Accountable Voting*

Takako Fujiwara-Greve† Yoko Kawada‡ Yuta Nakamura§
Noriaki Okamoto¶

April 1, 2023

Abstract

We consider social decision-making situations where some voters may have a conflict of interest in some social alternatives/applicants to be ranked. Then the accountability of the collective decision process is essential. We extend the standard social welfare function to include the interest relationships between the voters and the applicants. We introduce two accountability axioms: *Exclusion of Interested Party Evaluations* on the voters' individual preference manifests and the *No-Power-Game property* for changes in the interest relationships. While we obtain three impossibility theorems among the accountability axioms and the extensions of some standard axioms in voting theory, we also give two directions for remedies.

JEL classification number: D71; D63.
Key words: accountability, interested party, social welfare function, impossibility, plurality rule.

---

*We are grateful to Koji Yokote, Steffen Lippert, Salvador Barberà, Arkadii Slinko, Toru Hokari, Geir Asheim and seminar participants at the Summer Meeting on Game Theory 2020 (online) and Keio University for useful comments. This work was supported by JSPS KAKENHI Grant Numbers 19K23200 (Nakamura) and 20K13456 (Okamoto). Declarations of interest: none.

†Corresponding author. Department of Economics, Keio University, Tokyo 108-8345, Japan. Email: tgreve@keio.jp.

‡Faculty of Economics, Komazawa University, Tokyo 154-8525, Japan. Email: ykawada@komazawa-u.ac.jp.

§School of Economics and Business Administration, Yokohama City University, Kanagawa 236-0027. Email: y_naka@yokohama-cu.ac.jp.

¶Department of Economics, Meiji Gakuin University, Tokyo 108-8636, Japan. Email: nokamoto@eco.meijigakuin.ac.jp.

1. Introduction

1.1. Motivation and the outline of the results

One of the authors recently experienced the following problem at a national research fund committee. A committee consisting of eight members needed to rank about twenty applicants, but two of the members were closely related to two different applicants. They had either a co-author relationship or an advisor-advisee relationship with an applicant. The protocol of this committee prescribes that such a member must leave the room when others discuss the evaluation of the related applicant. Probably, the "leaving-the-room" rule is to address the accountability of the committee.[1] The committee members who have a related applicant cannot manipulate the discussion. Thus all applicants and the taxpayers (the sponsors) are assured that relationships with the committee members do not matter. Unfortunately, the protocol did not specify how to vote if the discussions do not lead to a collective decision. If each committee member is allowed to submit a ranking of all applicants, including the related ones, how the interest structure be taken care of? If some members give only a ranking over a subset of the applicants, how should the rankings over different sets be aggregated? This paper attempts to find accountable voting rules[2] that deter partisanship when some voters are biased on some social alternatives but allow all voters to express their opinions over unrelated alternatives.

The standard axiomatic social choice theory (e.g., Arrow, 1951) did not need to address accountability because the set of voters equals the society. Hence all affected parties are equally included in the social decision process. In many real-world problems, however, a subset of society, such as experts and referees (voters, henceforth), chooses an outcome that affects the entire society. Moreover, since the voters should be able to evaluate the alternatives/applicants (applicants, henceforth) well, they may be "close" to some applicants, as in the above example.[3]

[1] As Mulgan (2000) explains, the scope and meaning of "accountability" has been extended in a number of directions well beyond its core sense of being called to account for one's actions to some authority. In this paper, we consider accountability of collective-decision processes (not of individuals). Our notion of accountability is similar to the second extension in Mulgan (2000): a feature of institutional checks and balances (accountability as 'control').

[2] Another approach is the principal-agent model that designs contracts for each delegate to reward/punish actions to enforce accountability. We design collective decision rules, and we do not need to know the preferences of each delegate.

[3] Examples abound. In many organizations, the hiring or promotion decision of highly skilled positions

Then society should require accountability of the collective decision process that warrants fair treatment of all applicants regardless of how they are related to the voters.

The social welfare functions (Arrow, 1951) collect the individual preference orders of the voters and rank the applicants. We extend this notion to have two domains, the preference profile domain and the set of possible interest structures. Each interest structure specifies which voter has an interest in which applicant or not and can be identified by a matrix of 0 and 1. We call the extended notion *social welfare functions with interest structure* (SWFI).

We introduce two accountability axioms for the two domains of the SWFIs. First, if a voter changes her/his preference order of the applicant(s) of her/his interest relative to other applicants, that should not affect the social outcome (*Exclusion of Interested Party Evaluations*; henceforth *Interest-Exclusion*[4]). This is an inter-preference-profile requirement and incorporates the idea of the "leaving-the-room" protocol of the first example.

Second, suppose that a voter changes her/his status from neutral to an interested party of some applicant. In that case, it should not improve the social ranking of the applicant (*No-Power-Game property*). This is an inter-interest-structure requirement.

The benefit of our accountability axioms includes the following effects. First, under *Interest-Exclusion*, the outcome is not a direct consequence of the interested voter's preference revelation. Hence the related parties can maintain a good relationship regardless of the collective decision outcome. This feature makes it easy to solicit volunteer voters/referees as well. Second, under the *No-Power-Game property*, the applicants do not need to expend effort to increase interested voters and can concentrate on the correct attempt to improve their qualities. Third, competitions with an accountable decision-making process can attract entrants to become applicants/contestants because they are assured not to be disadvantaged by being unknown persons to the voters.

Note that we focus on situations where voters may have verifiable biased relationships

requires a small number of experts who may be closely related to some of the applicants/candidates because they are trained in the same field. Universities regularly encounter problems that advisors are included in the committee to evaluate students for scholarships, etc. Conflict of interest also exists when the social alternatives are not humans, e.g., when new public construction projects need to be ranked and the voters/experts live in the affected regions. Our formulation includes such cases.

[4]The axiom is a generalization of the "self-exclusion" property in the peer rating literature (e.g., Ng and Sun, 2003).

(interest) with some applicants. Our focus is because it is impossible to make voters reveal unobservable interest relationships in the simple one-shot voting context. Nevertheless, our model still covers many essential problems. In the academic world, advisor-advisee and co-author relationships are verifiable and positively biased. Referees and contestants from the same country are similar examples in international competitions (e.g., sports, music). Our model includes verifiable negative interest relationships, such as people related to archrival sports teams and countries. The interest structure of the peer rating/selection models (e.g., Ng and Sun (2003), Ohseto (2012), Holzman and Moulin (2013), Tamura and Ohseto (2014), Mackenzie (2015), Aziz et al. (2019), Edelman and Por (2021), and Alcalde-Unzu et al. (2022)) is trivially verifiable. The underlying assumption in this literature is that the set of applicants and the set of voters are the same, and the only interested party of each applicant is herself/himself.

Our main results include both possibility and impossibility theorems. Many score-based SWFIs satisfy both of the accountability axioms. However, no SWFI satisfies the *No-Power-Game property* and a variant of the axiom of Unanimity (Arrow, 1951), which we call *A-Unanimity* (Theorem 1). A much weaker axiom than *A-Unanimity* is still not compatible with the two accountability axioms and some standard axioms in voting theory, e.g., *No-Pairwise-Loser property*, a.k.a. *Condorcet-Loser criterion*[5] (Theorem 2). Therefore, accountability and some standard axioms are hard to sustain jointly.

A possible cause of the impossibilities is that an SWFI must take any *non-trivial* interest structure[6] as an input. Careful choice of voters can control the set of possible interest structures. Thus we consider restricting the domain of the set of feasible interest structures of an SWFI. Theorem 1 continues to hold even if we require that, for any two distinct applicants, at least k voters must exist who have no interest in them, as long as $k \leqq [\frac{n}{2}]$ (where n is the number of voters). However, when $k \geqq [\frac{n}{2}] + 1$, there is an SWFI that satisfies both the *No-Power-Game property* and *A-Unanimity*. Hence a solution to the impossibility result is to keep the set of voters to those with few related applicants. (However, one impossibility result remains even if most voters have no conflict of interest. Therefore, the impossibility results are not dependent

[5]The name is based on Borda (1784) and Condorcet (1784). McLean and Hewitt (1994) and McLean and Urken (1995) give the English translations of these articles.

[6]For any two distinct applicants, at least one voter does not have a conflict of interest with them.

on the large domain of all non-trivial interest-structures.)

