CAES Working Paper Series

Informational requirements of social choice rules to avoid the Condorcet loser

Shin Sato

Faculty of Economics

Fukuoka University

WP-2011-007

Center for Advanced Economic Study

Fukuoka University

(CAES)

8-19-1 Nanakuma, Jonan-ku, Fukuoka,
JAPAN 814-0180

Informational requirements of social choice rules to avoid the Condorcet loser

Shin Sato

Faculty of Economics, Fukuoka University

8-19-1 Nanakuma, Jonan-ku, Fukuoka, JAPAN 814-0180

E-mail: `shinsato@adm.fukuoka-u.ac.jp`

June 8, 2011

Abstract

I consider informational requirements of social choice rules satisfying anonymity, neutrality, monotonicity, and efficiency, and never choosing the Condorcet loser. Depending on the number of agents and the number of alternatives, either the plurality rule or a plurality with a runoff is characterized. In one case, the plurality rule is the most selective rule among the rules operating on the minimal informational requirement. In the other case, each rule operating on the minimal informational requirement is a two-stage rule, and among them, a plurality with a runoff is the rule whose choice at the first stage is most selective. These results not only clarify properties of the plurality rule and a plurality with a runoff, but also explain why they are widely used in real societies.

Keywords: Condorcet loser, informational requirement, plurality rule, plurality with a runoff.

JEL classification: D71.

1 Introduction

Imagine that you are the rule designer. Your intention is to choose the best alternatives. (Ties are allowed.) Along with standard axioms such as *anonymity* and *neutrality*, assume that you want to avoid the Condorcet loser and to minimize the cost of informational processing. Moreover, you do not want to depend on a tie-breaking rule as much as possible. Then, what is the “right” rule? This paper gives the answer to this question. The answer is either a plurality rule or a plurality with a runoff; it is the plurality rule if and only if $n = 2$ or $m = 2$ or $(n, m) = (4, 3)$, where n is the number of agents, and m is the number of alternatives; in all the other cases, the answer is a plurality with a runoff. Thus, in most cases, we have a plurality with a runoff as the answer to the designer’s question.

I believe that the informational requirement is one of the most important features of each rule. For example, if we ignore the informational aspects of rules, it would be difficult to explain why the plurality rule is most dominant in our lives. Sato (2009) shows that among the one-stage rules satisfying *anonymity*, *neutrality*, *monotonicity*, and *efficiency*, the plurality rule operates on the minimal informational requirement, and the plurality rule is more selective than each other such rule operating on the minimal informational requirement. Sato (2009)’s result theoretically supports our intuition and experience that, compared with other “reasonable” social choice rules, a main advantage of the plurality rule lies in its simplicity and selectivity.

However, it is well known that the plurality rule can choose “bad” alternatives.[1] Thus, in this paper, in addition to *anonymity*, *neutrality*, *monotonicity*, and *efficiency*, I require each rule not to choose the *Condorcet loser*. The Condorcet loser is the counterpart of the *Condorcet winner* which wins a strict majority against each alternative in a pairwise comparison. The Condorcet loser is defeated by each

[1]For example, see Nurmi (1999, Section 3.1).

other alternative. If you consider that the Condorcet winner is the best alternative when it exists, then you should be ready to accept the view that the Condorcet loser is the worst alternative. Even if you do not consider that the Condorcet winner and the loser are the best and the worst alternative,[2] it is reasonable to exclude the Condorcet loser from the social choice; it would be widely accepted that the Condorcet loser is not the best alternative, and hence should not belong to the social choice.

As I mentioned earlier, in most cases, the answer to the the designer's question is a plurality with a runoff, which is another widely used rule, especially in "serious" social choice problems such as elections of the president in some nations, the dean of faculties, and so on.[3] This result is consistent with our intuition that although a plurality with a runoff uses more amount of information, it is still sufficiently simple, and more "correctly" reflects preferences than the plurality rule.

Finally, I discuss related literature. Fishburn (1977) is closest to my work in the sense that we both study informational requirements and discriminability (or selectivity) as they would give us important insights on social choice rules. Conitzer and Sandholm (2005) study *communication complexities* of eleven widely used voting rules. Among many differences, the most significant and essential one between my approach and Conitzer and Sandholm (2005) is that I consider a minimal informational requirement as an "axiom" and hence measuring the informational size of specific social choice rules is not my objective while it is in Conitzer and Sandholm (2005). The plurality rule is axiomatically characterized by Richelson (1978); Roberts (1991); Ching (1996); Yeh (2008), among others, but I do not know any characterization of a plurality with a runoff.

[2] I believe that choosing the Condorcet winner whenever it exists is not necessarily an appealing property of a rule.

[3] Precisely speaking, a plurality with a runoff in this paper is slightly different from some of them. In this paper, if an alternative winning a majority in a weak sense (50 out of 100) in the first stage, then there is no second stage, and the alternative is the winner. In real societies, an alternative is often required to win a *strict* majority (51 out of 100) to be the winner without the second stage.

This paper is organized as follows. Section 2 gives basic notation and definitions. Section 3 gives main results. Section 4 concludes the paper. Proofs are collected in the Appendix.

2 Basic notation and definitions

For each real number a, let $\lceil a \rceil$ denote the smallest integer greater than or equal to a. Let $N = \{1, \dots, n\}$ be a finite set of *agents* with $n \geq 2$, and X be a finite set of *alternatives* with $|X| = m \geq 2$. Let $\mathcal{X}$ be the set of all nonempty subsets of X. For each $x \in X$, I write x for $\{x\}$ when the omission of the braces does not cause confusion. For each $A \in \mathcal{X}$, let $\mathcal{L}(A)$ denote the set of all *linear orders* (complete, transitive, and antisymmetric binary relations) on A. Let $\mathcal{L}$ denote $\mathcal{L}(X)$. Each linear order can be represented as a ranking of the alternatives without ties. I use linear orders to represent agents' preferences over each $A \in \mathcal{X}$. Let R be our generic notation for a linear order. When a preference relation belongs to a particular agent $i \in N$, we write it as R_i. The strict parts of R and R_i are denoted by P and P_i, respectively. An n-tuple $\boldsymbol{R} = (R_1, \dots, R_n)$ is a *preference profile*. For each $R \in \mathcal{L}(A)$ and each $k \in \{1, \dots, |A|\}$, let $r_k(R)$ denote the kth ranked alternative according to R in A. For each $R \in \mathcal{L}$ and each $A \in \mathcal{X}$, let $R|A \in \mathcal{L}(A)$ be the restriction of R to A. Let $\boldsymbol{R}|A = (R_i|A)_{i \in N}$. For each $A \in \mathcal{X}$, each $x \in A$, and each $\boldsymbol{R} \in \mathcal{L}(A)^N$, let $N(\boldsymbol{R}, x)$ be the set of agents who most prefer x at $\boldsymbol{R}$, and let $n(\boldsymbol{R}, x) = |N(\boldsymbol{R}, x)|$. I often write $N(x)$ and $n(x)$ for $N(\boldsymbol{R}, x)$ and $n(\boldsymbol{R}, x)$, respectively, when the preference profile under consideration is obvious.

I consider a rule consisting of at most two stages.[4] In our definition of a rule, I use partitions of the set of linear orders as domains of social choice rules in each stage. I explain an idea behind the formulation. Because we are interested in informational requirements, we have to explicitly account for a kind of information

[4]This is for simplicity. With more than two stages, my results do not change.

each social choice rule uses. As an example, consider the plurality rule. It chooses the most preferred alternatives by the largest number of agents. Then, each pair R_i and R'_i such that $r_1(R_i) = r_1(R'_i)$ are "equivalent" from the viewpoint of the plurality rule. This is because the plurality rule uses information only on what are the top ranked alternatives. In other words, R_i and R'_i send the same "message" to the plurality rule. Based on this observation, for each $x \in X$, let $M(x)$ be the collection of all $R \in \mathcal{L}$ such that $r_1(R) = x$. Let $\mathcal{M}_i = \{M(x) \mid x \in X\}$ for each $i \in N$. Then, $\mathcal{M}_i$ is a partition of $\mathcal{L}$, and each $M_i \in \mathcal{M}_i$ tells us what is the top ranked alternative by agent i. Such M_i and $\mathcal{M}_i$ are called a *message* and a *message space*, respectively. For the plurality rule, each *message profile* $\boldsymbol{M} = (M_1, \ldots, M_n)$ contains enough information to make a social choice.

Next, consider the Borda rule. Intuitively speaking, the Borda rule needs much more information on agents' preferences than the plurality rule. A finer message space reflects this fact; for the Borda rule, no R_i and R'_i with $R_i \neq R'_i$ are "equivalent". Thus, the message space for the Borda rule is $\mathcal{M}'_i = \{\{R_i\} \mid R_i \in \mathcal{L}\}$. Two distinct preference relations serve as distinct messages.

Generally speaking, the further information is required, the more pairs of preference relations should be distinguished from the viewpoint of the rule. This observation motivates our definition of an informational size of a rule. What is important at this point is that coarseness of each message space tells us the informational requirement of each rule, and message spaces are explicitly incorporated into the definition of a rule.

