Games and Economic Behavior 104 (2017) 350–371

Contents lists available at ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

Ordinal versus cardinal voting rules: A mechanism design approach ☆

Semin Kim

School of Economics, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, Republic of Korea

ARTICLE INFO

Article history:
Received 10 February 2015
Available online 8 May 2017

JEL classification:
C72
D01
D02
D72
D82

Keywords:
Ordinal rule
Pareto efficiency
Incentive compatibility
Bayesian mechanism design

ABSTRACT

We consider the performance and incentive compatibility of voting rules in a Bayesian environment: agents have independent private values, there are at least three alternatives, and monetary transfers are prohibited. First, we show that in a neutral environment, meaning alternatives are symmetric ex-ante, essentially any ex-post Pareto efficient ordinal rule is incentive compatible. Importantly, however, we can improve upon ordinal rules. We show that we can design an incentive compatible cardinal rule which achieves higher utilitarian social welfare than any ordinal rule. Finally, we provide numerical findings about incentive compatible cardinal rules that maximize utilitarian social welfare.

1. Introduction

For the problem of aggregation of individual preferences into a group decision, the vast majority of social choice literature uses a purely ordinal model, meaning that agents can rank the candidates but intensity of preferences is discarded. Similarly, the voting literature mainly studies ordinal voting rules which depend only on ranking information in agents' preferences. Furthermore, attention is usually restricted to strategy-proof rules,[1] i.e., rules in which truth-telling is best for each agent regardless of other agents' reports.

These two assumptions greatly restrict the types of aggregation procedures that can be explored. Thus, relaxing these assumptions proves valuable. First, preference intensity information needs to be considered for the problem. A simple example is when a slight majority of people weakly prefers one alternative and a slight minority strongly prefers another alternative. Which alternative is most desirable for the society? A more concrete example is a representative democracy. Representatives need to choose one of the alternatives. A representative knows how much his district is willing to pay for

☆ This work is based on a chapter from my Ph.D. dissertation written at the department of Economics of The Ohio State University under the supervision of Prof. Yaron Azrieli. I am greatly indebted to Yaron Azrieli for his thoughtful guidance. I thank two anonymous referees for their feedback and suggestions that improved the quality of this article. I am also grateful to Paul Healy, James Peck, Katherine Baldiga Coffman, Christian Hellwig, Dan Levin, Caleb Cox, Jinwoo Kim, Ian Krajbich, Tilman Börgers, Peter Postl, Roberto Serrano, Hye Lin Choi, Jack Rouzer, Robert Munk, Dimitry Mezhvinsky, and audiences at several seminars and conferences for their helpful comments and suggestions. Mu Yang provided valuable research assistance. All errors are my own responsibility.

E-mail address: seminkim@yonsei.ac.kr.

[1] Strategy-proof rules are often called dominant incentive compatible rules in the mechanism design literature.

http://dx.doi.org/10.1016/j.geb.2017.04.012
0899-8256/

the alternatives to be chosen, which may represent his district's preference intensities over the alternatives. He needs to consider not only the ordinal preference of his group, but also the preference intensity of his group. However, if a voting rule that aggregates the votes of the representatives is ordinal, it may not sufficiently satisfy the whole society. Moreover, requiring strategy-proofness may be too demanding. It is well known from the Gibbard–Satterthwaite theorem (Gibbard, 1973 and Satterthwaite, 1975) that any deterministic and strategy-proof rule is dictatorial under mild assumptions.

Our approach to the problem differs from past literature in that, in our model, agents assign cardinal values to the alternatives to indicate their relative desirability. Rather than requiring strategy-proofness, we consider the weaker concept of *incentive compatibility*. It only requires agents to be truth-telling if other agents are also truth-telling when there is incomplete information about other agents' preferences. Our model, therefore, closely resembles the classic mechanism design problem with the crucial difference that monetary transfers cannot be used to remedy incentive problems. To the best of the author's knowledge, only a few papers have investigated incentive compatible cardinal voting rules with preference intensity information.[2]

Our objective is to apply a mechanism design approach, without monetary transfers, to examine the efficiency and incentive compatibility of voting rules. We consider a neutral environment, i.e., when alternatives are symmetric ex-ante, and show that essentially any ex-post Pareto efficient ordinal rule is Incentive Compatible (IC). Further, we prove that there exists an IC cardinal rule which outperforms any ordinal rule. We also report numerical results on IC cardinal rules which maximize utilitarian social welfare, called second-best rules.

