(R)) = x_p * x_q.$$

It is obvious that if  $D_{\{x_p, x_q\}}(t_i(R), s_{-i}^*(R)) = x_p x_q$ , then  $U_i^{pq}(t_i(R), s_{-i}^*(R)) = U_i^{pq}(s^*(R))$ .

And if  $D_{\{x_p, x_q\}}(t_i(R), s_{-i}^*(R)) = x_p * x_q$ , then  $U_i^{pq}(t_i(R), s_{-i}^*(R)) = 1 - \lambda$ ; and from the hypothesis that  $\alpha = \lambda$ , we obtain  $1 - \alpha = 1 - \lambda$ , that is,  $U_i^{pq}(s^*(R)) \geq U_i^{pq}(t_i(R), s_{-i}^*(R))$ .

**Second case:**  $x_q x_p \in D(s^*(R))$ .

The proof is similar to that of the first case.

**Third case:**  $x_p * x_q \in D(s^*(R))$ .

$$\text{In this case, } U_i^{pq}(s^*(R)) = \begin{cases} 1 - \beta & \text{if } R_{\{x_p, x_q\}}^i = x_p x_q \\ 1 - \beta & \text{if } R_{\{x_p, x_q\}}^i = x_q x_p \\ 1 - \lambda & \text{if } R_{\{x_p, x_q\}}^i = (x_p x_q). \end{cases}$$

If  $R_{\{x_p, x_q\}}^i = x_p x_q$ , then  $s_i^{*pq}(R) = 1$  and  $U_i^{pq}(s^*(R)) = 1 - \beta$ .

If  $t_i(R) = 0$ , then

$$(N^1(t_i(R), s_{-i}^*(R)), N^0(t_i(R), s_{-i}^*(R)), N^{-1}(t_i(R), s_{-i}^*(R))) = (N^1(s^*(R)) \setminus \{i\}, N^0(s^*(R) \cup \{i\}), N^{-1}(s^*(R))) \notin W.$$

We know that  $x_p * x_q \in D(s^*(R))$  implies that  $(N^{-1}(s^*(R)), N^0(s^*(R)), N^1(s^*(R))) \notin W$ , but

$(N^{-1}(s^*(R)), N^0(s^*(R) \cup \{i\}), N^1(s^*(R)) \setminus \{i\})$  might belong to  $W$ , which means that it is possible that

$x_q x_p \in D(t_i(R), s_{-i}^*(R))$ . Therefore, for  $t_i^{pq}(R) = 0$ ,  $D_{\{x_p, x_q\}}(t_i(R), s_{-i}^*(R)) \in \{x_q x_p, x_p * x_q\}$ , which leads

to  $U_i^{\{x_p, x_q\}}(t_i(R), s_{-i}^*(R)) \in \{0, 1 - \beta\}$ . Finally, we obtain  $U_i^{pq}(t_i(R), s_{-i}^*(R)) \leq U_i^{pq}(s^*(R)) = 1 - \beta$ .

If  $t_i^{pq}(R) = -1$ , then  $D_{\{x_p, x_q\}}(t_i(R), s_{-i}^*(R)) \in \{x_q x_p, x_p * x_q\}$  and, again, we obtain

$$U_i^{pq}(t_i(R), s_{-i}^*(R)) \leq U_i^{pq}(s^*(R)).$$

When  $R_{\{x_p, x_q\}}^i = x_q x_p$ , the reasoning is similar to the case where  $R_{\{x_p, x_q\}}^i = x_p x_q$ .

Now assume that  $R_{\{x_p, x_q\}}^i = (x_p x_q)$ . In this case,  $s_i^*(R) = 0$  and  $U_i^{pq}(s^*(R)) = 1 - \lambda$ .

If  $t_i^{pq}(R) = 1$ , then either  $D_{\{x_p, x_q\}}(t_i(R), s_{-i}^*(R)) = x_p * x_q$  or  $D_{\{x_p, x_q\}}(t_i(R), s_{-i}^*(R)) = x_p x_q$ , which

implies that  $U_i^{pq}(t_i(R), s_{-i}^*(R)) = 1 - \lambda$  or  $U_i^{pq}(t_i(R), s_{-i}^*(R)) = 1 - \alpha$ .

From the assumption that  $\lambda = \alpha$ , we have  $1 - \alpha = 1 - \lambda$ , that is,  $U_i^{pq}(t_i(R), s_{-i}^*(R)) = U_i^{pq}(s^*(R))$ .

2) Conversely, let us assume that  $\lambda \neq \alpha$ , that is,  $\lambda < \alpha$  or  $\alpha < \lambda$ .

Assume that  $\alpha < \lambda$ . Consider simple majority rule, in which a tripartition  $(S_1, S_2, S_3)$  is winning if and only if  $|S_1| > |S_3|$ . Let  $R$  be the preference profile defined as follows :

$$R^1 : (x_1 x_2) x_3 \dots x_m;$$

$$R^2 : x_1 x_2 x_3 \dots x_m;$$

$$R^3 : x_2 x_1 x_3 \dots x_m;$$

$$R^k : (x_1 x_2) x_3 \dots x_m \text{ for all } k \notin \{1, 2, 3\}.$$

The truthful vote profile  $s^*(R)$  is given by:  $s_1^*(R) = (0, 1, 1, \dots, 1)$ ,  $s_2^*(R) = (1, 1, \dots, 1)$ ,  $s_3^*(R) = (-1, 1, \dots, 1)$  and  $s_k^*(R) = (0, 1, 1, \dots, 1)$  for all  $k \notin \{1, 2, 3\}$ .

