A Characterization for Procedural Choice Based on Dichotomous Preferences Over Criteria

Takahiro Suzuki[(✉)] and Masahide Horita

Department of International Studies, Graduate School of Frontier Sciences,
The University of Tokyo, 5-1-5, Kashiwanoha, Kashiwashi, Chiba, Japan
ts.takahiro.suzuki@gmail.com, horita@k.u-tokyo.ac.jp

Abstract. Many lessons for procedural choice have been provided by axiomatic studies of decision procedures. However, there appears to be a gap between these axiomatic studies and the actual determination of appropriate procedures, as an axiomatic characterization does not directly answer which axiom should be appropriate—particularly when there is no agreement on the relative desirability of criteria. The present study proposes a formal model of procedural choice based on preferences over criteria (PCBPC). Specifically, we focus on the aggregation rule that maps a dichotomous preference profile over criteria for decision procedures to a nonempty set of decision procedures. We prove that the counting rule, which chooses the decision procedures with greatest supports, is the unique aggregation rule that satisfies anonymity (A), neutrality (N), strict monotonicity (SM), and partition consistency (PC), where PC is proposed based on the idea that representations of equivalent criteria in different ways should not affect the results. Two distinct standpoints for PCBPC are highlighted: one is to regard criteria as atomic, i.e., inseparable, objects and the other as composite, i.e., separable, objects. The difference between them is made explicit with two impossibility theorems showing the inconsistency between unanimity in the former standpoint and A (or PC) in the latter standpoint.

Keywords: Procedural choice · Preferences over criteria · Counting rule

1 Introduction

Axiomatic studies of decision procedures (social welfare functions, social choice rules, multiple criteria decision-making methods (MCDM), etc.) have provided many lessons for aggregating voters' preferences over alternatives. The choice of an appropriate procedure is, however, still a challenging problem and there seems no agreement on the best procedure, even among social choice theorists (Nurmi [1]).

Part of the reason for this difficulty may be that axiomatic characterizations do not, at least in some cases, directly answer which is the appropriate decision procedure. For instance, theorems such as "f is the unique social choice rule that satisfies criteria A and B", or "g is the unique social choice rule that satisfies criteria A and C" deepen our

understanding of these rules. But the choice between f and g is still controversial when there is no agreement on the relative desirability of B and C.

This observation implies the necessity of a decision support system that deals with voters' preferences over criteria. To this end, the present paper provides a formal model for procedural choice based on preferences over criteria (PCBPC). Although each of the axiomatic studies has provided the justifications for particular axioms, Nurmi [1] was the first to study the system for PCBPC (see also de Almeida et al. [2]; application of this method for business contexts is found in de Almeida and Nurmi [3]). Based on the performance matrix of voting rules (Table 1), Nurmi [1] proposes two different methods to create collective rankings/choices of voting rules based on the voters' preferences over criteria. One method is to first determine the rankings of criteria with an ordinary decision procedure (e.g., Borda's or Kemeny's rule) and then determine the rankings of decision procedures (p. 247–248 in Nurmi [1]). The other method is to determine the rankings of decision procedures directly using MCDM methods (p. 250–251 in Nurmi [1]). As Nurmi points out, these methods are designed for situations where voters have linear preferences over criteria and some other methods can be used instead. The present study first aims to provide a formal model for PCBPC in order to make clear the foundations and implicit assumptions of PCBPC. Next, we characterize the counting rule which chooses those decision procedures with the greatest support when voters have dichotomous preferences over criteria.

Table 1. Whether a SCR satisfies a given criterion (1) or not (0) (from de Almeida et al. [2])

	Criterion								
	a	b	c	d	e	f	g	h	i
Amendment	1	1	1	1	0	0	0	0	0
Copeland	1	1	1	1	1	0		0	0
Dodgson	1	0	1	0	1	0	0	0	0
Maximin	1	0	1	1	1	0	0	0	0
Kemeny	1	1	1	1	1	0	0	0	0
Plurality	0	0	1	1	1	1	0	0	1
Borda	0	1	0	1	1	1	0	0	1
Approval	0	0	0	1	0	1	1	0	1
Black	1	1	1	1	1	0	0	0	0
Pl. runoff	0	1	1	0	1	0	0	0	0
Nanson	1	1	1	0	1	0	0	0	0
Hare	0	1	1	0	1	0	0	0	0
Coombs	0	1	1	0	1	0	0	0	0

a: The Condorcet winner criterion, b: The Condorcet loser criterion, c: The strong Condorcet criterion, d: Monotonicity, e: Pareto, f: Consistency, g: Chernoff property, h: Independence of irrelevant alternatives, i: Invulnerability to the no-show paradox. In the table, 1 (0) means that the voting rule satisfies (fails to satisfy) the criterion.

Let us review the literature related to procedural choice. Some authors study the desirable way of procedural choice when the preferences over decision procedures are consequentially induced from the preferences over alternatives (Rae [4]; Koray [5]; Barbera and Jackson [6]; Suzuki and Horita [7]). Others begin with the preferences over decision procedures (Dietrich [8]; Diss and Merlin [9]). Compared with these studies, PCBPC is new in that it is based on the preferences over criteria.

