

Pareto-efficiency of ordinal multiwinner voting rules¹

Jean Lainé, Jérôme Lang, İpek Özkal-Sanver, Remzi Sanver

Abstract

We investigate the Pareto efficiency of ordinal multiwinner voting rules. Defining the Pareto-optimality of a committee requires relating the voters' rankings over individual candidates to their preferences over committees. We consider two well-known extension principles that extend rankings over candidates to preferences over committees: responsive and lexicographic. As the responsive extension outputs partial orders, we consider two associated Pareto-optimality notions: a committee is possibly (respectively, necessary) Pareto-optimal if it is Pareto-optimal for some (respectively, every) completion of these partial orders. We define several Pareto-efficiency notions for multiwinner rules, depending on whether some (respectively, all) committees in the output are Pareto-optimal for one of the latter notions. We review what we believe to be a complete list of ordinal multiwinner rules that have been studied in the literature, and identify which Pareto-efficiency notions they satisfy. We find that, somewhat surprisingly, these rules show a huge diversity: some satisfy the strongest notion, some do not even satisfy the weakest one, with many other rules at various intermediate levels.

1 Introduction

Multiwinner voting rules (or *committee rules*) are natural generalizations of single-winner voting rules. They are useful in a variety of situations, from shortlisting to proportional representation and group recommendation. See [14] for a survey. A multiwinner voting rule outputs a set of k candidates, also called a *committee*, for some integer k . The literature distinguishes two important families of such rules, depending on the format of the input: those based on approval votes, and those based on ordinal votes (each voter ranking the candidates). *We focus on the latter family.*

The recent literature extensively discusses axiomatic and computational properties of multiwinner voting rules. We focus on *Pareto-efficiency*. For defining Pareto-efficiency we first need to define the satisfaction of a voter by a committee. If the input of the rule consists of approval ballots, a way doing so is to count how many approved candidates it contains: this is the path followed by Lackner and Skowron [21], who show that most well-studied approval-based committee rules are Pareto-efficient in this sense, with the noticeable exception of sequential rules. If the input of the rule is ordinal, then things are more difficult: while the input allows to say how each voter ranks single candidates, *it does not generally allow to say how they rank committees.*²

We could think of going around this difficulty by saying that a committee S is Pareto-optimal if whenever candidate x Pareto-dominates candidate y , then y is in S only if x is there too. Although this is a plausible necessary condition for Pareto-optimality to hold, as a definition it is too weak: consider the profile P with candidates a, b, x, y and four votes $a \succ x \succ b \succ y$, $b \succ x \succ a \succ y$, $a \succ y \succ b \succ x$ and $b \succ y \succ a \succ x$. All four candidates are Pareto-efficient, yet $\{a, b\}$ should clearly dominate $\{x, y\}$.

Another way consists in considering a *preference extension principle* mapping a ranking over candidates to a partial order over committees, and then applying classical Pareto dominance directly on committees. Such extension principles have been frequently used in various subdomains of social choice, including irresolute rules, fair division of indivisible items, many-to-one matching, and hedonic games. Three

¹This is an abridged version of a paper with the same title that has been accepted for publication in the *Review of Economic Design* (special issue in the honour of Arkadii Slinko).

²As there are exponentially many committees if k is variable, asking voters to rank committees explicitly is not usually considered an option.

different interpretations of sets of candidates have been considered [6]: final outcomes (all elements in the set are jointly obtained), complete uncertainty (only one of the elements in the set is obtained in the end, nature will decide which one), or opportunities (only one of the elements in the set is obtained in the end, and the concerned agent can choose which one). The first of these three interpretations is *conjunctive* while the other two are *disjunctive*. In the committee election setting (as in fair division, matching and hedonic games), the interpretation that prevails is the conjunctive one: a subset of candidates S is seen as a *joint set of candidates* (as opposed to the choice, by nature or by the agent, of one alternative within S).

Aziz *et al.* [3, 1] consider several extension principles. For each of them, they study the computational complexity of determining whether a committee is Pareto-optimal, of computing *some* Pareto-optimal committee, and when possible, they give simple characterizations of Pareto-optimal committees. They do not, however, consider the following question: given a multiwinner rule f and a preference extension E , does f always output committees that are Pareto-optimal with respect to E ?

Some of the common preference extension principles extend rankings over singletons to *partial* orders over committees. In such a case, we are sometimes not able to say whether a committee is preferred by a voter to another one. It is however possible to consider, for such extension principles, two modal notions: *possible* and *necessary Pareto-optimality* and *efficiency*.³ A committee is possibly Pareto-optimal (PPO) if it is Pareto-optimal for *some* completion of these partial preferences, and necessarily Pareto-optimal (NPO) if it is Pareto-optimal for *all* completions of these partial preferences. (Obviously, when an extension principle outputs a complete preference relation then both notions coincide.) These PPO and NPO notions with respect to an extension principle carry on to multiwinner voting rules: a voting rule is necessarily (respectively, possibly) Pareto-efficient if all the committees it outputs are NPO (respectively, PPO). As the rules we consider are irresolute, we also introduce a weaker notion: a rule is weakly PPE if at least one of its output committees is PPO.

Our aim is to study the Pareto-efficiency of most well-studied ordinal multiwinner voting rules under two classical preference extension principles that are especially relevant for our setting.

The central preference extension principle we consider is the *responsive extension*, which is particularly suitable for the context of multiwinner elections, since it assumes that voters have additively decomposable preferences over committees. The responsive extension can be seen as the ordinal counterpart of additivity. It has been introduced for the first time by [27], in the context of one-to-many matching, and studied further in [26, 8]. It is arguably the most suitable preference extension principle under the conjunctive interpretation, and has been used several times in this context, especially in matching [19, 7], fair division [2, 9, 5, 28], committee selection [3, 1] and coalition formation [25, 18]. As it produces a partial order, we will consider its “possible” and “necessary” versions, defined by quantifying over complete extensions of these partial preferences.

On the profile P introduced above, $\{a, b\}$ necessarily Pareto-dominates $\{x, y\}$: if every voter has an additively decomposable preference consistent with their ordinal preferences over single candidates, then whatever the choice of the utility values, $\{a, b\}$ Pareto-dominates $\{x, y\}$. If we replace the fourth vote by $x \succ y \succ b \succ a$, then $\{b, y\}$ is not necessarily Pareto-optimal; however, it is possibly Pareto-optimal (see details in the long version).

Beyond the responsive extension, we also consider the *lexicographic* (or *leximax*) extension principle, introduced in [8], and used in voting contexts in [20, 22, 5, 3, 1]. As it is complete, Pareto-efficiency is directly applicable (we do not need to distinguish between possible and necessary Pareto-optimality) and as it is a refinement of the responsive extension, Pareto-efficiency for the lexicographic extension is stronger than possible Pareto-efficiency and weaker than necessary Pareto-efficiency.

We do not consider other extension principles; we explain how below.

³We use the term *optimality* for committees and the term *efficiency* for rules.

The lexicographic extension principle is a refinement of the ‘best’ extension principle [3], which orders committees only according to their best element. This extension principle is rougher, less sensitive to voters’ preferences, than the lexicographic extension principle. As it is a coarsening of it, Pareto-efficiency for the lexicographic extension implies Pareto-efficiency for the ‘best’ extension, so all our positive results about lexicographic extension implies Pareto-efficiency for the ‘best’ extension.

Instead of focusing on the best element we could focus on the worst element and define a *leximin* extension principle, defined exactly as the (or *leximax*) principle, but starting from the worst committee members instead of the best ones; this is a refinement of the ‘worst’ extension principle [3]. It has been argued in several places (see, e.g., [30] and the references therein) that in multiwinner voting, focusing on the best element (the best representative of a voter in a committee, her best item or nearest facility in a set) generally makes more sense than focusing on the worst element. Also, beyond focusing on the best or the worst elements, there’s a continuum of possibilities ([30]; see in particular the discussion in Section 3); as they are less common we will not consider them here.

Finally, we do not consider extension principles that make sense for the disjunctive interpretation of sets (used, for instance, for studying the axiomatic properties of irresolute rules), but that make little or no sense for the “sets of final outcomes” interpretation: Fishburn’s, Gärdenfors’ and Kelly’s [6].

Our results allow us to classify rules into six classes, according to their level of Pareto-efficiency:

- *Class 6*: those for which all winning committees are necessarily Pareto-optimal for the responsive extension (from now on we will simply say “necessarily Pareto-optimal”): they are said to be *necessarily Pareto efficient* (NPE). Such rules are rare; still, the *perfectionist* rule [15] is NPE.
- *Class 5*: those for which winning committees are lexicographically Pareto-optimal (they are said to be *lexicographically Pareto efficient*, or LPE) but not always necessarily Pareto-optimal. This is the case for some committee scoring rules, as well as for sequential rules such as Single Transferable Vote (STV), sequential plurality and sequential Chamberlin-Courant.
- *Class 4*: those that may output committees that are not lexicographically Pareto-optimal, but ensure that for any input profile, at least one of the committees in the output is lexicographically Pareto-optimal (these rules are said to be *weakly lexicographically Pareto-efficient*, WLPE).
- *Class 3*: those that are not LPE, but for which all winning committees are possibly Pareto-optimal for the responsive extension (from now on we will simply say “possibly Pareto-optimal”); they are said to be *possibly Pareto efficient* (PPE). This class contains compromise rules, all committee scoring rules with a strictly decreasing scoring function that fail LPE, and more generally a large fraction of committee scoring rules.
- *Class 2*: those that may output committees that are not possibly Pareto-optimal, but ensure that for any input profile, at least one of the committees in the output is possibly Pareto-optimal; these rules are said to be *weakly possibly Pareto-efficient* (WPPE). This is notably the case for committee scoring rules that are not PPE, including Single Non-Transferable Vote (SNTV), Bloc, and most variants of Chamberlin-Courant rules.
- *Class 1*: finally, some rules are not even WPPE. This class includes, typically, Condorcetian rules such as Number of External Defeats (NED) or Minimum Size of External Opposition (SEO).

Section 2 gives the necessary background on ordinal multiwinner voting rules and preference extensions. Section 3 defines possible, lexicographic and necessary Pareto-optimality (respectively efficiency) of committees (respectively rules). Section 4 identifies the “level of Pareto-efficiency” of all our (families of) rules. The Appendix contains further results, examples, as well as all proofs that are not included in the 12-page extended abstract.

2 Background

2.1 Basic notions and notation

Let N be a set of voters with $|N| = n \geq 2$ and A a set of candidates with $|A| = m \geq 3$. $\mathcal{L}(A)$ is the set of linear orders (or rankings) over A . A profile P is a collection of *votes*, each vote being a linear order over candidates: $P = (\succ_1, \dots, \succ_n) \in \mathcal{L}(A)^n$. To avoid overloaded notation, when writing votes we often omit the $\succ$ symbol: for instance, the vote $a \succ b \succ c \succ d$ is simply written $abcd$. Also, when several votes in a profile are identical, we use the following notation: $P = (2 \times abcd, 3 \times dcba)$ is the profile containing 3 votes $a \succ b \succ c \succ d$ and 2 votes $d \succ c \succ b \succ a$.

If $k \in \{1, \dots, m-1\}$ then $\mathcal{S}_k(A) = \{X \subset A : |X| = k\}$. A member of $\mathcal{S}_k(A)$ is called a *committee* of size k , or k -committee. We generally omit curly brackets when writing committees: for $A = \{a, b, c, d\}$ we write $\mathcal{S}_2(A) = \{ab, ac, ad, bc, bd, cd\}$ instead of $\{\{a, b\}, \{a, c\}, \{a, d\}, \{b, c\}, \{b, d\}, \{c, d\}\}$.