Consider the following deviation  $t_1(R)$  of player 1:  $t_1(R) = (1, 1, 1, \dots, 1)$ . We have:

$$\begin{aligned} D(s^*(R)) &= \{x_1 * x_2\} \cup \{x_p x_q : p < q, (p, q) \neq (1, 2)\} \\ U_1(s^*(R)) &= [(1 - \lambda) + (m - 2)] + (m - 2) + \dots + 2 + 1 = \frac{m(m - 1)}{2} - \lambda \\ D(t_1(R), s_{-1}^*(R)) &= \{x_p x_q : p < q\} \\ U_1(t_1(R), s_{-1}^*(R)) &= [(1 - \alpha) + (m - 2)] + (m - 2) + \dots + 2 + 1 = \frac{m(m - 1)}{2} - \alpha \end{aligned}$$

Furthermore,  $\alpha < \lambda$  implies  $\frac{m(m-1)}{2} - \alpha > \frac{m(m-1)}{2} - \lambda$ , that is,  $U_1(t_1(R), s_{-1}^*(R)) > U_1(s^*(R))$ ; this means that  $s^*(R)$  is not a Nash equilibrium for the game  $\Gamma_R$ .

Assume that  $\alpha > \lambda$ . Consider again simple majority rule with abstention  $W$ , where a tripartition  $(S_1, S_2, S_3)$  is winning if and only if  $|S_1| > |S_3|$ . Let  $R$  be the preference profile defined as follows:

$$\begin{aligned} R^1 &: (x_1 x_2) x_3 \dots x_m; \\ R^2 &: x_1 x_2 x_3 \dots x_m; \\ R^k &: (x_1 x_2) x_3 \dots x_m \text{ for all } k \notin \{1, 2\}. \end{aligned}$$

The sincere voting profile  $s^*(R)$  is given by:  $s_1^*(R) = (0, 1, 1, \dots, 1)$ ,  $s_2^*(R) = (1, 1, 1, \dots, 1)$  and  $s_k^*(R) = (0, 1, 1, \dots, 1)$  for all  $k \notin \{1, 2, 3\}$ .

Consider the following deviation of player 1:  $t_1(R) = (-1, 1, 1, \dots, 1)$ . We have:

$$\begin{aligned} D(s^*(R)) &= \{x_p x_q : p < q\} \\ U_1(s^*(R)) &= [(1 - \alpha) + (m - 2)] + (m - 2) + \dots + 2 + 1 = \frac{m(m - 1)}{2} - \alpha \\ D(t_1(R), s_{-1}^*(R)) &= \{x_p x_q : p < q, (p, q) \neq (1, 2)\} \cup \{x_1 * x_2\} \\ U_1(t_1(R), s_{-1}^*(R)) &= [(1 - \lambda) + (m - 2)] + (m - 2) + \dots + 2 + 1 = \frac{m(m - 1)}{2} - \lambda \end{aligned}$$

Furthermore,  $\alpha > \lambda$  implies  $\frac{m(m-1)}{2} - \lambda > \frac{m(m-1)}{2} - \alpha$ , that is,  $U_1(t_1(R), s_{-1}^*(R)) > U_1(s^*(R))$ ; again, we conclude that  $s^*(R)$  is not a Nash equilibrium for the game  $\Gamma_R$ . ■

### Proof of Theorem 2

The proof is similar to that of Theorem 1 since the proof of this theorem does not use the argument that  $d_{\alpha, \beta, \lambda}$  is a metric. ■

### Proof of Theorem 3

The proof is similar to that of Theorem 1 since this proof does not rely on the assumption that all the pairs of alternatives are presented simultaneously. ■

### Proof of Theorem 4

Let  $\alpha, \beta, \lambda \in ]0, 1]$  such that  $\alpha = \lambda$ . The proof will proceed in two parts.

**First part:** We assume that  $\beta < 1$  and prove that  $\geq_T$  and  $\geq_{P_{\alpha, \beta, \lambda}}$  are equivalent in all societies  $(N, A, V)$  (or  $(N, A, W)$ ).

Let  $(N, A, W)$  be a society, and  $i, j \in N$  two voters.

a) Assume that  $i \geq_T j$ .

We want to show that  $i \geq_{P_{\alpha, \beta, \lambda}} j$ ; that is, for all preference profile  $R \in \mathcal{U}^N$ ,  $d_{\alpha, \beta, \lambda}(R^j, D(R_{ij})) \leq d_{\alpha, \beta, \lambda}(R^j, D(R))$ .

Let  $R \in \mathcal{U}^N$  be a preference profile. It suffices to prove that for all  $\{x, y\} \in \mathcal{P}_2(A)$ ,  $d_{\alpha, \beta, \lambda}^{\{x, y\}}(R^j, D(R_{ij})) \leq d_{\alpha, \beta, \lambda}^{\{x, y\}}(R^j, D(R))$ . We will distinguish four cases corresponding to the value of  $d_{\alpha, \beta, \lambda}^{\{x, y\}}(R^j, D(R_{ij})) \in \{0, \alpha, \beta, 1\}$ .

- If  $d_{\alpha, \beta, \lambda}^{\{x, y\}}(R^j, D(R_{ij})) = 0$ , the proof is obvious.
- If  $d_{\alpha, \beta, \lambda}^{\{x, y\}}(R^j, D(R_{ij})) = \alpha$ , then two subcases are possible.