Following the definition, I explain each component.

Definition 2.1

A *rule* consists of the following components:

- (**A set of message profiles**) For each $i \in N$, $\mathcal{M}_i$ is a partition of $\mathcal{L}$ and $\mathcal{M} = \mathcal{M}_1 \times \cdots \times \mathcal{M}_n$,

- **(A social choice rule in the first stage)** f is a function from $\mathcal{M}$ into $\mathcal{X} \times (\mathcal{X} \cup \{\emptyset\} \cup \{0\})$ such that $f_1(\boldsymbol{M}) \supset f_2(\boldsymbol{M})$ for each $\boldsymbol{M} \in \mathcal{M}$ with $f_2(\boldsymbol{M}) \neq 0$, where $f_1(\boldsymbol{M})$ and $f_2(\boldsymbol{M})$ are the first and the second component of f, respectively.

$\star$ The following components are necessary if and only if $\mathcal{I} \equiv \{A \mid A = f(\boldsymbol{M})$ with $f_2(\boldsymbol{M}) \neq 0$ for some $\boldsymbol{M} \in \mathcal{M}\} \neq \emptyset$.

- **(A tie-breaking rule)** T is a function from $\mathcal{I}$ into $\mathcal{X}$ such that $A_1 \setminus A_2 \subset T(A) \subset A_1$ for each $A \in \mathcal{I}$. Let $\mathcal{A} \equiv \{T(A) \mid A \in \mathcal{I}\}$.

- **(A set of message profiles in the second stage)** For each $i \in N$ and each $A \in \mathcal{A}$, $\mathcal{M}(A)_i$ is a partition of $\mathcal{L}(A)$, and $\mathcal{M}(A) = \mathcal{M}(A)_1 \times \cdots \times \mathcal{M}(A)_n$.

- **(A social choice rule in the second stage)** For each $A \in \mathcal{A}$, f_A is a function from $\mathcal{M}(A)$ into $\mathcal{X}$ such that $f_A(\boldsymbol{M}) \subset A$ for each $\boldsymbol{M} \in \mathcal{M}(A)$.

I explain a procedure of making a social choice according to a rule. First, agents send a message profile $\boldsymbol{M} = (M_1, \ldots, M_n)$ to the social choice rule f. The message profile contains information on preferences over X. Although I call f a social choice rule, it does more than usual ones. The value of f at $\boldsymbol{M}$ consists of two components $f_1(\boldsymbol{M})$ and $f_2(\boldsymbol{M})$. The first component $f_1(\boldsymbol{M})$ is a social choice in the first stage. If the second component $f_2(\boldsymbol{M})$ is 0, then we do not proceed to the second stage. In this case, the social choice as a whole is $f_1(\boldsymbol{M})$. Other than 0, $f_2(\boldsymbol{M})$ can take a value in $\mathcal{X} \cup \{\emptyset\}$. In this case, we proceed to the second stage, and we interpret $f_2(\boldsymbol{M})$ as the set of alternatives subject to tie-breaking before the second stage. To be consistent with this interpretation, we assume $f_2(\boldsymbol{M}) \subset f_1(\boldsymbol{M})$. When $f_2(\boldsymbol{M}) = \emptyset$, then each alternative in $f_1(\boldsymbol{M})$ proceed to the second stage. The other extreme is the case $f_2(\boldsymbol{M}) = f_1(\boldsymbol{M})$. In this case, each alternative in $f_1(\boldsymbol{M})$ is subject to tie-breaking before the second stage. Generally, each alternative in $f_1(\boldsymbol{M}) \setminus f_2(\boldsymbol{M})$ proceed to the second stage, but some alternatives in $f_2(\boldsymbol{M})$ might be dropped by a tie-breaking rule T. $T[f(\boldsymbol{M})]$ is the set of alternatives at the beginning of the second stage. Then, because each alternative in $f_1(\boldsymbol{M}) \setminus f_2(\boldsymbol{M})$ proceeds to the second stage, $f_1(\boldsymbol{M}) \setminus f_2(\boldsymbol{M}) \subset T[f(\boldsymbol{M})]$. Because the tie-breaking rule does not add alternatives to $f_1(\boldsymbol{M})$, $T[f(\boldsymbol{M})] \subset f_1(\boldsymbol{M})$. Let $A = T[f(\boldsymbol{M})]$. In the second stage, agents send a message profile in $\mathcal{M}(A)$ to a social choice rule f_A. The message profile contains information on preferences over A. Finally, f_A makes a social choice based on the message profile.

If $\mathcal{I} = \emptyset$, then the rule is called a *one-stage rule*. "$\mathcal{I} = \emptyset$" is equivalent to "$f_2(\boldsymbol{M}) = 0$ for each $\boldsymbol{M} \in \mathcal{M}$". To describe a one-stage rule, it suffices to specify $\mathcal{M}$ and f_1. If $\mathcal{I} \neq \emptyset$, then the rule is called a *two-stage rule*. Note that it is possible that $f_2(\boldsymbol{M}) = 0$ for some $\boldsymbol{M} \in \mathcal{M}$ in a two-stage rule.

I often use F to represent a rule. Let $(\mathcal{M}, f, T, (\mathcal{M}(A))_{A\in\mathcal{A}}, (f_A)_{A\in\mathcal{A}})$ be our generic notation for a two stage rule, and $(\mathcal{M}, f_1)$ for a one-stage rule. For convenience, unless clearly specified, $F = (\mathcal{M}, f, T, \mathcal{M}(A), f_A)$ can be either a one-stage or two-stage rule.[5] (Whenever it does not cause confusion, I drop the subscript "$_{A\in\mathcal{A}}$" from the presentation.)

Essentially, social choice rules f and f_A in a rule are functions on agents' preferences. Each domain of social choice rules such as $\mathcal{M}$ and $\mathcal{M}(A)$ represents an *invariance condition* of f and f_A. Formally, for each $\boldsymbol{R} \in \mathcal{L}^N$, let $\varphi(\boldsymbol{R}) \in \mathcal{M}$ be the message profile such that $R_i \in \varphi(\boldsymbol{R})_i$ for each $i \in N$. Then, $f[\varphi(\boldsymbol{R})] = f[\varphi(\boldsymbol{R}')]$ for each pair $\boldsymbol{R}, \boldsymbol{R}' \in \mathcal{L}^N$ such that $\varphi(\boldsymbol{R}) = \varphi(\boldsymbol{R}')$, i.e., $\boldsymbol{R}$ and $\boldsymbol{R}'$ belong to the same message profile. In this sense, $\mathcal{M}$ can be considered as an invariance condition. I often consider a relationship between preferences and social choices. In such cases, for simplicity, I drop "φ", and write $f(\boldsymbol{R})$ for $f[\varphi(\boldsymbol{R})]$. Similar arguments apply to f_A on $\mathcal{M}(A)$. In other words, I frequently describe

[5] Precisely speaking, this is an abuse of notation because T, $\mathcal{M}(A)$, and f_A should be absent in a one-stage rule.

a rule f on $\mathcal{M}$ as if it were a function on $\mathcal{L}^N$ respecting the invariance condition $\mathcal{M}$, and f_A on $\mathcal{M}(A)$ as if it were a function on $\mathcal{L}(A)^N$ respecting the invariance condition $\mathcal{M}(A)$.

Let $\boldsymbol{R} \in \mathcal{L}^N$. For each one-stage rule F, $f_1(\boldsymbol{R})$ is the *social choice of F at $\boldsymbol{R}$*. For each two-stage rule F, the *social choice of F at $\boldsymbol{R}$* is

- $f_1(\boldsymbol{R})$ if $f_2(\boldsymbol{R}) = 0$, and
- $f_A(\boldsymbol{R}|A)$ if $f_2(\boldsymbol{R}) \neq 0$, where $A = T[f(\boldsymbol{R})]$.

Several remarks are in order.

First, a social choice as a whole can be either single-valued or multi-valued.

Second, in a two-stage rule, by setting $f_2(\boldsymbol{M}) = \emptyset$ for each $\boldsymbol{M} \in \mathcal{M}$, we have a rule without any tie-breaking.

Third, I explicitly consider tie-breaking between the first and the second stage. The purpose of this is to concentrate on an "essence" of social choice rules.[6]

Fourth, I do not consider evolution of preferences from the first to the second stage.

I define two rules. Each rule play an important role in later sections.

Definition 2.2 (Plurality rule)

For each $x \in X$, let $M(x) = \{R \in \mathcal{L} \mid r_1(R) = x\}$. For each $i \in N$, let $\mathcal{M}_i = \{M(x) \mid x \in X\}$. For each $\boldsymbol{M} \in \mathcal{M}$ and each $x \in X$, let $n(x) = |\{i \in$

[6]Some problem is caused due to tie breaking. For example, consider the *plurality with a runoff*. As I define later, the social choice proceeding to the second stage should contain at most two alternatives. Assume $N = \{1, 2, 3\}$, $X = \{x, y, z\}$, xR_1yR_1z, yR_2zR_2x, and zR_3xR_3y. This is a typical situation leading to the *voting paradox*. As long as a "reasonable" social choice rule (without tie breaking) is used in the first stage, $\{x, y, z\}$ is chosen. However, to proceed to the second stage, one alternative should be dropped. Then, regardless of which alternative is dropped and how that alternative is chosen, the choice before the second stage as a whole does not satisfy *anonymity* and *neutrality*. This violation is due to tie-breaking, and not an essential problem. By dividing the procedure into a choice based on preferences and tie-breaking, we can consider *anonymous* and *neutral* social choice rules.