$W(\mathcal{S}_k(A))$ is the set of weak orders over $\mathcal{S}_k(A)$ and $\Pi(\mathcal{S}_k(A))$ is the set of strict partial orders over $\mathcal{S}_k(A)$. The strict partial order associated with $\sqsubseteq_i \in W(\mathcal{S}_k(A))$ is defined as usual by $x \sqsubseteq_i y$ if $x \sqsubseteq_i y$ and not $y \sqsubseteq_i x$. If $\sqsubseteq_1, \dots, \sqsubseteq_n$ are weak orders over $\mathcal{S}_k(A)$, $Q = (\sqsubseteq_1, \dots, \sqsubseteq_n)$ is called a weak order profile over $\mathcal{S}_k(A)$. Given a strict partial order $\sqsubseteq_i$ over $\mathcal{S}_k(A)$, a *completion* of $\sqsubseteq_i$ is a complete weak order $\sqsubseteq_i$ over $\mathcal{S}_k(A)$ such that for all $x, y \in A$, if $x \sqsubseteq_i y$ then $x \sqsubseteq_i y$.

2.2 Ordinal multi-winner rules

An (ordinal) multiwinner rule is a function f that, given a profile P and an integer $k \leq m$, outputs a nonempty subset of $\mathcal{S}_k(A)$. We list below several (most?) prominent ordinal multiwinner rules that have been well-studied in the literature. The most recent survey on ordinal multiwinner rules is [14]. See the Appendix for a detailed description of the rules and detailed examples.

Committee scoring rules (CSR) [13] Given a vote $\succ_i$ and $S \in \mathcal{S}_k(A)$, we denote by $pos(S, \succ_i)$ the sequence of positions of the members of S sorted increasingly: for example, $pos(\{b, d\}, abcd) = (2, 4)$. A *committee scoring function* γ maps every such sequence of positions into a number, and satisfies a monotonicity condition. Each vote $\succ_i$ gives a score $\gamma(pos(S, \succ_i))$ to each committee S . The winning committee(s) is (are) the one(s) maximizing the sum of scores for all votes $\sum_{i=1}^n \gamma(pos(S, \succ_i))$. Some particular CSRs: *best- k* rules (such as k -Borda and Single Nontransferrable Vote) correspond to additively decomposable scoring functions; the *Bloc* rule, with $\gamma(pos(S, \succ_i))$ being the number of candidates in S ranked in the first k positions of $\succ_i$; the *perfectionist* rule, that outputs the set S of k candidates with the largest number of voters whose set of preferred k candidates is exactly S (in any order); and the family of *Chamberlin-Courant* rules, where $\gamma(S, \succ_i)$ is the score of the best candidate in S for $\succ_i$.

Condorcetian rules Two common ways of extending the Condorcet criterion from single candidates to committees: Gehrlein stability (a committee S is Gehrlein stable if every $x \in S$ majority defeats every $y \in A \setminus S$) [16, 17], and local stability for quota q (S is locally stable for q if for any $y \in A \setminus S$, at least qn voters prefer some candidate in X to y [12]). A rule is Gehrlein-consistent if it elects the Gehrlein stable committee whenever there exists one. Two specific Gehrlein-consistent rules are NED (for “number of external defeats”) and SEO (for “size of external opposition”) [11], that can be seen as respective multiwinner counterparts of the Copeland and maximin single-winner rules. We also consider the *locally stable extension of maximin* (LSE-maximin) [4]: it outputs the sets S that are locally stable for the maximal possible quota q .

Compromise rules Let $\alpha \in [0, 1]$. For a profile P and a candidate x , $\lambda(\alpha, P, x)$, is the smallest integer j such that x appears in the first j positions in more than αn votes. The compromise rule MC_k^α [29, 31] identifies the smallest integer j such that there exist at least k alternatives with $\lambda(\alpha, P, x) \leq j$, and

then outputs the k alternatives with the smallest values of $\lambda(\alpha, P, x)$; ties are broken according to the number of voters who rank them in the first j positions. Note that $MC_k^{1/2}$, called *majoritarian compromise*, is a multiwinner version of the *Bucklin* rule.

Sequential rules There are several variants of multiwinner *single transferable vote* (STV). We present the most common one: let $q = \lceil \frac{n}{k} \rceil$ (quota). If some candidate x has a plurality score $S(x) \geq q$, then x is elected, and each of the votes for x becomes a fractional vote with weight $1 - \frac{q}{S(x)}$, with x removed; otherwise the candidate with the lowest plurality score is eliminated from all votes. This operation is repeated until k candidates have been elected. *Sequential plurality* elects first the plurality winner, removes it from the list of candidates, then elects the plurality winner from the obtained profile, and so on until k candidates have been elected. *Greedy s-Chamberlin-Courant* rules select candidates in sequence according to their marginal contribution to the Chamberlin-Courant score of the committee with respect to a scoring vector s .

2.3 Preference extensions

An *extension principle* is a function $E : \mathcal{L}(A) \rightarrow \Pi(\mathcal{S}_k(A))$: E maps a linear order over candidates to a strict order over k -committees. We note $\succ_i^E$ for $E(\succ_i)$. The implicit assumption is that i 's actual preference $\sqsubseteq_i$ is compatible with $\succ_i^E$, i.e., that it is one of its completions. We write, for each $\succ_i \in \mathcal{L}(A)$,

$$\kappa^E(\succ_i) = \{\sqsubseteq_i \in W(\mathcal{S}_k(A)) : \sqsubseteq_i \text{ is a completion of } \succ_i^E\}$$

and, by a slight abuse of notation, for each profile $P = (\succ_1, \dots, \succ_n)$,

$$\kappa^E(P) = \kappa^E(\succ_1) \times \dots \times \kappa^E(\succ_n).$$

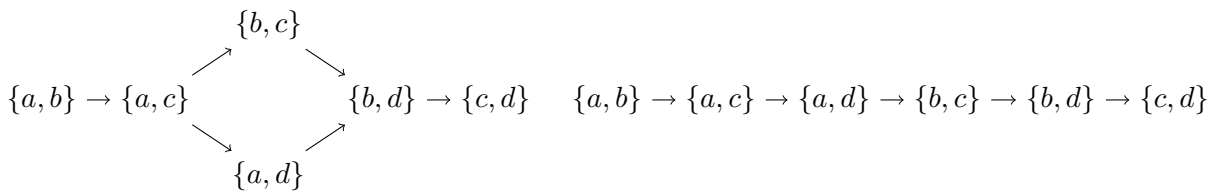
The *responsive extension principle* ρ [27] says that for any subset A of candidates containing x and not containing y , if B is obtained from A by replacing x by y , then B is preferred to A if and only if y is preferred to x . Formally, given any $X, Y \in \mathcal{S}_k(A)$ and any $\succ_i \in \mathcal{L}(A)$, we say that Y is an elementary improvement for X at $\succ_i$ if and only if $Y = (X \setminus \{x\}) \cup \{y\}$ for some $x \in X$ and $y \in A \setminus X$ with $y \succ_i x$. The responsive extension ρ of $\succ_i$ is then defined as the inclusion-wise smallest transitive relation satisfying $Y \succ_i^\rho X$ whenever Y is an elementary improvement for X at $\succ_i$.

The responsive extension $\succ_i^\rho$ can be characterised equivalently by stochastic dominance. For any $h \in \{1, \dots, |X|\}$, we write $r_h(X; \succ_i) \in X$ for the h th ranked alternative in $X \subseteq A$ at $\succ_i \in \mathcal{L}(A)$. At each $\succ_i \in \mathcal{L}(A)$ and for any distinct $X, Y \in \mathcal{S}_k(A)$, we define the stochastic dominance relation $\sigma^k(\succ_i)$ over $\mathcal{S}_k(A)$ as $X \sigma^k(\succ_i) Y$ iff $r_h(X; \succ_i) = r_h(Y; \succ_i)$ or $r_h(X; \succ_i) \succ_i r_h(Y; \succ_i)$ for all $h \in \{1, \dots, k\}$.

The *lexicographic extension principle* lex maps every $\succ_i \in \mathcal{L}(A)$ to the linear order $lex(\succ_i) = \succ_i^{lex}$ in $\mathcal{L}(\mathcal{S}_k(A))$ defined by: for $X \in \mathcal{S}_k(A)$ and $Y \in \mathcal{S}_k(A)$, $X \succ_i^{lex} Y$ if and only if for some $h^* \in \{1, \dots, k\}$, the following two conditions hold: (1) $r_{h^*}(X; \succ_i) < r_{h^*}(Y; \succ_i)$, and (2) $r_h(X; \succ_i) = r_h(Y; \succ_i)$ for all $h \in \{1, \dots, h^* - 1\}$.

Unlike the responsive extension, the lexicographic extension outputs a total order on $\mathcal{S}_k(A)$, which is one of the completions of the responsive extension. Therefore, $X \succ_i^{lex} Y$ implies $X \succ_i^\rho Y$.

Example 2.1. Let $A = \{a, b, c, d\}$, $k = 2$, and $a \succ_1 b \succ_1 c \succ_1 d$. The responsive extension $\kappa^\rho(\succ_1)$ is the partial order on $\mathcal{S}_2(A)$ and the lexicographic extension $\succ_1^{lex}$ of $\succ_1$ on the figure below.



Responsive (left) and lexicographic (right) extensions of $a \succ_1 b \succ_1 c \succ_1 d$.

3 Pareto-optimality and Pareto-efficiency

As already mentioned, Pareto-optimality is a property conditional to the choice of an extension principle.

As the lexicographic extension principle generates a linear order over $\mathcal{S}_k(A)$, Pareto-optimality is defined in a natural way.

Definition 3.1. Given a profile $P = (\succ_i)_{i \in N} \in \mathcal{L}(A)^N$, and two committees $X, Y \in \mathcal{S}_k(A)$, Y lexicographically Pareto-dominates X at P if $Y \succ_i^{lex} X$ holds for every $i \in N$, and X is lexicographically Pareto optimal at P if it is not lexicographically Pareto-dominated by any other committee in $\mathcal{S}_k(A)$.

As the responsive extension generates only a partial order over committees, we cannot directly apply Pareto-optimality. One classical way of extending to collections of partial orders a notion that usually applies to collections of total orders consists in quantifying over completions. Possible and necessary Pareto-efficiency correspond respectively to existential and universal quantification (see [10, 9, 2, 5]).

We start by defining Pareto-dominance. We recall that ρ is the responsive extension principle.

Definition 3.2. Let $X, Y \in \mathcal{S}_k(A)$.

- Y Pareto-dominates X at $Q = (\sqsupseteq_1, \dots, \sqsupseteq_n) \in W(\mathcal{S}_k(A))^N$ if $Y \sqsupseteq_j X$ for all $j \in N$ and $Y \sqsupseteq_i X$ for some $j \in N$.
- Y possibly Pareto-dominates X at $P \in \mathcal{L}(A)^N$ for ρ if Y Pareto-dominates X at Q for some $Q \in \rho(P)$.
- Y necessarily Pareto-dominates X at $P \in \mathcal{L}(A)^N$ for ρ if Y Pareto-dominates X at Q for every $Q \in \rho(P)$.

We now define Pareto optimality.

Definition 3.3. Let $X \in \mathcal{S}_k(A)$.

- X is Pareto optimal at $Q \in W(\mathcal{S}_k(A))^N$ if and only if there is no $Y \in \mathcal{S}_k(A)$ such that Y Pareto-dominates X at Q .
- X is necessarily Pareto optimal at $P \in \mathcal{L}(A)^N$ for ρ if and only if X is Pareto optimal at every $Q \in \rho(P)$.
- X is possibly Pareto optimal at $P \in \mathcal{L}(A)^N$ for ρ if and only if X is Pareto optimal at some $Q \in \rho(P)$.

The following is immediate:

Proposition 3.1.