**Subcase 1:**  $x \simeq_{R^j} y$  and  $x *_{D(R_{ij})} y$ .

From the definition of a political tournament and thanks to  $x \simeq_{R^j} y$ ,  $x \simeq_{D(R)} y$  is impossible; thus

$d_{\alpha,\beta,\lambda}^{\{x,y\}}(R^j, D(R)) \neq 0$ . Furthermore, we have  $d_{\alpha,\beta,\lambda}^{\{x,y\}}(R^j, D(R)) \neq \beta$  because  $x \simeq_{R^j} y$ . Finally we have

either  $d_{\alpha,\beta,\lambda}^{\{x,y\}}(R^j, D(R)) = \alpha$  or  $d_{\alpha,\beta,\lambda}^{\{x,y\}}(R^j, D(R)) = 1$ , and hence,  $d_{\alpha,\beta,\lambda}^{\{x,y\}}(R^j, D(R)) \geq d_{\alpha,\beta,\lambda}^{\{x,y\}}(R^j, D(R_{ij}))$ .

**Subcase 2:**  $x \simeq_{R^j} y$  and  $x \succ_{D(R_{ij})} y$ .

Again, we have  $d_{\alpha,\beta,\lambda}^{\{x,y\}}(R^j, D(R)) \neq 0$  and  $d_{\alpha,\beta,\lambda}^{\{x,y\}}(R^j, D(R)) \neq \beta$ ; thus, either  $d_{\alpha,\beta,\lambda}^{\{x,y\}}(R^j, D(R)) = \alpha$  or

$d_{\alpha,\beta,\lambda}^{\{x,y\}}(R^j, D(R)) = 1$ .

• If  $d_{\alpha,\beta,\lambda}^{\{x,y\}}(R^j, D(R_{ij})) = \beta$ , then  $x \succ_{R^j} y$  and  $x *_{D(R_{ij})} y$ . Since  $x \succ_{R^j} y$  and we cannot have  $x \simeq_{D(R)} y$ ,

it is the case that  $d_{\alpha,\beta,\lambda}^{\{x,y\}}(R^j, D(R)) \neq \alpha$ . Now, we will show that  $d_{\alpha,\beta,\lambda}^{\{x,y\}}(R^j, D(R)) \neq 0$ . Assume on the

contrary that  $d_{\alpha,\beta,\lambda}^{\{x,y\}}(R^j, D(R)) = 0$ , that is,  $x \succ_{D(R)} y$ . Then,  $(N(x, y, R), N^0(x, y, R), N(y, x, R)) \in W$ .

- If  $i \in N(x, y, R)$ , then

$$(N(x, y, R_{ij}), N^0(x, y, R_{ij}), N(y, x, R_{ij})) = (N(x, y, R), N^0(x, y, R), N(y, x, R)) \in W$$

and therefore,  $x \succ_{D(R_{ij})} y$ , which is a contradiction.

- If  $i \in N^0(x, y, R)$ , then

$$(N(x, y, R_{ij}), N^0(x, y, R_{ij}), N(y, x, R_{ij})) = (N(x, y, R) \setminus \{j\} \cup \{i\}, N^0(x, y, R) \setminus \{i\} \cup \{j\}, N(y, x, R)).$$

From the assumption that  $i \geq_T j$ , and from  $(N(x, y, R), N^0(x, y, R), N(y, x, R)) \in W$ , we have

$(N(x, y, R_{ij}), N^0(x, y, R_{ij}), N(y, x, R_{ij})) \in W$ , implying again that  $x \succ_{D(R_{ij})} y$ , which is a contradiction.

- If  $i \in N(y, x, R)$ , then

$$(N(x, y, R_{ij}), N^0(x, y, R_{ij}), N(y, x, R_{ij})) = (N(x, y, R) \setminus \{j\} \cup \{i\}, N^0(x, y, R), N(y, x, R) \setminus \{i\} \cup \{j\}).$$

As above, we obtain  $x \succ_{D(R_{ij})} y$ , which is a contradiction because  $x *_{D(R_{ij})} y$ .

As  $d_{\alpha,\beta,\lambda}^{\{x,y\}}(R^j, D(R)) \notin \{0, \alpha\}$ , we deduce that  $d_{\alpha,\beta,\lambda}^{\{x,y\}}(R^j, D(R)) \geq d_{\alpha,\beta,\lambda}^{\{x,y\}}(R^j, D(R_{ij}))$ .

• If  $d_{\alpha,\beta,\lambda}^{\{x,y\}}(R^j, D(R_{ij})) = 1$ , without loss of generality, assume that  $x \succ_{R^j} y$  and  $y \succ_{D(R_{ij})} x$ .

$y \succ_{D(R_{ij})} x$  implies that  $(N(y, x, R_{ij}), N^0(y, x, R_{ij}), N(x, y, R_{ij})) \in W$ , and  $x \succ_{R^j} y$  implies  $x \succ_{R_{ij}^i} y$ , which implies that  $i \in N(x, y, R_{ij})$ .

- If  $j \in N(x, y, R_{ij})$ , then

$$(N(y, x, R), N^0(y, x, R_{ij}), N(x, y, R)) = (N(y, x, R_{ij}), N^0(y, x, R_{ij}), N(x, y, R_{ij})) \in W,$$

implying that  $y \succ_{D(R)} x$  and  $d_{\alpha,\beta,\lambda}^{\{x,y\}}(R^j, D(R)) = 1$ .