As Nurmi [1] points out, the PCBPC model looks similar to a MCDM model such as PROMETHEE at first glance, because both of them deal with a performance matrix such as Table 1. In our view, the biggest difference between them is that our PCBPC model is based on the social choice theoretic approach: preferences are supposed to be binary relations rather than numerical utility functions; each concept (criterion, aggregation rule, etc.) is defined through set theory; and axioms in social choice theory turn out to be important (as we will argue).

2 Designing a Formal Model for PCBPC

Let $N = \{1, 2, \ldots, n\}$ be the set of voters $(n \geq 2)$ and $\mathcal{F}$ be the set of feasible decision procedures. The elements of $\mathcal{F}$ can be social choice rules, social welfare functions, social decision functions, MCDM methods, etc. A criterion C (for $\mathcal{F}$) is a proposition on the elements of $\mathcal{F}$. More formally, a criterion C is a function that returns "true" or "false" for each $f \in \mathcal{F}$ ($C(f) = $ true means that $f \in \mathcal{F}$ satisfies the criterion).

The most straightforward approach to PCBPC would be simply to use ordinary decision procedures (e.g., Borda's or Kemeny's rule) by regarding criteria as alternatives and then use another procedure to yield the rankings/choices of decision procedures:
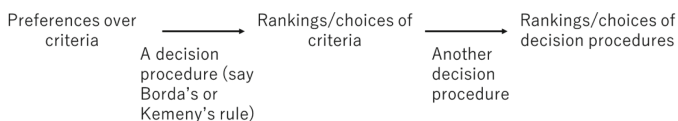

Fig. 1. A sketch of a two-step PCBPC

Although this approach is intuitive, it can be controversial because of the differences between criteria and alternatives in the ordinary sense. The most essential difference is in the logical relationships between criteria, i.e., some sets of criteria are consistent, but others are not. As is well known, Arrow [10] 's impossibility theorem says that Pareto, independence of irrelevant alternatives, and non-dictatorship reach a contradiction in the design of social welfare functions (with universal domains). Therefore, to approve of all three criteria equally is logically impossible even though the society members may unanimously agree about them all. On the other hand, the combination of anonymity, neutrality, and positive responsiveness is consistent (May [11]) and so these three can be approved simultaneously. This implies that

unsatisfactory results can happen if the two steps are independently determined. In other words, the set of admissible collective judgments of criteria should be restricted by the logical relationships of decision procedures at hand. In order to make matters simple, our model assumes that the PCBPC method returns the collective choices on decision procedures directly from the set of preference profiles over criteria (and not via the collective judgments of criteria).

Assumption 1 (Outline of a PCBPC Method)
A PCBPC method is a single-step procedure[1] that returns a nonempty subset of $\mathcal{F}$ for any profile of voters' preferences over criteria.

When examining voters' preferences over criteria, however, to require such logical consistency seems too demanding. Suppose $\mathcal{F} = \{\text{Amendment}, \text{Copeland}\}$. Then, Table 1 says that criteria a-d are logically equivalent in the sense that f satisfies criterion x if and only if f satisfies criterion y, for all $f \in \mathcal{F}$ and $x, y \in \{a, b, c, d\}$. In this case, logical consistency of voters' preferences over criteria would demand that a: the Condorcet winner criterion and d: Monotonicity must be preferred indifferently. This sounds too demanding especially when the voter is not familiar with those criteria. Assumption 2 allows voters to have any preference irrespective of the consistency:

Assumption 2 (Universal Domain)
The domain, or the set of admissible preference profiles, of the PCBPC method is universal.

Another justification of Assumption 2 is that voters' preferences over criteria can be independent of $\mathcal{F}$. For instance, a voter who favors plurality might prefer the criterion of monotonicity to the Condorcet winner criterion because plurality satisfies only the former. Assumption 2 is necessary for allowing preferences caused by such external factors. In summary, Assumption 2 says that voters may approve even an inconsistent set of criteria (e.g., Pareto, independence of irrelevant alternatives, and non-dictatorship for social welfare functions), and they may approve several criteria independently that are coincidentally equivalent under $\mathcal{F}$.

In the present article, we restrict our attention to dichotomous preferences:

Assumption 3 (Dichotomous Preferences)
Each voter is presumed to have a dichotomous preference over criteria, i.e., each voter either approves or disapproves of each criterion[2].

According to Assumption 3, i's preference over criteria can be expressed by arraying those criteria that i approves, e.g., $(C_{i,1}, C_{i,2}, \ldots, C_{i,a_i})$. For any criterion C, let $\mathcal{F}_C = \{f \in \mathcal{F} / f \text{ satisfies } C\}$.

---

[1] Mathematically speaking, even though the PCBPC method is made up of two separate steps as in Fig. 1, composition of the procedures may turn it into a single-step procedure. In this sense, the phrase "a single-step procedure" in Assumption 1 is not a rigorous mathematical statement, but an intuitive expression of our model.