Another direction for a remedy is to weaken one of the accountability axioms. *Interest-Exclusion* is easy to satisfy, and the peer rating/selection literature considers it a fundamental requirement. Therefore, we weaken the *No-Power-Game property.* We show that the *Weak No-Power-Game property* is compatible with all other axioms of Theorem 2 by explicitly constructing an extension of the Borda rule (which we call *the winning rate rule*). The *Weak No-Power-Game property* requires that if a voter changes the status from neutral to an interested party of an applicant and if this voter's most preferred alternative was this applicant, it would not improve the social ranking of the relevant applicant. This property is still a desirable requirement, and thus the winning rate rule is a reasonable solution.[7]

1.2. Related literature

Since we introduce a new framework, there is no directly comparable previous work. Our motivation is close to that of the peer rating/selection literature cited above. However, our model and our impossibility theorem structure differ substantially from those of the literature. In the peer rating/selection literature, the set of voters is the set of social alternatives, and each voter is the unique interested party of one social alternative, herself/himself. By contrast, we deal with arbitrary sets of social alternatives and arbitrary interest structures between the voters and the alternatives. The axiom configurations in the impossibility results of this paper and the literature are not directly comparable because we require the *No-Power-Game property*, which is defined on the set of possible interest structures. This set is degenerate in the peer rating/selection models.[8]

The peer rating models consider social ranking problems by the voters who are also the candidates to be ranked. Ng and Sun (2003) and Ohseto (2012) assume that voters submit cardinal rankings of themselves, and the social decision is a cardinal ranking. These papers

[7]In the working paper version (Fujiwara-Greve et al., 2021), we examine the properties of other concrete SWFIs in terms of the axioms introduced in this paper. We show that taking the average for the 0-1 matrix representing the interest structure is the key to mitigating the unequal influence of interested voters.

[8]This comparison resembles the contrast between the Arrovian social welfare function (Arrow, 1951) and the Bergson-Samuelson social welfare function (e.g., Kemp and Ng, 1976). Our remedy is in the same spirit as the Arrovian literature to restrict the domain (e.g., Kalai and Muller, 1977) or to weaken the requirements. For a survey, see Suzumura (1983) Chapter 3.

require "self-exclusion" such that the social ranking does not change if the evaluation marks over other candidates (who are not oneself) do not change for all voters. In the peer rating model, this requirement is the same as our *Interest-Exclusion*. The peer selection models (e.g., Holzman and Moulin (2013), Mackenzie (2015), and Edelman and Por (2021)) consider social choice problems, i.e., the voters choose one of them as the winner. The voters are not allowed to nominate themselves. Thus our *Interest-Exclusion* is embedded in their model.

Holzman and Moulin (2013) showed an impossibility between impartiality (similar to strategyproofness) and anonymity or some kind of unanimity.[9] Mackenzie (2015) resolves the impossibility by allowing stochastic winners and characterizes the recommended uniform random dictator rule. Edelman and Por (2021) allow variable populations and stochastic winners to characterize the same rule. Instead of stochasticity, Tamura and Ohseto (2014) allow a set of voters to become winners to resolve the impossibility. Recall that our aim is not to make voting rules strategyproof. Instead, we want to deter biases on the social rankings by interested voters. Hence we propose rating rules to dilute the inputs from voters interested in some applicants (see Sections 4 and 5).

Biased voter problems are analyzed in other contexts as well. For example, Amorós (2020) deals with a dual-preference problem such that voters (experts) may have biased preferences over the outcomes in addition to their true evaluations of the alternatives. While we aim at the elimination of biases on the social rankings, Amorós (2020) investigated whether a mechanism that always results in the "true" majoritarian (or unanimity) outcome is implementable.[10]

Accountability is also addressed in political economy literature. This literature usually considers the principal-agent model, and the issue is to control individual politicians/delegates to act correctly for society. Instead of trying to manage individual voters, we design accountable collective decision rules that work for arbitrary preference orders and interest relationships.

This paper is organized as follows. Section 2 gives the basic setup of the voting problems

[9] Alcalde-Unzu et al. (2022) generalize the impossibility result of Holzman and Moulin (2013) to the peer rating model and allow arbitrary message spaces.

[10] There are also game-theoretic papers on vote transparency (see Name-Correa and Yildirim, 2019, and the references therein), which explicitly consider strategic voting games by biased voters (experts). This literature fixes the voting rule (e.g., the unanimity rule) and investigates the effects of external factors such as social pressure and reputation to prevent biased outcomes. By contrast, we investigate the existence of voting rules with desirable properties.

with interest structure and introduces axioms. Section 3 gives impossibility theorems. In Section 4, we find a restriction of the domain of the feasible interest structures to resolve the impossibility between the *No-Power-Game property* and *A-Unanimity.* Section 5 weakens the *No-Power-Game property* as another remedy. Section 6 concludes. Important proofs are given in the text, while analogous proofs are in Appendix.

2. Model and axioms

2.1. Voting problems

Let $N = \{1, 2, \ldots, n\}$ be a finite set of *voters* or *committee members* and assume $n \geqq 2$. There is a finite set $X = \{x_1, x_2, \ldots, x_j, \ldots, x_m\}$ of *applicants* (alternatives) to be ranked, and we sometimes use $x, y, z \in X$ as typical elements as well. We assume $m \geqq 3$. Denote by $\mathcal{L}$ (resp. $\mathcal{R}$) the set of all strict orders (resp. weak orders) on X. We assume that each voter $i \in N$ has some strict order $\succ_i \in \mathcal{L}$.

In this paper, an interest relationship is a verifiable connection between a voter and an applicant such that society worries that the voter has a biased opinion of the applicant. For example, in an academic context, a voter i can be said to have an interest in an applicant x_j if i and x_j are co-authors or i was an advisor of x_j. We do not consider the elicitation of unverifiable interests. Instead, we want to eliminate biases on the social ranking when having some voters with connections to some applicants may be unavoidable. We do not restrict our attention to positive interest (such that a voter likes to see a high social ranking of her/his related applicant(s)). If a negative interest relationship (such as archrival team members in a sports context) is verifiable, our model includes it.

An *interest structure* is a matrix $A \in \{0, 1\}^{n \times m}$ such that, if $a_{ij} = 1$, then there is no interest between the voter i and the applicant x_j, and, if $a_{ij} = 0$, then the voter i has some interest in x_j (see Figure 1). The 1-0 structure is for mathematical convenience when we want to take only relationships with no conflict of interest into account (see Section 5, Definition 11).

Given an interest structure $A \in \{0, 1\}^{n \times m}$, for any $i \in N$, let

$$A(i) := \{x_j \in X : a_{ij} = 1\}$$

$$A = \begin{array}{c} \text{voters}\backslash\text{applicants} \\ 1 \\ 2 \\ \vdots \\ n \end{array} \begin{array}{c} \begin{array}{cccc} x_1 & x_2 & \cdots & x_m \end{array} \\ \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} \end{array}$$

Figure 1: Interest structure matrix

be the set of "admissible" applicants whom i has no interest in and it is safe to allow i to express her/his opinion freely on $A(i)$. We also say that voter i is "unrelated" (or neutral) to an applicant x_j if $a_{ij} = 1$. From the voters' point of view, for any $A \in \{0,1\}^{n\times m}$ and any $x_j \in X$, let

$$I(x_j; A) := \{i \in N : a_{ij} = 0\}$$

be the set of x_j's "interested voters" under the interest structure A.

We focus on interest structures such that any pair of applicants have at least one voter who can compare them without worrying about the conflict of interest.

Definition 1. An interest structure $A \in \{0,1\}^{n\times m}$ is *non-trivial* if, for any distinct $x_j, x_k \in X$, there exists $i \in N$ such that $\{x_j, x_k\} \subset A(i)$. Let $\mathcal{A} \subset \{0,1\}^{n\times m}$ be the set of all non-trivial A's.

We use the notation $\mathbf{1}_{n,m}$ for the all-ones, $n \times m$ matrix. This is the "ordinary" situation, and clearly $\mathbf{1}_{n,m} \in \mathcal{A}$. We extend social welfare functions to the domain $\mathcal{L}^N \times \mathcal{A}$.

Definition 2. A *social welfare function with interest structure (SWFI)* is a function f such that

$$f : \mathcal{L}^N \times \mathcal{A} \to \mathcal{R}.$$

Our model has the following features to compare with the peer rating/selection models. First, we only require that each voter has an ordinal preference relation over X. Second, we do not restrict our attention to the case that the set of the voters is the set of applicants, nor the interest structure such that each voter has a single applicant (herself/himself) of interest. Third, we do not need a predetermined number of winners since a social welfare function ranks all applicants.

2.2. Accountability axioms

Since a social welfare function has two sets $\mathcal{L}^N \times \mathcal{A}$ in the domain, we define accountability axioms for each. Our first axiom for accountability is on the preference domain such that the social ranking does not change as long as each voter's preferences on the admissible/unrelated applicants $A(i)$ do not change. In other words, no matter how hard a voter pushes her/his "related" applicants, that should not affect the social ranking.[11] If this axiom is violated, a voter changes the social decision by manipulating her/his ranking of the applicant of the interest.

For any $i \in N$, any $\succ_i \in \mathcal{L}$, and any nonempty $S \subset X$, let $\succ_i|_S \subset S \times S$ be the restricted strict order on S induced by $\succ_i$. Then, $\succ_i|_S = \succ'_i|_S$ means that $\succ_i$ and $\succ'_i$ coincide on S.