- If  $j \in N^0(x, y, R_{ij})$ , then

$$(N(y, x, R), N^0(y, x, R), N(x, y, R)) = (N(y, x, R_{ij}), N^0(y, x, R_{ij}) \setminus \{j\} \cup \{i\}, N(x, y, R_{ij}) \setminus \{i\} \cup \{j\}).$$

But  $(N(y, x, R_{ij}), N^0(y, x, R_{ij}), N(x, y, R_{ij})) \in W$  and  $i \geq_T j$ ; hence

$$(N(y, x, R), N^0(y, x, R), N(x, y, R)) \in W, \text{ implying that } y \succ_{D(R)} x, \text{ and hence } d_{\alpha, \beta, \lambda}^{\{x, y\}}(R^j, D(R)) = 1.$$

- If  $j \in N(y, x, R_{ij})$ , then

$$(N(y, x, R), N^0(y, x, R), N(x, y, R)) = (N(y, x, R_{ij}) \setminus \{j\} \cup \{i\}, N^0(y, x, R_{ij}), N(x, y, R_{ij}) \setminus \{i\} \cup \{j\});$$

and for the same reason as above, we obtain  $(N(y, x, R), N^0(y, x, R), N(x, y, R)) \in W$ , which implies that

$$y \succ_{D(R)} x, \text{ and hence } d_{\alpha, \beta, \lambda}^{\{x, y\}}(R^j, D(R)) = 1.$$

For each of these scenarios, we therefore conclude that  $d_{\alpha, \beta, \lambda}^{\{x, y\}}(R^j, D(R)) = 1$ , which implies

$$d_{\alpha, \beta, \lambda}^{\{x, y\}}(R^j, D(R)) = 1 \geq d_{\alpha, \beta, \lambda}^{\{x, y\}}(R^j, D(R_{ij})).$$

b) We now assume that  $i \geq_{P_{\alpha, \beta, \lambda}} j$  and would like to show that  $i \geq_T j$ . By contradiction, let us assume that  $\text{non}(i \geq_T j)$ . We should then show that  $\text{non}(i \geq_{P_{\alpha, \beta, \lambda}} j)$ . Following the definition of  $\geq_{P_{\alpha, \beta, \lambda}}$ , we have to show that there exists a preference profile  $R \in \mathcal{U}^N$  such that  $d_{\alpha, \beta, \lambda}(R^j, D(R_{ij})) > d_{\alpha, \beta, \lambda}(R^j, D(R))$ . Let  $A = \{x_1, x_2, \dots, x_m\}$ . From  $\text{non}(i \geq_T j)$ , three subcases are possible:

**Subcase 1 :** There exists  $(S_1, S_2, S_3) \in 3^N$  such that: 
$$\begin{cases} (S_1, S_2, S_3) \in W \\ j \in S_1, i \in S_3 \text{ and} \\ ((S_1 \setminus \{j\}) \cup \{i\}, S_2, (S_3 \setminus \{i\}) \cup \{j\}) \notin W. \end{cases}$$

Consider the following preference profile  $R$ :

$$\begin{cases} \forall k \in S_1, x_1 \succ_{R^k} x_2 \succ_{R^k} x_3 \succ_{R^k} \dots \succ_{R^k} x_m \\ \forall k \in S_2, x_1 \simeq_{R^k} x_2 \succ_{R^k} x_3 \succ_{R^k} \dots \succ_{R^k} x_m \\ \forall k \in S_3, x_2 \succ_{R^k} x_1 \succ_{R^k} x_3 \succ_{R^k} \dots \succ_{R^k} x_m \end{cases}$$

The computation of  $D(R)$  gives :

$$x_1 \succ_{D(R)} x_2 \succ_{D(R)} x_3 \succ_{D(R)} \dots \succ_{D(R)} x_m, \text{ which implies } d_{\alpha, \beta, \lambda}(R^j, D(R)) = 0.$$

But  $\text{non}(x_1 D(R_{ij}) x_2)$  since

$$(N(x_1, x_2, R_{ij}), N^0(x, y, R_{ij}), N(x_2, x_1, R_{ij})) = (S_1 \setminus \{j\} \cup \{i\}, S_2, S_3 \setminus \{i\} \cup \{j\}) \notin W.$$

Therefore,  $d_{\alpha, \beta, \lambda}(R^j, D(R_{ij})) \geq \beta > 0 = d_{\alpha, \beta, \lambda}(R^j, D(R))$ .

**Subcase 2 :** There exists  $(S_1, S_2, S_3) \in 3^N$  such that: 
$$\begin{cases} (S_1, S_2, S_3) \in W \\ j \in S_1, i \in S_2 \text{ and} \\ ((S_1 \setminus \{j\}) \cup \{i\}, (S_2 \setminus \{i\}) \cup \{j\}, S_3) \notin W \end{cases}$$

We consider the same profile defined in subcase 1.

Again we obtain :  $x_1 \succ_{D(R)} x_2 \succ_{D(R)} x_3 \succ_{D(R)} \dots \succ_{D(R)} x_m$ , which implies  $d_{\alpha, \beta, \lambda}(R^j, D(R)) = 0$ ; meanwhile, we have  $\text{non}(x_1 D(R_{ij}) x_2)$  since

$$(N(x_1, x_2, R_{ij}), N^0(x, y, R_{ij}), N(x_2, x_1, R_{ij})) = (S_1 \setminus \{j\} \cup \{i\}, S_2 \setminus \{i\} \cup \{j\}, S_3) \notin W.$$

It follows that  $d_{\alpha,\beta,\lambda}(R^j, D(R_{ij})) \geq \beta > 0 = d_{\alpha,\beta,\lambda}(R^j, D(R))$ .