[2] To make matters simple, we assume that each voter casts their ballot sincerely, i.e., strategic manipulation is assumed not to exist. Thus, we do not distinguish between ballots and preferences.

Assumption 4 (Input to the PCBPC Method: Preferences over Criteria)
The input to the PCBPC method is the preference profile over criteria. A voter's preference over criteria is expressed as a multiset of $\mathfrak{P}(\mathcal{F})$ such as

$$\mathcal{X}_i = \left[\mathcal{F}_{C_{i,1}}, \mathcal{F}_{C_{i,2}}, \ldots, \mathcal{F}_{C_{i,a_i}} \right], \tag{1}$$

where $C_{i,1}, C_{i,2}, \ldots, C_{i,a_i}$ are interpreted as criteria that the voter i approves.
   Let us give some comments on this assumption.

(1) Roughly speaking, Assumption 4 means that the names of the criteria do not matter. Let us explain why. According to Assumption 4, the PCBPC method looks at not the criterion C itself but $\mathcal{F}_C$. In the choices of Copeland and Dodgson, for example, there is supposed to be no difference between criteria b and d in Table 1 because for the two voting rules, satisfying b is equivalent to satisfying d. In this example, Assumption 4 demands that "a voter approves criterion b" is equivalent to "a voter approves criterion d". The point is that Copeland is supported by one reason, no matter which of b or d is approved. In this sense, Assumption 4 is an axiom that treats equivalent criteria equally.

(2) It is also noteworthy that the meaning of a criterion C is measured within the set of feasible decision procedures $\mathcal{F}$. Consequently, the "same" criterion, say Condorcet winner criterion (CW), may be translated into different sets if different $\mathcal{F}$'s are given: for instance, $\mathcal{F}_{CW} = \mathcal{F}$ when $\mathcal{F}$ is made up of amendment, Copeland, and Dodgson, while $\mathcal{F}'_{CW} = \phi$ if $\mathcal{F}'$ is made up of plurality, Borda, and approval (see Table 1). The difference reflects the logical strength of CW in the two cases (CW is extremely weak in the former case and extremely strong in the latter).

   In this sense, Assumption 4 demands that the meaning of criteria be argued with respect to $\mathcal{F}$, which can be interpreted as the domain of discourse (what kind of decision procedures are considered). This view is based on the usual practice of social choice theory. One can find many examples where impossibility theorems are conquered by expanding the domain $\mathcal{F}$. One of the most famous cases is that the combination of Pareto, independence of irrelevant alternatives, and non-dictatorship (as well as unrestricted domain) is found to be consistent if social decision functions are considered (Sen [12]). This shows that we cannot argue the consistency without specifying the domain of discourse.

(3) There may be some objections to this assumption. The strongest of them would be "Doesn't it pay little attention to other important features of criteria, such as logical relationships between criteria?" The answer is "yes" and "no". For one thing, the assumption neglects the relationship outside $\mathcal{F}$. The fact that criterion a is stronger than b in the choice of $\mathcal{F}' = \{\text{Amendment, Copeland, Dodgson}\}$ is not considered in the choice of $\mathcal{F} = \{\text{Amendment, Copeland}\}$ (the two criteria are equivalent under $\mathcal{F}$). At the same time, however, the relationships inside $\mathcal{F}$ are fully included.

This is because the logical relationships between C and C' under $\mathcal{F}$ can be fully expressed by $\mathcal{F}_C$ and $\mathcal{F}_{C'}$.[3]

(4) A multiset $[A, B, C, \ldots, Z]$ is a set admitting repetition. Note that the ordering does not matter: $[A, B] \neq [A, B, B] = [B, A, B] = [B, B, A]$.

More formally, we say that $\mathcal{M}$ is a multiset of a set S if $\mathcal{M}$ is a function from S to $\mathbb{Z}_{\geq 0}$ (the set of all nonnegative integers), where $\mathcal{M}(A) = k$ means that A appears k times[4]. For distinction, square brackets are used to express multisets.

(5) We use multisets instead of ordinary sets so that we can admit the approval of some criteria that are coincidentally equivalent under $\mathcal{F}$ (recall Assumption 2).

(6) Note also that our model does not restrict the set of criteria. Each voter can select his/her own set of criteria (e.g., those rules that he/she knows). This is one of the main differences between the present model and Nurmi [1] 's, which considers a fixed set of criteria.

Let $\mathfrak{D}$ be the set of all multisets of $\mathfrak{P}(\mathcal{F})$. Based on Assumption 1-4, we obtain the definition of the PCBPC method (from now on, we call it the aggregation rule).

Definition 1
An *aggregation rule* is a function from $\mathfrak{D}^n$ to the set of all nonempty subsets of $\mathcal{F}$.