Let us be more explicit regarding our environment. We consider a Bayesian environment where an agent's preference over at least three alternatives is private information. To capture the intensity of preferences, each agent has cardinal values for the alternatives. These values are random variables which, we assume, are drawn identically and independently across agents. We also assume that values are neutral between alternatives, which implies that the value distributions of each alternative are symmetric. A voting rule f is a mapping from the reported value profiles to lotteries over the set of alternatives. Social welfare is measured in terms of expected utilities induced by the rule.

We first focus on the class of ordinal rules and consider ex-post Pareto efficiency. An ordinal rule f is Ordinally Paretian Ex-post (OPE) if when every agent prefers one alternative to another, the rule must choose the inferior alternative with zero probability.[3] We show that for any OPE rule f, there exists an IC ordinal rule g that delivers the same expected utility for every agent as f (Theorem 1). This is somewhat surprising, because even in the class of ordinal rules, incentive compatibility has often proven problematic in the strategic voting and mechanism design literature. In other words, an agent often has an incentive to misrepresent his preference in order to prevent an undesirable outcome. Theorem 1, however, guarantees that there is no incentive compatibility problem in our set-up, as long as the rule satisfies the extremely weak criterion of ex-post Pareto efficiency.[4]

The next question deals with whether it is possible to design a rule that is superior to ordinal rules. If we consider the class of cardinal rules, it is known that there exists a trade-off between Pareto efficiency and incentive compatibility.[5] However, we show that an IC cardinal rule can be designed which achieves higher utilitarian social welfare than any ordinal rule (Theorem 2). With three alternatives, this rule is closely related to the (A, B)-scoring rule in Myerson (2002), which is similar to the standard scoring rule, but with two possible score vectors, $(1, A, 0)$ or $(1, B, 0)$ where $(0 \leq A \leq B \leq 1)$. It is the general form of well-known voting rules: *plurality, negative, Borda count, and approval* voting rules.[6] We show that our rule is an IC (A, B)-scoring rule that maximizes utilitarian social welfare in an environment with three alternatives (Proposition 2). Since (A, B)-scoring rules are simple to implement and widely used in reality, Proposition 2 could inform voting rules in a variety of institutional settings.

Finally, we attempt to find the second-best rule with a numerical approach. We first consider the simplest non-trivial environment with two agents, three alternatives, and a uniform distribution of values. We find that the second-best rule loses efficiency only when two agents have the opposite ordinal preference over alternatives. We calculate the utilitarian social welfare of the first-best rule, the second-best rule, the best-ordinal rule, and the rule suggested in Theorem 2. The welfare loss for incentive constraints is much smaller than the loss for the restriction of ordinal rules. Then, we vary the distribution to a beta distribution to check the robustness of our findings under a uniform distribution.

The remainder of this paper is organized as follows. The next section reviews related literature. In Section 3, we introduce the environment and define ordinal rules and scoring rules. Section 4 describes a motivating example. Section 5 discusses the notion of ordinal Pareto efficiency and incentive compatibility of voting rules. In Section 6, we construct an IC cardinal rule superior to any ordinal rule, and show the connection between our rule and the (A, B)-scoring rules. Section 7 reports the numerical findings about the second-best rule. Section 8 concludes. The Appendix contains all the proofs and the beta distribution analysis.

---

[2] See the related literature section.

[3] This notion of Pareto efficiency is widely used in social choice literature (for example, see Pattanaik and Peleg (1986) or McLennan (1980)).