1. if X is necessarily Pareto at P for ρ then X is lexicographically Pareto optimal at P .
2. if X is lexicographically Pareto optimal at P then X is possibly Pareto at P for ρ .

Proposition 3.2. For any $X \in \mathcal{S}_k(A)$ and $P \in \mathcal{L}(A)^N$:

1. X is necessarily Pareto optimal at P for ρ if and only if there is no $Y \in \mathcal{S}_k(A)$ that possibly Pareto-dominates X at P for ρ
2. X is possibly Pareto optimal at P for ρ if and only if there is no $Y \in \mathcal{S}_k(A)$ that necessarily Pareto-dominates X at P for ρ .

Example 3.1. Let $P = (efbdca, abcdef)$. At this profile,

- $\{a, b\}$ is LPO, because it is the most preferred committee for voter 2.
- $\{a, e\}$ is LPO: voter 1 lexicographically prefers $\{a, e\}$ to any committee that does not contain e , and voter 2 lexicographically prefers $\{a, e\}$ to any committee that does not contain a .
- $\{b, e\}$ is LPO: the only committee that voter 1 lexicographically prefers to $\{b, e\}$ is $\{e, f\}$, but voter 2 lexicographically prefers $\{b, e\}$ to $\{e, f\}$.
- $\{c, e\}$ is not LPO, as it is lexicographically Pareto-dominated by $\{b, e\}$.
- $\{b, f\}$ and $\{b, d\}$ are not LPO at P , as they are lexicographically Pareto-dominated by $\{a, e\}$.
- $\{a, b\}$ is NPO, because it is the most preferred committee for voter 2.
- $\{e, b\}$ is NPO: because of voter 1, the only committee that can possibly Pareto-dominate it is $\{e, f\}$; but voter 2 necessarily prefers $\{e, b\}$ to $\{e, f\}$;
- $\{a, e\}$ is not NPO, as it is possibly Pareto-dominated by $\{b, c\}$. As it is lexicographically Pareto optimal it is *a fortiori* possibly Pareto optimal.
- $\{b, d\}$ is not LPO, because it is lexicographically Pareto-dominated by $\{a, e\}$. It is PPO: if it was necessarily Pareto-dominated by another committee, this would be $\{b, e\}$, $\{b, f\}$, $\{d, e\}$, $\{d, f\}$ or $\{e, f\}$ because of voter 1; but voter 2 necessarily prefers $\{b, d\}$ to all of these.
- $\{d, f\}$ is not PPO, as it is necessarily Pareto-dominated by $\{b, e\}$.

We have defined so far possible and necessary Pareto optimality of a *committee* with respect to some extension principle. For multiwinner rules, we have to take irresoluteness into account. We define the following five levels of efficiency:⁴

Definition 3.4. Given an extension principle E , and a multiwinner voting rule f , we say that

- f is necessarily (resp. possibly, lexicographically) Pareto-efficient for E if for any profile P over A , every $S \in f(P)$ is necessarily (resp. possibly, lexicographically) Pareto-optimal for E .
- f is weakly possibly Pareto-efficient for E if for any profile P , some $S \in f(P)$ is possibly Pareto-optimal for E .
- f is lexicographically Pareto-efficient if for any profile P , every $S \in f(P)$ is lexicographically Pareto-optimal.
- f is weakly lexicographically Pareto-efficient if for any profile P , some $S \in f(P)$ is lexicographically Pareto-optimal.

Figure 1 shows the logical relations between the five levels of Pareto-efficiency.

As we see further, for some rules, weak possible Pareto-efficiency serves to guarantee possible Pareto-efficiency even for pathological profiles.

When $k = 1$, f becomes a single-winner rule. Then possible and necessary Pareto-efficiency reduce to standard Pareto-efficiency: if f is an irresolute single-winner voting rule f with ordinal input, f is Pareto-efficient if for every profile P , every $x \in f(P)$ is Pareto-optimal. Let us say that f is *weakly*

⁴We do not define a weak version of necessary Pareto-efficiency, which would have little interest: we do not know any interesting rule that is guaranteed to output *some* necessarily Pareto-optimal committee, but that fails necessary Pareto-efficiency. (Of course, such rules exist: for instance, the rule that outputs all possible committees.)

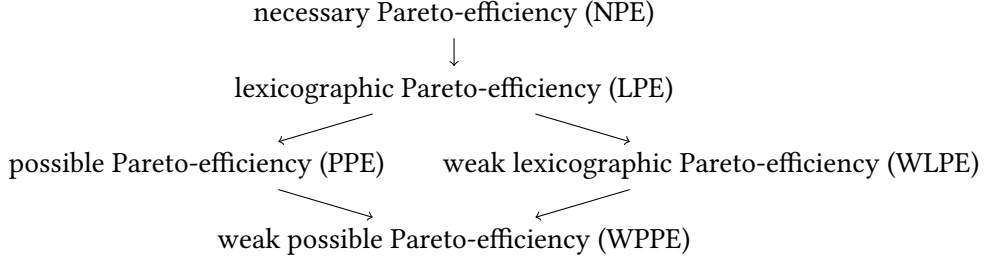


Figure 1: Relations between the five levels of Pareto-efficiency

Pareto-efficient if for every profile P , some $x \in f(P)$ is Pareto-optimal. As far as we can tell, all irresolute single-winner rules that have received some attention in the literature satisfy at least weak Pareto-efficiency. Most of them satisfy the stronger Pareto-efficiency property; a few exceptions are some positional scoring rules with a scoring vector that is not strictly decreasing (such as k -approval for $k \geq 2$),⁵ as well as maximin, and tournament solutions such as the Top Cycle and the Banks set [23].

4 Detailed analysis of rules

Now we proceed and identify the Pareto-efficiency levels of all the rules we introduced in Section 2. The presentation structure of the results differs from that in the journal paper.

4.1 Committee scoring rules

Now we proceed and identify the Pareto-efficiency levels of all the rules we introduced in Section 2.

Proposition 4.1. *Every CSR f_γ is WPPE. If γ is strict then f_γ is PPE.*

When γ is not strict, PPE can fail indeed: for instance, for SNTV, $k = 2$, $P = (abc, abc)$, SNTV outputs $\{\{a, b\}, \{a, c\}\}$; the latter is not PPO. This failure of PPE, however, occurs only in the case where at least k alternatives are ranked first by some voter is LPE.

The question of identifying nonstrict CSRs that are PPE is nontrivial, even for best- k rules. (See a discussion in the Appendix.)

We can say something stronger about SNTV.

Proposition 4.2. *SNTV is WLPE but not LPE. Its restriction to profiles where at least k alternatives are ranked first by some voter is LPE.*

Proposition 4.3. *Bloc fails PPE and WLPE.*

Proof. For PPE, consider the profile $P = (zyxuv, zyxuv, xuzyv, xvzyu)$: $\{x, y\}$ is winning but necessarily Pareto-dominated by $\{x, z\}$, therefore not PPO. For WLPE, let $k = 3$, $m = 6$, and $P = (axyzbc, byzxca, czxyab)$. The winning committee $\{x, y, z\}$ is lexicographically Pareto-dominated by $\{a, b, c\}$. ■

Proposition 4.4. *β -CC is WLPE and not PPE.*

Proof. If $S \in \beta\text{-CC}(P)$ is lexicographically Pareto-dominated, then some S' lexicographically dominates S : then, for each voter i , her best candidate in S' is at least as good as her best candidate in S , therefore

⁵For an exact characterization of Pareto-efficient positional scoring rules see [24].

the β -CC score of S' is no smaller than the β -CC score of S . We iterate this process until we reach a lexicographically optimal committee S^* , which is also in $\beta\text{-CC}(P)$. This implies that $\beta\text{-CC}$ is WLPE.

It fails PPE: take any profile such that the top alternative is the same (x^*) in all votes, and take two alternatives z, y such that $y \succ z$ in all votes. Any committee containing x^* and z , but not y , is a winning committee although it is necessarily Pareto-dominated. ■

In the Appendix we define a lexicographic refinement of $\beta\text{-CC}$, called $\beta\text{-CC}^*$, and prove that it is LPE.

k -Borda is PPE, as it is based on a strict γ , but fails WLPE: for the profile $(abcde-, ebcda-)$, the unique winning committee for $k = 2$ is $\{b, c\}$, which is lexicographically Pareto-dominated by $\{a, e\}$.⁶ However, the failure of PPE for k -Borda, occurs only for pathological profiles for which the number of candidates appearing in top position in some votes is less than k .

Proposition 4.5. *The perfectionist rule is NPE.*

Proof. For any committee S output by the perfectionist rule there is at least one voter i such that the set of i 's top k candidates is S . Therefore i necessarily prefers S any other committee S' , and S is necessarily Pareto-optimal. ■

On the negative side, we have this general result (whose proof is rather long):

Proposition 4.6. *For any $k \geq 2$ and $m \geq 2k + 1$, no best- k rule satisfies necessary Pareto-efficiency.*

Characterizing LPE and WLPE committee scoring rules appears to be difficult in the general case. As we already observed in Section 3, when $k = 1$, LPE coincides with standard Pareto-efficiency, and characterizing Pareto-efficient single-winner positional scoring rules is already not trivial [24]. Therefore, we should not expect to obtain an easy characterization of lexicographic Pareto-efficiency of multiwinner CSRs in the general case. But somewhat surprisingly, obtaining such a characterization is difficult even for the simple case of $k = 2$ and “best- k ” committee scoring rules.

To give an idea of the difficulty, let us restrict to additive CSRs and consider the case $k = 2, m = 4$. An additive CSR for $k = 2$ is associated with an additive scoring function $\gamma_{m,2}$: there exists a non-increasing scoring vector $(s_1, \dots, s_m)$, with $s_1 > s_m$, such that $\gamma_{m,2}(i, j) = s_i + s_j$. Without loss of generality, we assume $s_m = 0$. Given profile $\succ = (\succ_1, \dots, \succ_n)$, recall that $\text{Score}(x, \succ) = \sum_{i=1}^n s_{\text{pos}(x, \succ_i)}$. We denote by f_2^s the corresponding additive CSR.

Proposition 4.7. *If $m = 4$ and $k = 2$, then f_2^s is LPE if and only if $s_1 > s_2 + s_3$ and $s_2 > s_3$.*

4.2 Sequential rules

Given any $P_i \in \mathcal{L}(A)$ together with $B \subseteq A$, define $1(\succ_i, B)$ as the top candidate in B for $\succ_i$. We introduce below a property of multiwinner rules which has its own interest: *top-sequentiality* expresses that candidates are selected in a sequence, and at each step, the selected candidate is the most preferred candidate, among those that have not been selected yet, for at least one voter. For instance, sequential dictatorship — where a voter picks her preferred candidate, then a second voter (who can be the same one) picks her preferred candidate among those who remain, etc. — is obviously top-sequential. As we see below, many more interesting rules are top-sequential as well.

Definition 4.1. *A multiwinner rule f is top-sequential if for all $P \in \mathcal{L}(A)^n$, $S = \{a_1, \dots, a_k\} \in f(P)$, one can order candidates $a_1, \dots, a_k$ so that $\forall h \in \{1, \dots, k\}$, $a_h = 1(\succ_{i_h}, A \setminus \{a_1, \dots, a_{h-1}\})$ for some $i_h \in N$.*

⁶ k -Borda is however WLPE for $m = 4$ and $k = 2$.