Definition 3. An SWFI $f : \mathcal{L}^N \times \mathcal{A} \to \mathcal{R}$ satisfies *Exclusion of Interested Party Evaluations* (henceforth *Interest-Exclusion*) if, for any $\boldsymbol{\succ} = (\succ_1, \ldots, \succ_n) \in \mathcal{L}^N$, any $A \in \mathcal{A}$, any $i \in N$, and any $\succ'_i \in \mathcal{L}$ such that $\succ_i|_{A(i)} = \succ'_i|_{A(i)}$,

$$f(\boldsymbol{\succ}, A) = f(\succ'_i, \succ_{-i}, A).$$

Interest-Exclusion is an inter-preference-profile requirement.[12] In many real-world situations, voters are often asked not to express their opinions on the applicants of their interest, which is the spirit of this axiom. Peer selection literature also requires that a voter should not nominate oneself.

Our second axiom is an inter-interest-structure requirement such that applicants should not improve their social ranks by increasing voters who are interested in them.

Definition 4. An SWFI $f : \mathcal{L}^N \times \mathcal{A} \to \mathcal{R}$ satisfies the *No-Power-Game property* if, for any $A, A' \in \mathcal{A}$, any $\boldsymbol{\succ} \in \mathcal{L}^N$, and any $x \in X$,

---

[11] Even if a voter denounces the related applicants, that does not affect social ranking, either. Hence this axiom deals with both favorable and hostile interests.

[12] *Interest-Exclusion* may be viewed as a too stringent requirement because it discards any information regarding preferences on $X \setminus A(i)$ if the set is not a singleton. However, our point is that many rules satisfy this strong axiom; hence we keep it this way. An interesting future research is to weaken *Interest-Exclusion* to utilize relative rankings within $X \setminus A(i)$.

if $I(z;A) = I(z;A')$ for all $z \neq x$ and $I(x;A) \cup \{i\} = I(x;A')$ for some $i \in N$, then for any $y \in X$,

$$[yf(\succ, A)x \Rightarrow yf(\succ, A')x] \text{ and } [[yf(\succ, A)x \text{ and } \neg xf(\succ, A)y] \Rightarrow [yf(\succ, A')x \text{ and } \neg xf(\succ, A')y]].$$

To interpret the condition of this axiom, as the interest structure A changes to A', the applicant x has one more interested voter i ($I(x;A) \cup \{i\} = I(x;A')$), while the interest structure of the other applicants remains the same ($I(z;A) = I(z;A')$ for all $z \neq x$). The requirement is that, in this case, the ranking between x and a socially preferred y should not reverse. Otherwise, favoritism is suspected, and the decision process is not accountable.[13] If the new relationship is negative, this axiom cannot prevent that x declines in the social ranking. This axiom only prevents a power-game-type competition which encourages applicants to increase the number of positively interested voters.

A simple way to satisfy *Interest-Exclusion* is to utilize only the preferences on $A(i)$, as in the peer rating/selection models. This may be the underlying idea of the protocol mentioned earlier at a funding committee that a voter must leave the room when her/his related person's application is on the table.

Proposition 1. *The two accountability requirements are compatible.*

Proof. We construct an extension of the scoring rule (Young, 1975) based on the plurality rule. A similar construction based on the Borda rule is also possible (see Fujiwara-Greve et al., 2021). For any $(x, \succ_i, A(i)) \in X \times \mathcal{L} \times 2^X$, let

$$s^{plu}(x, \succ_i, A(i)) = \begin{cases} 1 & \text{if } x \in A(i) \text{ and } \left|\{y \in A(i) : y \succ_i x\}\right| = 0 \\ 0 & \text{otherwise.} \end{cases}$$

This function gives point 1 to the top applicant in $A(i)$ at $\succ_i$ and 0 to all other applicants. Define the *extended plurality rule with interest structure* f^{plu} as

$$xf^{plu}(\succ_1, \ldots, \succ_n, A)y \iff \sum_{i \in N} s^{plu}(x, \succ_i, A(i)) \geqq \sum_{i \in N} s^{plu}(y, \succ_i, A(i)).$$

[13] We think verifiable relationships are not easy to nullify (e.g., advisor-advisee), and thus we do not consider the decrease in related voters.

That is, for each applicant x, we count the voters who are not interested in x and put x at the top. *Interest-Exclusion* is satisfied by definition. The *No-Power-Game property* is also straightforward because when an applicant x has one more interested voter, the points of x either decrease by one or stay the same. In contrast, the points of all other applicants either increase by one or remain the same. Hence the social ranking of x does not improve. $\square$

2.3. Extensions of standard axioms in voting theory

We adapt some of the standard axioms in the Arrovian literature to our framework while allowing arbitrary non-trivial interest structures. When no voter has an interest in any applicant ($A = \mathbf{1}_{n,m}$), our axioms reduce to the standard axioms. When there are interested voters, our axioms require that, for each applicant, at least the preferences of voters who do not have an interest in that applicant should be taken into account.

As a preparation, for any $S \subseteq X$ and $\succ \in \mathcal{L}$, let $\text{top}(\succ, S) \in S$ be the top alternative in the set S according to the strict order $\succ$, i.e., $\text{top}(\succ, S) \succ y$ for all $y \in S \setminus \{\text{top}(\succ, S)\}$.

Definition 5. An SWFI $f : \mathcal{L}^N \times \mathcal{A} \to \mathcal{R}$ satisfies *A-Unanimity* if, for any $(\boldsymbol{\succ}, A) \in \mathcal{L}^N \times \mathcal{A}$ and any $x \in X$,

$$[\text{top}(\succ_i, A(i)) = x, \ \ \forall\, i \in N \setminus I(x; A)] \text{ implies that } [x f(\boldsymbol{\succ}, A) z \text{ and } \neg z f(\boldsymbol{\succ}, A) x, \ \ \forall z \in X].$$

This is a unanimity axiom that includes accountability (hence A is added in the name). It requires that if there is an x such that all unrelated voters consider it the best, then it should be ranked as the social best. Voters with a positive interest in x would not mind the outcome. The accountability property of this axiom is that we prevent voters who have a hostile interest in x from denouncing x as long as all neutral voters agree to make x the best.

The following remark guarantees that *A-Unanimity* is not self-contradictory.

Remark 1. *For any $(\boldsymbol{\succ}, A) \in \mathcal{L}^N \times \mathcal{A}$, if there exists $x \in X$ such that [$top(\succ_i, A(i)) = x, \ \ \forall\, i \in N \setminus I(x; A)$] holds, then it is unique.*

Proof. Suppose that there are distinct $x, x' \in X$ such that

for any $i \in N \setminus I(x; A)$, $\text{top}(\succ_i, A(i)) = x$, and for any $i \in N \setminus I(x'; A)$, $\text{top}(\succ_i, A(i)) = x'$.

Since $A \in \mathcal{A}$ is a non-trivial interest structure, $\{x, x'\} \subset A(j)$ for some $j \in N$, but $\succ_j$ is a strict order, so it cannot be that both $x \succ_j x'$ and $x' \succ_j x$ hold. $\square$

It is also important not to put a pairwise-majority loser (in some literature, it is called a Condorcet-loser) at the top of the social ranking. This requirement was originally due to Borda (1784).[14]

Definition 6. For any $(\boldsymbol{\succ}, A) \in \mathcal{L}^N \times \mathcal{A}$, $x \in X$ is a *pairwise-majority-loser with respect to* $(\boldsymbol{\succ}, A)$ if

$$\forall y \in X \setminus \{x\}, \quad \big|\{i \in N : y \succ_i|_{A(i)} x\}\big| > \big|\{i \in N : x \succ_i|_{A(i)} y\}\big|.$$

Definition 7. An SWFI $f : \mathcal{L}^N \times \mathcal{A} \to \mathcal{R}$ satisfies the *No-Pairwise-Loser property* if, for any $(\boldsymbol{\succ}, A) \in \mathcal{L}^N \times \mathcal{A}$ and any $x \in X$ such that x is a pairwise-majority-loser with respect to $(\boldsymbol{\succ}, A)$, there exists $y \in X \setminus \{x\}$ such that $yf(\boldsymbol{\succ}, A)x$ and $\neg xf(\boldsymbol{\succ}, A)y$.

As a much weaker requirement than *A-Unanimity*, we introduce an axiom such that voters should be able to put any applicant at the top of a social ranking.

Definition 8. An SWFI $f : \mathcal{L}^N \times \mathcal{A} \to \mathcal{R}$ satisfies *Non-Restrictedness* if, for any $A \in \mathcal{A}$ and any $x \in X$, there exists $\boldsymbol{\succ} \in \mathcal{L}^N$ such that $xf(\boldsymbol{\succ}, A)y$ for all $y \in X$.