[4] Our notion of ex-post Pareto efficiency is based on ranking information in agents' preferences. On the other hand, Holmström and Myerson (1983) defined sets of ex-ante, interim, and ex-post Pareto efficient rules based on cardinal utility. The set of ex-post Pareto efficient rules is the largest group among them. According to their definition, we can define the set of ex-post Pareto efficient rule in the class of ordinal rule, but it is smaller than the set of our OPE rules.

[5] See Ehlers et al. (2016) and Börgers and Postl (2009).

[6] The detailed explanation is in Section 6.2.

2. Related literature

There is extensive literature examining voting (decision) rules subject to incentive constraints. The seminal Gibbard–Satterthwaite theorem (Gibbard, 1973 and Satterthwaite, 1975) investigates strategy-proof voting rules, showing that under mild assumptions, a deterministic rule that is strategy-proof is dictatorial. Much of the subsequent literature (Gibbard, 1978; Moulin, 1980; Freixas, 1984, etc.) has either confirmed the robustness of the theorem, or established a positive result by relaxing its assumptions (e.g., allowing randomized rules or cardinal rules, or restricting the domain of preferences). Unfortunately, since strategy-proofness may be too demanding, using preference intensity information does not aid in designing strategy-proof voting rules. Thus, we consider (Bayesian) incentive compatibility which is weaker than strategy-proofness, but still requires agents to be truthful.

Some recent papers adopted a Bayesian mechanism design approach. The usefulness of this approach results from the availability of information about agents' value distributions. Jackson and Sonnenschein (2007) showed that incentive constraints can be overcome by linking decision problems and identifying the cumulative value reports of agents with the value distributions of agents. Barberà and Jackson (2006) derived the optimal weights assigned to representatives based on the value distributions in a model of indirect democracy. We also adopt this approach.

Apesteguia et al. (2011) derived the form of the utilitarian, maximin, and maximax ordinal rules. These rules are the optimal ordinal rules based on different social welfare functions, such as the utilitarian, maximin, and maximax welfare functions. They do not consider incentive compatibility, assuming that agents reveal their true preferences. Thus, a natural question is whether their rules are IC. Our Theorem 1 addresses the issue of incentive compatibility of the set of efficient rules, which includes all of their rules.

Majumdar and Sen (2004) analyzed the implications of weakening incentive constraints from strategy-proofness to Ordinal Bayesian Incentive Compatibility similarly to our approach. They show that under uniform priors, a wide class of rules satisfies Ordinal Bayesian Incentive Compatibility, which is related to our Theorem 1. In contrast to their work, however, we address the relationship between incentive compatibility and Pareto efficiency. We also, allow for randomized rules, and our Theorem 1 does not require their assumptions of neutrality or elementary monotonicity.[7] Furthermore, we characterize the ordinally Pareto ex-post rules (Proposition 1).

Recently, many papers in voting and mechanism design literature have considered cardinal preferences. Apesteguia et al. (2011) and Gershkov et al. (2017) used cardinal utilities to evaluate the welfare under voting rules and Majumdar and Sen (2004) used cardinal utilities to define the concept of their incentive compatibility. However, they restricted their attention to ordinal rules similarly to our Section 5. On the other hand, several papers, similar to our Section 6 and 7, have attempted to investigate mechanisms that take into account cardinal preferences and improve ordinal mechanisms. Hortala-Vallve (2012) proposed qualitative voting on the simultaneous-binary decision on two issues. Casella (2005) suggested storable votes on dynamic-binary decisions on the identical issue. Börgers and Postl (2009) numerically found the second-best rule with two agents, and three alternatives, fixing oppositely ordinal preferences.[8] Miralles (2012) found an ex-ante utilitarian IC allocation rule with two ex-ante symmetric objects. Abdulkadiroğlu et al. (2011) argued that Boston Mechanism, which uses cardinal preferences on tie cases, could perform better than Deferred Acceptance.[9]

Ehlers et al. (2016) recently showed that every cardinal Bayesian incentive compatible mechanism satisfying uc-continuity property must be ordinal in terms of interim allocation of probability of alternatives. Their concept of uc-continuity requires the interim allocation of the probability of alternatives under the mechanism to be uniformly continuous if an agent varies his cardinal preferences without changing his ordinal preferences over alternatives. We can interpret their result such that if we want to design a mechanism which can utilize the information about the intensity of preferences, we need to look at mechanisms that are not uc-continuous. It is consistent with that the rule we suggest in Section 5 is not uc-continuous.