Lemma 4.1. *Sequential plurality and STV are top-sequential.*

Proof. For Sequential plurality: at each step, the selected candidate maximizes the plurality score among the remaining candidates; therefore it is ranked first by at least one voter. For STV: at each step where a candidate is selected, it reaches the quota, therefore it is ranked first by at least one voter. ■

Proposition 4.8. *Every top-sequential rule is LPE.*

Proof. Let f be top-sequential and pick $P \in \mathcal{L}(A)^n$ and $X \in f(P)$. Writing $S = \{a_1, \dots, a_k\}$, and assuming elements of S are selected w.r.t. order $a_1 > \dots > a_k$, the definition of top-sequentiality implies that for all $h \in \{1, \dots, k\}$, $a_h = 1(\succ_{i_h}, A \setminus \{a_1, \dots, a_{h-1}\})$ for some $i_h \in N$. If $S' \succ_i^{lex} S$ for all $i \in N$, one must have $a_1 \in S'$ (otherwise, by definition of ρ^{lex} , $S \succ_{i_1}^{lex} S'$). Replicating this argument for $a_2, \dots, a_k$ shows that $S = S'$, which is impossible. Hence, S is lex-Pareto optimal at P , which shows that f satisfies LPE. ■

As an immediate consequence of Proposition 4.8 and Lemma 4.1 we have

Proposition 4.9. *Sequential plurality and STV are LPE.*

Proposition 4.8 can also be used to prove that sequential dictatorships are LPE, which can also be obtained as a by-product of Theorem 2 in [1].

Proposition 4.10. *STV and sequential plurality fail necessary Pareto-efficiency.*

The last remaining sequential rule is greedy β -CC, for which we have a negative result.

Proposition 4.11. *Greedy β -CC is not WLPE.*

4.3 Compromise rules

We have these general results, one positive and one negative (proofs in Appendix).

Proposition 4.12.

1. For any $\alpha \in [0, 1)$, MC_k^α is PPE.
2. For any $\alpha \in (0, 1)$ and $k \geq 2$, MC_k^α is not WLPE.

4.4 Condorcetian rules

When $k = 1$, SEO is the *maximin* rule, known to fail Pareto-efficiency, but to satisfy weak Pareto-efficiency. It turns out that for $k > 1$, SEO does not even satisfy weak possible Pareto-efficiency.

Proposition 4.13. *SEO is not WPPE.*

Proof. We give a counterexample for $n = 20$, $k = 2$, $m = 12$, $C = \{b, d\} \cup A^* \cup C^*$ where $A^* = \{a_1, \dots, a_5\}$ and $C^* = \{c_1, \dots, c_5\}$.

P is the following profile with 20 voters grouped in four types:

- 5 $A^* \succ b \succ C^* \succ d$
- 5 $A^* \succ d \succ C^* \succ b$
- 5 $C^* \succ b \succ A^* \succ d$
- 5 $C^* \succ d \succ A^* \succ b$

The 5 voters of a given type have cyclic preferences over the clone sets A^* and C^* : for instance, those of type $A^* \succ b \succ C^* \succ d$ are

- $a_1 \succ a_2 \succ a_3 \succ a_4 \succ a_5 \succ b \succ c_1 \succ c_2 \succ c_3 \succ c_4 \succ c_5 \succ d$
- $a_2 \succ a_3 \succ a_4 \succ a_5 \succ a_1 \succ b \succ c_2 \succ c_3 \succ c_4 \succ c_5 \succ c_1 \succ d$
- $a_3 \succ a_4 \succ a_5 \succ a_1 \succ a_2 \succ b \succ c_3 \succ c_4 \succ c_5 \succ c_1 \succ c_2 \succ d$
- $a_4 \succ a_5 \succ a_1 \succ a_2 \succ a_3 \succ b \succ c_4 \succ c_5 \succ c_1 \succ c_2 \succ c_3 \succ d$
- $a_5 \succ a_1 \succ a_2 \succ a_3 \succ a_4 \succ b \succ c_5 \succ c_1 \succ c_2 \succ c_3 \succ c_4 \succ d$

For $S \in \mathcal{S}_2(C)$, let $SEO(S) = \min_{x \in S, y \in C \setminus S} |\{i : x \succ_i y\}|$. For each $i \in \{1, \dots, 5\}$, 16 voters out of 20 prefer a_i to $a_{i+1[5]}$, therefore, if $S \cap A^* \neq \emptyset$ then $SEO(S) \leq 4$. Similarly, if $S \cap C^* \neq \emptyset$ then $SEO(S) \leq 4$. Finally, for each i , 15 voters out of 20 prefer the a_i 's to b , 15 prefer the a_i 's to d , 15 prefer the c_i 's to b , and 15 prefer the c_i 's to d ; therefore, $SEO(\{b, d\}) = 5$. The only winning committee is $\{b, d\}$; however, it is necessarily Pareto-dominated by $\{a_1, c_1\}$. ■

We do not have any better news with NED. While, for $k = 1$, NED is the *Copeland* rule, which is Pareto-efficient, for $k > 1$ we do not even have weak possible Pareto-efficiency.

Proposition 4.14. *NED is not WPPE.*

The proof is in the Appendix.

Given that NED and SEO are Gehrlein-consistent but fail possible Pareto-efficiency, we may wonder whether Gehrlein-consistency and possible Pareto-efficiency are compatible. The answer is positive. We even have this more general result:

Proposition 4.15. *A Gehrlein stable committee is possibly Pareto-optimal.*

Proof. For any profile P and $x, y \in C$, let $W(x, y, P) = |\{i : x \succ_i y\}|$. For any committees $S, S' \in \mathcal{S}_k(A)$, define $G(S, S', P) = \sum_{x \in S} \sum_{y \in S'} W(x, y, P)$. Assume $S' \in \mathcal{S}_k(A)$ necessarily Pareto-dominates $S \in \mathcal{S}_k(A)$. Then $S' \setminus S$ necessarily Pareto-dominates $S \setminus S'$. Let $|S' \setminus S| = |S \setminus S'| = r$. For every voter i , let $S' \setminus S = \{x_1^i, \dots, x_r^i\}$, with $x_1^i \succ_i x_2^i \succ_i \dots \succ_i x_r^i$. Then, for each $j = 1, \dots, r$, i prefers x_j^i to at most $r - j + 1$ candidates in $S \setminus S'$. This implies $G(S \setminus S', S' \setminus S, P) \leq n(1 + \dots + (r - 1)) = \frac{r(r-1)}{2}n < r^2 \frac{n}{2}$. Now, assume S is Gehrlein stable for P ; then for each $x \in S \setminus S'$ and $y \in S' \setminus S$, $W(x, y, P) > \frac{n}{2}$, therefore $G(S \setminus S', S' \setminus S, P) > r^2 \frac{n}{2}$. The contradiction between the two inequalities imply that S cannot be both Gehrlein stable and necessarily Pareto-dominated. ■

It is easy to construct a rule that is both Gehrlein-consistent and PPE: given any PPE rule f , the rule that outputs the unique Gehrlein stable k -committee if there is one, and the winner of f otherwise, is Gehrlein stable and PPE.

Proposition 4.16. *LSE-maximin fails possible Pareto-efficiency, but satisfies weak possible Pareto-efficiency.*

The failure of LPE, and even WLPE, for Condorcetian rule, is a consequence of their failing possible Pareto-efficiency. We give here a stronger result:

Proposition 4.17. *When $k \geq 3$, no Gehrlein-consistent rule is weakly lexicographically Pareto-efficient.*

Proof. Let $k \geq 3$ and $P = (axyzbc-, byzxca-, czxyab-)$. This profile has a Gehrlein-stable committee $\{x, y, z\}$, which however is lexicographically Pareto-dominated by $\{a, b, c\}$. ■

For $k = 2$, however, lexicographically Pareto-efficiency and Gehrlein-consistency are compatible. (See the Appendix.)

5 Discussion

Table 1 summarizes all results.

At one extreme, we have a class (Class 1) of rules that fail even the weakest notion of Pareto-efficiency (WLPE): two Gehrlein-consistent rules (SEO and NED), and Greedy β -CC. Failing such a weak property

	WPPE	PPE	WLPE	LPE	NPE	class
CSRs	+	some	some	some	some	2, 3, 4, 5, 6
strict CSRs	+	+	some	some	some	3, 5, 6
best- k	+	some	some	some	-	2,3,4,5
SNTV	+	-	+	-	-	4
k -Borda	+	+	-	-	-	3
β -CC	+	-	+	-	-	4
β -CC*	+	+	+	+	-	5
perfectionist	+	+	+	+	+	6
Bloc	+	-	-	-	-	2
MC_k^α	+	+	-	-	-	3
STV	+	+	+	+	-	5
sequential plurality	+	+	+	+	-	5
greedy β -CC	-	-	-	-	-	1
LSE-maximin	+	-	-	-	-	2
Gehrlein-consistent rules	some	some	-	-	-	1,2,3
SEO, NED	-	-	-	-	-	1

Table 1: Multiwinner rules and degrees of Pareto-efficiency

sends a negative signal about these rules: they should be selected with care, and for good reasons that counterbalance this failure.

Just above, we find Class 2, containing LSE-Maximin and Bloc, who satisfy WPPE but nothing above. LSE-Maximin, SEO and NED are the three Condorcetian rules (extending Condorcet-consistent rules) we considered: the message is that Gehrlein stability does not fit well with Pareto (which perhaps does not come as a surprise). As for Bloc, this is one more slightly negative signal, which should contribute to be cautious about using it.

Two classes of rules are above this class. Class 3 contains compromise rules, and some committee scoring rules, including k -Borda, that satisfy PPE, but fail WLPE. WLPE is rather strong, so we can consider that some safety test is passed as to what concerns Pareto-efficiency, in the sense that this should not be a reason to exclude these rules.

The other class above Class 2 is Class 4, which is incomparable with Class 3. It contains some committee scoring rules, including β -CC and SNTV: they satisfy WPPE and WLPE but fail PPE, but this is mostly because of pathological profiles, so again they should probably not be excluded on this ground.

Now we move towards classes of rules that behave very well regarding Pareto-efficiency. Class 5 contains rules that satisfy LPE (but fail NPE). It contains some CSRs, including the new rule β -CC*, as well as STV and sequential plurality. Satisfying LPE is a good arguments to choose one of these rules in any context where the lexicographic extension makes sense, that is, when an agent pays more attention to her preferred alternative in a committee than on the other ones.

Finally, Class 6 is composed of rules that satisfy the strongest property (NPE). We have identified only one known rule that satisfies it (the perfectionist rule); on the other hand, this rule has so many other drawbacks that it should be chosen with extreme care. These rather negative finding about NPE sends the signal that this property is too strong, rather than the signal that we should select rules that satisfy it.

Acknowledgements

We are grateful to the two anonymous reviewers for helpful comments. This work was supported in part by project ANR-22-CE26-0019 “Citizens” and project PR[AI]RIE-PSAI ANR-23-IACL-0008, both funded by the Agence Nationale de la Recherche, as well as by the BILGI Research Development Innovation Programme, POLarization viewed from SOcial choice Perspective (POSOP).