Non-Restrictedness is weaker than the non-imposition condition in the standard Arrovian framework, given an A. If an SWFI satisfies *A-Unanimity*, then it satisfies *Non-Restrictedness*. For an impossibility result, this weaker axiom is sufficient.

3. Impossibility theorems

In this section, we show various impossibility theorems when the second domain of SWFI is the set of all non-trivial interest structures. On the surface, the *No-Power-Game property* and *A-Unanimity* seem orthogonal in the deterrence of biases: the *No-Power-Game property* reduces biases by an increase of (positively) interested parties, while *A-Unanimity* reduces

[14]See, for example, Fishburn and Gerhlein (1976) and Okamoto and Sakai (2019) for surveys on the *No-Pairwise-Loser property.*

biases by (negatively) interested parties. However, our first result is that these two axioms are incompatible.

Theorem 1. *There is no SWFI $f : \mathcal{L}^N \times \mathcal{A} \to \mathcal{R}$ that satisfies the* No-Power-Game property *and* A-Unanimity.

Proof. Suppose that an SWFI $f : \mathcal{L}^N \times \mathcal{A} \to \mathcal{R}$ satisfies A-*Unanimity* and the *No-Power-Game property.*

Claim. For any $(\succ, A) \in \mathcal{L}^N \times \mathcal{A}$ and any $x \in X$, if there exists $i \in N$ such that $A(i) = X$ and $\text{top}(\succ_i, A(i)) = x$, then

$$x f(\succ, A) y \text{ and } \neg y f(\succ, A) x, \quad \forall y \in X \setminus \{x\}. \tag{1}$$

Proof of the Claim.

Take any $(\succ, A) \in \mathcal{L}^N \times \mathcal{A}$ and any $x \in X$ such that there exists $i \in N$ for which $A(i) = X$ and $\text{top}(\succ_i, A(i)) = x$. Suppose that there exists $y \neq x$ such that $y f(\succ, A) x$, negating (1).

If each voter $j \in N \setminus I(x; A)$ (who does not have an interest in x under A) ranks x as her/his best, $y f(\succ, A) x$ contradicts A-*Unanimity.* Hence there must be one voter with the property that she/he is not an interested party of x and her/his best among admissible applicants is not x. Let this voter be $j_1 \in N \setminus I(x; A)$. Her/his top alternative in $A(j_1)$ is not x.

Change A so that j_1 becomes an interested party of x. Let $A^1 \in \mathcal{A}$ be such that $I(x; A^1) = \{j_1\} \cup I(x; A)$ and $I(z; A^1) = I(z; A)$ for all $z \neq x$. Then, by the *No-Power-Game property* of f, the social ranking of x should not improve, and we have $y f(\succ, A^1) x$.

If each voter $j \in N \setminus I(x; A^1)$ who does not have an interest in x under A^1 ranks x as her/his best, $y f(\succ, A^1) x$ contradicts A-*Unanimity.* Hence there must be $j_2 \in N \setminus I(x; A^1)$ such that x is not her/his top in $A(j_2)$. Change A^1 to A^2 by making j_2 an interested party of x.

Repeating this operation leads to the interest structure A^k (for some $k \leqq n-1$) in which all voters except i becomes an interested party of x. Then the unique neutral voter i considers x the best. By A-*Unanimity*, x must be the social best, which contradicts that $y f(\succ, A^k) x$. This completes the proof of the Claim. //

We derive a contradiction using the above Claim for two voters with opposing preferences. Formally, consider $A = \mathbf{1}_{n,m} \in \mathcal{A}$, and fix arbitrary distinct x and y. There exists $\boldsymbol{\succ} \in \mathcal{L}^N$ such that $x \succ_1 z$ for all $z \neq x$ and $y \succ_2 z$ for all $z \neq y$. Then, by Claim,

$$x f(\boldsymbol{\succ}, \mathbf{1}_{n,m}) y \text{ and } \neg y f(\boldsymbol{\succ}, \mathbf{1}_{n,m}) x,$$

$$y f(\boldsymbol{\succ}, \mathbf{1}_{n,m}) x \text{ and } \neg x f(\boldsymbol{\succ}, \mathbf{1}_{n,m}) y,$$

a contradiction. □

Theorem 1 is shocking because *A-Unanimity* is a natural extension of the standard Unanimity: if all neutral voters agree on the top applicant, that should also be the social top. The above impossibility may be due to the strength of the *No-Power-Game property* because even if we weaken *A-Unanimity* to *Non-Restrictedness*, we have another impossibility result as follows.

Theorem 2. *Assume $n \geqq 3$. There is no SWFI $f : \mathcal{L}^N \times \mathcal{A} \to \mathcal{R}$ that satisfies the* No-Power-Game property*,* Interest-Exclusion*,* Non-Restrictedness*, and the* No-Pairwise-Loser property*.*

Proof. Suppose that an SWFI $f : \mathcal{L}^N \times \mathcal{A} \to \mathcal{R}$ satisfies all properties. Fix any $x \in X$ and let $A^* \in \mathcal{A}$ be such that $A^*(n) = X$ and $A^*(i) = X \setminus \{x\}$ for all $i \in \{1, \ldots, n-1\}$. By *Non-Restrictedness*, there exists $\boldsymbol{\succ} \in \mathcal{L}^N$ such that $x f(\boldsymbol{\succ}, A^*) y$, for all $y \in X$.

Let $\boldsymbol{\succ}' \in \mathcal{L}^N$ be such that

- $\succ'_n = \succ_n$,
- $\succ'_i |_{X \setminus \{x\}} = \succ_i |_{X \setminus \{x\}}$ and $y \succ'_i x$ for all $i \neq n$ and all $y \neq x$.

That is, $\succ_i$ and $\succ'_i$ coincide on $A^*(i)$ for all $i \in N$, but $\succ'_i$ for $i \neq n$ puts x to the bottom. Since $n \geqq 3$, if all voters are unrelated to all applicants, i.e., if $A = \mathbf{1}_{n,m}$, x is a pairwise-majority-loser for the new preference profile $\boldsymbol{\succ}'$. The *No-Pairwise-Loser property* then implies that there exists $y \in X$ such that $y f(\boldsymbol{\succ}', \mathbf{1}_{n,m}) x$ and $\neg x f(\boldsymbol{\succ}', \mathbf{1}_{n,m}) y$.

Change $A = \mathbf{1}_{n,m}$ into $A^1 \in \mathcal{A}$ such that the voter 1 becomes an interested party of x, i.e., $A^1(1) = X \setminus \{x\}$ and $A^1(i) = X$ for all $i \in \{2, \ldots, n\}$. Then, by the *No-Power-Game property*, we have $y f(\boldsymbol{\succ}', A^1) x$ and $\neg x f(\boldsymbol{\succ}', A^1) y$. Next, let $A^2 \in \mathcal{A}$ be such that $A^2(1) = A^2(2) = X \setminus \{x\}$

and $A^2(i) = X$ for all $i \in \{3, \ldots, n\}$. Then, again by the *No-Power-Game property*, we have $yf(\boldsymbol{\succ}', A^2)x$ and $\neg xf(\boldsymbol{\succ}', A^2)y$. Repeating this operation for $n-1$ times, we have that $yf(\boldsymbol{\succ}', A^*)x$ and $\neg xf(\boldsymbol{\succ}', A^*)y$, where A^* is defined above. Recall that $\succ_i$ and $\succ_i'$ coincide on $A^*(i)$ for all $i \in N$ and $xf(\boldsymbol{\succ}, A^*)y$ for all $y \in X$. Thus we obtain a contradiction to *Interest-Exclusion*. □

Theorems 1 and 2 imply that accountability and some standard axioms in voting theory are hard to sustain jointly.

Although in a smaller environment, we show that the *No-Power-Game property* is strong because it is incompatible with the *No-Pairwise-Loser property* alone. Recall that m is the number of the applicants in X, and we can write $X = \{x_1, x_2, \ldots, x_m\}$.

Theorem 3. *Suppose that $n \geqq m+2$. There is no SWFI $f : \mathcal{L}^N \times \mathcal{A} \to \mathcal{R}$ that satisfies the* No-Power-Game property *and the* No-Pairwise-Loser property*.*

Proof. Suppose that an SWFI $f : \mathcal{L}^N \times \mathcal{A} \to \mathcal{R}$ satisfies the *No-Power-Game property* and the *No-Pairwise-Loser property.* There exists $A \in \mathcal{A}$ such that

- for any $i \in \{1, \ldots, m\}$, $A(i) = \{x_i\}$,
- for any $i \in \{m+1, m+2\}$, $A(i) = X$,
- for any $i \in N \setminus \{1, \ldots, m+2\}$, $A(i) = \emptyset$.

That is, each of the first m voters has a different and unique admissible applicant x_i. To guarantee the non-triviality of A, two voters, $m+1$ and $m+2$, have no interest in any applicant. If there are more than $m+2$ voters, they are not relevant in A.