**Subcase 3 :** There exists  $(S_1, S_2, S_3) \in 3^N$  such that: 
$$\begin{cases} (S_1, S_2, S_3) \in W \\ j \in S_2, i \in S_3 \text{ and} \\ (S_1, (S_2 \setminus \{j\}) \cup \{i\}, (S_3 \setminus \{i\}) \cup \{j\}) \notin W. \end{cases}$$

- If  $((S_3 \setminus \{i\}) \cup \{j\}, (S_2 \setminus \{j\}) \cup \{i\}, S_1) \in W$ , then let  $T_1 = (S_3 \setminus \{i\}) \cup \{j\}$ ,  $T_2 = (S_2 \setminus \{j\}) \cup \{i\}$  and

$T_3 = S_1$ . From the definition of a voting rule with abstention, we have  $(T_1, T_2, T_3) \in W$ ,  $j \in T_1, i \in T_2$

and  $((T_1 \setminus \{j\}) \cup \{i\}, (T_2 \setminus \{i\}) \cup \{j\}, T_3) = (S_3, S_2, S_1) \notin W$ . Therefore,  $(T_1, T_2, T_3)$  satisfies the assumption of subcase 2 that is already solved.

- Assume that  $(S_3 \setminus \{i\} \cup \{j\}, S_2 \setminus \{j\} \cup \{i\}, S_1) \notin W$ . Consider the following preference profile  $Q$ :

$$\begin{cases} \forall k \in S_1 : & x_2 \succ_{Q^k} x_1 \succ_{Q^k} x_3 \succ_{Q^k} \dots \succ_{Q^k} x_m \\ \forall k \in S_2 \setminus \{j\} \cup \{i\} : & x_1 \simeq_{Q^k} x_2 \succ_{Q^k} x_3 \succ_{Q^k} \dots \succ_{Q^k} x_m \\ \forall k \in S_3 \setminus \{i\} \cup \{j\} : & x_1 \succ_{Q^k} x_2 \succ_{Q^k} x_3 \succ_{Q^k} \dots \succ_{Q^k} x_m \end{cases}$$

Clearly, we have  $(N(x_1, x_2, Q), N^0(x_1, x_2, Q), N(x_2, x_1, Q)) = ((S_3 \setminus \{i\}) \cup \{j\}, (S_2 \setminus \{j\}) \cup \{i\}, S_1) \notin W$

and  $(N(x_2, x_1, Q), N^0(x_2, x_1, Q), N(x_1, x_2, Q)) = (S_1, (S_2 \setminus \{j\}) \cup \{i\}, (S_3 \setminus \{i\}) \cup \{j\}) \notin W$ . This proves

that  $x_2 *_{D(Q)} x_1$  (that is,  $x_2$  and  $x_1$  are socially incomparable). We therefore obtain:

$$x_2 *_{D(Q)} x_1 \succ_{D(Q)} x_3 \succ_{D(Q)} \dots \succ_{D(Q)} x_m; \text{ hence } d_{\alpha,\beta,\lambda}(Q^j, D(Q)) = \beta.$$

Considering  $Q_{ij}$ , we have  $x_2 \succ_{D(Q_{ij})} x_1$  because  $(S_1, S_2, S_3) \in W$ . More generally,

$$x_2 \succ_{D(Q_{ij})} x_1 \succ_{D(Q_{ij})} x_3 \succ_{D(Q_{ij})} \dots \succ_{D(Q_{ij})} x_m, \text{ which implies } d_{\alpha,\beta,\lambda}(Q^j, D(Q_{ij})) = 1 > \beta = d_{\alpha,\beta,\lambda}(Q^j, D(Q)).$$

**Conclusion :** If  $\beta < 1$ , then  $\geq_T$  and  $\geq_{P_{\alpha,\beta,\lambda}}$  are equivalent for all society  $(N, A, W)$ .

**Second part :** We now assume that  $\geq_T$  and  $\geq_{P_{\alpha,\beta,\lambda}}$  are equivalent for all society  $(N, A, V)$  (or  $(N, A, W)$ ) and would like to show that  $\beta < 1$ . We proceed by contradiction. Assume that  $\beta = 1$ , and show that there exists at least one society  $(N, A, W)$  for which  $\geq_{P_{\alpha,\beta,\lambda}}$  and  $\geq_T$  are not equivalent.

Consider  $N = \{1, 2, 3, 4, 5\}$ ,  $A = \{x, y\}$  and  $W_m = \{(1, 25, 34), (14, 2, 35), (12, 4, 35)\}$ .

In this game, we have  $non(4 \geq_T 5)$  because  $(1, 25, 34) \in W$  but  $(1, 24, 35) \notin W$  (in fact,  $5 >_T 4$ ). We show below that  $4 \geq_{P_{\alpha,\beta,\lambda}} 5$ .

Let  $R$  be a preference profile. Let us prove that  $d_{\alpha,\beta,\lambda}(R^5, D(R_{45})) \leq d_{\alpha,\beta,\lambda}(R^5, D(R))$ . Without loss of generality, two cases are possible for player 5: either  $x \succ_{R^5} y$  or  $x \simeq_{R^5} y$ .

**First case:**  $x \succ_{R^5} y$ .

- If  $y \succ_{D(R^N)} x$ , then the conclusion is straightforward.

- If  $x *_{D(R)} y$ , then the conclusion follows since  $d_{\alpha,\beta,\lambda}(R^5, D(R)) = \beta = 1$ .

- If  $x \succ_{D(R)} y$ , then  $d_{\alpha,\beta,\lambda}(R^5, D(R)) = 0$  and  $(N(x, y, R), N^0(x, y, R), N(y, x, R)) \in W$ . By the definition of  $W$ , we have  $1 \in N(x, y, R)$ . So  $5 \in N(x, y, R)$ .