This paper is also related to voluminous literature on scoring rules. Myerson (2002) introduced (A, B)-scoring rules in a model of three-candidate elections. Giles and Postl (2014) studied symmetric Bayes Nash equilibria under (A, B)-scoring rules and the welfare implication of the rules. They showed that some (A, B)-scoring rules can increase social welfare relative to some voting rules such as the plurality, negative, Borda count, and approval voting rules. Their result is closely related to our Theorem 2. While they used numerical methods for the welfare comparison with some ordinal rules in a model of three alternatives, we analytically prove the existence of a superior rule to any ordinal rule with at least three alternatives (Theorem 2).

There are several papers that have investigated voting rules only with two alternatives. Azrieli and Kim (2014) showed that there is no incentive compatible rule which outperforms optimal ordinal rules, and they characterized the set of interim incentive efficient rules as the set of weighted majority rules. Schmitz and Tröger (2012) proved that the optimal rule among all anonymous and neutral rules in both symmetric and neutral environments is the majority rule, which is indeed

[7] We discuss this point in Section 5.

[8] The crucial difference compared with our Section 7 is that they fix the preferences, which is known to agents. Our set-up is to allow any ordinal preferences, which are privately observed by agents. Some readers may find it interesting to compare their numerical findings with ours.

[9] Some papers have found an optimal (utilitarian or Pareto efficient) rule subject to incentive compatibility in their set-up. However, to the author's knowledge, no one has yet analytically found the second-best voting rule with at least three alternatives or issues (or objects for allocation rules).

ordinal. However, we showed that the investigation with at least three alternatives leads to greatly different results regarding efficiency and incentive compatibility.

Finally, many papers, including this one, do not allow monetary transfers despite applying the mechanism design approach. For example, Gershkov et al. (2017) dealt with a voting mechanism, Miralles (2012) examined an allocation mechanism, and Carrasco and Fuchs (2009)[10] studied a communication protocol. Monetary transfers are excluded because they are infeasible in many cases for ethical or institutional reasons (e.g., child placement in public schools, organ transplants, collusion in markets, voting).

3. The model

3.1. Environment

Consider a standard Bayesian environment with private values. The set of agents is $N = \{1, 2, ..., n\}$, and the set of alternatives is L with $\mid L \mid = m$.[11] The typical elements of L will be denoted by l, l', l'', and so on. Each agent $i \in N$ has a von Neumann–Morgenstern value (utility) function for alternative l, $v_i^l : L \to \mathbb{R}$. We denote by $v_i = \left(v_i^l\right)_{l \in L} \in \mathbb{R}^m$ the vector of values that agent i assigns to the alternatives. We assume that no agent is ever indifferent between alternatives, that is $v_i^l \neq v_i^{l'}$ for each agent $i \in N$ and for any alternatives $l, l' \in L$.[12] Let $V_i \subseteq \mathbb{R}^m$ denote agent i's value space, and $V = V_1 \times \cdots \times V_n$ be the set of value profiles. The set of orderings of alternatives is T^{ORD}, t denotes an element of T^{ORD}, and τ denotes an element of $\left[T^{ORD}\right]^n$. We define an ordinal type function $t_i : V_i \to T^{ORD}$ such that if $v_i^l > v_i^{l'} > \cdots > v_i^{l''}$, then $t_i(v_i) = ll' \cdots l''$, which shows the ordinal preferences of agent i. We allow any possible ordinal preferences over the alternatives, such that $\mid \{t : t_i(v_i) = t\} = T^{ORD} \mid = m!$. A subscript $-i$ on a vector means that the ith coordinate is excluded.