Let $\boldsymbol{\succ} \in \mathcal{L}^N$ be an arbitrary profile such that

- for any $i \in \{1, \ldots, m\}$, $\text{top}(\succ_i, X) = x_i$,
- $x_1 \succ_{m+1} x_2 \succ_{m+1} \cdots \succ_{m+1} x_m$,
- $x_m \succ_{m+2} x_{m-1} \succ_{m+2} \cdots \succ_{m+2} x_1$.

Note that the voter $m+1$ and the voter $m+2$ have opposite preferences.

Since $f(\succ, A)$ is a weak order and $|X|$ is finite, there exits $x \in X$ such that

$$x\, f(\succ, A)\, y \ \text{ for all } y \in X. \tag{2}$$

Let $A' \in \mathcal{A}$ be such that

- for any $i \in \{1, \ldots, m\}$, $A'(i) = \{x_i, x\}$ (not necessarily a two-element set),
- for any $i \in \{m+1, m+2\}$, $A'(i) = X$,
- for any $i \in N \setminus \{1, \ldots, m+2\}$, $A'(i) = \emptyset$.

Under A', for any $x_i \neq x$, there are two voters (voter i and some $j \in \{m+1, m+2\}$) who prefers x_i to x and $A'(i)$ and $A'(j)$ include x_i and x, but there is only one voter $k \in \{m+1, m+2\} \setminus \{j\}$ who prefers x to x_i and $A'(k)$ includes x_i and x. Therefore x is a pairwise-majority-loser with respect to $(\succ, A')$. By the *No-Pairwise-Loser property,* there exists $y \neq x$ such that $yf(\succ, A')x$ and $\neg\, xf(\succ, A')y$.

Let us change one voter at a time to become an interested party of x starting from voter 1 to m. That is, change A' to $A^1 \in \mathcal{A}$ such that $A^1(1) = \{x_1\}$ and $A^1(j) = A'(j)$ for all $j \neq 1$. Then change A^1 to $A^2 \in \mathcal{A}$ such that $A^2(2) = \{x_2\}$ and $A^2(j) = A^1(j)$ for all $j \neq 2$, and so on. This operation is a finite-step process ending with $A^m = A$. In each step, the social rank of x should not improve by the *No-Power-Game property.* In the end, we have that

$$yf(\succ, A)x \text{ and } \neg\, xf(\succ, A)y,$$

which is a contradiction to (2). □

Possible causes of the impossibility results include the large size of $\mathcal{A}$. In many situations, the organizers can control the voter configurations; hence we may restrict the interest-structure domain of SWFIs to a smaller set. Section 4 considers restrictions of $\mathcal{A}$. Another possible cause is the strength of the *No-Power-Game property.* Section 5 weakens the *No-Power-Game property* to recover from the impossibility.

4. Restrictions of the feasible interest structures

The impossibility results explicitly use the large domain of the interest structures. Realistic prevention of accountability problems is to have a two-step procedure. First, when forming a committee of voters, reduce the possibility of conflict of interest as much as possible. Second, use an accountable voting rule. The first step allows us to focus on a smaller domain of the interest structures than $\mathcal{A}$. In many situations, however, it is impossible always to form a committee of neutral voters. (A committee is often gathered before the set of alternatives/applicants is known. Expert voters should be in a field close enough to those of the applicants so that some relationships may not be avoidable.) In this section, we consider restrictions of $\mathcal{A}$ to find the borderline of the impossibility of Theorem 1. Then, if we can always form committees within that set[15], there is an accountable and unanimous voting rule and vice versa.

For any $k = 1, 2, \dots, n$, define

$$\mathcal{A}(k) := \{A \in \{0,1\}^{n \times m} : \forall x, y \in X \text{ such that } x \neq y,\ |\{i \in N : \{x, y\} \subset A(i)\}| \geq k\}.$$

In this set, for any distinct pair, at least k voters have no conflict of interest. Note that $\mathcal{A}(1)$ coincides with the set $\mathcal{A}$ of all non-trivial interest structures. We can adapt all axioms to $\mathcal{A}(k)$.

Proposition 2. *For any $k = 1, \dots, [\frac{n}{2}]$, there is no $f : \mathcal{L}^N \times \mathcal{A}(k) \to \mathcal{R}$ that satisfies the* No-Power-Game property *and* A-Unanimity*.*

Proof. See Appendix. □

Hence, the logic similar to the proof of Theorem 1 holds even if, for each pair of applicants, nearly half of the voters have no verifiable connection to them.

Next, we show that Theorem 1 does not hold when $k \geqq [\frac{n}{2}] + 1$, i.e., when each pair of applicants is admissible for more than half of the voters. Namely, *the extended plurality rule with interest structure* in the proof of Proposition 1 satisfies both axioms. If all neutral voters of some x consider it the best and if the number of the neutral voters (of any pair including x) is more than half of the voters, then the plurality-based SWFI places x at the social top, satisfying

[15] Many universities require doctoral degree evaluation committees to have some external members outside of the candidate's field. This rule can be an attempt to restrict the possible interest structures.

A-Unanimity. The *No-Power-Game property* (and moreover *Interest-Exclusion*) holds for the plurality-based SWFI by construction.

Proposition 3. *When* $k \geqq [\frac{n}{2}] + 1$*, there exists* $f : \mathcal{L}^N \times \mathcal{A}(k) \to \mathcal{R}$ *satisfying* Interest-Exclusion*, the* No-Power-Game property*, and* A-Unanimity*.*

Proof. Let $k \geqq [\frac{n}{2}]+1$ and let $f_k^{plu} : \mathcal{L}^N \times \mathcal{A}(k) \to \mathcal{R}$ be *the extended plurality rule with interest structure* f^{plu} (see the proof of Proposition 1) adapted to $\mathcal{A}(k)$.

Interest-Excluson and the *No-Power-Game property* hold by the same argument as that in the proof of Proposition 1.

A-Unanimity: Take any $(\succ, A) \in \mathcal{L}^N \times \mathcal{A}(k)$ and any $x \in X$ such that x is the top for all voters who have no interest in x; for any $i \in N \setminus I(x; A)$,

$$\text{top}(\succ_i, A(i)) = x.$$

Since $A \in \mathcal{A}(k)$, for any $y \in X \setminus \{x\}$, the number of voters who have no interest in the pair (x, y) is at least k. Hence,

$$\left|\{i \in N : \text{top}(\succ_i, A(i)) = x\}\right| \geqq \left[\frac{n}{2}\right] + 1.$$

Then, for any $y \in X \setminus \{x\}$, $x f_k^{plu}(\succ, A) y$ and $\neg\, y f_k^{plu}(\succ, A) x$. Hence, f_k^{plu} satisfies *A-Unanimity.* □

Therefore a solution to the impossibility results is to keep the set of voters within an appropriate range such that each pair of applicants is admissible for more than half of the voters.

However, the incompatibility between the *No-Power-Game property* and the *No-Pairwise-Loser property* remains even if most voters have no conflict of interest in any pair.

Proposition 4. *Assume that* $n \geqq 2m$*. For any* $k \leqq n-4$*, there is no SWFI* $f : \mathcal{L}^N \times \mathcal{A}(k) \to \mathcal{R}$ *that satisfies the* No-Power-Game property *and the* No-Pairwise-Loser property*.*

Proof. See Appendix. □

The proof uses a similar construction of $(\succ, A)$ and repeated applications of the *No-Power-Game property* as in the proof of Theorem 3. If the number of voters is large, then $n-4$ is close to n. Hence Proposition 4 covers the case when nearly all voters have no conflict of interest in any pair, and thus Theorem 3 does not hinge on the largest interest-structure domain. This finding is another striking result.

5. Weakening of the No-Power-Game Property

Another remedy is to weaken one of the accountability axioms. *Interest-Exclusion* is easy to satisfy, and the peer rating/selection literature considers it a fundamental requirement. Therefore, we weaken the *No-Power-Game Property.*

Definition 9. An SWFI $f : \mathcal{L}^N \times \mathcal{A} \to \mathcal{R}$ satisfies the *Weak No-Power-Game property* if, for any $A, A' \in \mathcal{A}$, any $\succ \in \mathcal{L}^N$, and any $x \in X$,
if $I(z; A) = I(z; A')$ for all $z \neq x$, $I(x; A) \cup \{i\} = I(x; A')$ for some $i \in N$ with $\text{top}(\succ_i, A) = x$, then for any $y \in X$,

$$[y f(\succ, A) x \Rightarrow y f(\succ, A') x] \text{ and } [[y f(\succ, A) x \wedge \neg x f(\succ, A) y] \Rightarrow [y f(\succ, A') x \wedge \neg x f(\succ, A') y]] .$$

This is a satisfactory weakening of the *No-Power-Game property*: if a voter who puts x at the top changes the status from neutral to an interested party of x, then x's social rank should not improve. By contrast, the *No-Power-Game property* requires that changing any voter into an interested party should not improve the ranking.