$N \mid M(x) = M_i\}|$. The integer $n(x)$ is the number of agents who rank x at the top of their preferences. Then, for each $\boldsymbol{M} \in \mathcal{M}$, let $f_1^p(\boldsymbol{M}) = \{x \in X \mid n(x) \geq n(y),\ \forall y \in X\}$. The one-stage rule $(\mathcal{M}, f_1^p)$ is called the *plurality rule*.

For each $\boldsymbol{R} \in \mathcal{L}^N$ and each $x \in f_1^p(\boldsymbol{R})$,[7]

$$m \times n(\boldsymbol{R}, x) \geq \sum_{y \in X} n(\boldsymbol{R}, y) = n,$$

and hence $n(\boldsymbol{R}, x) \geq \lceil \frac{n}{m} \rceil$. This fact is crucial for our results because when $\lceil \frac{n}{m} \rceil \geq \lceil \frac{n}{2} \rceil$, for each $\boldsymbol{R} \in \mathcal{L}^N$, each $x \in f_1^p(\boldsymbol{R})$ is not the Condorcet loser. See Section 3 for more discussion on this condition.

Definition 2.3 (Plurality with a runoff)

Before a formal definition, I explain an idea. In the first stage, each agent reports the most preferred alternative. If some alternative wins a majority in a weak sense (henceforth, simply "a majority"),[8] the set of the alternatives winning a majority is the social choice as a whole. (This set contains at most two alternatives.) Otherwise, we choose two alternatives most preferred by the agents in the first stage, and proceed to the second stage. If we cannot choose exactly two alternatives based on the number of votes by the agents, we use a tie-breaking rule. In the second stage, the social outcome is determined by a simple majority between the two alternatives.

I define formally. For each $i \in N$, let $\mathcal{M}_i$ be the one in the plurality rule. For each $\boldsymbol{M} \in \mathcal{M}$, let $V_1(\boldsymbol{M}) = f_1^p(\boldsymbol{M})$, and $V_2(\boldsymbol{M}) = \{x \in X \setminus V_1(\boldsymbol{M}) \mid n(x) \geq n(y),\ \forall y \in X \setminus V_1(\boldsymbol{M})\}$, where $n(x)$ is the number of the agents whose message is $M(x)$. Then, let f^r be a social choice rule in the first stage defined by for each

[7] Remember that $f_1^p(\boldsymbol{R})$ stands for $f_1^p(\varphi(\boldsymbol{R}))$, where $\varphi(\boldsymbol{R})$ is the message profile such that $\boldsymbol{R} \in \varphi(\boldsymbol{R})$.

[8] An alternative wins a majority *in a weak sense* if he gets 50 out of 100, *in a strict sense* if he gets 51 out of 100.

$\boldsymbol{M} \in \mathcal{M}$,

$$f^r(\boldsymbol{M}) = \begin{cases} (V_1(\boldsymbol{M}), 0), & \text{if } n(x) \geq \frac{n}{2} \text{ for each } x \in V_1(\boldsymbol{M}) \\ (V_1(\boldsymbol{M}), V_1(\boldsymbol{M})), & \text{if } |V_1(\boldsymbol{M})| \geq 2 \text{ and } n(x) < \frac{n}{2} \text{ for each } x \in V_1(\boldsymbol{M}) \\ (V_1(\boldsymbol{M}) \cup V_2(\boldsymbol{M}), V_2(\boldsymbol{M})), & \text{if } |V_1(\boldsymbol{M})| = 1 \text{ and } n(x) < \frac{n}{2} \text{ for each } x \in V_1(\boldsymbol{M}). \end{cases}$$

The above three cases are exhaustive, but some cases do not occur under some environments. For example, when $n = 2$, only the first case is relevant. When $n = 3$, the third case never occurs.

Let T be a tie-breaking rule on $\mathcal{I}$ such that $|T(A)| = 2$ for each $A \in \mathcal{I}$. For each $A \in \mathcal{A}$, let $\mathcal{M}(A)_i = \{\{R\} \mid R \in \mathcal{L}(A)\}$, and f_A^r be a simple majority rule, i.e., for each $\boldsymbol{M} \in \mathcal{M}(A)$, $x \in f_A^r(\boldsymbol{M})$ if and only if at least half of the agents prefer x to the alternative in $A \setminus x$. The rule $(\mathcal{M}, f^r, T, \mathcal{M}(A), f_A^r)$ is a *plurality with a runoff*.

In the above definition, the only requirement for a tie-breaking rule is that the set of alternatives at the beginning of the second stage contains two alternatives. Ties might be broken alphabetically such as $T(\{x, y, z\}, \{y, z\}) = \{x, y\}$ and $T(\{x, y, z\}, \{x, y\}) = \{x, z\}$. Ties might be broken "inconsistently" such as $T(\{x, y, z\}, \{x, y\}) = \{x, z\}$ and $T(\{x, y, z\}, \{x, y, z\}) = \{y, z\}$.

I define properties of a rule. For each rule F, a social choice rule f (for a one-stage rule, consider only f_1) in the first stage satisfies

- *anonymity* if for each permutation σ of N and each $\boldsymbol{R} \in \mathcal{L}^N$, $f(\boldsymbol{R}) = f(\boldsymbol{R}^\sigma)$, where $\boldsymbol{R}^\sigma \in \mathcal{L}^N$ is defined by for each $i \in N$, $R_i^\sigma = R_{\sigma(i)}$.

- *neutrality* if for each permutation ρ of X and each $\boldsymbol{R} \in \mathcal{L}^N$, $\rho[f_1(\boldsymbol{R})] = f_1[\rho(\boldsymbol{R})]$[9] and $\rho[f_2(\boldsymbol{R})] = f_2[\rho(\boldsymbol{R})]$,[10] where $\rho(\boldsymbol{R}) = (\rho(R_1), \ldots, \rho(R_n)) \in$

[9]For each $A \in \mathcal{X}$, $\rho(A) \equiv \{\rho(x) \mid x \in A\}$.

[10]Define $\rho(\emptyset) = \emptyset$ and $\rho(0) = 0$. (Remember that $f_2(\boldsymbol{R})$ can be $\emptyset$ and 0.)

$\mathcal{L}^N$ is defined by for each $i \in N$, $\rho(R_i) = \{(x, y) \in X^2 \mid (\rho^{-1}(x), \rho^{-1}(y)) \in R_i\}$.

- *monotonicity* if for each pair $\boldsymbol{R}, \boldsymbol{R}' \in \mathcal{L}^N$ such that $x \in f_1(\boldsymbol{R})$, $\boldsymbol{R}$ and $\boldsymbol{R}'$ coincide on $(X \setminus x)^2$, and $x = r_k(R_i) = r_{k'}(R'_i)$ with $k' \leq k$ for each $i \in N$, we have $x \in f_1(\boldsymbol{R}')$.

- *efficiency* if for each $\boldsymbol{R} \in \mathcal{L}$ and each pair $x, y \in X$ such that xP_iy for each $i \in N$, $y \notin f_1(\boldsymbol{R})$.

Anonymity and *neutrality* means a symmetric treatment of agents and alternatives, respectively. The symmetric treatments are not only at the choice of alternatives before tie-breaking, but also at the choice of alternatives subject to tie-breaking. *Monotonicity* means that when $x \in f_1(\boldsymbol{R})$ and the position of x improves through the change from $\boldsymbol{R}$ to $\boldsymbol{R}'$ while the relative comparison of each other pair of alternatives is unchanged, then $x \in f_1(\boldsymbol{R}')$. *Efficiency* means that when $y \in X$ is dominated by some alternative, then y cannot be chosen.

For each $A \in \mathcal{A}$, *anonymity*, *neutrality*, *monotonicity*, and *efficiency* of f_A for each $A \in \mathcal{A}$ are defined in a similar manner. (In the above definition, ignore the condition on f_2, replace f and f_1 by f_A, $\mathcal{L}^N$ by $\mathcal{L}(A)^N$, and X by A.)

A one-stage rule satisfies *anonymity* if f_1 satisfies *anonymity*. A two-stage rule satisfies *anonymity* if f and f_A for each $A \in \mathcal{A}$ satisfy *anonymity*. I use *neutrality*, *monotonicity*, and *efficiency* of a rule in a similar manner.

For each $\boldsymbol{R} \in \mathcal{L}^N$, an alternative $x \in X$ is the *Condorcet loser at* $\boldsymbol{R}$ if $|\{i \in N \mid xR_iy\}| < |\{i \in N \mid yR_ix\}|$ for each $y \in X \setminus x$. A rule *avoids the Condorcet loser* if for each $\boldsymbol{R} \in \mathcal{L}^N$, the Condorcet loser at $\boldsymbol{R}$ does not belong to the social choice of a rule at $\boldsymbol{R}$.