To capture the preference intensity, it is useful to make a linear order and normalize as follows. For a vector $x \in \mathbb{R}^m$ and an integer $k \in \{1, .., m\}$, we denote by $x^{[k]}$ the k th-highest value among the coordinates of x. Now, let $1 = v_i^{[1]} > v_i^{[2]} > ... > v_i^{[m]} = 0$ for every $i \in N$.[13]The value of the first ranked alternative is normalized to 1 and the value of the lowest to 0, which can be interpreted as the base of the preference intensities. Then, the preference intensities of the intermediate alternatives are measured between 0 and 1.

The information structure is as follows. We assume that the vector of values of agent i is an identical and independent random variable $\hat{v}_i$ across agents with values in V_i. The assumption of strict preference and the normalization establish agent i's value space $V_i = \left\{v_i : 1 = v_i^{[1]} > v_i^{[2]} > ... > v_i^{[m]} = 0\right\}$. We consider the conditional distribution of $v_i^{[k]}$ given type $t \in T^{ORD}$ which is denoted by $\mu^{k,t}$. Let μ^t be the joint distribution of $\hat{v}_i^{[k]}$ for $k = \{1, ..., m\}$ given type $t \in T^{ORD}$. We mainly study a neutral environment where the type distribution and the joint distribution are symmetric across ordinal types, that is

for any $t, t' \in T^{ORD}$,

$$Pr\left[t_i\left(\hat{v}_i\right) = t\right] = Pr\left[t_i\left(\hat{v}_i\right) = t'\right] \text{ and } \mu^t = \mu^{t'}$$

As a consequence, the structure is simplified such that $\mu^t = \mu$ and $Pr\left[t_i\left(\hat{v}_i\right) = t\right] = \frac{1}{m!}$ for every $t \in T^{ORD}$ and for each agent $i \in N$. The above features of the environment are common knowledge among agents. However, the realized value v_i is assumed to be observed only by agent i.

A voting rule is a measurable mapping $f : V \to \triangle(L)$,[14] meaning that randomized rules are allowed. Let F be the set of all rules. For every agent i and rule f, $U_i(f) = \mathbb{E}\left(\hat{v}_i \cdot f\left(\hat{v}\right)\right)$ denotes the expected utility of agent i under the rule f.[15] To illustrate the features of our environment, we present the following example.[16]

Example 1. Consider three alternatives a, b, and c, implying that $L = \{a, b, c\}$, $\hat{v}_i = (\hat{v}_i^a, \hat{v}_i^b, \hat{v}_i^c)$, and $T^{ORD} = \{abc, acb, bac, bca, cab, cba\}$. Suppose $t_i(v_i) = abc$. Then, the normalization implies that $\hat{v}_i^a = 1$, $\hat{v}_i^c = 0$ and $0 < \hat{v}_i^b < 1$. In a neutral

[10] Their procedure to derive the optimal communication protocol has a similar flavor to our Theorem 2 in the sense that each type is divided into two types. However, unlike our paper, they cover two players, infinite action spaces and, more importantly, develop a dynamic mechanism.

[11] For every finite set X, $\mid X \mid$ denotes the number of elements in X.

[12] This assumption is not essential for our results, but simplifies some definitions and notations.

[13] Our results still hold without this normalization except Proposition 2. Without the normalization, the superior rule we suggest may not be an $(A, B)-$scoring rule in a certain environment. It implies that $(A, B)-$scoring rules are not sufficient to find an IC rule superior to any ordinal rule. Then, our suggested rule may have more originality in terms of the structure of rules. However, we think Proposition 2 is meaningful with the normalization because it shows the clear connection between the rule and $(A, B)-$scoring rules, which are simple and widely used in practice. Since the normalization also simplifies some arguments, we choose to use it. The similar normalization appears in Börgers and Postl (2009). On the other hand, it is arguable whether this normalization can capture the preference intensity and validate the interpersonal comparisons of utility. In fact, the issue of cardinal utility and interpersonal comparisons of utility has been controversial in the literature (for example, see Strotz, 1953; Harsanyi, 1955; Roberts, 1980). This is beyond the scope of our paper.