Proposition 5. *There exists an SWFI $f : \mathcal{L}^N \times \mathcal{A} \to \mathcal{R}$ that satisfies the* Weak No-Power-Game property*,* Interest-Exclusion*,* A-Unanimity *(and thus* Non-Restrictedness*), and the* No-Pairwise-Loser property*.*

To prove this, we construct a new rule, the *winning-rate rule*, based on the Borda rule. First, we introduce the *net-winning points.*

Definition 10. For any $(\succ, A) \in \mathcal{L}^N \times \mathcal{A}$, any $i \in N$, and any $x \in X$, define the *net-winning* *points* of x for voter i at $(\succ_i, A(i))$ by

$$w(x, \succ_i, A(i)) := \begin{cases} 0 & \text{if } x \notin A(i) \\ |\{y \in A(i) : x \succ_i y\}| - |\{y \in A(i) : y \succ_i x\}| & \text{otherwise.} \end{cases}$$

For example, if $|A(i)| = 2$, then the *net-winning points* of the top applicant in $A(i)$ is 1 and that of the bottom applicant in $A(i)$ is -1, while that of any $x \notin A(i)$ is 0. These scores are different from the usual Borda score, which counts the number of alternatives ranked lower than the relevant one.[16]

The problem with comparing the total *net-winning points* is that applicants with many neutral voters may get disproportionally many (positive or negative) points. Therefore, we consider an averaging system with respect to **the number of comparisons** between admissible applicants to milden such effect. In practice, such averages are often used to measure the performances of sports players/teams whose number of plays may differ significantly.[17]

Definition 11. For any $j \in \{1, 2, \ldots, m\}$ and any $(\succ, A) \in \mathcal{L}^N \times \mathcal{A}$, the *winning rate* of applicant x_j at $(\succ, A)$ is

$$WR(x_j, \succ, A) := \frac{\sum_{i \in N} w(x_j, \succ_i, A(i))}{\sum_{i \notin I(x_j, A)} (|A(i)| - 1)} \left(= \frac{\sum_{i \in N} w(x_j, \succ_i, A(i))}{\sum_{i \in N} a_{ij} (|A(i)| - 1)} \right).$$

The *winning-rate rule* f^{WR} is an SWFI which satisfies, for any $j, k \in \{1, 2, \ldots, m\}$ and any $(\succ, A) \in \mathcal{L}^N \times \mathcal{A}$,

$$x_j f^{WR}(\succ, A) x_k \iff WR(x_j, \succ, A) \geqq WR(x_k, \succ, A).$$

Recall that the *net-winning points* count the "winnings" of x_j minus the "loss" of x_j. The *winning rate* divides this difference by the number of pairwise comparisons that x_j faces under A, which is $\sum_{i \notin I(x_j, A)} (|A(i)| - 1)$. The 0-1 structure of the A matrix formulation makes the summation easy to compute.

[16] For why putting related applicants in the center instead of the bottom is better, see our working paper Fujiwara-Greve et al. (2021).

[17] For example, the definition of the winning-rate rule coincides with that of "the gravity center rating" given in the formula 9.4 in Langville and Meyer (2013, Chapter 9), assuming that $k_{ij} = 1$ if team i beats j and $k_{ij} = -1$ if team j beats team i. They provide ratings and rankings of American football teams using data of match results in NFL 2009–2010 based on the gravity center rating or the winning-rate rule.

To illustrate how comparing the total *net-winning points* and the *winning-rate rule* differ, consider the following example where $X = \{x_1, x_2, x_3\}$ and $N = \{1, 2, \ldots, 10\}$.

Let $(\boldsymbol{\succ}, A) \in \mathcal{L}^N \times \mathcal{A}$ be such that

- $A(1) = X$ and $A(i) = \{x_2, x_3\}$ for all $i \in \{2, \ldots, 10\}$,
- $x_1 \succ_1 x_2 \succ_1 x_3$ and $x_2 \succ_i x_3$ for all $i \in \{2, \ldots, 10\}$.

Then

$$\begin{aligned}
&w(x_1, \succ_1, A(1)) = 2, \quad w(x_1, \succ_i, A(i)) = 0, \ \forall i \in \{2, \ldots, 10\},\\
&w(x_2, \succ_1, A(1)) = 0, \quad w(x_2, \succ_i, A(i)) = 1, \ \forall i \in \{2, \ldots, 10\},\\
&w(x_3, \succ_1, A(1)) = -2, \quad w(x_3, \succ_i, A(i)) = -1, \ \forall i \in \{2, \ldots, 10\},
\end{aligned}$$

so that x_2 has the largest total *net-winning points*. However,

$$\begin{aligned}
WR(x_1, \boldsymbol{\succ}, A) &= \frac{w(x_1, \succ_1, A(1))}{|A(1)| - 1} = \frac{2}{2},\\
WR(x_2, \boldsymbol{\succ}, A) &= \frac{\sum_{i \in N} w(x_2, \succ_i, A(i))}{\sum_{i \in N} (|A(i)| - 1)} = \frac{9}{11},\\
WR(x_3, \boldsymbol{\succ}, A) &= \frac{\sum_{i \in N} w(x_3, \succ_i, A(i))}{\sum_{i \in N} (|A(i)| - 1)} = \frac{-11}{11}.
\end{aligned}$$

Hence, the *winning-rate rule* ranks x_1 as the top. In this example, the applicant x_1 is admissible only for the voter 1, while the applicant x_2 is admissible for all voters. The *winning-rate rule* is better if a big difference in the number of neutral voters should not affect the social ranking.

Proof of Proposition 5. We show that f^{WR} satisfies all axioms. *Interest-Exclusion* holds by definition.

No-Pairwise-Loser property: Take any $(\boldsymbol{\succ}, A) \in \mathcal{L}^N \times \mathcal{A}$. By the definition of w, the total *net-winning points* over all applicants is 0:

$$\sum_{y \in X} \sum_{i \in N} w(y, \succ_i, A(i)) = \sum_{i \in N} \sum_{y \in X} w(y, \succ_i, A(i)) = 0.$$

Now, let x be the highest ranked alternative with respect to $f^{WR}(\boldsymbol{\succ}, A)$. Since the denominator of $WR(x, \succ_i, A)$ is positive, the numerator must satisfy $\sum_{i \in N} w(x, \succ_i, A(i)) \geqq 0$.

By Lemma 1 in Appendix, for any $z \in X$ which is a pairwise-majority-loser with respect to A, we have

$$\sum_{i \in N} w(z, \succ_i, A(i)) < 0.$$

Hence x is not a pairwise-majority-loser with respect to $(\succ, A)$.

A-Unanimity: Take any $(\succ, A) \in \mathcal{L}^N \times \mathcal{A}$, and assume that there is $x \in X$ such that $top(\succ_i, A) = x$ for any $i \in N \setminus I(x; A)$.

Then the *net-winning points* of x is $w(x, \succ_i, A(i)) = |A(i)| - 1$ for any $i \in N \setminus I(x; A)$, and thus its *winning rate* is 1; $WR(x, \succ, A) = 1$. Since this achieves the maximum value of the WR, we have $x f^{WR}(\succ, A) y$ for any $y \in X$.

Next, consider any $y \neq x$. Since A is non-trivial, there exists $j \in N$ such that $\{x, y\} \subset A(j)$. By the assumption of x, we must have $x \succ_j y$. Hence the winning rate of y cannot be 1. That is, $\neg y f^{WR}(\succ, A) x$ holds.

Weak No-Power-Game property: Let $A, A' \in \mathcal{A}$, $\succ \in \mathcal{L}^N$ and $x \in X$ be such that the interested party of $z \neq x$ are the same across A and A', but the interested party of x weakly expands from A to A' in such a way that a voter who puts x at the top joins the interested party under A':

$$I(z; A) = I(z; A'),\ \forall z \neq x,\ \exists i \in N; \mathrm{top}(\succ_i, A) = x,\ I(x; A) \cup \{i\} = I(x; A').$$

Take any $y \in X$. We want to show

$$[y f(\succ, A) x \Rightarrow y f(\succ, A') x] \text{ and } [[y f(\succ, A) x \wedge \neg x f(\succ, A) y] \Rightarrow [y f(\succ, A') x \wedge \neg x f(\succ, A') y]]. \quad (3)$$

If $y = x$, then clearly (3) holds. Hence assume that $y \neq x$. Also if $I(x; A) = I(x; A')$, (3) holds. Hence assume that $i \notin I(x; A)$. Because i is not counted at A' and other voters do not change their interest structures ($A(j) = A'(j)$), the *winning rate* of x at A' is

$$\sum_{j \in N} \frac{w(x, \succ_j, A'(j))}{\sum_{k \notin I(x, A')} (|A'(k)| - 1)} = \frac{0 + \sum_{j \neq i} w(x, \succ_j, A(j))}{0 + \sum_{k \notin I(x, A) \cup \{i\}} (|A(k)| - 1)},$$

while the *winning rate* of x at A is

$$\sum_{j \in N} \frac{w(x, \succ_j, A(j))}{\sum_{k \notin I(x, A)} (|A(k)| - 1)} = \frac{|A(i)| - 1 + \sum_{j \neq i} w(x, \succ_j, A(j))}{|A(i)| - 1 + \sum_{k \notin I(x, A) \cup \{i\}} (|A(k)| - 1)},$$

because i puts x as the top in $A(i)$. This means that the *winning rate* of x weakly decreases when the interest structure changes from A to A'.