Let $\mathcal{R}$ denote the set of rules avoiding the Condorcet loser and satisfying *anonymity*, *neutrality*, *monotonicity*, and *efficiency*.

It is easy to see that a plurality with a runoff belongs to $\mathcal{R}$. Whether the plurality rule belongs to $\mathcal{R}$ or not depends on n and m. For example, if $n = 2$, then the plurality rule avoids the Condorcet loser. On the other hand, if $n = m = 3$, there is $\boldsymbol{R} \in \mathcal{L}^N$ such that $f_1^p(\boldsymbol{R})$ contains the Condorcet loser. Our theorems in the next section specify the condition on n and m under which the plurality rule belongs to $\mathcal{R}$.

As I discussed before Definition 2.1, the number of messages in each message space $\mathcal{M}_i$ is considered as the informational requirement of a rule. For each rule F and each $\boldsymbol{R} \in \mathcal{L}$, the *informational size of F at $\boldsymbol{R}$* is defined by

- $\sum_{i \in N} |\mathcal{M}_i|$, if $f_2(\boldsymbol{R}) = 0$,
- $\sum_{i \in N} |\mathcal{M}_i| + \sum_{i \in N} |\mathcal{M}(T[f(\boldsymbol{R})])_i|$, if $f_2(\boldsymbol{R}) \neq 0$.

The integer $\sum_{i \in N} |\mathcal{M}_i|$ is the informational size in the first stage. Thus, if there is no second stage, i.e., $f_2(\boldsymbol{R}) = 0$, then $\sum_{i \in N} |\mathcal{M}_i|$ is the informational size of a rule at $\boldsymbol{R}$. If $f_2(\boldsymbol{R}) \neq 0$, the set of alternatives at the beginning of the second stage is $T[f(\boldsymbol{R})]$, and the message space of agent i in the second stage is $\mathcal{M}(T[f(\boldsymbol{R})])_i$. Thus, $\sum_{i \in N} |\mathcal{M}(T[f(\boldsymbol{R})])_i|$ is the informational size in the second stage. The sum of the informational size of the first and the second stage is the informational size of a rule at $\boldsymbol{R}$.

For a one-stage rule, the informational size at $\boldsymbol{R}$ and that at $\boldsymbol{R}'$ are the same for each pair $\boldsymbol{R}, \boldsymbol{R}' \in \mathcal{L}^N$, while not necessarily the same for a two-stage rule. Thus, for a two-stage rule, we have some degree of freedom in the choice of the informational size. Fortunately, for our purpose, the following definition of the minimal informational requirement, which seems uncontroversial, is sufficient.

Definition 2.4

A rule F *operates on the minimal informational requirement in* $\mathcal{R}$ if for each rule F' in $\mathcal{R}$ and each $\boldsymbol{R} \in \mathcal{L}^N$, the informational size of F at $\boldsymbol{R}$ is less than or equal to the informational size of F' at $\boldsymbol{R}$.

When a rule F operates on the minimal informational requirement in $\mathcal{R}$, for each rule $F' \in \mathcal{R}$, whichever preference profile we have, the informational size of F at the profile is not greater than the informational size of F' at the profile.

3 Results

In this section, let $F^p = (\mathcal{M}^p, f_1^p)$ denote the plurality rule, and let $F^r = (\mathcal{M}^r, f^r, T^r,$ $(\mathcal{M}(A)^r)_{A \in \mathcal{A}^r}, (f_A^r)_{A \in \mathcal{A}^r})$ denote a plurality with a runoff.

Theorem 3.1
Assume $\lceil \frac{n}{m} \rceil = \lceil \frac{n}{2} \rceil$.

(i) The plurality rule operates on the minimal informational requirement in $\mathcal{R}$.

(ii) If a rule F *operates on the minimal informational requirement in* $\mathcal{R}$*, then* F *is a one-stage rule* $(\mathcal{M}, f_1)$ *such that*

(a) $\mathcal{M}_i = \mathcal{M}_i^p$ *for each* $i \in N$*, and*

(b) $f_1(\boldsymbol{R}) \supset f_1^p(\boldsymbol{R})$ *for each* $\boldsymbol{R} \in \mathcal{L}^N$.

Two remarks are in order.

First, $\lceil \frac{n}{m} \rceil = \lceil \frac{n}{2} \rceil$ if and only if

- $n = 2$, or
- $m = 2$, or
- $(n, m) = (4, 3)$.

Second, the theorem says that each rule operating on the minimal informational requirement in $\mathcal{R}$ is necessarily a one-stage rule, and it is a *supercorrespondence* of the plurality rule. This can be considered as a characterization of the plurality rule; it is the most selective rule among the ones operating on the minimal informational requirement in $\mathcal{R}$.

Theorem 3.2

Assume $\lceil \frac{n}{m} \rceil < \lceil \frac{n}{2} \rceil$.

(i) *A plurality with a runoff operates on the minimal informational requirement in* $\mathcal{R}$.

(ii) *If a rule* F *operates on the minimal informational requirement in* $\mathcal{R}$*, then* F *is a two-stage rule* $(\mathcal{M}, f, T, (\mathcal{M}(A))_{A\in\mathcal{A}}, (f_A)_{A\in\mathcal{A}})$ *such that*

(a) $\mathcal{M}_i = \mathcal{M}_i^r$ *for each* $i \in N$,

(b) *for each* $\boldsymbol{R} \in \mathcal{L}^N$, $f_2(\boldsymbol{R}) = 0$ *if and only if* $f_2^r(\boldsymbol{R}) = 0$,

(c) $f_1(\boldsymbol{R}) \supset f_1^r(\boldsymbol{R})$ *for each* $\boldsymbol{R} \in \mathcal{L}^N$,

(d) *each* $A \in \mathcal{A}$ *contains two alternatives,*

(e) *for each* $i \in N$ *and each* $A \in \mathcal{A}$, $\mathcal{M}(A)_i = \{\{R\} \mid R \in \mathcal{L}(A)\}$*, and*

(f) *for each* $A \in \mathcal{A}$, f_A *is the simple majority rule.*

Several remarks are in order.

First, because $m \geq 2$, the two cases $\lceil \frac{n}{m} \rceil = \lceil \frac{n}{2} \rceil$ and $\lceil \frac{n}{m} \rceil < \lceil \frac{n}{2} \rceil$, considered in Theorems 3.1 and 3.2, respectively, are exhaustive.

Second, Theorem 3.2 says that each rule operating on the minimal informational requirement in $\mathcal{R}$ is necessarily a two-stage rule, and the social choice rule in the first stage is a *supercorrespondence* of the one in a plurality with a runoff. Moreover, only two alternatives proceed to the second stage, and the second stage is the simple majority rule between them.

Third, it is more common to require a majority in a strict sense instead of a weak sense for an alternative to be a winner without the second stage. Let F^{r+1} denote such a rule. If n is odd, there is no difference between F^r and F^{r+1}. Assume n is even. Then, at some preference profile, the informational size of F^{r+1} is larger than F^r. Nevertheless, it would be safe to say that the difference is practically negligible in such "big" social choice problems (for example, presidential elections by a people) that the cost of information processing matters.

Fourth, we have a characterization of a plurality with a runoff. Assume that you are the rule designer, and that you want to choose a rule in $\mathcal{R}$. Also, you want a rule with the least informational requirements. Moreover, in the first stage, you want the rule to be as selective as possible. In other words, you do not want to depend on a tie-breaking rule as much as possible. Then, the answer is a plurality with a runoff.

4 Concluding remarks

What we understand through this paper and Sato (2009) concerning the cost of information processing can be summarized as follows:

$$\begin{bmatrix} \text{Anonymity} \\ \text{Neutrality} \\ \text{Monotonicity} \\ \text{Efficiency} \\ \text{No Condorcet loser} \end{bmatrix} + [\text{minimal information}] + [\text{selectivity}] = \begin{cases} \text{plurality}, \\ \text{or} \\ \text{plurality with a runoff}, \end{cases}$$

and

$$\begin{bmatrix} \text{Anonymity} \\ \text{Neutrality} \\ \text{Monotonicity} \\ \text{Efficiency} \end{bmatrix} + [\text{minimal information}] + [\text{selectivity}] = \text{plurality}.$$

Unlike usual "$+$", the above "$+$" are not commutative. If we first imposed minimality of informational requirements, then constant rules would be the answer. Thus, the above formulas should be read from left to right.

The difference between the left sides of the two formulas is "No Condorcet loser". If $\lceil \frac{n}{m} \rceil = \lceil \frac{n}{2} \rceil$, this difference does not make a difference in the "answer".

However, if $\lceil \frac{n}{m} \rceil < \lceil \frac{n}{2} \rceil$, the existence of "No Condorcet loser" is crucial and changes the conclusion.

Like other axiomatic analyses in social choice theory, these results clarify properties of the plurality rule and a plurality with a runoff. In addition to this, I believe that the results explain whey these two rules are widely used in real societies. This is possible because we consider a practically important feature of rules, i.e., the amount of information each rule uses to make a social choice.