References

- [1] Haris Aziz and Jérôme Monnot. Computing and testing Pareto optimal committees. *Auton. Agents Multi Agent Syst.*, 34(1):24, 2020.
- [2] Haris Aziz, Serge Gaspers, Simon Mackenzie, and Toby Walsh. Fair assignment of indivisible objects under ordinal preferences. *Artif. Intell.*, 227:71–92, 2015.
- [3] Haris Aziz, Jérôme Lang, and Jérôme Monnot. Computing Pareto optimal committees. In Subbarao Kambhampati, editor, *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI 2016, New York, NY, USA, 9-15 July 2016*, pages 60–66, 2016.
- [4] Haris Aziz, Edith Elkind, Piotr Faliszewski, Martin Lackner, and Piotr Skowron. The Condorcet principle for multiwinner elections: From shortlisting to proportionality. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI 2017, Melbourne, Australia, August 19-25, 2017*, pages 84–90, 2017.
- [5] Haris Aziz, Péter Biró, Jérôme Lang, Julien Lesca, and Jérôme Monnot. Efficient reallocation under additive and responsive preferences. *Theor. Comput. Sci.*, 790:1–15, 2019.
- [6] S. Barbera, W. Bossert, and P. K. Pattanaik. Ranking sets of objects. In *Handbook of Utility Theory, volume II*, chapter 17, page 893–977. Kluwer Academic Publishers, 2004.
- [7] Khaled Belahcène, Vincent Mousseau, and Anaëlle Wilczynski. Combining fairness and optimality when selecting and allocating projects. In *IJCAI*, pages 38–44, 2021.
- [8] W. Bossert. Preference extension rules for ranking sets of alternatives with a fixed cardinality. *Theory and Decision*, 39:301–317, 1995.
- [9] Sylvain Bouveret, Ulle Endriss, and Jérôme Lang. Fair division under ordinal preferences: Computing envy-free allocations of indivisible goods. In *ECAI 2010 - 19th European Conference on Artificial Intelligence, Lisbon, Portugal, August 16-20, 2010, Proceedings*, pages 387–392, 2010.
- [10] Steven Brams, Paul H. Edelman, and Peter C. Fishburn. Fair division of indivisible items. *Theory and Decision*, 55:147–180, 2003.
- [11] Danilo Coelho. *Understanding, Evaluating and Selecting Voting Rules Through Games and Axioms*. PhD thesis, 2004.
- [12] Edith Elkind, Jérôme Lang, and Abdallah Saffidine. Condorcet winning sets. *Soc. Choice Welf.*, 44(3):493–517, 2015.
- [13] Edith Elkind, Piotr Faliszewski, Piotr Skowron, and Arkadii Slinko. Properties of multiwinner voting rules. *Soc. Choice Welf.*, 48(3):599–632, 2017.
- [14] Piotr Faliszewski, Piotr Skowron, Arkadii Slinko, and Nimrod Talmon. Multiwinner voting: A new challenge for social choice theory. In Ulle Endriss, editor, *Trends in Computational Social Choice*. 2017.
- [15] Piotr Faliszewski, Piotr Skowron, Arkadii Slinko, and Nimrod Talmon. Multiwinner analogues of the plurality rule: axiomatic and algorithmic perspectives. *Soc. Choice Welf.*, 51(3):513–550, 2018.
- [16] W. Gehrlein. The Condorcet criterion and committee selection. *Mathematical Social Sciences*, 10(3):199–209, 1985.
- [17] Eric Kamwa. Stable rules for electing committees and divergence on outcomes. *Group Decision and Negotiation*, 26(3):547–564, 2017.
- [18] Anna Maria Kerkmann, Jérôme Lang, Anja Rey, Jörg Rothe, Hilmar Schadrack, and Lena Schend. Hedonic games with ordinal preferences and thresholds. *J. Artif. Intell. Res.*, 67:705–756, 2020.

- [19] R. Khare, S. Roy, and T. Storcken. Stability in matching with couples having responsive preferences. Technical report, 2021.
- [20] Christian Klamler, Ulrich Pferschy, and Stefan Ruzika. Committee selection under weight constraints. *Math. Soc. Sci.*, 64(1):48–56, 2012. doi: 10.1016/j.mathsocsci.2011.11.006. URL <https://doi.org/10.1016/j.mathsocsci.2011.11.006>.
- [21] Martin Lackner and Piotr Skowron. Utilitarian welfare and representation guarantees of approval-based multiwinner rules. *Artif. Intell.*, 288:103366, 2020.
- [22] Jérôme Lang, Jérôme Mengin, and Lirong Xia. Voting on multi-issue domains with conditionally lexicographic preferences. *Artif. Intell.*, 265:18–44, 2018. doi: 10.1016/j.artint.2018.05.004. URL <https://doi.org/10.1016/j.artint.2018.05.004>.
- [23] Jean-Francois Laslier. *Tournament Solutions and Majority Voting*. Springer, 1997.
- [24] Bonifacio Llamazares and Teresa Peña. Scoring rules and social choice properties: some characterizations. *Theory and Decision*, 78(3):429–450, 2015.
- [25] Roberto Lucchetti, Stefano Moretti, and Tommaso Rea. Coalition formation games and social ranking solutions. In *AAMAS*, pages 862–870, 2022.
- [26] A. Roth and M. Sotomayor. Two-sided matching: A study in game-theoretic modeling and analysis. 1990.
- [27] Alvin E. Roth. The college admissions problem is not equivalent to the marriage problem. *Journal of economic Theory*, 36(2):277–288.
- [28] Erel Segal-Halevi, Avinatan Hassidim, and Haris Aziz. Fair allocation with diminishing differences. *J. Artif. Intell. Res.*, 67:471–507, 2020.
- [29] Murat R. Sertel. Lecture notes in microeconomics. 1986.
- [30] Piotr Skowron, Piotr Faliszewski, and Jérôme Lang. Finding a collective set of items: From proportional multirepresentation to group recommendation. *Artif. Intell.*, 241:191–216, 2016.
- [31] Bilge Yilmaz and Murat R. Sertel. The majoritarian compromise is majoritarian-optimal and subgame-perfect implementable. *Social Choice and Welfare*, 16(4):615–627, 1999.

Jean Lainé
 CNAM, Paris, France
 Email: jeanlainelirsa@gmail.com

Jérôme Lang
 Université Paris-Dauphine, Université PSL, CNRS, LAMSADE
 Paris, France
 Email: lang@lamsade.dauphine.fr

İpek Özkal-Sanver
 Bilgi University
 Istanbul, Turkey
 Email: ipek.sanver@bilgi.edu.tr

Remzi Sanver
 Université Paris-Dauphine, Université PSL, CNRS, LAMSADE
 Paris, France Email: sanver@lamsade.dauphine.fr

Appendix

A1. A zoo of ordinal multiwinner rules

A multiwinner rule is a function f that, given a profile P and an integer $k \leq m$, outputs a nonempty subset of $\mathcal{S}_k(A)$. We list here a few prominent multiwinner rules that have been well-studied in the literature.

Committee scoring rules (CSR)

Committee scoring rules were first defined in [13]. Given a vote $\succ_i$ and a candidate c , we denote by $\text{pos}(c, \succ_i)$ the position of c in $\succ_i$ (the top-ranked candidate has position 1, the one ranked last has position m). Given $S \in \mathcal{S}_k(A)$, the position of S in $\succ_i$, denoted by $\text{pos}(S, \succ_i)$ is the sequence of positions of the members of S sorted increasingly. We denote by $[m]_k$ the set of all size- k increasing sequences of elements from $\{1, \dots, m\}$. For $I = (i_1, \dots, i_k)$ and $J = (j_1, \dots, j_k)$ in $[m]_k$, we say that $I \succeq J$ if for each $t = 1, \dots, k$ we have $i_t \leq j_t$. We write $I \succ J$ for $I \succeq J$ and not $J \succeq I$.

A *committee scoring function* $\gamma_{m,k} : [m]_k \rightarrow \mathbb{R}$ associates each committee position with a score and satisfies *monotonicity*: if $I \succeq J$ then $\gamma_{m,k}(I) \geq \gamma_{m,k}(J)$. Moreover, $\gamma_{m,k}$ is *strict* if $\gamma_{m,k}(I) > \gamma_{m,k}(J)$ whenever $I, J \in [m]_k$ are such that $I \succ J$.

Given $\gamma_{m,k}$ and profile P , the committee scoring rule (CSR) $f_{\gamma_{m,k}}$, which we will write simply f_γ by abuse of notation, outputs committees S maximising

$$\text{score}(S, P) = \sum_{i=1}^n \gamma_{m,k}(\text{pos}(S, \succ_i))$$

A CSR f_γ is strict if γ is strict. A few well-known particular CSRs are listed below:

- A committee scoring function γ is *additively decomposable* if

$$\gamma_{m,k}(i_1, \dots, i_k) = \sum_{j=1}^k \gamma_m(i_j)$$

for some function $\gamma_m = \{1, \dots, m\} \rightarrow \mathbb{R}$. If γ is additively decomposable then f_γ outputs the k candidates x maximizing $\text{score}(x, P) = \sum_{i=1}^n \gamma_m(\text{pos}_{\succ_i}(x))$. Such a CSR is the natural multiwinner extension of a single-winner positional scoring rule; it is called a *best- k* CSR. Well-known examples of best- k CSR are Single Non-Transferable Vote (SNTV), defined by $\gamma_m(1) = 1$ and $\gamma_m(j) = 0$ for each $j > 1$, and k -Borda, defined by $\gamma_m(j) = m - j$ for all $j = 1, \dots, m$.

- if $\gamma_{m,k}(i_1, \dots, i_k) = |\{j : i_j \leq k\}|$ then f_γ is the *Bloc* rule. In words, the Bloc rule outputs the candidates listed most often in the top k candidates of the votes. Note that, although $\gamma_{m,k}$ is additively decomposable, Bloc is not a best- k rule, because $\gamma_{m,k}$ depends on k .
- if $\gamma_{m,k}(1, \dots, k) = 1$ and $\gamma_{m,k}(I) = 0$ for all $I \neq (1, \dots, k)$ then f_γ is called the *perfectionist rule* [15].
- let $s = (s_1, \dots, s_m)$ with $s_1 \geq \dots \geq s_m$ and $s_1 > s_m$. The Chamberlin-Courant k -multiwinner rule associated with scoring vector s , denoted by s -CC, is the CSR defined by $\gamma_{m,k}(i_1, \dots, i_k) = s_{i_1}$. If s is the Borda vector, defined by $s_i = m - i + 1$ for every i then f_γ is the *Borda-Chamberlin-Courant* rule (β -CC for short).

- A family of rules, which contains both Chamberlin-Courant and best- k rules, is obtained by using an ordered weighted average (OWA) to compute the satisfaction of an agent: the score of her j th best candidate in the selection is weighted by w_j [30]. For a reason that will become clear in Section ??, we consider a specific rule in this family, a lexicographic refinement of β -CC, which we denote by β -CC*: let $\varepsilon < \frac{1}{nm}$, then

$$\gamma_{m,k}(i_1, \dots, i_k) = \sum_{j=1}^k \varepsilon^{j-1} (m - i_j).$$

It can be checked easily that the winning committees do not depend on ε and that β -CC* is a refinement of β -CC.

Example 5.1. Let $A = \{a, b, c, d, e, f\}$, $n = 10$, $k = 2$, and $P = (4 \times fedbca, 3 \times abcdef, 2 \times bcaedf, 1 \times dcabef)$.

- $SNTV(P) = \{af\}$: f and a , in this order, are the two candidates ranked first in the largest number of votes.
- $2\text{-Borda}(P) = \{bc, bd\}$: b has the highest Borda score, followed by c and d (tied).
- $Bloc(P) = \{be, bf\}$: b is ranked in the top 2 positions in 5 votes; e and f , in 4 votes (and other candidates, in at most 3 votes).
- $\beta\text{-CC}(P) = \{af, bf\}$: 7 votes have a or f in first position, and for the other 3, the better candidate among a and f is in third position: the β -CC score of $\{a, f\}$ is $7 \times 5 + 3 \times 3 = 44$. Next, 6 votes have b or f in first position, 3 in second position, in one on fourth position: the β -CC score of $\{b, f\}$ is $6 \times 5 + 3 \times 4 + 2 = 44$. It can be checked that all other committees of size 2 have a smaller β -CC score.
- $\beta\text{-CC}^*(P) = \{bf\}$: the tie between the tied winning committees for β -CC is resolved by looking at the position of the second best (that is: worst!) committee member in all votes. For $\{a, f\}$, this second best candidate appears in the last position in all votes, while for $\{b, f\}$, it appears in the last position in 6 votes and in position 4 in 4 votes.
- the perfectionist rule applied to P outputs $\{ef\}$: 4 votes have $\{e, f\}$ as their top two elements, and no set of two candidates does better.