\* If  $4 \in N(x, y, R)$ , then  $x \succ_{D(R_{45})} y$  and therefore,  $d_{\alpha, \beta, \lambda}(R^5, D(R_{45})) = 0$ .

\* If  $4 \in N^0(x, y, R)$ , then

$$(N(x, y, R_{45}), N^0(x, y, R_{45}), N(y, x, R_{45})) = ((N(x, y, R_{45}) \setminus \{5\}) \cup \{4\}, (N^0(x, y, R_{45}) \setminus \{4\}) \cup \{5\}, N(y, x, R_{45})).$$

Since  $(N(x, y, R), N^0(x, y, R), N(y, x, R))$  contains a tripartition  $(S_1, S_2, S_3) \in W_m$ , from the definition of  $W$ ,

$$((N(x, y, R_{45}) \setminus \{5\}) \cup \{4\}, (N^0(x, y, R_{45}) \setminus \{4\}) \cup \{5\}, N(y, x, R_{45})) \text{ contains the same tripartition } (S_1, S_2, S_3);$$

this means that  $x \succ_{D(R_{45})} y$  (for example, if  $(N(x, y, R), N^0(x, y, R), N(y, x, R)) = (15, 24, 3)$ , then it

$$\text{contains } (1, 25, 34). \text{ We have } ((N(x, y, R_{45}) \setminus \{5\}) \cup \{4\}, (N^0(x, y, R_{45}) \setminus \{4\}) \cup \{5\}, N(y, x, R_{45})) = (14, 25, 3)$$

which still contains the same tripartition  $(1, 25, 34)$  and consequently,  $d_{\alpha, \beta, \lambda}(R^5, D(R_{45})) = 0$ ).

\*If  $4 \in N(y, x, R)$ , then

$$(N(x, y, R_{45}), N^0(x, y, R_{45}), N(y, x, R_{45})) = ((N(x, y, R_{45}) \setminus \{5\}) \cup \{4\}, N^0(x, y, R_{45}), (N(y, x, R_{45}) \setminus \{4\}) \cup \{5\}).$$

Again, similarly to the argument above, we obtain  $x \succ_{D(R_{45})} y$  and thus  $d_{\alpha, \beta, \lambda}(R^5, D(R_{45})) = 0$ .

We conclude that in this case,  $d_{\alpha, \beta, \lambda}(R^5, D(R_{45})) \leq d_{\alpha, \beta, \lambda}(R^5, D(R))$ .

**Second case:**  $x \simeq_{R^5} y$ .

In this case, depending on the comparison of  $x$  and  $y$  by  $D(R)$ , we have either  $d_{\alpha, \beta, \lambda}(R^5, D(R)) = \alpha$  or

$d_{\alpha, \beta, \lambda}(R^5, D(R)) = \lambda$ . But since  $\lambda = \alpha$  by assumption, it is the case that  $d_{\alpha, \beta, \lambda}(R^5, D(R)) = \alpha$ . No matter how  $D(R_{45})$  compares  $x$  and  $y$ , we get  $d_{\alpha, \beta, \lambda}(R^5, D(R_{45})) = \alpha$ .

In both cases, we obtain  $d_{\alpha, \beta, \lambda}(R^5, D(R_{45})) \leq d_{\alpha, \beta, \lambda}(R^5, D(R))$ , that is  $4 \geq_{P_{\alpha, \beta, \lambda}} 5$ . ■

### Proof of Proposition 1

Let  $\alpha, \beta, \lambda \in ]0, 1]$ . Let us prove that for any political tournament,  $\geq_T \subseteq \geq_{P_{\alpha, \beta, \lambda}}$ . Assume a tournament  $(N, A, \mathcal{P}_2(A), R, V, D)$  and  $i, j \in N$  such that  $i \geq_T j$ . The proof that  $i \geq_{P_{\alpha, \beta, \lambda}} j$  directly follows from part **a**) of the **first case** of the proof of Theorem 4. In that part, we show that if for any society  $(N, A, V)$ ,  $i \geq_T j$ , then  $\geq_T \subseteq \geq_{P_{\alpha, \beta, \lambda}}$  at any preference profile  $R$  defined on  $A$ . (Note that in that part of the proof, we did not need to assume that  $\beta < 1$ .)

Now, we would like to show a tournament  $(N, A, \mathcal{P}_2(A), R, V, D)$  such that there are two players  $i, j \in N$  such that  $i \geq_{P_{\alpha, \beta, \lambda}} j$  but  $\text{not}(i \geq_T j)$  and  $\text{not}(j \geq_T i)$ . Consider the following tournament where  $N = \{1, 2, 3\}$ ,  $A = \{x, y\}$ ,  $R = (xy, yx, (xy))$ , and  $W$  is such that  $W_m = \{(1, 3, 2), (3, 2, 1)\}$ . Consider players 1 and 2. We have:

$\text{not}(2 \geq_T 1)$  because  $(1, 3, 2) \in W$  and  $(2, 3, 1) \notin W$ ;

$\text{not}(1 \geq_T 2)$  because  $(3, 2, 1) \in W$  and  $(3, 1, 2) \notin W$ .

Also,  $D(R) = xy$  and  $D(R_{12}) = yx$ , which yields:

$$d(R^2, D(R)) = 1 \text{ and } d(R^2, D(R_{12})) = 0, \text{ and therefore } d(R^2, D(R_{12})) < d(R^2, D(R)).$$

It follows that  $1 >_{P_{\alpha, \beta, \gamma}} 2$ .