Appendix Proofs

I use many permutations of N and X. For simplicity, when I describe a permutation, I do not specify the part on which the permutation is the identity function. For example, when I say that σ is the permutation of N exchanging i and j, then it should be understood that σ is the identity function on $N \setminus \{i, j\}$.

Let $F \equiv (\mathcal{M}, f, T, (\mathcal{M}(A))_{A \in \mathcal{A}}, (f_A)_{A \in \mathcal{A}})$ be a rule which operates on the minimal informational requirement in $\mathcal{R}$. This is either a one-stage or two-stage rule. (See footnote 5.) However, whenever I talk about $\mathcal{A}$, $\mathcal{M}(A)$, and f_A, then it should be understood that the rule under consideration is a two-stage rule.

Among the following lemmas, Lemmas A.1, A.2, and A.5 are based on Sato (2009)'s Lemma 4.1, Lemma 4.2, and arguments in Sato (2009, p.196–197), respectively. I give proofs of them for completeness.

Lemma A.1

(i) For each pair $i, j \in N$, $\mathcal{M}_i = \mathcal{M}_j$.

(ii) For each pair $i, j \in N$ and for each $A \in \mathcal{A}$, $\mathcal{M}(A)_i = \mathcal{M}(A)_j$.

Proof. (i) Suppose $\mathcal{M}_i \neq \mathcal{M}_j$ for some pair $i, j \in N$.

CLAIM 1: Either

(a) there is $\bar{M}_i \in \mathcal{M}_i$ such that $\bar{M}_i \cap M_j^1$ and $\bar{M}_i \cap M_j^2$ for some distinct $M_j^1, M_j^2 \in$

$\mathcal{M}_j$, or

(b) there is $\bar{M}_j \in \mathcal{M}_j$ such that $M_i^1 \cap \bar{M}_j$ and $M_i^2 \cap \bar{M}_j$ for some distinct $M_i^1, M_i^2 \in \mathcal{M}_i$.

Proof of Claim 1. Suppose neither one of (a) and (b) holds. First, I prove $\mathcal{M}_i \subset \mathcal{M}_j$. Let $M_i \in \mathcal{M}_i$. Then, there is $M_j \in \mathcal{M}_j$ such that $M_i \subset M_j$. For M_j, there is $M_i' \in \mathcal{M}_i$ such that $M_j \subset M_i'$. Thus, $M_i \subset M_j \subset M_i'$. Because $\mathcal{M}_i$ is a partition of $\mathcal{L}$, $M_i = M_i'$. This implies $M_i = M_j$, and hence $M_i \in \mathcal{M}_j$. Because M_i was arbitrary in $\mathcal{M}_i$, we have $\mathcal{M}_i \subset \mathcal{M}_j$. Symmetric arguments show $\mathcal{M}_i \supset \mathcal{M}_j$. Thus, $\mathcal{M}_i = \mathcal{M}_j$. This is a contradiction to the assumption $\mathcal{M}_i \neq \mathcal{M}_j$. $\square$

Without loss of generality, assume (a).

CLAIM 2: For each $\boldsymbol{M}_{-j} \in \mathcal{M}_{-j}$, $f(M_j^1, \boldsymbol{M}_{-j}) = f(M_j^2, \boldsymbol{M}_{-j})$.

Proof of Claim 2. Let $R_j^1 \in \bar{M}_i \cap M_j^1$ and $R_j^2 \in \bar{M}_i \cap \mathcal{M}_j^2$. Let $\boldsymbol{M}_{-j} \in \mathcal{M}_{-j}$ and $\boldsymbol{R}_{-j} \in \boldsymbol{M}_{-j}$. Let σ be the permutation of N exchanging i and j. Because $(R_j^1, \boldsymbol{R}_{-j})^\sigma$ and $(R_j^2, \boldsymbol{R}_{-j})^\sigma$ belong to the same message profile, we have $f[(R_j^1, \boldsymbol{R}_{-j})^\sigma] = f[(R_j^2, \boldsymbol{R}_{-j})^\sigma]$. Because f satisfies *anonymity*, $f([(R_j^1, \boldsymbol{R}_{-j})^\sigma]^\sigma) = f([(R_j^2, \boldsymbol{R}_{-j})^\sigma]^\sigma)$. Because $[(R_j^1, \boldsymbol{R}_{-j})^\sigma]^\sigma = (R_j^1, \boldsymbol{R}_{-j})$ and $[(R_j^2, \boldsymbol{R}_{-j})^\sigma]^\sigma = (R_j^2, \boldsymbol{R}_{-j})$, we have $f(R_j^1, \boldsymbol{R}_{-j}) = f(R_j^2, \boldsymbol{R}_{-j})$. This implies that $f(M_j^1, \boldsymbol{M}_{-j}) = f(M_j^2, \boldsymbol{M}_{-j})$. $\square$

Claim 2 shows that M_j^1 and M_j^2 in $\mathcal{M}_j$ do not serve as "distinct" messages. Let $M_j^* = M_j^1 \cup M_j^2$. By merging M_j^1 and M_j^2 into one message M_j^*, we have the message space $\mathcal{M}_j^* = \{M_j \mid M_j \in \mathcal{M}_j \setminus \{M_j^1, M_j^2\} \text{ or } M_j = M_j^*\}$. Replace $\mathcal{M}_j$ in $\mathcal{M}$ by $\mathcal{M}_j^*$. Let $\mathcal{M}^*$ denote the resulting profile of message spaces. Define f^* on $\mathcal{M}^*$ by $f^*(\boldsymbol{R}) = f(\boldsymbol{R})$ for each $\boldsymbol{R} \in \mathcal{L}^N$.[11] Then, the rule $F^* = (\mathcal{M}^*, f^*, T, (\mathcal{M}(A))_{A \in \mathcal{A}}, (f_A)_{A \in \mathcal{A}})$ is in $\mathcal{R}$. Also, at each $\boldsymbol{R} \in \mathcal{L}^N$,

[11] The function f^* on $\mathcal{M}^*$ can be defined in this manner because f respects an invariance condition $\mathcal{M}^*$.

F^* operates on less informational requirements than F. This is a contradiction to the assumption that F operates on the minimal informational requirement in $\mathcal{R}$.

(ii) Let $A \in \mathcal{A}$. Suppose $\mathcal{M}(A)_i \neq \mathcal{M}(A)_j$ for some $i, j \in N$. By similar arguments in (i), we have $\mathcal{M}(A)^*$ and f_A^* on $\mathcal{M}(A)^*$ such that $f_A^*(\boldsymbol{R}) = f_A(\boldsymbol{R})$ for each $\boldsymbol{R} \in \mathcal{L}(A)^N$ and $\sum_{i \in N} |\mathcal{M}(A)_i| > \sum_{i \in N} |\mathcal{M}(A)_i^*|$. In F, replace $\mathcal{M}(A)$ by $\mathcal{M}(A)^*$, and f_A by f_A^*. Let F^* be the resulting rule. Then, F^* is in $\mathcal{R}$. Let $\boldsymbol{R} \in \mathcal{L}^N$ be such that $T[f(\boldsymbol{R})] = A$. The informational size of F at $\boldsymbol{R}$ is greater than that of F^* at $\boldsymbol{R}$. This is a contradiction to the assumption that F operates on the minimal informational requirement in $\mathcal{R}$. ■

Lemma A.2

For each $i \in N$, each $M \in \mathcal{M}_i$, and each permutation ρ of X, $\rho(M) \equiv \{\rho(R) \mid R \in M\} \in \mathcal{M}_i$.

Proof. Suppose $\rho(M) \notin \mathcal{M}_i$. There are two cases.

CASE 1: $\rho(M) \subsetneq M'$ for some $M' \in \mathcal{M}_i$. Because $\rho(M) \subsetneq M'$, we have $M \subsetneq \rho^{-1}(M')$.[12] Because $\mathcal{M}_i$ is a partition of $\mathcal{L}$, there is $M^* \in \mathcal{M}_i$ such that $M^* \neq M$ and $M^* \cap \rho^{-1}(M') \neq \emptyset$. Let $R_i^1 \in M$ and $R_i^2 \in M^* \cap \rho^{-1}(M')$. I claim that for each $\boldsymbol{R}_{-i} \in \mathcal{L}^{N \setminus \{i\}}$, $f(R_i^1, \boldsymbol{R}_{-i}) = f(R_i^2, \boldsymbol{R}_{-i})$. Suppose not. Then, $\rho[f(R_i^1, \boldsymbol{R}_{-i})] \neq \rho[f(R_i^2, \boldsymbol{R}_{-i})]$, where $\rho[f(R_i^1, \boldsymbol{R}_{-i})] = (\rho[f_1(R_i^1, \boldsymbol{R}_{-i}), \rho[f_2(R_i^1, \boldsymbol{R}_{-i})])$ and $\rho[f(R_i^2, \boldsymbol{R}_{-i})] = (\rho[f_1(R_i^2, \boldsymbol{R}_{-i}), \rho[f_2(R_i^2, \boldsymbol{R}_{-i})])$. By *neutrality*, $f[\rho(R_i^1, \boldsymbol{R}_{-i})] \neq f[\rho(R_i^2, \boldsymbol{R}_{-i})]$. However, because $\rho(R_i^1)$ and $\rho(R_i^2)$ belong to the same message M', we have $f[\rho(R_i^1, \boldsymbol{R}_{-i})] = f[\rho(R_i^2, \boldsymbol{R}_{-i})]$.