Condorcetian rules

Two ways of extending the Condorcet criterion from single winners to candidates are discussed in [4]: Gehrlein stability (a committee S is Gehrlein stable if every $x \in S$ majority defeats every $y \in A \setminus S$) [16, 17], and local stability for quota q (S is locally stable for quota q if for any $y \in A \setminus S$, at least qn voters prefer some candidate in S to y [12]).

A rule is Gehrlein-consistent if it elects the (unique) Gehrlein stable committee whenever there exists one. Two specific Gehrlein-consistent rules are NED (for “number of external defeats”) and SEO (for “size of external opposition”) [11], that can be seen as the respective multiwinner counterparts of the Copeland and maximin single-winner rules. The NED rule outputs committees S that maximize the number of pairs $(x, y) \in S \times A \setminus S$ such that x majority-beats y in P . The SEO rule outputs committees S that maximise $\min_{x \in S, y \in A \setminus S} |\{i : x \succ_i y\}|$. We give only one locally stable rule: the maximal θ -winning sets rule [12], also called the *locally stable extension of maximin* (LSE-maximin) in [4]: it outputs the sets S that are locally stable for the maximal possible quota q .

Continuing Example 5.1:

- there is no Gehrlein-stable committee; there is however a unique *weak* Gehrlein stable committee: each of b and c defeats a , e and f , and weakly defeats d . This leads to $\text{NED}(P) = \text{SEO}(P) = \{bc\}$;
- $\text{LSE-maximin}(P) = \{af\}$: $\{a, f\}$ is locally stable for $q = \frac{7}{10}$, because 8 voters prefer either a or f to b , 7 voters prefer either a or f to c , 9 voters prefer either a or f to d , and all voters prefer either a or f to e ; and no committee does better or equally good.

Compromise rules

Let $\alpha \in [0, 1)$. For a given profile P , for each alternative x , the *compromise index of x with respect to α and P* , $\lambda(\alpha, P, x)$, is the smallest integer j such that x appears in the first j positions in more than αn votes: that is, x appears in the first $\lambda(\alpha, P, x)$ positions in more than αn votes but in the first $\lambda(\alpha, P, x) - 1$ positions in at most αn votes.

The compromise rule MC_k^α [29, 31] identifies the smallest integer j such that there exist at least k alternatives with $\lambda(\alpha, P, x) \leq j$, and then outputs the k alternatives with the smallest values of $\lambda(\alpha, P, x)$; in case of a tie, meaning that there are more alternatives with $\lambda(\alpha, P, x) = j$ than necessary, the tie is broken according to the number of voters who rank them in the first j positions.

Note that $MC_k^{1/2}$, called *majoritarian compromise*, is a multiwinner version of the *Bucklin* rule.

Continuing Example 5.1: let us first take $\alpha = 1/3$. We have

$$\begin{array}{lll} \lambda(1/3, P, a) = 3 & \lambda(1/3, P, b) = 2 & \lambda(1/3, P, c) = 3 \\ \lambda(1/3, P, d) = 3 & \lambda(1/3, P, e) = 2 & \lambda(1/3, P, f) = 1 \end{array}$$

We have $j = 2$. There is a tie between b and e , resolved in favour of b , since b and e are ranked in the first 3 positions by respectively 5 and 4 voters. Therefore, $MC_k^{1/3}(P) = \{bf\}$.

Let us now take $\alpha = 1/2$. We have $j = 3$ and $MC_k^{1/2}(P) = \{ac\}$ (no tie-breaking is needed).

Sequential rules

There are several variants of the multiwinner version of *single transferable vote* (STV). We present the most common one: let $q = \lceil \frac{n}{k} \rceil$ (quota). If some candidate x has a plurality score $S(x) \geq q$, then x is elected, and each of the votes for x becomes a fractional vote with weight $1 - \frac{q}{S(x)}$, with x removed; otherwise the candidate with the lowest plurality score is eliminated from all votes (using tie-breaking if necessary). This operation is repeated until k candidates have been elected.

Sequential plurality elects first the plurality winner (using tie-breaking if necessary), removes it from the list of candidates, then elects the plurality winner from the obtained profile, and so on until k candidates have been elected.

Given a scoring vector $s = (s_1, \dots, s_m)$ with $s_1 \geq \dots \geq s_m$ and $s_1 > s_m$, and a subset of candidates T with $|T| \leq k$, let

$$\text{score}_{CC}^s(T, \succ) = \sum_{i=1}^n \max_{y \in T} s_{\text{pos}_{\succ_i}(y)}(y).$$

If $|T| < k$ and $x \notin T$, $\text{score}_{CC}^s(x|T, \succ) = \text{score}_{CC}^s(x \cup T, \succ) - \text{score}_{CC}^s(T, \succ)$ is the marginal score of x with respect to T and $\succ$. *Greedy s -Chamberlin-Courant* (s -GCC) elects the k winning candidates in sequence, including at each step the candidate with the largest marginal score with respect to the candidates already included.

- it first elects y_1 maximising $\text{score}_{CC}^s(\{y_1\}, \succ)$

- then, for each $j \in \{2, \dots, k\}$, it elects y_j maximising

$$\text{score}_{CC}^s(\{y_j | \{y_1, \dots, y_{j-1}\}, \succ).$$

Note that sequential plurality coincides with s -GCC for $s = (1, 0, \dots, 0)$.

For all these sequential rules, if ties occur, then all possibilities for resolving them are taken into account (which is sometimes called the “parallel universe” assumption).

Continuing Example 5.1:

- $STV(P) = \{af\}$: the quota is 5, no candidate reaches it; c and e are eliminated, still no candidate reaches the quota; d is eliminated, then b .
- $\text{SeqPlu}(P) = \{ef\}$: f is elected first, then e .
- $\beta\text{-GCC}(P) = \{be\}$: the Borda winner b is selected first, and e gives the highest marginal score given that b has been selected.

A2. Missing proofs

The following Theorem implies point 1 of Proposition 3.2.

Theorem A2.3 *Let $X \in \mathcal{S}_k(A)$ and $P \in \mathcal{L}(A)^N$. The following statements are equivalent:*

1. X is necessarily Pareto optimal at P for ρ .
2. There is no $Y \in \mathcal{S}_k(A)$ that possibly Pareto-dominates X at P for ρ .
3. For every $Y \in \mathcal{S}_k(A) \setminus \{X\}$, there exists $i \in N$ such that $X \succ_i^\rho Y$ holds.

Proof. Let $X \in \mathcal{S}_k(A)$ and $P \in \mathcal{L}(A)^N$.

- We first show that 1 implies 2. Assume there exists $Y \in \mathcal{S}_k(A)$ which possibly Pareto-dominates X at P for ρ . By definition, this means that there exists $\tilde{Q} = (\tilde{\sqsupset}_1, \dots, \tilde{\sqsupset}_n) \in \kappa^\rho(P)$ at which Y Pareto-dominates X . Hence, X is not Pareto optimal at every $Q \in \kappa^\rho(P)$, which shows that X is not necessarily Pareto optimal at P for ρ .
- We show that 2 implies 3. Assume there exists $Y \in \mathcal{S}_k(A) \setminus \{X\}$ such that $X \succ_i^\rho Y$ fails for all $i \in N$. This implies the existence of $Q = (\sqsupset_1, \dots, \sqsupset_n) \in \kappa^\rho(P)$ such that $Y \sqsupseteq_i X$ for all $i \in N$. By Remark 3.1, this is equivalent to saying that there exists $Q = (\sqsupset_1, \dots, \sqsupset_n) \in \kappa^\rho(P)$ such that $Y \sqsupseteq_i X$ for all i . Thus, Y Pareto-dominates X at some $Q \in \kappa^\rho(P)$, hence Y possibly Pareto-dominates X .
- Finally, we show that 3 implies 1. Assume that for every $Y \in \mathcal{S}_k(A) \setminus \{X\}$ there exists $i(Y) \in N$ such that $X \succ_{i(Y)}^\rho Y$. By Remark 3.1, for every $Y \in \mathcal{S}_k(A) \setminus \{X\}$ there exists $i(Y) \in N$ such that $X \sqsupseteq_{i(Y)} Y$ for every $\sqsupset_{i(Y)} \in \kappa^\rho(\succ_{i(Y)})$. Pick any $Y \in \mathcal{S}_k(A) \setminus \{X\}$ and any $Q = (\sqsupset_1, \dots, \sqsupset_n) \in \kappa^\rho(P)$. Since $X \sqsupseteq_{i(Y)} Y$ for some $i(Y) \in N$, Y does not Pareto-dominates X at Q . As this holds for any $Y \in \mathcal{S}_k(A) \setminus \{X\}$, X is Pareto optimal at Q for ρ . Finally, as the argument applies to any $Q \in \kappa^\rho(P)$, X is necessarily Pareto optimal at P for ρ .

■

Now, the following Theorem implies point 2 of Proposition 3.2

Theorem A2.4 *Let $X \in \mathcal{S}_k(A)$ and $P \in \mathcal{L}(A)^N$. The following statements are equivalent:*

1. *X is possibly Pareto optimal at P for ρ .*
2. *There is no $Y \in \mathcal{S}_k(A)$ that necessarily Pareto-dominates X at P for ρ .*
3. *For every $Y \in \mathcal{S}_k(A)$, there exists $i \in N$ such that $Y \succ_i^\rho X$ fails.*

Before proving Theorem A2.4, we introduce a definition and several simple lemmas. Given $\succ_i \in \mathcal{L}(A)$, $X \in \mathcal{S}_k(A)$, and $\sqsubseteq_i \in \kappa(\succ_i^\rho)$, we say that $\sqsubseteq_i^{X+}$ is an X -best completion of $\succ_i^\rho$ if for any $Y \in \mathcal{S}_k(A)$, if $Y \succ_i^\rho X$ does not hold then $X \sqsubseteq_i^{X+} Y$. Moreover, we say that $Q^{X+} = (\sqsubseteq_1^{X+}, \dots, \sqsubseteq_n^{X+})$ is an X -best completion of $P^\rho = (\succ_1^\rho, \dots, \succ_n^\rho)$ if for every i , $\sqsubseteq_i^{X+}$ is an X -best completion of $\succ_i^\rho$.