3   Results

We use script letters $\mathcal{X}, \mathcal{Y}, \ldots$ to represent preference profiles and script letters with subscripts $\mathcal{X}_i, \mathcal{X}_j, \ldots$ to represent preferences. Subsets of $\mathcal{F}$ are denoted by capital letters $X, Y, \ldots$

Definition 2 (Axioms[5] for Aggregation Rules)
An aggregation rule F is said to satisfy

1) *anonymity* (A) if for all $i, j \in N$ and $\mathcal{X}_i, \mathcal{X}_j \in \mathfrak{D}$ such that $C \in \mathcal{X}_j$, if $\mathcal{X}'_i = \mathcal{X}_i \cup [C]$ and $\mathcal{X}'_j = \mathcal{X}_j \backslash [C]$, then[6] $F(\mathcal{X}_i, \mathcal{X}_j, \mathcal{X}_{-i,j}) = F(\mathcal{X}'_i, \mathcal{X}'_j, \mathcal{X}_{-i,j})$.

2) *neutrality* (N) if for all $f, g \in \mathcal{F}$, if $\mathcal{X}' \in \mathfrak{D}^n$ is a profile obtained from $\mathcal{X} \in \mathfrak{D}$ by swapping the positions of f and g, then $f \in F(\mathcal{X}) \Leftrightarrow g \in F(\mathcal{X}')$.

---

[3] For instance, "Under $\mathcal{F}$, C is logically stronger than C'" is by definition equivalent to $\mathcal{F}_C \subseteq \mathcal{F}_{C'}$. More complicated statements such as "Under $\mathcal{F}$, if C and not D then E" can be also expressed as $(\mathcal{F}_C \backslash \mathcal{F}_D) \subseteq \mathcal{F}_E$.

[4] Throughout the paper, we use both interpretations of a multiset (i.e., a set with repetition or a function from the base set to $\mathbb{Z}_{\geq 0}$.

[5] Although these are in fact "criteria" for selecting aggregation rules in our sense, we use the word "axiom" in this context to make the distinction from criteria that make up the elements of $\mathcal{F}$.

[6] As usual, preferences other than i and j are expressed as $\mathcal{X}_{-i,j}$.

3) *strict monotonicity* (SM) if for all $i \in N$, $\mathcal{X} = (\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_n) \in \mathfrak{D}^n$, and $X \subseteq \mathcal{F}$, if $F(\mathcal{X}) \cap X \neq \phi$, then $F(\mathcal{X}') = F(\mathcal{X}) \cap X$, where $\mathcal{X}'$ is a profile obtained from $\mathcal{X}$ by adding X to someone's preference.

4) *partition consistency* (PC) if for all $i \in N$, disjoint $X, Y \subseteq \mathcal{F}$ such that $X \cup Y \in \mathcal{X}_i$, $F(\mathcal{X}_i, \mathcal{X}_{-i}) = F(\mathcal{X}'_i, \mathcal{X}_{-i})$ where $\mathcal{X}'_i = (\mathcal{X} \backslash [X \cup Y]) \cup [X, Y]$.

As usual, A and N demand that the results are not susceptible to the names of the voters or decision procedures in $\mathcal{F}$, respectively.

SM is also a straightforward modification of strict monotonicity for social choice rules. It demands that if a criterion C is additionally approved by someone, the results will change in favor of the decision procedures supported by the criterion.

The basic idea of PC is that representations of equivalent criteria in different ways should not affect the results. For instance, consider a voter who is in favor of those decision procedures satisfying at least one of either the Condorcet winner criterion (CW) or the Condorcet loser criterion (CL). The union of them can be expressed in several ways[7]:

$$CW \vee CL = (CW \wedge \neg CL) \vee CL = CW \wedge (CL \wedge \neg CW). \qquad (2)$$

For this rephrasing, PC demands that each of the following is regarded as the same in the eyes of the aggregation rule:

1) Approve the united axiom $CW \vee CL$;
2) Approve $(CW \wedge \neg CL)$ and CL; and
3) Approve CW and $(CL \wedge \neg CW)$.

Another justification of PC is as follows. Suppose that criterion d in Table 1 (monotonicity) is favored by the society members. At one time, however, the importance of criterion i (invulnerability to the no-show paradox) is argued and the group is divided into three subgroups: [i] those who prefer $d \wedge i$, [ii] those who prefer $d \wedge \neg i$ (maybe those who do not approve criterion i), and [iii] those who prefer both (maybe those who do not care about criterion i). In this case, we can say either that voters in [iii] approve d or that they approve both $(d \wedge i)$ and $(d \wedge \neg i)$. PC demands that these two ways of preference revelation yield the same outcomes.

For a preference profile $\mathcal{X} = (\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_n) \in \mathfrak{D}^n$, we define the appearance of $f \in \mathcal{F}$ at $\mathcal{X}_i$ as

$$a(f : \mathcal{X}_i) := \sum_{X \subseteq \mathcal{F} \text{ s.t. } f \in X} \mathcal{X}_i(X). \qquad (3)$$

Also, the appearance of $f \in \mathcal{F}$ at $\mathcal{X} = (\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_n)$ is

$$a(f : \mathcal{X}) := \sum_{i \in N} a(f : \mathcal{X}_i). \qquad (4)$$

---

[7] As usual, $\vee$ means "or" and $\wedge$ means "and".