[12] I prove $M \subsetneq \rho^{-1}(M')$. First, I prove $M \subset \rho^{-1}(M')$. Let $R \in M$. Because $\rho(M) \subset M'$, $\rho(R) \in M'$. Then, by definition of $\rho^{-1}(M')$, $\rho^{-1}[\rho(R)] \in \rho^{-1}(M')$. Because $\rho^{-1}[\rho(R)] = R$, we have $R \in \rho^{-1}(M')$. Thus, $M \subset \rho^{-1}(M')$. Next, I prove $M \neq \rho^{-1}(M')$. Because $\rho(M) \subsetneq M'$, there exists $R' \in M'$ such that $R' \notin \rho(M)$. By definition, $\rho^{-1}(R') \in \rho^{-1}(M')$. I claim $\rho^{-1}(R') \notin M$. Suppose $\rho^{-1}(R') \in M$. Then, $\rho[\rho^{-1}(R')] \in \rho(M)$. Because $\rho[\rho^{-1}(R')] = R'$, we have $R' \in \rho(M)$, which is a contradiction.

This is a contradiction. Therefore, $f(R_i^1, \boldsymbol{R}_{-i}) = f(R_i^2, \boldsymbol{R}_{-i})$. This implies that M and M^* can be integrated into one message without any substantial change. (See the arguments after Claim 2 in the proof of Lemma A.1.) This is a contradiction to the assumption that F operates on the minimal informational requirement in $\mathcal{R}$.

CASE 2: $\rho(M) \cap M^1 \neq \emptyset$ and $\rho(M) \cap M^2 \neq \emptyset$ for some $M^1, M^2 \in \mathcal{M}_i$ with $M^1 \neq M^2$. Let $R_i^1 \in \rho(M) \cap M^1$ and $R_i^2 \in \rho(M) \cap M^2$. In this case, it can be seen that $f(R_i^1, \boldsymbol{R}_{-i}) = f(R_i^2, \boldsymbol{R}_{-i})$ for each $\boldsymbol{R}_{-i} \in \mathcal{L}^{N\setminus\{i\}}$. (See arguments in Case 1.) Then, two messages M^1 and M^2 can be integrated into one message without any substantial change. This is a contradiction to minimality of the informational requirement of F. ■

Lemma A.3

(i) For each $i \in N$ and each $M \in \mathcal{M}_i$, there is $x \in X$ such that $M \subset M(x) \equiv \{R \in \mathcal{L} \mid r_1(R) = x\}$.

(ii) For each $i \in N$, each $A \in \mathcal{A}$, and each $M \in \mathcal{M}(A)_i$, there is $x \in A$ such that $M \subset M_A(x) \equiv \{R \in \mathcal{L}(A) \mid r_1(R) = x\}$.

Proof. (i) Let $i \in N$ and $M \in \mathcal{M}_i$. Suppose $M \not\subset M(x)$ for each $x \in X$. Let $\boldsymbol{M} \in \mathcal{M}$ be such that $M_j = M$ for each $j \in N$. (By Lemma A.1, such $\boldsymbol{M}$ exists in $\mathcal{M}$.) Then, there are $R, R' \in M$ such that $r_1(R) \neq r_1(R')$. Let $x = r_1(R)$ and $y = r_1(R')$. Let $\boldsymbol{R} \in \boldsymbol{M}$ be such that $R_i = R$ for each $i \in N$. Let $\boldsymbol{R}' \in \boldsymbol{M}$ be such that $R'_i = R'$ for each $i \in N$. Because $\boldsymbol{R}, \boldsymbol{R}' \in \boldsymbol{M}$, $f_1(\boldsymbol{R}) = f_1(\boldsymbol{R}')$. Because f satisfies *efficiency*, for each $z \in X \setminus x$, we have $z \notin f_1(\boldsymbol{R})$. By *efficiency*, we also have $x \notin f_1(\boldsymbol{R}')$. Thus, $f_1(\boldsymbol{R}) = f_1(\boldsymbol{R}') = \emptyset$, which is a contradiction.

(ii) In the proof of (i), replace $\mathcal{M}_i$ by $\mathcal{M}(A)_i$, X by A, $\mathcal{M}$ by $\mathcal{M}(A)$, and f_1 by f_A. ■

Lemma A.4

For each $i \in N$, $\mathcal{M}_i = \mathcal{M}_i^p = \mathcal{M}_i^r$.

Proof. The second equality holds by definition. Thus, I prove $\mathcal{M}_i = \mathcal{M}_i^r$.

Let $\boldsymbol{R} \in \mathcal{L}^N$ be such that $R_i = R_j$ for each $i \in N$. Then, the informational size of F^r at $\boldsymbol{R}$ is nm. Because F operates on the minimal informational requirement in $\mathcal{R}$ and $F^r \in \mathcal{R}$, the informational size of F at $\boldsymbol{R}$ is less than or equal to nm. By Lemma A.3, for each pair $R, R' \in \mathcal{L}$ such that $r_1(R) \neq r_1(R')$, R and R' belong to different messages. This implies that $|\mathcal{M}_i| \geq m = |\mathcal{M}_i^r|$ for each $i \in N$. Therefore, the informational size of F at $\boldsymbol{R}$ is greater than or equal to nm. Thus, the informational size of F at $\boldsymbol{R}$ is nm.

STEP 1: $\mathcal{M}_i \subset \mathcal{M}_i^r$.

Let $M \in \mathcal{M}_i$. Let $x \in X$ be such that $M \subset M(x)$. (By Lemma A.3, such x exists.) I claim $M = M(x)$. Suppose $M \subsetneq M(x)$. Because $\mathcal{M}_i$ and $\mathcal{M}_i^r$ are partitions of $\mathcal{L}$, there is $M' \in \mathcal{M}_i$ such that $M' \subset M(x)$ and $M \neq M'$. Let $y \in X$, and ρ be the permutation of X exchanging x and y. Then, $\rho(M)$ and $\rho(M')$ are subsets of $M(y)$. By Lemma A.2, $\rho(M), \rho(M') \in \mathcal{M}_i$. Because y was arbitrary, for each $z \in X$, $M(z)$ contains at least two messages in $\mathcal{M}_i$. Thus, $|\mathcal{M}_i| \geq 2m$. By Lemma A.1, $\sum_{i \in N} |\mathcal{M}_i| \geq 2mn$. Therefore, the informational size of F at $\boldsymbol{R}$ is at least $2mn$, which is a contradiction. Thus, $M = M(x)$ and $M \in \mathcal{M}_i^r$.

STEP 2: $\mathcal{M}_i \supset \mathcal{M}_i^r$.

Let $M(x) \in \mathcal{M}_i^r$. Because $\mathcal{M}_i^r$ and $\mathcal{M}_i$ are partitions of $\mathcal{L}$, there is $M \in \mathcal{M}_i$ such that $M(x) \cap M \neq \emptyset$. By Lemma A.3, $M(x) \supset M$. Then, $M(x) = M$ can be established by the same arguments in Step 1. ■

Lemma A.5

For each $\boldsymbol{R} \in \mathcal{L}^N$, $f_1(\boldsymbol{R}) \supset f_1^p(\boldsymbol{R})$.

Proof. Suppose $f_1(\boldsymbol{R}) \not\supset f_1^p(\boldsymbol{R})$ for some $\boldsymbol{R} \in \mathcal{L}^N$. Let $x \in f_1^p(\boldsymbol{R}) \setminus f_1(\boldsymbol{R})$.

First, I claim $f_1^p(\boldsymbol{R}) \cap f_1(\boldsymbol{R}) = \emptyset$. Suppose that there is $y \in f_1^p(\boldsymbol{R}) \cap f_1(\boldsymbol{R})$. Let σ be a permutation of N such that $\sigma(\{i \in N \mid r_1(R_i) = x\}) = \{i \in N \mid r_1(R_i) = y\}$ and $\sigma(\{i \in N \mid r_1(R_i) = y\}) = \{i \in N \mid r_1(R_i) = x\}$. By *anonymity*, $y \in f_1(\boldsymbol{R}^\sigma)$. Let ρ be the permutation of X exchanging x and y. By *neutrality*, $x \in f_1[\rho(\boldsymbol{R}^\sigma)]$. Note that the two n-tuples of top ranked alternatives in $\boldsymbol{R}$ and $\rho(\boldsymbol{R}^\sigma)$ are the same. Therefore, by Lemma A.4, $\boldsymbol{R}$ and $\rho(\boldsymbol{R}^\sigma)$ belong to the same message profile. Thus, $x \in f_1(\boldsymbol{R})$, which is a contradiction.