Lemma A2.5 *There exists an X -best completion of $\rho(P)$.*

Proof. For each i , consider the relation $\triangleright_i = \succ_i^\rho \cup \{(X, Y) \mid Y \succ_i^\rho X \text{ does not hold}\}$. $\triangleright_i$ is acyclic: if it had a cycle, since $\succ_i^\rho$ is acyclic, the cycle would contain a pair (X, Y) such that $Y \succ_i^\rho X$ does not hold, and thus would contain a path from X to Y in $\succ_i^\rho$; because $\succ_i^\rho$ is transitive, this would contradict the fact that $Y \succ_i^\rho X$ does not hold. Therefore the transitive closure $\triangleright_i$ is a strict partial order, and any of its completions is a X -best completion of $\succ_i^\rho$. This being true for every i , there exists an X -best completion of $\rho(P)$. ■

Lemma A2.6 *For any $i \in N$ and $X, Y \in \mathcal{S}_k(A)$, $Y \sqsubseteq_i^{X+} X$ holds for some X -best completion of $\succ_i^\rho$ if and only if $Y \sqsubseteq_i X$ for all completions $\sqsubseteq_i$ of $\succ_i^\rho$.*

Proof. The right-to-left direction is trivial. From left to right, assume that $Y \sqsubseteq_i X$ fails in some completion $\sqsubseteq_i$ of $\succ_i^\rho$. Then $X \sqsubseteq_i Y$, which implies that $Y \succ_i^\rho X$ does not hold; by definition of an X -best completion $\sqsubseteq_i^{X+}$, we have $X \sqsubseteq_i^{X+} Y$, and therefore $Y \sqsubseteq_i^{X+} X$ does not hold. ■

Lemma A2.7 *For any $i \in N$ and $X, Y \in \mathcal{S}_k(A)$, $X \sqsubseteq_i Y$ holds for some completion of $\succ_i^\rho$ if and only if $X \sqsubseteq_i^{X+} Y$ holds for some X -best completion of $\succ_i^\rho$.*

Proof. The right-to-left direction is trivial. From left to right, assume that for some X -best completion of $\succ_i^\rho$, $X \sqsubseteq_i^{X+} Y$ does not hold. Then, by definition of an X -best completion, we have $Y \succ_i^\rho X$, therefore $X \sqsubseteq_i Y$ does not hold in any completion of $\succ_i^\rho$. ■

Lemma A2.8 *$X \in \mathcal{S}_k(A)$ is Pareto-dominated at some X -best completion Q^{X+} of $\rho(P)$ if and only if X is not possibly Pareto-optimal.*

Proof. The right-to-left direction is a direct consequence of Lemma A2.5. From left to right, assume $X \in \mathcal{S}_k(A)$ is Pareto-dominated at some X -best completion $Q = (\sqsubseteq_1^{X+}, \dots, \sqsubseteq_n^{X+})$ of $\rho(P)$. Then there is an $Y \in \mathcal{S}_k(A)$ such that $Y \sqsubseteq_i^{X+} X$ for all i , and $Y \sqsubseteq_i^{X+} X$ for some i . By Lemma A2.6, $Y \sqsubseteq_i^{X+} X$ implies $Y \sqsubseteq_i X$ for all completions $\sqsubseteq_i$ of $\succ_i^\rho$. $Y \sqsubseteq_i^{X+} X$ implies that $X \sqsubseteq_i^{X+} Y$ does not hold, and by Lemma A2.7, $X \sqsubseteq_i Y$ holds for no completion of $\succ_i^\rho$, therefore $Y \sqsubseteq_i X$ holds for any completion of $\succ_i^\rho$. This allows us to conclude that if X is Pareto-dominated at all $Q \in \kappa^\rho(P)$, which is equivalent to saying that X is not possibly Pareto-optimal. ■

Now we are ready to prove Theorem A2.4.

Proof.

- We show by contradiction that 1 implies 2. Assume there is an $Y \in \mathcal{S}_k(A)$ that necessarily Pareto-dominates X at P for ρ . Then for any completion Q of $\rho(P)$, Y Pareto-dominates X at Q , which implies that X is not Pareto-optimal at Q . This being true for all Q , X is not possibly Pareto-optimal at P .
- We show by contradiction that 2 implies 3. Assume there is $Y \in \mathcal{S}_k(A)$ such that $Y \succ_i^\rho X$ holds for all i : then for all i , and for any extension $\sqsupseteq_i$ of $\succ_i^\rho$, we have $Y \sqsupseteq_i X$. By Remark A2.1, this is equivalent to saying that for all i , and for any extension $\sqsupseteq_i$ of $\succ_i^\rho$, we have $Y \sqsupseteq_i X$. Therefore X is necessarily Pareto-dominated by Y .
- We show by contradiction that 3 implies 1. Assume X is not possibly Pareto-optimal: then by Lemma A2.8, X is Pareto-dominated by some $Y \in \mathcal{S}_k(A)$ at some X -best completion $Q^{X+} = (\sqsupseteq_1^{X+}, \dots, \sqsupseteq_n^{X+})$ of P^ρ , which implies that for all i we have $Y \sqsupseteq_i^{X+} X$. By Lemmas A2.7 and A2.8, for all i we have $Y \sqsupseteq_i X$ and all completions $\sqsupseteq_i$ of $\succ_i^\rho$, which implies that $Y \succ_i^\rho X$ holds.

■

Proposition A2.9 *Given any $P \in \mathcal{L}(A)^N$ and any $i \in N$:*

1. $\text{top}(k, \succ_i)$ is necessarily Pareto optimal at P ;
2. let $X = \text{top}(k+1, \succ_i) \setminus \{x\}$ for some $x = \{r_t(A; \succ_i)\}$ with $t \in \{1, \dots, k\}$. Then X is necessarily Pareto optimal at P if and only if $y \succ_j x$ for some $j \in N$ and $y \in \text{top}(k+1, \succ_i) \setminus \text{top}(t, \succ_i)$.

Proof. For 1, observe that $\text{top}(k, \succ_i)$ is the most preferred committee by i : there is no $Y \in \mathcal{S}_k(A)$ such that $Y \neq X$ and $Y \sqsupseteq_i X$: $\text{top}(k, \succ_i)$ cannot be possibly Pareto-dominated, and is therefore necessarily Pareto optimal at P .

For 2, let $X = \text{top}(k+1, \succ_i) \setminus \{x\}$ for some $x = \{r_t(A; \succ_i)\}$ with $t \in \{1, \dots, k\}$. Assume X is necessarily Pareto optimal at P : in particular, for any i , it is not possibly Pareto-dominated by $\text{top}(k, \succ_i)$, which means that there is a j such that $X \sqsupseteq_j \text{top}(k, \succ_i)$. Since X is obtained from $\text{top}(k, \succ_i)$ by replacing $\{r_t(A; \succ_i)\}$ by $\{r_{k+1}(A; \succ_i)\}$, this means that $\{r_t(A; \succ_i)\} \succ_j x$ for some $t \in \{1, \dots, k\}$. ■

Proposition 4.1 *Every CSR f_γ is WPPE. If γ is strict then f_γ is PPE.*

Proof. The monotonicity of γ implies that if S necessarily Pareto-dominates S' for some profile P , then $\text{score}(S', P) \geq \text{score}(S, P)$, therefore f_γ is WPPE.

Assume γ is strict. Let $S \in f_\gamma(P)$ and let $S' \in \mathcal{S}_k(A)$ such that S' necessarily Pareto-dominates S . Then $\text{pos}(S, \succ_i) \succ \text{pos}(S', \succ_i)$ for each i , which implies $\text{score}(S', P) > \text{score}(S, P)$, contradicting $S \in f_\gamma(P)$. ■

Proposition 4.2 *SNTV is WLPE but not LPE. Its restriction to profiles where at least k alternatives are ranked first by some voter is LPE.*

Proof. For any profile P let $\text{Top}(P)$ be the set of candidates that are ranked on top by at least one voter. Let $q = |\text{Top}(P)|$. If $q \geq k$, then SNTV is equivalent to the top-sequential rule that selects the candidates with the highest k plurality scores, and therefore, by Proposition 4.8, the restriction of SNTV to such profiles is LPE. If $q < k$, consider the following top-sequential rule: all q candidates with strictly positive plurality score are selected, and then the remaining $k - q$ candidates are voters 1's top $k - q$ candidates among those remaining. This rule is top-sequential by definition, and its outcome belongs to $\text{SNTV}(P)$, therefore SNTV is WLPE. SNTV is however not LPE, because of pathological profiles with less than k alternatives ranked first by some voter. For instance, if $k = 2$ and $P = (abcd, acdb)$, then

SNTV outputs $\{ab, ac, ad\}$; while ab and ac are lexicographically Pareto-optimal, ad is lexicographically Pareto-dominated. ■

Proposition 4.6

For any $k \geq 2$ and $m \geq 2k + 1$, no best- k rule satisfies necessary Pareto-efficiency.

Proof. Let $k \geq 2$ and $m \geq 2k + 1$. Consider the following m -candidate, k -voter profile:

$x_1 z_1 \dots z_k$	---	$x_2 \dots x_k$
$x_2 z_2 \dots z_1$	---	$x_3 \dots x_2$
$x_3 z_3 \dots z_2$	---	$x_4 \dots x_3$
...		
$x_k z_k \dots z_{k-1}$	---	$x_1 \dots x_{k-1}$

Let $X = \{x_1, \dots, x_k\}$, $Z = \{z_1, \dots, z_k\}$ and $Y = A \setminus (X \cup Z)$. Every candidate in Y is Pareto-dominated by every candidate in Z , therefore a necessarily Pareto optimal committee must be contained in $X \cup Z$.

Let $s = (s_1, \dots, s_m)$ and consider the best- k rule f_s . The score of each $x \in X$ is $s_1 + s_{m-k+2} + \dots + s_m$; the score of each $z \in Z$ is $s_2 + \dots + s_{k+1}$. The score of every $y \in Y$ candidates is at most $s_{k+2} + \dots + s_{2k+1}$. Therefore:

- If $s_1 + s_{m-k+2} + \dots + s_m > s_2 + \dots + s_{k+1}$ then X is the only winning committee. It is not necessarily Pareto optimal, as it is possibly Pareto-dominated by Z .
- If $s_2 + \dots + s_{k+1} > s_1 + s_{m-k+2} + \dots + s_m$ then Z is the only winning committee contained in $X \cup Z$ (recall that all other committees cannot be necessarily Pareto optimal). It is not necessarily Pareto optimal, as it is possibly Pareto-dominated by X .

The only remaining cases are when $s_1 + s_{m-k+2} + \dots + s_m = s_2 + \dots + s_{k+1}$, which we now assume.

- Assume $s_1 + s_{m-k+2} + \dots + s_m = s_2 + \dots + s_{k+1}$ and $s_{k+2} > 0$. Assume also $k \geq 3$. We consider the following profile:

$x_1 z_1 \dots z_k x_2 \dots x_k$	---
$x_2 z_2 \dots z_1 x_3 \dots x_2$	---
$x_3 z_3 \dots z_2 x_4 \dots x_3$	---
...	
$x_k z_k \dots z_{k-1} x_1 \dots x_{k-1}$	---

The only winning committee is X (which is not necessarily Pareto optimal).

- Assume $k \geq 3$, $s_1 + s_{m-k+2} + \dots + s_m = s_2 + \dots + s_{k+1}$ and $s_{k+2} = 0$, which means that $s = (s_1, s_2, \dots, s_{k+1}, 0, \dots, 0)$ with $s_1 = s_2 + \dots + s_{k+1}$. This also implies that $s_2 > 0$. Consider this profile:

$x_1 x_2 z_1 \dots z_{k-1}$	z_k	---
$x_2 x_3 z_2 \dots z_k$	z_1	---
...		
$x_k x_1 z_k \dots z_{k-2}$	z_{k-1}	---

The score of every x_i is $s_1 + s_2$ and that of every z_i is $s_3 + \dots + s_{k+1}$. Since $s_2 > 0$, $s_1 + s_2 > s_3 + \dots + s_{k+1}$, therefore the winning committee is X , which is not necessarily Pareto optimal because it is possible Pareto-dominated by Z .

- The only remaining case is $k = 2$ and $s = (s_1, s_2, s_3, 0, \dots, 0)$ with $s_1 = s_2 + s_3$. Assume first $s_3 > 0$. Consider the profile

$$\begin{array}{l} x_1 z_1 z_3 \text{ --- } x_2 \\ x_2 z_2 z_3 \text{ --- } x_1 \end{array}$$

The winning committee is $\{x_1, x_2\}$; it is not necessarily Pareto optimal. Finally, assume $s_3 = 0$, that is, $s = (s_1, s_1, 0, \dots, 0)$. Consider the profile

$$\begin{array}{l} x_1 z_1 \text{ --- } x_2 \\ x_1 z_2 \text{ --- } x_2 \\ x_1 z_3 \text{ --- } x_2 \\ x_2 z_1 \text{ --- } x_1 \\ x_2 z_2 \text{ --- } x_1 \\ x_2 z_3 \text{ --- } x_1 \end{array}$$

The only winning committee is $\{x_1, x_2\}$; it is not necessarily Pareto optimal.