In words, $a(f : \mathcal{X}_i)$ and $a(f : \mathcal{X})$ represent how many criteria are approved that support f (as noted in (4) of Assumption 4, $\mathcal{X}_i(X)$ denotes how many times X appears in $\mathcal{X}_i$). An aggregation rule F is called the *counting rule*, denoted by F_c, if it selects the decision procedures with the highest appearances:

$$F(\mathcal{X}) = \mathrm{argmax}_{f \in \mathcal{F}} \, a(f : \mathcal{X}). \tag{5}$$

In the later arguments, we use

$$A(\mathcal{X}) = \sum_{f \in \mathcal{F}} a(f : \mathcal{X}). \tag{6}$$

Example 1
Consider a group of three voters, denoted by 1-3. Suppose voter 1 approves criteria a, b, and c in Table 1 (maybe a person who prefers the Condorcet-related criteria) and voter 2 approves e and h (maybe a person who prefers Arrow's framework), and voter 3 approves d and i. Then, among the 13 procedures in Table 1, Copeland, Kemeny, and Black get the highest appearances of 5 pts and thus selected by the counting rule.

Theorem 1
An aggregation rule F satisfies A, N, SM, and PC if and only if it is the counting rule.

[Proof of Theorem 1]
The 'if' part is straightforward. We prove the 'only if' part. Let F be an aggregation rule satisfying A, N, SM, and PC. The proof is made by an induction on $A(\mathcal{X})$.

Suppose that $A(\mathcal{X}) = 0$. For any $f, g \in \mathcal{F}$, $\mathcal{X}$ remains the same when the positions of f and g are swapped. So, N demands that $f \in F(\mathcal{X}) \Leftrightarrow g \in F(\mathcal{X})$ for all $f, g \in \mathcal{F}$. Since $F(\mathcal{X}) \neq \phi$ by definition, it follows that $F(\mathcal{X}) = \mathcal{F}$.

Suppose that the "if" part holds when $A(\mathcal{X}') \leq k$. Suppose $A(\mathcal{X}) = k + 1$. Let $\mathcal{X}^0 = \mathcal{X} = (\mathcal{X}_1, \mathcal{X}_2 \dots, \mathcal{X}_n)$, $\mathcal{X}_i = [X_{i,1}, X_{i,2}, \dots, X_{i,a_i}]$, and $E := \{f \in \mathcal{F} | a(f : \mathcal{X}) = k + 1\}$. The proof is completed if we show $F(\mathcal{X}) = E$.

For each $e \in E$, we choose $i(e) \in N$, $j(e) \in \{1, 2, \dots, a_{i(e)}\}$, and $c(e) = X_{i(e),j(e)}$ such that $e \in X_{i(e),j(e)}$ (such $X_{i(e),j(e)}$ exists because $a(e : \mathcal{X}) = k + 1 > 0$). Let $\mathcal{X}^1$ be a profile obtained from $\mathcal{X}$ by deleting each $e \in E$ from $c(e)$. By the assumption of induction, we have

$$F(\mathcal{X}^1) = \{f \in \mathcal{F} \mid a(f : \mathcal{X}^1) = k\}. \tag{7}$$

Note that $E \subseteq F(\mathcal{X}^1)$. Next, we construct $\mathcal{X}^2$ by adding E to $\mathcal{X}_1^1$ (voter 1's preference at $\mathcal{X}^1$). With SM, we have $F(\mathcal{X}^2) = F(\mathcal{X}^1) \cap E = E$. Let $E = \{e_1, e_2, \dots, e_k\}$. Divide the added E to singletons in $\mathcal{X}_1^2$ to make $\mathcal{X}_1^3$:

$$\mathcal{X}_1^2 = [\dots, E] \mapsto \mathcal{X}_1^3 = [\dots, \{e_1\}, \{e_2\}, \dots, \{e_k\}]. \tag{8}$$

By repeating PC, we have that $F(\mathcal{X}^3) = F(\mathcal{X}^2) = E$. Construct $\mathcal{X}^4$ by shifting each $\{e\}$ ($e = e_1, e_2, \ldots, e_k$) from $\mathcal{X}_1^3$ to $\mathcal{X}_{i(e)}^3$. With A, it follows that $F(\mathcal{X}^3) = F(\mathcal{X}^4)$. With PC, we can restore each $X_{i(e),j(e)}$ without changing the results: $F(\mathcal{X}^4) = F(\mathcal{X}^0)$. So, we have $F(\mathcal{X}^0) = E$.   ∎

[Independence of the axioms]

To verify the independence of each axiom used in Theorem 1, we provide several examples. The proof is omitted because they are straightforward.