We have shown that for any political tournament  $(N, A, \mathcal{P}_2(A), R, V, D)$ ,  $\geq_T \subseteq \geq_{P_{\alpha, \beta, \lambda}}$ , and that this inclusion may be strict. ■

### Proof of Theorem 5

Let  $\alpha, \beta, \lambda \in ]0, 1]$  such that  $\beta < 1$  and  $(N, A, V)$  be a society.

If  $V$  is swap robust, then it follows from Lemma 1 that the influence relation  $\geq_T$  is complete and transitive. Given that  $\beta < 1$ , it follows from Theorem 4 that  $\geq_T$  and  $\geq_{P_{\alpha, \beta, \lambda}}$  are equivalent. We then conclude that  $\geq_{P_{\alpha, \beta, \lambda}}$  is complete and transitive.

We prove the converse implication by contradiction. Assume that  $V$  is not swap-robust. It follows from Lemma 1 that  $\geq_T$  is not complete. Since  $\beta < 1$ , it follows from Theorem 4 that  $\geq_{P_{\alpha, \beta, \lambda}}$  and  $\geq_T$  are equivalent, which implies that  $\geq_{P_{\alpha, \beta, \lambda}}$  is not complete, which in turn logically implies that  $\geq_{P_{\alpha, \beta, \lambda}}$  is not complete or transitive. ■

### Proof of Theorem 6

Let  $\alpha, \beta, \lambda \in ]0, 1]$  be such that  $\alpha = \lambda$ , and let  $(N, A, V)$  be a society.

1- Assume that  $\beta < 1$  and  $V$  is swap-robust.

Thanks to Theorem 4,  $\geq_{P_{\alpha, \beta, \lambda}}$  and  $\geq_T$  are equivalent because  $\beta < 1$ .

From Lemma 2,  $\geq_T$ ,  $\geq_S$  and  $\geq_B$  are equivalent because  $V$  is swap-robust. It follows that  $\geq_{P_{\alpha, \beta, \lambda}}$ ,  $\geq_S$  and  $\geq_B$  are equivalent.

2- Assume that  $\beta = 1$  or that  $V$  is not swap-robust. We want to show that  $\geq_{P_{\alpha, \beta, \lambda}}$ ,  $\geq_S$  and  $\geq_B$  are not equivalent in general. We will show that there exists at least one society for which they do not coincide. We distinguish two cases below.

**First case:** Assume  $\beta = 1$ .

Consider the following society taken from the proof of Theorem 4:

$N = \{1, 2, 3, 4, 5\}$ ,  $A = \{x, y\}$ , and  $W$  is such that  $W_m = \{(1, 25, 34), (14, 2, 35), (12, 4, 35)\}$  (the set of minimal winning tripartitions).

It can be easily checked that the influence relation gives the following ranking of voters:

$$1 >_T 2 >_T 5 >_T 4 >_T 3.$$

It follows that  $W$  is swap-robust (Lemma 1). From Lemma 2, this implies that the Shapley-Shubik and Banzhaf power indices give the following rankings:

$$1 >_S 2 >_S 5 >_S 4 >_S 3 \text{ and } 1 >_B 2 >_B 5 >_B 4 >_B 3.$$

However  $4 \geq_{P_{\alpha, \beta, \lambda}} 5$ . Therefore,  $\geq_{P_{\alpha, \beta, \lambda}}$ ,  $\geq_S$  and  $\geq_B$  are not equivalent.

**Second case:** Assume that  $\beta < 1$  and  $V$  is not swap-robust.

It follows from Theorem 5 that  $\geq_{P_{\alpha,\beta,\lambda}}$  is not complete or transitive. Therefore,  $\geq_{P_{\alpha,\beta,\lambda}}$  cannot be equivalent to  $\geq_S$  or  $\geq_B$  since  $\geq_S$  and  $\geq_B$  are complete and transitive. This also proves the second sentence of item 2) of Theorem 6. ■

### **Proof of Corollary 1**

Let  $\alpha, \beta, \lambda \in ]0, 1]$  such that  $\alpha = \lambda$ . Assume that  $\beta < 1$  and show that  $\geq_{P_{\alpha,\beta,\lambda}}, \geq_S$  and  $\geq_B$  are equivalent for all linear society  $(N, A, V)$ .

Let  $(N, A, V)$  be a linear society. By definition,  $V$  is swap-robust. It follows from item 1) of Theorem 6 that  $\geq_{P_{\alpha,\beta,\lambda}}, \geq_S$  and  $\geq_B$  are equivalent.

Conversely, assume by contradiction that  $\beta < 1$ . Then we need to show that there exists at least one linear society for which  $\geq_{P_{\alpha,\beta,\lambda}}, \geq_S$  and  $\geq_B$  are not equivalent. This is immediate from part 2) and the first case of the proof of Theorem 6. ■