Let $z \in f_1(\boldsymbol{R})$. By the above argument, $z \notin f_1^p(\boldsymbol{R})$. Let N_x be a subset of $\{i \in N \mid r_1(R_i) = x\}$ such that $|N_x| = |\{i \in N \mid r_1(R_i) = z\}|$. Let σ' be a permutation of N such that $\sigma'(N_x) = \{i \in N \mid r_1(R_i) = z\}$ and $\sigma'(\{i \in N \mid r_1(R_i) = z\}) = N_x$. By *anonymity*, $z \in f_1(\boldsymbol{R}^{\sigma'})$. Let ρ' be the permutation exchanging x and z. By *neutrality*, $x \in f_1[\rho'(\boldsymbol{R}^{\sigma'})]$. For each $i \in \{j \in N \mid r_1(R_j) = x\} \setminus N_x$, lift x to the top in $\rho'(R_i^{\sigma'})$. Let $\boldsymbol{R}'' \in \mathcal{L}^N$ be the resulting preference profile. By *monotonicity*, $x \in f_1(\boldsymbol{R}'')$. The two n-tuples of top ranked alternatives in $\boldsymbol{R}$ and $\boldsymbol{R}''$ are the same. Therefore, $\boldsymbol{R}$ and $\boldsymbol{R}''$ belong to the same message profile. Because $x \in f_1(\boldsymbol{R}'')$, it follows that $x \in f_1(\boldsymbol{R})$. This is a contradiction to the choice of $x \in f_1^p(\boldsymbol{R}) \setminus f_1(\boldsymbol{R})$. ■

A.1 A proof of Theorem 3.1

STEP 1: The plurality rule is in $\mathcal{R}$.

I prove that the plurality rule avoids the Condorcet loser. Let $\boldsymbol{R} \in \mathcal{L}^N$ and $x \in f_1^p(\boldsymbol{R})$. Then, as I showed below Definition 2.2, $n(\boldsymbol{R}, x) \geq \lceil \frac{n}{m} \rceil \geq \lceil \frac{n}{2} \rceil$. Thus, for each $y \in X$, at least half of the agents prefer x to y. Therefore, x is not the Condorcet loser.

It is easy to see that the plurality rule satisfies *anonymity*, *neutrality*, *monotonicity*, and *efficiency*. Therefore, the plurality rule is in $\mathcal{R}$.

STEP 2: For each $i \in N$, $\mathcal{M}_i = \mathcal{M}_i^p$.

This is established in Lemma A.4.

STEP 3: F is a one-stage rule, and its informational size at each $\boldsymbol{R} \in \mathcal{L}^N$ is nm.

Because F operates on the minimal informational requirement in $\mathcal{R}$ and the informational size of the plurality rule at each $\boldsymbol{R} \in \mathcal{L}^N$ is nm, the informational size of F at each $\boldsymbol{R} \in \mathcal{L}^N$ is not greater than nm. By definition, the informational size of the second stage of each rule is positive. Thus, there is no second stage in F. Therefore, F is a one-stage rule, and by Step 2, its informational size at each $\boldsymbol{R} \in \mathcal{L}^N$ is nm.

STEP 4: The plurality rule operates on the minimal informational requirement in $\mathcal{R}$.

By Step 3, at each $\boldsymbol{R} \in \mathcal{L}^N$, the informational size of F and that of the plurality rule are the same. Because F operates on the minimal informational requirement in $\mathcal{R}$, the plurality rule also operates on the minimal informational requirement in $\mathcal{R}$.

STEP 5: For each $\boldsymbol{R} \in \mathcal{L}^N$, $f_1(\boldsymbol{R}) \supset f_1^p(\boldsymbol{R})$.

This is established in LemmaA.5.

A.2 A proof of Theorem 3.2

Let $F = (\mathcal{M}, f, (\mathcal{M}(A))_{A\in\mathcal{A}}, (f_A)_{A\in\mathcal{A}})$ be a rule which operates on the minimal informational requirement in $\mathcal{R}$. It is clear that $F^r \in \mathcal{R}$. Because the informational size of F^r at each $\boldsymbol{R} \in \mathcal{L}^N$ is at most $nm + 2n$, the information size of F at each $\boldsymbol{R} \in \mathcal{L}^N$ is also at most $nm + 2n$.

STEP 1: For each $i \in N$, $\mathcal{M}_i = \mathcal{M}_i^r$.

This is established in Lemma A.4.

STEP 2 F is not a 1-stage rule.

Suppose that F is a 1-stage rule. Let $x \in X$. There is $\boldsymbol{R} \in \mathcal{L}^N$ such that $x \in f_1^p(\boldsymbol{R})$ and $n(\boldsymbol{R}, x) = \lceil \frac{n}{m} \rceil$.[13] Because $f_1(\boldsymbol{R}) \supset f_1^p(\boldsymbol{R})$ (by Lemma A.5), we have $x \in f_1(\boldsymbol{R})$. For each $i \in N \setminus N(\boldsymbol{R}, x)$, take x to the bottom of R_i. Let $\boldsymbol{R}'$ be the resulting preference profile. Because $\mathcal{M}_i = \mathcal{M}_i^r$ for each $i \in N$ (by Step 1), $\boldsymbol{R}$ and $\boldsymbol{R}'$ belong to the same message profile. Thus, $x \in f_1(\boldsymbol{R}')$. Because $n(\boldsymbol{R}, x) = \lceil \frac{n}{m} \rceil < \lceil \frac{n}{2} \rceil$, a strict majority puts x at the bottom of preferences at $\boldsymbol{R}'$. Therefore, $x \in f_1(\boldsymbol{R}')$ is a contradiction to the assumption that F avoids the Condorcet loser.

STEP 3: For each $\boldsymbol{R} \in \mathcal{L}^N$, $f_2(\boldsymbol{R}) = 0$ if and only if $f_2^r(\boldsymbol{R}) = 0$.

Let $\boldsymbol{R} \in \mathcal{L}^N$.

First, assume $f_2(\boldsymbol{R}) = 0$. By Lemma A.5, $f_1(\boldsymbol{R})$ contains $f_1^p(\boldsymbol{R})$. Also, each alternative in $f_1^p(\boldsymbol{R})$ wins a majority.[14] By definition, $f_2^r(\boldsymbol{R}) = 0$.

---

[13] I explain formally. Let α and β be nonnegative integers such that $n = \alpha m + \beta$ with $\beta < m$. If $\beta = 0$, then let $\boldsymbol{R} \in \mathcal{L}^N$ be such that for each $y \in X$, $n(\boldsymbol{R}, y) = \alpha$. If $\beta > 0$, let $A \subset X$ be such that $|A| = \beta$ and $x \in A$, and let $\boldsymbol{R} \in \mathcal{L}^N$ be such that for each $y \in X$,

$$n(\boldsymbol{R}, y) = \begin{cases} \alpha, & \text{if } y \in X \setminus A \\ \alpha + 1, & \text{if } y \in A. \end{cases}$$

In either case ($\beta = 0$ or $\beta > 0$), $x \in f_1^p(\boldsymbol{R})$ and $n(\boldsymbol{R}, x) = \lceil \frac{n}{m} \rceil$.

[14] To see this, suppose that some alternative in $f_1^p(\boldsymbol{R})$ does not win a majority. Let $x \in f_1^p(\boldsymbol{R})$ be such an alternative. For each $j \in N \setminus N(\boldsymbol{R}, x)$, take x to the bottom of R_j. Let $\boldsymbol{R}'$ be the resulting preference profile. Because $|N \setminus N(\boldsymbol{R}, x)| > n/2$, a strict majority puts x at the bottom of preferences. Thus, x is the Condorcet loser. By Step 1, $\boldsymbol{R}$ and $\boldsymbol{R}'$ belong to the same message profile. Thus, $f(\boldsymbol{R}) = f(\boldsymbol{R}')$. Therefore, $x \in f_1(\boldsymbol{R}')$. Because $f_2(\boldsymbol{R}') = 0$, x belongs to the social

Next, assume $f_2^r(\boldsymbol{R}) = 0$. Then, the informational size of F^r at $\boldsymbol{R}$ is nm. By Step 1, the informational size of the first stage of F is nm. Because F operates on the minimal informational requirement in $\mathcal{R}$, there is no second stage at $\boldsymbol{R}$, i.e., $f_2(\boldsymbol{R}) = 0$.

STEP 4: For each $\boldsymbol{R} \in \mathcal{L}^N$, $f_1(\boldsymbol{R}) \supset f_1^r(\boldsymbol{R})$.

Let $\boldsymbol{R} \in \mathcal{L}^N$. By Lemma A.5, $f_1(\boldsymbol{R}) \supset f_1^p(\boldsymbol{R})$. By definition of F^p, $f_1^p(\boldsymbol{R}) = f_1^r(\boldsymbol{R})$ in the first and the second case in the definition of f^r. (See Definition 2.3.) Thus, it suffices to consider the remaining case where $|f_1^p(\boldsymbol{R})| = 1$ and $n(\boldsymbol{R}, x) < \frac{n}{2}$ for each $x \in f_1^p(\boldsymbol{R})$. In other words, the set of winners under the plurality rule is a singleton, and the alternative does not win a majority. Let $\{x\} \equiv f_1^p(\boldsymbol{R})$.