– A, SM, PC but not N: constant rule $F(\mathcal{X}) = \{f\}$ for some fixed $f \in \mathcal{F}$.
– N, SM, PC, but not A: fix a voter $i^* \in N$. With a slight abuse of notation, let

$$F_C(\mathcal{X}) \cap \mathcal{X}_{i^*} = \{f \in F_C(\mathcal{X}) \mid \exists X \in \mathcal{X}_{i^*} \text{ s.t. } f \in X\}. \tag{9}$$

Define an aggregation rule F as:

$$F(\mathcal{X}) = \begin{cases} F_c(\mathcal{X}) & \text{if } F_C(\mathcal{X}) \cap \mathcal{X}_{i^*} = \phi \\ F_C(\mathcal{X}) \cap \mathcal{X}_{i^*} & \text{otherwise.} \end{cases} \tag{10}$$

– A, N, SM, but not PC: An alternative definition of F_c is as follows. Given a profile $\mathcal{X} = (\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_n)$ with $\mathcal{X}_i = [X_{i,1}, \ldots, X_{i,a_i}]$, F_c searches for each f among $X_{i,1}, X_{i,2}, \ldots, X_{i,a_i}$ and assigns one point to f every time it is found in some $X_{i,j}$. Those decision procedures with the highest points are selected.

Slightly modifying the algorithm above, assign $1/|X_{i,j}|$ points every time f is found in some $X_{i,j}$. Such a rule satisfies A, N, SM, but not PC.

– A, N, PC, but not SM: constant rule $F(\mathcal{X}) = \mathcal{F}$. Another example is the aggregation rule that chooses the decision procedures with the greatest and the second greatest number of appearances.

4 Discussion

In Sect. 3, we provide a characterization of counting rule F_C from the four axioms introduced. In other words, we justify the choice of those decision procedures with highest supports. The significance of the theorem depends on the plausibility of the assumptions as well as the present model. This section discusses what happens if the foundations of Theorem 1 is slightly changed, which clarifies the limitations and the significance of the present model.

(1) *What is a criterion – an atomic element or a composite?*

First, let us discuss the foundation of PC. Let $u(\mathcal{X})$ be the set of criteria that all individuals agree at profile $\mathcal{X}$: $u(\mathcal{X}) = \{X \subseteq \mathcal{F}/X \in \mathcal{X}_i \text{ for all } i \in N\}$.

We say that an aggregation rule satisfies *unanimity on criteria* (UC) if $\mathcal{F}(\mathcal{X}) \subseteq \left(\bigcap_{X \in u(\mathcal{X})} X \right)$ holds whenever $\bigcap_{X \in u(\mathcal{X})} X \neq \phi$. In words, if there exist unanimously agreed criteria and they are consistent in the sense that some decision procedure (s) satisfy them all, the winners must be among such decision procedures. Though it looks plausible, this axiom is inconsistent with PC.

Theorem 2 (Impossibility of PC and UC)
If $|\mathcal{F}| \geq 3$, there is no aggregation rule that satisfies both PC and UC.

(Proof of Theorem 2) Assume to the contrary that F satisfies UC and PC. Let f, g, h be distinct elements in $\mathcal{F}$. Let $\mathcal{X}_1 = [\{f, h\}, \{g, h\}]$ and $\mathcal{X}_i = [\{f\}, \{g\}]$ for all $i \geq 2$. Let $\mathcal{X}_1^{'} = [\{f\}, \{h\}, \{g, h\}]$ and $\mathcal{X}_1^{''} = [\{f, h\}, \{g\}, \{h\}]$. With PC, we have $F(\mathcal{X}) = F(\mathcal{X}_1^{'}, \mathcal{X}_{-1}) = F(\mathcal{X}_1^{''}, \mathcal{X}_{-1})$. But UC implies that $F(\mathcal{X}_1^{'}, \mathcal{X}_{-1}) \subseteq \{f\}$ and $F(\mathcal{X}_1^{''}, \mathcal{X}_{-1}) \subseteq \{g\}$. Contradiction. ∎

Readers might find this result somewhat bewildering, because unanimity is one of the most plausible axioms in social choice theory. Thus, one might become concerned with the soundness of PC. Our answer is that these two criteria are based on totally different points of view, which should be called atomic or composite standpoints. Recall that PC demands that different ways to represent a criterion, e.g., $C_1 \vee C_2 = (C_1 \wedge \neg C_2) \vee C_2$, do not affect the results. Consequently, it is possible to partition criteria into pieces without changing the results (see the proof of Theorem 1). Put another way, PC sees each criterion as a composite object that can be decomposed (i.e., a composite standpoint). On the other hand, UC focuses on the form of each criterion as it is, implicitly distinguishing a criterion from its decomposition. In this sense, UC is based on an atomic standpoint: a criterion is indivisible. The impossibility shown in Theorem 2 can be interpreted as the (huge) gap between these two standpoints.

Indeed, the impossibility soon disappears if we think of unanimity based on a composite point of view instead. Let us say that an aggregation rule satisfies *unanimity on procedures* (UP) if whenever there exists a decision procedure f that is unanimously supported by every criterion in the profile, i.e., $f \in X$ for all $X \in \mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_n$, then the winners must be found among such f's. One can easily see that UP and PC are consistent, because F_c satisfies them both.