It suffices to prove that the *winning rate* of y weakly increases when the interest structure changes from A to A'.

Case 1: $y \notin A(i)$.

Since the interest structure of other applicants than x does not change, $y \notin A'(i)$ also holds, so that $w(y, \succ_i, A(i)) = w(y, \succ_i, A'(i)) = 0$. The *net-wining points* and the number of pairwise comparisons of y for other voters do not change. Hence the *winning rate* of y does not change from A to A'.

Case 2: $y \in A(i)$.

Since the interested parties of y do not change between A and A', i is not an interested party of y at A'. By moving from A to A', the total number that the voter i compares y with something else has decreases by 1, i.e., becomes $|A(i)| - 2$. Recall that $\text{top}(\succ_i, A) = x$, but i does not count for x under A'. Hence the *net-winning points* of y for voter i at $(\succ_i, A')$ increases by one point, and the *winning rate* of y at A' is

$$\sum_{j \in N} \frac{w(y, \succ_j, A'(j))}{\sum_{k \notin I(y,A')} (|A'(k)| - 1)} = \frac{w(y, \succ_i, A(i)) + 1 + \sum_{j \neq i} w(y, \succ_j, A(j))}{|A(i)| - 2 + \sum_{k \notin I(y,A) \cup \{i\}} (|A(k)| - 1)},$$

which is greater than the *winning rate* of y at A. Therefore, we have (3). □

The logic of how the weakening of the *No-Power-Game property* helps resolve the impossibility is as follows. When the most favorable voter becomes an additional interested party, this voter can no longer push the applicant under *Interest-Exclusion*. Therefore, the *Weak No-Power-Game property* is easily satisfied. On the other hand, if a voter who is not that favorable to an applicant becomes an additional interested party, *Interest-Exclusion* may work oppositely. Previously, this voter may have harmed this applicant but can no longer do that by becoming an interested party. Hence the change may improve the ranking of the applicant, and the *No-Power-Game property* is more difficult to satisfy.

6. Concluding remarks

Our main contribution is introducing a new framework to analyze voting rules under an arbitrary conflict of interest situation. We formalized axioms of accountability to deter partisanship for any non-trivial (and verifiable) interest structure. Although we show impossibility results, we also provide two directions of remedies.

Among many future directions, let us mention some of them. First, our formulation may be too extreme. *Interest-Exclusion* threw away the information of one's related applicants $X \setminus A(i)$. A better-informed and still accountable voting rule would utilize the relative rankings within $A(i)$ and $X \setminus A(i)$ as independently as possible. Although we have restricted our attention to 0 or 1 matrix A, mathematically, other weights are possible.

Second, we have yet to investigate strategyproof-related concerns. We can adapt the impartiality of Holzman and Moulin (2013) in our framework with general interest structures. It is an important next question what kind of impossibility or possibility results hold with the (adapted) impartiality.

Third, there are many real-world applications. For example, checking whether the actual judging protocols in some sports (e.g., gymnastics, snowboarding, figure skating) satisfy our axioms is useful for fairness. Rankings are important not only for fairness but also for resource allocations. For example, university rankings may affect government funding and student applications, hence the universities' revenue.[18] The calculation rules of the actual university rankings are not often revealed, but the operations may include interested parties' inputs, such as peer evaluations (university professors' votes) and alums questionnaires. Therefore, real-world ranking systems should be examined with respect to accountability. We hope this paper provides a first step to designing accountable voting systems.

[18] The recent phenomenon of university withdrawals from a ranking (e.g., "Yale, Harvard and UC Berkeley law schools withdraw from US News rankings," *The Guardian*, Nov. 18, 2022, and "Harvard leads an exodus of medical schools withdrawing from US News rankings," *The Observer,* Jan. 24, 2023) is evidence that rankings affect their benefits.

References

Alcalde-Unzu, J., Berga, D., and Gjorgjiev, R. (2022) “Impartial Social Rankings: Some Impossibilities,” mimeo. Universidad Pública de Navarra.

Amorós, P. (2020) “Aggregating Experts’ Opinions to Select the Winner of a Competition,” *International Journal of Game Theory*, 49, 833-849.

Arrow, K. J. (1951) *Social Choice and Individual Values,* Wiley, New York. (3rd ed., 2012, Yale University Press.)

Aziz, H., Lev, O., Matti, N., Rosenschein, J.S., and Walsh, T. (2019) “Strategyproof peer selection using randomization, partitioning, and apportionment,” *Artificial Intelligence*, 275, 295-309.

Borda, J.-C. de (1784) “Mémoire sur les élections au scrutin,” *Histoire de l’Académie Royal des Sciences*, Year 1781, 657–664.

Condorcet, M. de (1784) “Sur les élections au scrutin,” *Histoire de l’Académie Royal des Sciences*, Year 1781, 31–34.

Edelman, P. and Por, A. (2021) “A New Axiomatic Approach to the Impartial Nomination Problem,” *Games and Economic Behavior*, 130, 443-451.

Fishburn, P. C. and Gehrlein, W. V. (1976) “Borda’s Rule, Positional Voting, and Condorcet’s Simple Majority Principle,” *Public Choice*, 28, 79–88.

Fujiwara-Greve, T., Kawada, Y., Nakamura, Y. and Okamoto, N. (2021) “Accountable Voting: Impossibility and an Analysis of Concrete Rules,” SSRN 3786674.

Holzman, R. and Moulin, H. (2013) ”Impartial Nominations for a Prize,” *Econometrica*, 81, 173–196.

Kalai, E. and Muller, E. (1977) “Characterization of domains admitting nondictatorial social welfare functions and nonmanipulable voting procedures,” *Journal of Economic Theory*, 16(2), 457-469.

Kemp, M. and Ng, Y-K (1976) "On the Existence of Social Welfare Functions, Social Orderings and Social Decision Functions," *Economica*, 43, 59-66.

Langville, A. and Meyer, C. (2013) *Who's # 1?: The Science of Rating and Ranking*, Princeton University Press.

Mackenzie, A. (2015) "Symmetry and Impartial Lotteries," *Games and Economic Behavior*, 94, 15-28.

McLean, I. and Hewitt, F. (1994) *Condorcet: Foundations of Social Choice and Political Theory*, Edward Elgar Publishing Limited.

McLean, I. and Urken, A. B. (1995) *Classics of Social Choice*, The University of Michigan Press.

Mulgan, R. (2000) "'Accountability': An Ever-Expanding Concept?," *Public Administration*, 78, 555-573.

Name-Correa A and Yildirim H. (2019) "Social Pressure, Transparency, and voting in committees," *Journal of Economic Theory*, 184, 104943.

Ng, Y-K. and Sun G-Z. (2003) "Exclusion of Self Evaluations in Peer Ratings: An Impossibility and Some Proposals," *Social Choice and Welfare*, 20, 443–456.

Ohseto, S. (2012) "Exclusion of Self Evaluations in Peer Ratings: Monotonicity versus Unanimity on Finitely Restricted Domains," *Social Choice and Welfare*, 38, 109–119.

Okamoto, N. and Sakai, T. (2019) "The Borda Rule and the Pairwise-Majority-Loser Revisited," *Review of Economic Design,* 23, 75–89.

Suzumura, K. (1983) *Rational Choice, Collective Decisions, and Social Welfare,* Cambridge University Press.

Tamura, S. and Ohseto, S. (2014) "Impartial Nomination Correspondences," *Social Choice and Welfare*, 43, 47–54.

Young, H. P. (1975) "Social Choice Scoring Functions," *SIAM Journal of Applied Mathematics*, 28, 824–838.

Appendix: Additional Proofs

Proof of Proposition 2. This proof is similar to that of Theorem 1. Take any $k \in \{1, \ldots, [\frac{n}{2}]\}$. Suppose that an SWFI $f : \mathcal{L}^N \times \mathcal{A}(k) \to \mathcal{R}$ satisfies *A-Unanimity* and the *No-Power-Game property*.

Claim. For any $(\succ, A) \in \mathcal{L}^N \times \mathcal{A}(k)$ and any $x \in X$, if there exists $K \subset N$ such that $|K| = k$ and, for any $i \in K$, $A(i) = X$ and $\text{top}(\succ_i, X) = x$ hold, then

$$\forall y \neq x, \quad xf(\succ, A)y \text{ and } \neg yf(\succ, A)x. \tag{4}$$

Proof of the Claim.