## References

- [1] Aguiar, V. H., Pongou, R., Serrano, R., and Tondji, J. B. (2020). An Index of Unfairness. in Handbook of the Shapley Value, E. Algava et al. (eds.), CRC Press Taylor and Francis.
- [2] Aguiar, V. H., Pongou, R., and Tondji, J. B. (2018). A non-parametric approach to testing the axioms of the Shapley value with limited data. *Games and Economic Behavior*, 111, 41-63.
- [3] Arrow. K.J. (1951): Social Choice and Individual Values. John Wiley.
- [4] Bahel, E., and Y. Sprumont (2020): Strategyproof choice of social acts. *American Economic Review*, 110, 596-627.
- [5] Banzhaf, J. F. (1965) : Weighted voting doesn't work : A mathematical Analysis. *Rutgers Law Review*, 19, 317-343.
- [6] Barbie, M., Puppe, C., and Tasnadi, A. (2006): Non-manipulable domains for the Borda count. *Economic Theory*, 27(2), 411-430.
- [7] Can, B., and T. Storcken (2018): A re-characterization of the Kemeny distance. *Journal of Mathematical Economics*, 79, 112-116.
- [8] Cramer, G. (1750): Introduction à l'analyse des lignes courbes algébriques. Frères Cramer et C. Philibert, Genève.
- [9] Dasgupta, P., and E. Maskin (2008): On the robustness of majority rule. *Journal of the European Economic Association*, 6(5), 949-973.
- [10] Dasgupta, P., and E. Maskin (2020): Strategy-proofness, IIA, and majority rule. *American Economic Review: Insights*, forthcoming.
- [11] Dowding, K. M. (1991): *Rational choice and political power*. Aldershot, Hampshire, England Brookfield, Vermont, USA: Edward Elgar.
- [12] Difo Lambo., L. and Moulen J. (2000): Quel pouvoir mesure-t-on dans un jeu de vote ? *Mathématiques & Sciences humaines*, 152, 27-47.
- [13] Difo Lambo. L., Tchantcho, B., and Moulen, J. (2012): Comparing influence theories in voting games under locally generated measures of dissatisfaction. *International Journal of Game Theory*, 41, 719-731.
- [14] Felsenthal, D. S., and Machover M. (1997): Voting games with abstention. *International Journal of Game Theory*, 26, 335-351.
- [15] Felsenthal, D. S., and Machover M. (1998): *The measurement of Voting Power. Theory and Practice, Problems and Paradoxes*. Edward Elgar Publishers, London.
- [16] Freixas, J., and Parker, C. (2015): Manipulation in games with multiple levels of output. *Journal of Mathematical Economics*, 61, 144-151.
- [17] Freixas, J., and Zwicker W.S. (2003): Weighted voting, abstention, and multiple levels of approval. *Social Choice and Welfare*, 21, 399-431.
- [18] Freixas, J., and Zwicker W.S. (2009): Anonymous yes-no voting with abstention and multiple levels of approval. *Games and Economic Behavior*, 67, 428-444.

- [19] Gibbard, A. (1973): Manipulation of voting scheme: a general result. *Econometrica*, 41, 587-601.
- [20] Guemmegne, J. and Pongou, R. (2014): A policy-based rationalization of collective rules: Dimensionality, specialized houses, and decentralized authority. *Journal of Mathematical Economics*, 52, 182-193.
- [21] Isbell, J. R., (1958): A class of simple games. *Duke Mathematical Journal*, 25, 423-439.
- [22] Hobbes, T. (1651): *Leviathan: or, the matter, forme and power of commonwealth, ecclesiasticall and civil*. Touchstone.
- [23] Kemeny, J. (1959): Mathematics without numbers. *Daedalus*, 88 (4), 577-591.
- [24] Kendall, M. (1938): A new measure of rank correlation. *Biometrika*, 30(1/2), 81-93.
- [25] Kruskal, W.H. (1958): Ordinal measures of association. *Journal of the American Statistical Association*, 53(284), 814-861.
- [26] Laruelle, A., and Valenciano, F. (2005): Assessing success and decisiveness in voting situations. *Social Choice and Welfare*, 24, 171-197.
- [27] Laruelle, A., Martinez, R., and Valenciano, F. (2006): Success Versus Decisiveness. *Journal of Theoretical Politics*, 18, 185-205.
- [28] Maskin, E. (2020): Arrow's IIA Condition, May's Axioms, and the Borda Count. Working paper.
- [29] Parker, C. (2012): The influence relation for ternary voting games. *Games and Economic Behavior*, 75, 867-881.
- [30] Pongou, R., Difo Lambo, L., and Tchantcho, B. (2008): Cooperation, stability and social welfare under majority rule. *Economic Theory*, 35, 555-574.
- [31] Pongou, R. Tchantcho, B. and Difo Lambo. L. (2011): Political influence in multi-choice institutions: Cyclicity, anonymity and transitivity. *Theory and Decision*, 70, 157-178.
- [32] Pongou, R., Tchantcho, B., and Tedjeugang, N. (2014): Power theories for multi-choice organizations and political rules: Rank-order equivalence. *Operations Research Perspectives*, 1, 42-49.
- [33] Pongou, R., Tchantcho, B., and Tedjeugang, N. (2017): Properties of Ladder Tournaments. *European Journal of Operational Research*, 263, 203-213.
- [34] Pongou, R., and Tondji, J. B. (2018). Valuing inputs under supply uncertainty: The Bayesian Shapley value. *Games and Economic Behavior*, 108, 206-224.
- [35] Rubinstein, R. (1980): Stability of decision systems under majority rule. *Journal of Economic Theory*, 23, 150-159.
- [36] Satterthwaite, M.A. (1975): Strategy-proofness and Arrow's conditions: existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10, 187-217.
- [37] Sen, A.K. (1974): Choice, orderings and morality. In: Körner, S. (ed) *Practical reason*. Oxford, 54-67.
- [38] Shapley, L. S. and Shubik, M. (1954): A model for Evaluating the Distribution of Power in a Committee System. *American Political Science Review*, 48, 787-792.
- [39] Taylor, A. T. and Zwicker, W. (1999): *Simple games: Desirability relations, trading, pseudoweightings*. Princeton University Press.

- [40] Tchantcho, B. Dikko Lambo, L. Pongou, R., and Mbama Engoulou, B. (2008): Voters' power in voting games with abstention: Influence relation and ordinal equivalence of power theories. *Games and Economic Behavior*, 64, 335-350.