In the rest of this discussion, we argue how PCBPC can be based on the atomic standpoint. To do this, we need to show another impossibility (Theorem 3 below).

(2) *To distinguish criteria that are logically equivalent under $\mathcal{F}$*

Theorem 3 (Impossibility of A and UC)
There is no aggregation rule that satisfies both A and UC.

(Proof of Theorem 3)

Assume to the contrary that such an aggregation rule F exists. Let $\mathcal{X}_1 = [\{f\}, \{f\}, \{g\}, \{g\}]$, $\mathcal{X}_2 = \phi$, and $\mathcal{X}_i = [\{f\}, \{g\}]$ for all $i \geq 3$. Let $\mathcal{X}_1^{'} = [\{f\}, \{g\}, \{g\}]$,

$\mathcal{X}'_2 = [\{f\}]$, $\mathcal{X}''_1 = [\{f\}, \{f\}, \{g\}]$, and $\mathcal{X}''_2 = [\{g\}]$. Then, A implies that $F(\mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_{-1,2}) = F(\mathcal{X}'_1, \mathcal{X}'_2, \mathcal{X}_{-1,2}) = F(\mathcal{X}''_1, \mathcal{X}''_2, \mathcal{X}_{-1,2})$. However, UC demands that $F(\mathcal{X}'_1, \mathcal{X}'_2, \mathcal{X}_{-1,2}) \subseteq \{f\}$ and $F(\mathcal{X}''_1, \mathcal{X}''_2, \mathcal{X}_{-1,2}) = \{g\}$. Contradiction.   ∎

In the proof, two assumptions—the explicit and the implicit—play key roles. The explicit assumption is the very heart of anonymity, i.e., who approves a certain criterion ($\{f\}$ and $\{g\}$ in the proof), must not affect the result. The implicit assumption is the repeated appearances of logically equivalent criteria (two $\{f\}$'s and two $\{g\}$'s are included in a voter's preference), which is in fact the consequence of Assumption 2. As one can see, the proof of Theorem 3 fails if such repetition is not allowed[8]. Therefore, in the last part of this discussion, we argue PCBPC can be based on the atomic standpoint by dropping Assumption 2.

(3) *What if we drop Assumption 2?*

Finally, we give an informal discussion of what happens if we drop Assumption 2. Specifically, we consider a modified model that prohibits us from claiming logically equivalent criteria in a voter's preference. This restriction can be expressed by substituting $\mathbb{Z}_{\geq 0}$ with $\{0,1\}$ in the co-domain of $\mathcal{X}_i$.

We define Nurmi's aggregation rule F_N below. Note that the rule is slightly modified from Nurmi [1] so that it fits our dichotomous preference model.

Let $\mathcal{X} = (\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_n)$ be a profile and $\mathcal{C}$ be the finite set of criteria at hand. Each voter's preference is supposed to be a subset of $\mathcal{C}$. Let $a(C : \mathcal{X})$ be the number of voters who approve criterion $C \in \mathcal{C}$.

[Step 1] *Determination of collective ranking over $\mathcal{C}$.*
Define an weak ordering[9] $\gtrsim$ over $\mathcal{C}$ such that for all $C, D \in \mathcal{C}$,

$$C \gtrsim D \Leftrightarrow a(C : \mathcal{X}) \geq a(D : \mathcal{X}). \tag{11}$$

Suppose that all elements in $\mathcal{C}$ are arrayed from the greatest to the least as $C_1 \sim C_2 \sim \cdots \sim C_k \gtrsim D_1 \sim D_2 \sim \cdots D_l \gtrsim E_1 \sim E_2 \sim \cdots \sim E_m \gtrsim F_1 \sim \cdots$.
[Step 2] *Collective choice over $\mathcal{F}$ based on the ranking generated in Step 1.*
Let $\mathcal{F}'_1 = \mathcal{F}_{C_1} \cap \mathcal{F}_{C_2} \cap \cdots \cap \mathcal{F}_{C_k}$. If $\mathcal{F}'_1$ is empty, $F_N(\mathcal{X}) = \mathcal{X}$. If $\mathcal{F}'_1$ is a singleton, $F(\mathcal{X}) = \mathcal{F}'_1$.
Otherwise, let $\mathcal{F}'_2 = \mathcal{F}'_1 \cap \mathcal{F}_{D_1} \cap \mathcal{F}_{D_2} \cap \cdots \cap \mathcal{F}_{D_l}$. If $\mathcal{F}'_2$ is empty, $F_N(\mathcal{X}) = \mathcal{F}'_1$. If $\mathcal{F}'_2$ is a singleton, $F(\mathcal{X}) = \mathcal{F}'_2$.
Otherwise, let $\mathcal{F}'_3 = \mathcal{F}'_2 \cap \mathcal{F}_{E_1} \cap \mathcal{F}_{E_2} \cap \cdots \cap \mathcal{F}_{E_m}$. If $\mathcal{F}'_3$ is empty, $F_N(\mathcal{X}) = \mathcal{F}'_2$. If $\mathcal{F}'_3$ is a singleton, $F(\mathcal{X}) = \mathcal{F}'_3$. Otherwise, consider $\mathcal{F}'_4$ and the process goes in this way until it finds a singleton (if the intersection of all elements in C includes more than one element, then let $F_N(\mathcal{X})$ be the intersection).