■

We note that in all of our profiles we needed at least $2k$ candidates, except in the last one where we needed at least $2k + 1$. Therefore, if $m \leq 2k$, some best- k rules may be NPE.⁷

Proposition 4.7 *If $m = 4$ and $k = 2$, then f_2^s is LPE if and only if $s_1 > s_2 + s_3$ and $s_2 > s_3$.*

Proof. If $s_1 \leq s_2 + s_3$, take $\succ = (abcd, dcba)$. We have $\text{Score}(a, \succ) = \text{Score}(d, \succ) = s_1 \leq \text{Score}(b, \succ) = \text{Score}(c, \succ) = s_2 + s_3$, therefore $\{b, c\}$ is a winning committee, which is lexicographically Pareto-dominated by $\{a, d\}$. If $s_2 = s_3$, then the single-voter profile $(abcd)$ has $\{a, c\}$ as a winning committee although it is dominated by $\{a, b\}$. This shows the necessary part.

For the sufficiency part, assume that $s_1 > s_2 + s_3$ and $s_2 > s_3$, and let a profile $\succ$ such that $\{y_1, y_2\}$ is lexicographically Pareto-dominated by $\{x_1, x_2\}$.

If $\{y_1, y_2\} \cap \{x_1, x_2\} \neq \emptyset$ then without loss of generality, $\{y_1, y_2\} = \{x_1, y_2\}$ with $y_2 \neq x_1$. Since $\{x_1, y_2\}$ is lexicographically Pareto-dominated by $\{x_1, x_2\}$, y_2 is Pareto-dominated by x_2 . If x_2 is ranked at least once in position 1 or 2, then $s_1 > s_2 + s_3$ and $s_2 > s_3$ imply that $\text{Score}(x_2, \succ) > \text{Score}(y_2, \succ)$. If x_2 is always ranked in position 3, then y_2 is always ranked in position 4 and the single winning committee consists of the other two candidates. In both cases, $\{y_1, y_2\}$ cannot be a winning committee.

Now, assume $\{y_1, y_2\} \cap \{x_1, x_2\} = \emptyset$. Because $\{y_1, y_2\}$ is lexicographically Pareto-dominated by $\{x_1, x_2\}$, each vote $\succ_i$ has the form $xyy'x'$ or $xyx'y'$ or $xx'y'y'$, where $\{x, x'\} = \{x_1, x_2\}$ and $\{y, y'\} = \{y_1, y_2\}$. Then, $s_1 > s_2 + s_3$ implies that $\text{Score}(x_1, \succ) + \text{Score}(x_2, \succ) > \text{Score}(y_1, \succ) + \text{Score}(y_2, \succ)$. Thus, $\{y_1, y_2\}$ cannot be a winning committee. ■

Replacing strict inequalities by weak inequalities in the proof of Proposition 4.8 leads to the characterization of WLPE best- k rules for $m = 4$ and $k = 2$: if $m = 4$, then f_2^s is WLPE if and only if $s_1 \geq s_2 + s_3$ and $s_2 \geq s_3$.

When m becomes larger, generalizing such a characterization becomes difficult.

⁷This is actually the case: take $k = 2$, $m = 3$, $A = \{x, y, z\}$, $s = (s_1, s_2, 0)$. Assume that for some profile P , $\{x, y\}$ is a winning committee but is not necessarily Pareto optimal; without loss of generality, assume $\{x, y\}$ is possibly Pareto-dominated by $\{x, z\}$. Because $\{x, y\}$ is possibly Pareto-dominated by $\{x, z\}$ P contains no vote where z is ranked last and no vote yzx . Therefore, P consists of α votes xzy , β votes zxy , and γ votes zyx . But now, the score of z is $(\beta + \gamma)s_1 + \alpha s_2$ and the score of y is γs_2 , so $\{x, y\}$ cannot be a winning committee unless $(\beta + \gamma)s_1 + \alpha s_2 = \gamma s_2$, that is, $\alpha = \beta = 0$ and $s_1 = s_2$; but then P contains only votes zyx and the winning committee is $\{y, z\}$.

As a corollary of Proposition 4.8, k -Borda is not LPE. It is WLPE for $m = 4$ and $k = 2$, but no longer if $m \geq 5$, as witnessed by the profile $(abcde-, ebcda-)$, for which the unique winning committee for $k = 2$ is $\{b, c\}$, which is lexicographically Pareto-dominated by $\{a, e\}$.

On the other hand, k -Harmonic, defined by the scoring vector $(1 - \frac{1}{4}, \frac{1}{2} - \frac{1}{4}, \frac{1}{3} - \frac{1}{4}, 0)$, is LPE for $k = 2$ and $m = 4$. But it is not WLPE in the general case: let $k = 3$, $m = 14$, and $P = (atuv \dots bc, buvt \dots ca, ctuv \dots ab)$: the winning committee $\{u, u, v\}$ is lexicographically Pareto-dominated by $\{a, b, c\}$.

Proposition 4.10 *STV and sequential plurality fail necessary Pareto-efficiency.*

Proof. Consider the profile $P = (axyb, bxya)$, and $k = 2$. The winning STV, and sequential plurality, committee is $\{a, b\}$; it is not necessarily Pareto optimal. ■

Proposition 4.11 *Greedy β -CC is not WLPE.*

Proof. Let $k = 2$ and $P = (axztyub, ayutxzb, bxztyua, byztuxa)$. At first step, β -CC selects x or y , both with optimal Borda scores. Suppose it selects x (respectively y), then at step 2 it selects y (respectively x) with maximal marginal contribution. Therefore, the output is $\{x, y\}$, which is lexicographically Pareto-dominated by $\{a, b\}$. ■

Proposition 4.12 *For any $\alpha \in (0, 1)$ and $k \geq 2$, MC_k^α is not WLPE.*

Proof. Assume $\frac{1}{2} \leq \alpha < 1$. Let P be the following two-voter profile:

$$\begin{array}{l} 1 : \quad z_1 \dots z_{k-2} \quad a \quad u \quad v \quad b \quad \dots \\ 1 : \quad z_1 \dots z_{k-2} \quad b \quad v \quad u \quad a \quad \dots \end{array}$$

We have $\lambda(z_i, \alpha, P) = i$ for all $i = 1, \dots, k-2$; $\lambda(u, \alpha, P) = \lambda(v, \alpha, P) = k+1$; and $\lambda(a, \alpha, P) = \lambda(b, \alpha, P) = k+2$.

So $MC_k^\alpha(P) = \{z_1 \dots z_{k-2} uv\}$ although $z_1 \dots z_{k-2} ab$ lexicographically dominates $z_1 \dots z_{k-2} uv$.

Now assume $0 < \alpha < \frac{1}{2}$. Let n be the smallest integer such that $n \geq \frac{1}{\alpha}$. (For instance, if $\frac{1}{3} \leq \alpha < \frac{1}{2}$ then $n = 3$.) From $\alpha < \frac{1}{2}$ we have $\frac{2}{\alpha} - \frac{1}{\alpha} = \frac{1}{\alpha} > 2$, therefore $n < \frac{2}{\alpha}$, so that $\frac{1}{n} < \alpha < \frac{2}{n}$. Let P be the following n -voter profile:

$$\begin{array}{l} 1 : \quad z_1 \dots z_{k-2} \quad a \quad u \quad v \quad \times \quad \dots \\ 1 : \quad z_1 \dots z_{k-2} \quad b \quad v \quad u \quad \times \quad \dots \\ n-2 : \quad z_1 \dots z_{k-2} \quad \times \quad \times \quad \times \quad a \quad \dots \end{array}$$

such that

1. in each of the last $n-2$ votes, none of the candidates ranked between $z_1, \dots, z_{k-2}$ and a is b, u or v
2. no candidate appears more than once above a in the last $n-2$ votes (note that for this to be possible we must have at least $3(n-2) + 4$ candidates).

We have $\lambda(z_i, \alpha, P) = i$ for all $i = 1, \dots, k-2$; $\lambda(u, \alpha, P) = \lambda(v, \alpha, P) = k+1$; $\lambda(a, \alpha, P) = k+2$; and $\lambda(b, \alpha, P) > k+2$.

So $MC_k^\alpha(P) = \{z_1, \dots, z_{k-2}, u, v\}$, although $\{z_1, \dots, z_{k-2}, a, b\}$ lexicographically dominates $\{z_1, \dots, z_{k-2}, u, v\}$. ■

Proposition 4.14 *NED is not WPPE.*

Proof. We give a counterexample with $n = 5$, $k = 2$, $m = 14$, and $C = \{a, b, c, d\} \cup E^* \cup F^*$ with $E^* = \{e_1, e_2, e_3, e_4, e_5\}$ and $F^* = \{f_1, f_2, f_3, f_4, f_5\}$. The five candidates $e_1, \dots, e_5$ are clones and are ranked in the five votes in such a way that they form a cycle $e_1 \rightarrow e_2 \rightarrow e_3 \rightarrow e_4 \rightarrow e_5 \rightarrow e_1$ (see the proof of 4.13 for an explanation), and similarly for $f_1, \dots, f_5$. P is a five-voter profile for whose votes are:

$$\begin{aligned} a &\succ b \succ F^* \succ c \succ d \succ E^* \\ c &\succ b \succ E^* \succ a \succ d \succ F^* \\ E^* &\succ a \succ d \succ F^* \succ c \succ b \\ a &\succ b \succ c \succ d \succ E^* \succ F^* \\ F^* &\succ c \succ d \succ E^* \succ a \succ b \end{aligned}$$

where $a \succ b \succ F^* \succ c \succ d \succ E^*$ means that a and b are preferred to all the f_i 's, all of them being preferred to c, d and all the e_i 's, and so on. The majority graph is as follows:

$$\begin{aligned} a &\rightarrow b, c, d, F^* \\ b &\rightarrow d, E^*, F^* \\ c &\rightarrow b, d, E^* \\ d &\rightarrow E^*, F^* \\ E^* &\rightarrow a, F^* \\ F^* &\rightarrow c \end{aligned}$$

b and d are the only candidates that beat the 10 candidates in $E \cup F$: the NED score of $\{b, d\}$ is 20 (as b and d both beat 10 candidates in $C \setminus \{b, d\}$). Any other candidate beats at most 8 candidates, therefore, any 2-committee different from $\{b, d\}$ has a NED score at most $11+8 = 19$. Therefore the only winning committee is $\{b, d\}$, and it is necessarily Pareto-dominated by $\{a, c\}$. ■

Proposition 4.17 *LSE-maximin fails possible Pareto-efficiency, but satisfies weak possible Pareto-efficiency.*

Proof. For the failure of PPE, just consider the single-voter profile (abc) : LSE-maximin outputs $\{a, b\}$ and $\{a, c\}$, the latter being necessarily Pareto-dominated.

For the satisfaction of WPPE, assume (1) S' necessarily dominates S and (2) S is locally stable for quota q . (1) implies that there is a bijection σ_i from S to S' such that for all $x \in S$, $\sigma_i(x) \succeq_i x$. (2) means that for all $y \in A \setminus S$, at least qn voters prefer some candidate in S to y . Let $I(S, y) \subseteq N$ be the set of voters who prefer some candidate $c(i, S, y) \in S$ to y .

We now show that S' is locally stable for quota q . Let $y \in A \setminus S'$; we have to show that at least qn voters prefer some candidate in S' to y .

Assume first that $y \in A \setminus S$. Let $i \in I(S, y)$. Therefore, $\sigma_i(c(i, S, y)) \succeq_i c(i, S, y) \succ_i y$, and so $i \in I(S', y)$. This implies that $I(S', y) \supseteq I(S, y)$ and so that $|I(S', y)| \geq qn$.

Assume now that $y \in S$. Because $y \notin S'$, (1) implies that $\sigma_i(x) \succ_i x$ for all $i \in N$. This implies that $I(S', y) = N$ and *a fortiori* that $|I(S', y)| \geq qn$. ■