CLAIM 4.1: $f_1(\boldsymbol{R}) \neq \{x\}$.

Proof of Claim 4.1. Because x does not win a majority at $\boldsymbol{R}$, the agents in $N \setminus N(\boldsymbol{R}, x)$ form a strict majority. Then, for each $i \in N \setminus N(\boldsymbol{R}, x)$, take x down to the bottom of R_i. Let $\boldsymbol{R}'$ be the resulting preference profile. Because $\boldsymbol{R}$ and $\boldsymbol{R}'$ belong to the same message profile, $f(\boldsymbol{R}) = f(\boldsymbol{R}')$. Thus, if $\{x\} = f_1(\boldsymbol{R})$, then $\{x\} = f_1(\boldsymbol{R}')$. Because the social choice at each preference profile cannot be $\emptyset$, x is the social choice of F at $\boldsymbol{R}'$. Because x is the Condorcet loser at $\boldsymbol{R}'$, this is a contradiction. $\square$

It suffices to show $f_1(\boldsymbol{R}) \supset V_2(\boldsymbol{R})$. (See Example 2.3 for the definition of $V_2(\boldsymbol{R})$.) Suppose $f_1(\boldsymbol{R}) \not\supset V_2(\boldsymbol{R})$. Let $y \in V_2(\boldsymbol{R}) \setminus f_1(\boldsymbol{R})$.

CLAIM 4.2: $f_1(\boldsymbol{R}) \cap V_2(\boldsymbol{R}) = \emptyset$.

Proof of Claim 4.2. Apply the arguments of the second paragraph in the proof of Lemma A.5. (Replace $f_1^p(\boldsymbol{R})$ by $V_2(\boldsymbol{R})$, y by w, and x by y.) $\square$

Let $z \in f_1(\boldsymbol{R}) \setminus x$. By Claim 4.1, such z exists. By Claim 4.2, $z \notin V_2(\boldsymbol{R})$.

---

choice of F at $\boldsymbol{R}'$. This is a contradiction to the assumption that F avoids the Condorcet loser.

Because $z \notin (V_1(\boldsymbol{R}) \cup V_2(\boldsymbol{R}))$, $n(\boldsymbol{R}, y) > n(\boldsymbol{R}, z)$. Then, by applying the arguments of the last paragraph in the proof of Lemma A.5, we have a contradiction. (Replace x by y, and $f_1^p(\boldsymbol{R})$ by $V_2(\boldsymbol{R})$.)

STEP 5: For each $A \in \mathcal{A}$, $|A| = 2$.

Let $A \in \mathcal{A}$.

First, assume $|A| \geq 3$. By Lemma A.3 (ii), $|\mathcal{M}(A)_i| \geq 3$ for each $i \in N$. Thus, the informational size of the second stage is at least $3n$. This is a contradiction to the assumption that F operates on the minimal informational requirement in $\mathcal{R}$.

Next, assume $|A| = 1$. Let $\{x\} \equiv A$. Let $\boldsymbol{R} \in \mathcal{L}^N$ be such that $T[f(\boldsymbol{R})] = x$. This implies $f_2(\boldsymbol{R}) \neq 0$. (If $f_2(\boldsymbol{R}) \neq 0$, then $f(\boldsymbol{R})$ does not belong to the domain of T.) Because $f_2(\boldsymbol{R}) = 0$ if and only if $f_2^r(\boldsymbol{R}) = 0$ (see Step 3), $f_2(\boldsymbol{R}) \neq 0$ implies that x does not wins a majority at $\boldsymbol{R}$. For each $i \in N \setminus N(\boldsymbol{R}, x)$, take x down to the bottom of R_i. Let $\boldsymbol{R}'$ be the resulting preference profile. At $\boldsymbol{R}'$, x is the Condorcet loser. Because $\boldsymbol{R}$ and $\boldsymbol{R}'$ belong to the same message profile, $T[f(\boldsymbol{R})] = T[f(\boldsymbol{R}')]$. Thus, $x = T[f(\boldsymbol{R}')]$. Because the social choice of F at $\boldsymbol{R}'$ cannot be $\emptyset$, x is the social choice of F at $\boldsymbol{R}'$. This is a contradiction.

STEP 6: For each $i \in N$ and each $A \in \mathcal{A}$, $\mathcal{M}(A)_i = \{\{R\} \mid R \in \mathcal{L}(A)\}$.

Let $A \in \mathcal{A}$. By Step 5, A consists of two alternatives. Let $A = \{x, y\}$. There are only two linear orders over A; x is preferred to y, or y is preferred to x. By Lemma A.3 (ii), these two preferences belong to different messages. Thus, $\mathcal{M}(A)_i = \{\{R\} \mid R \in \mathcal{L}(A)\}$.

STEP 7: For each $A \in \mathcal{A}$, f_A is the simple majority rule between the two alternatives in A.

Let $A \in \mathcal{A}$. By Step 5, A consists of two alternatives. Let f_A^p be the simple majority rule between the two alternatives. I prove $f_A(\boldsymbol{R}^A) = f^p_A(\boldsymbol{R}^A)$ for each $\boldsymbol{R}^A \in \mathcal{L}(A)^N$. (In this step, I put a superscript A to a preference profile over A.)

Let $\boldsymbol{R}^A \in \mathcal{L}^N(A)$.

CLAIM 7.1: $f_A(\boldsymbol{R}^A) \supset f^p_A(\boldsymbol{R}^A)$.

Proof of Claim 7.1. This can be established by similar arguments in Lemma A.5. $\square$

CLAIM 7.2: $f_A(\boldsymbol{R}^A) \subset f^p_A(\boldsymbol{R}^A)$.

Proof of Claim 7.2. I prove the contrapositive; for each $x \in A$, $x \notin f^p_A(\boldsymbol{R}^A)$ implies $x \notin f_A(\boldsymbol{R}^A)$. Thus, let $x \in A$ be such that $x \notin f^p_A(\boldsymbol{R}^A)$. Let $y = A \setminus x$. Let $\boldsymbol{R} \in \mathcal{L}^N$ be such that $T[f(\boldsymbol{R})] = A$. By Step 3, both x and y do not win a majority, i.e., $n(\boldsymbol{R}, x) < \lceil \frac{n}{2} \rceil$ and $n(\boldsymbol{R}, y) < \lceil \frac{n}{2} \rceil$.

Then, there is $\boldsymbol{R}' \in \mathcal{L}^N$ such that $r_1(R_i) = r_1(R'_i)$ for each $i \in N$, x is the Condorcet loser at $\boldsymbol{R}'$, and at $\boldsymbol{R}'|A$, x does not win a majority in A by one vote. Such $\boldsymbol{R}'$ can be obtained as follows; let $N' \subset N$ be such that $N' \supset N(\boldsymbol{R}, x)$, $N' \cap N(\boldsymbol{R}, y) = \emptyset$, and $|N'| = \lceil \frac{n}{2} \rceil - 1$. Such N' exists because neither $N(\boldsymbol{R}, x)$ nor $N(\boldsymbol{R}, y)$ forms a majority. Modify $\boldsymbol{R}$ as follows; for each $i \in N'$, take y to the bottom of R_i, and for each $i \in N \setminus N'$, take x to the bottom of R_i. Let $\boldsymbol{R}'$ be the resulting preference profile. Then, $\boldsymbol{R}'$ is the desired one.

Because $\boldsymbol{R}$ and $\boldsymbol{R}'$ belong to the same message profile, $T[f(\boldsymbol{R}')] = A$. Because x is the Condorcet loser at $\boldsymbol{R}'$, $x \notin f_A(\boldsymbol{R}'|A)$. At $\boldsymbol{R}'|A$, $\lceil \frac{n}{2} \rceil - 1$ agents prefer x to y. Remember that at $\boldsymbol{R}^A$, at most $\lceil \frac{n}{2} \rceil - 1$ agents prefer x to y. By *anonymity* and *monotonicity*, it can be seen that $x \notin f_A(\boldsymbol{R}^A)$. $\square$

By Claims 7.1 and 7.2, we have $f_A(\boldsymbol{R}^A) = f^p_A(\boldsymbol{R}^A)$.

References

Ching, Stephen. (1996) A simple characterization of plurality rule. Journal of Economic Theory 71, 298–302.

Conitzer, Vincent; Sandholm, Tuomas (2005) Communication complexity of common voting rules. Proceedings of the ACM Conference on Electronic Commerce, 78–87.

Fishburn, Peter C. (1977) Condorcet social choice functions. SIAM Journal on Applied Mathematics 33, 469–489.

Nurmi, Hannu (1999) Voting paradoxes and how to deal with them. Springer, Berlin.

Richelson, Jeffrey T. (1978) A characterization result for the plurality rule. Journal of Economic Theory 19, 548–550.

Roberts, Fred S. (1991) Characterization of the plurality function. Mathematical Social Sciences 21, 101–127.

Sato, Shin (2009) Informational requirements of social choice rules. Mathematical Social Sciences 57, 188–198.

Yeh, Chun-Hsien (2008) An efficiency characterization of plurality rule in collective choice problems. Economic Theory 34, 575–583.