---

[8] Note that the proof of Theorem 2 does not use repetition. Therefore, we may state that the gap between PC and UC is more essential than that between A and UC.

[9] Nurmi considers the domain of linear preferences. As a result, he proposes Borda's or Kemeny's rule for this step.

Note that all voters (decision procedures) are treated equally in the algorithm, which means that F_N satisfies A (N). Furthermore, it also satisfies UC. To see this, note that the basic idea of F_N is to search for a singleton subset of $\mathcal{F}$ by descending from "the greatest" criteria (i.e., the greatest elements of $\mathcal{C}$ with respect to $\gtrsim$) to "the least" criteria. Throughout this process, we have $\mathcal{F}'_1 \supseteq \mathcal{F}'_2 \supseteq \mathcal{F}'_3 \cdots$. As unanimously agreed criteria, if they existed, would belong to $\mathcal{F}'_1$, UC holds under F_N. As we noted earlier, this process is based on a different point of view from the counting rule F_c, or the present paper's framework (Assumptions 1-4). To characterize a multi-step aggregation rule like F_N could be an interesting future topic.

5   Concluding Remarks

The present article provides a formal model for PCBPC, or procedural choice based on preferences over criteria. Theorem 1 characterizes the counting rule, which chooses the decision procedures with the highest number of supports by approved criteria, with four axioms: anonymity, neutrality, strict monotonicity and partition consistency. The first three are well-known and essentially the same as those used in Dietrich [8] 's characterization of "counting rule" for a procedural choice based on preferences over "decision procedures"[10]. As in Dietrich's view, the use of anonymity and neutrality may be justified on the grounds that procedural choice is made only by the judgments within the decision body. Whether certain voters or decision procedures have competence against others should not be decided a priori; that is a matter of collective choice. SM is also important to exclude constant rules or such absurd rules that choose those decision procedures with minimal support.

Our model focuses only on dichotomous preferences (Assumption 3). While this is a typical assumption in voting theory, many studies on MCDM often consider cardinal preferences over criteria. Therefore, to look for anonymous, neutral, strict monotonic, and partition consistent aggregation rules under more complicated preferences over criteria (e.g., linear preferences or even fuzzy preferences) is an interesting topic for future research both from technical and practical points of view.

References

1. Nurmi, H.: The choice of voting rules based on preferences over criteria. In: Kamiński, B., Kersten, Gregory E., Szapiro, T. (eds.) GDN 2015. LNBIP, vol. 218, pp. 241–252. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-19515-5_19
2. de Almeida, A.T., Morais, D.C., Nurmi, H.: Criterion based choice of rules. Systems, Procedures and Voting Rules in Context. AGDN, vol. 9, pp. 57–66. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-30955-8_7

---

[10] Indeed, if the admissible preference is restricted to a unique singleton, our counting rule corresponds with (Dietrich [8])'s.

3. de Almeida, A.T., Nurmi, H.: A framework for aiding the choice of a voting procedure in a business decision context. In: Kamiński, B., Kersten, G.E., Szapiro, T. (eds.) GDN 2015. LNBIP, vol. 218, pp. 211–225. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-19515-5_17

4. Rae, D.W.: Decision-rules and individual values in constitutional choice. Am. Polit. Sci. Rev. **63**(1), 40–56 (1969)

5. Koray, S.: Self-selective social choice functions verify arrow and Gibbard-Satterthwaite theorems. Econometrica **68**(4), 981–996 (2000)

6. Barbera, S., Jackson, M.O.: Choosing how to choose: self-stable majority rules and constitutions. Q. J. Econ. **119**(3), 1011–1048 (2004)

7. Suzuki, T., Horita, M.: Plurality, borda count, or anti-plurality: regress convergence phenomenon in the procedural choice. In: Bajwa, D., Koeszegi, S.T., Vetschera, R. (eds.) GDN 2016. LNBIP, vol. 274, pp. 43–56. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-52624-9_4

8. Dietrich, F.: How to reach legitimate decisions when the procedure is controversial. Soc. Choice Welfare **24**(2), 363–393 (2005)

9. Diss, M., Merlin, V.: On the stability of a triplet of scoring rules. Theor. Decis. **69**(2), 289–316 (2010). https://doi.org/10.1007/s11238-009-9187-6

10. Arrow, K.: Social Choice and Individual Values. Wiley, New York (1951)

11. May, K.O.: A set of independent necessary and sufficient conditions for simple majority decision. Econometrica **20**(4), 680–684 (1952)

12. Sen, A.K.: Collective Choice and Social Welfare. Holden-Day, San Francisco (1970)