Take any $(\succ, A) \in \mathcal{L}^N \times \mathcal{A}(k)$ and any $x \in X$ such that there exists $K \subset N$ for which $|K| = k$ and, for all $i \in K$, $A(i) = X$ and $\text{top}(\succ_i, X) = x$ hold.

Suppose that there exists $y \neq x$ such that $yf(\succ, A)x$, negating (4). If $\text{top}(\succ_j, A(j)) = x$ holds for all $j \in N \setminus I(x; A)$, then $yf(\succ, A)x$ contradicts *A-Unanimity*. Therefore, there exists $j_1 \in N \setminus I(x; A)$ such that x is not the top for voter j_1 in $A(j_1)$. Note that $j_1 \notin K$ by the definition of K.

Let $A^1 \in \mathcal{A}(k)$ be such that $I(x; A^1) = \{j_1\} \cup I(x; A)$ and $I(z; A^1) = I(z; A)$ for all $z \neq x$. Then, by the *No-Power-Game property* of f, we have $yf(\succ, A^1)x$.

If $\text{top}(\succ_j, A^1(j)) = x$ holds for all $j \in N \setminus I(x; A^1)$, then $yf(\succ, A^1)x$ contradicts *A-Unanimity*. Therefore, there exists $j_2 \in N \setminus I(x; A^1)$ such that x is not the top for j_2 in $A^1(j_2)$. Again, $j_2 \notin K$.

Let $A^2 \in \mathcal{A}$ be such that $I(x; A^2) = \{j_2\} \cup I(x; A^1)$ and $I(z; A^2) = I(z; A^1)$ for all $z \neq x$. Then, again by the *No-Power-Game property* of f, we have $yf(\succ, A^2)x$. Repeating this process for t times where $t \leqq n - [\frac{n}{2}]$, we exhaust the voters in $N \setminus K$. Since all $i \in K$ put x at the top, we obtain a contradiction to *A-Unanimity*, and thus (4) holds. This completes the proof of the Claim. //

Consider $A = \mathbf{1}_{n,m} \in \mathcal{A}(k)$, and fix arbitrary distinct x and y. There exists $\succ \in \mathcal{L}^N$ such that

$$\begin{aligned} &\mathrm{top}(\succ_i, A(i)) = x, \text{ for any } i = 1, 2, \ldots, \left[\frac{n}{2}\right], \text{ and} \\ &\mathrm{top}(\succ_i, A(i)) = y, \text{ for any } i = \left[\frac{n}{2}\right] + 1, \ldots, n. \end{aligned}$$

Note that the cardinality of the second group of voters is at least $k (\leqq [\frac{n}{2}])$. By the Claim,

$$x f(\succ, \mathbf{1}_{n,m}) y \text{ and } \neg y f(\succ, \mathbf{1}_{n,m}) x,$$

$$y f(\succ, \mathbf{1}_{n,m}) x \text{ and } \neg x f(\succ, \mathbf{1}_{n,m}) y,$$

a contradiction. □

Proof of Proposition 4. The proof is similar to that of Theorem 3. Assume that $n \geqq 2m$. Suppose that for some $k \leqq n - 4$, there exits $f : \mathcal{L}^N \times \mathcal{A}(k) \to \mathcal{R}$ that satisfies the *No-Power-Game property* and the *No-Pairwise-Loser property.* Consider $(\succ, A) \in \mathcal{L}^N \times \mathcal{A}(k)$ as in Table 1, where the applicants in brackets in a voter's ranking are the ones who are not admissible (that is, $A(1) = A(2) = X \setminus \{x_1\}$ and so on) and the worst for the relevant voters. Note that, for any pair of applicants, at least $n - 4$ voters have no interest in them.

									Replicate $[\frac{(n-2m)}{2}] - 1$ times				
	$\succ_1$	$\succ_2$	$\cdots$	$\succ_{2j-1}$	$\succ_{2j}$	$\cdots$	$\succ_{2m-1}$	$\succ_{2m}$	$\succ_{2m+1}$	$\succ_{2m+2}$	$\cdots$	$\succ_n$ (n even)	$\succ_n$ (odd)
top	x_2	x_m	$\cdots$	x_{j+1}	x_{j-1}	$\cdots$	x_1	x_{m-1}	x_1	x_m	$\cdots$	x_m	x_1
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	x_m	x_1	$\cdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	x_1	x_m	$\cdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
m-1-th	x_m	x_2	$\cdots$	x_{j-1}	x_{j+1}	$\cdots$	x_{m-1}	x_1	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
m-th	(x_1)	(x_1)	$\cdots$	(x_j)	(x_j)	$\cdots$	(x_m)	(x_m)	x_m	x_1	$\cdots$	x_1	x_m

Table 1: $(\succ, A)$ for Proposition 4

The preference profile in Table 1 is constructed in such a way that (i) if the number n of voters is even, then the preferences between any distinct pair of applicants x_j and x_h cancel out, and (ii) if n is odd, for any x_j and x_h such that $j < h$, the number of voters with $x_j \succ_i x_h$ is one more than the number of those with $x_h \succ_i x_j$.

Since $f(\succ, A)$ is a weak order and $|X|$ is finite, there exits $x_j \in \{x_1, \ldots, x_m\}$ such that

$$x_j \, f(\succ, A) \, y, \quad \text{for all } y \in X. \tag{5}$$

Change A to $A' \in \mathcal{A}(k)$ so that voters $2j-1$ and $2j$ are no longer interested parties of x_j, i.e., $A'(i) = A(i)$ for any $i \in N \setminus \{2j-1, 2j\}$ and $A'(i) = A(i) \cup \{x_j\}$ for any $i \in \{2j-1, 2j\}$.

To see that x_j is a pairwise-majority-loser with respect to $(\succ, A')$, take an arbitrary $x_h \neq x_j$. Note that in all pairs of voters other than $\{2j-1, 2j\}$, the two voters have opposite preferences over any pair of applicants. If n is even, then this exhausts all pairs except $\{2j-1, 2j\}$ and hence x_j has two less votes than that of x_h by the preferences of $\{2j-1, 2j\}$. If n is odd, voter n may prefer x_j over x_h, but still the two votes under A' against x_j makes x_j a pairwise-majority-loser.

Therefore, by the *No-Pairwise-Loser property*, there exists $y \neq x_j$ such that $y f(\succ, A') x_j$ and $\neg\, x_j f(\succ, A') y$.

Change A' to A'' by making the voter $2j-1$ an interested party of x_j (and all other interest structures unchanged). By the *No-Power-Game property*, $y f(\succ, A'') x_j$ and $\neg\, x_j f(\succ, A'') y$ must hold.

Finally, change A'' by making the voter $2j$ an interested party of x_j, and we are back to A. By the *No-Power-Game property* again, we have that $y f(\succ, A) x_j$ and $\neg\, x_j f(\succ, A) y$, contradicting (5). □

Lemma 1. *For any* $(\succ, A) \in \mathcal{L}^N \times \mathcal{A}$*, if* $x \in X$ *is a pairwise-majority-loser with respect to* $(\succ, A)$*, then its total* net-winning points *is negative;*

$$\sum_{i \in N} w(x, \succ_i, A(i)) < 0.$$

Proof of Lemma 1. Let $x \in X$ be a *pairwise-majority-loser* with respect to $(\succ, A)$, that is,

$$\forall y \neq x, \quad \left|\{i \in N : y \succ_i|_{A(i)} x\}\right| > \left|\{i \in N : x \succ_i|_{A(i)} y\}\right|.$$

Taking the sum over $y \neq x$ and adding $\left|\{i \in N : x \succ_i|_{A(i)} x\}\right|$ artificially to both sides, we have

$$\left|\{i \in N : x \succ_i|_{A(i)} x\}\right| + \sum_{y \neq x} \left|\{i \in N : y \succ_i|_{A(i)} x\}\right| > \left|\{i \in N : x \succ_i|_{A(i)} x\}\right| + \sum_{y \neq x} \left|\{i \in N : x \succ_i|_{A(i)} y\}\right|.$$

Therefore

$$\begin{aligned}
0 &> \sum_{y \in X} \left|\{i \in N : x \succ_i|_{A(i)} y\}\right| - \sum_{y \in X} \left|\{i \in N : y \succ_i|_{A(i)} x\}\right| \\
&= \sum_{i \in N} \left|\{y \in X : x \succ_i|_{A(i)} y\}\right| - \sum_{i \in N} \left|\{y \in X : y \succ_i|_{A(i)} x\}\right| \qquad (6) \\
&= \sum_{i \in N} w(x, \succ_i, A(i)),
\end{aligned}$$

where (6) is due to the alternative counting: for each i, count the number of admissible applicants with the relevant property and then take the sum over the voters. □