[14] For every finite set X, $\triangle(X)$ denotes the set of probability measures on X.

[15] $x \cdot y$ denotes the inner product of the vector x and y.

[16] The simplest case with three alternatives will be frequently used to help readers understand the structure and intuitions of the model.

environment, $Pr\left[t_i\left(\hat{v}_i\right)=t\right]=\frac{1}{6}$. For example, if $f(v)=(\frac{1}{2},\frac{1}{2},0)$, then at value profile v, alternatives a and b are each chosen with probability $\frac{1}{2}$ and c is never chosen.

In a neutral environment, it is convenient to define a permutation function and a neutral rule. Let $\sigma : L \to L$ be a permutation of L, and let ϕ be the set of all permutations with respect to L. Then, for $x=\left(x^l\right)_{l\in L}\in\mathbb{R}^m$ and $y=\left(\left(y_i^l\right)_{l\in L}\right)_{i\in N}\in\mathbb{R}^{m\times n}$, we denote by x^σ the vector $\left(x^{\sigma(l)}\right)_{l\in L}\in\mathbb{R}^m$ and by y^σ the vector $\left(\left(y_i^{\sigma(l)}\right)_{l\in L}\right)_{i\in N}\in\mathbb{R}^{m\times n}$.

Definition 1. A rule f is *neutral* if for any profile $v\in V$ and any permutation $\sigma\in\phi$, we have $f\left(v^\sigma\right)=f(v)^\sigma$.

A neutral rule treats all alternatives symmetrically.

Finally in this section, we introduce the standard definition of incentive compatibility from the mechanism design literature:

Definition 2. A rule f is *Incentive Compatible (IC)* if truth-telling is a Bayesian equilibrium of the direct revelation mechanism associated with f. In our environment, this means that for all $i\in N$ and all $v_i, v_i'\in V_i$,

$$v_i\cdot\left(\mathbb{E}\left(f(v_i,\hat{v}_{-i})\right)-\mathbb{E}\left(f(v_i',\hat{v}_{-i})\right)\right)\geq 0.$$

3.2. Ordinal rules and scoring rules

An ordinal rule acts only on the ordinal preference information in the reported value profile. To define an ordinal rule, let P_i^{ORD} be the ordinal partition of V_i into $|\,T^{ORD}\,|$ sets. That is,

for $t\in T^{ORD}$

$$V_i^t=\{v_i\in V_i\mid t_i(v_i)=t\}$$

The partition P_i^{ORD} reflects the ordinal types of agent i. Let $P^{ORD}=P_1^{ORD}\times\cdots\times P_n^{ORD}$ be the corresponding product partition of V where v and v' are in the same element of P^{ORD} if and only if v_i and v_i' are in the same element of P_i^{ORD} for every agent $i\in N$. Let $P^{ORD}(v)$ be the element of P^{ORD} that contains the value profile v.

Definition 3. A rule f is *ordinal* if it is P^{ORD}-measurable, i.e., if $f(v)=f\left(v'\right)$ whenever $P^{ORD}(v)=P^{ORD}\left(v'\right)$. The set of all ordinal rules is denoted by F^{ORD}.

Thus, an ordinal rule depends only on the ranking information in the reported value profile, and is unaffected by changes in the expressed intensity of preferences. We define a useful function related to the ranking of alternatives: for $i\in N$, let $r_l : V_i\to\{1,\ldots,m\}$ denote the ranking of alternative l in v_i.

Next, we consider an important class of ordinal rules: *scoring* rules. Let $s\equiv\left(s^1,\ldots,s^m\right)$ be a score vector in $\mathbb{R}^m$ where $s^1>\cdots>s^m$.

Definition 4. A rule f is a *scoring rule* if there exists a score vector $s\in\mathbb{R}^m$ such that

$$Supp\,(f(v))\subseteq \underset{l\in L}{argmax}\sum_{i\in N}s^{r_l(v_i)},$$

where $Supp\,(f(v))=\left\{l\in L : f(v)^l>0\right\}$.

We can regard $s^{r_l(v_i)}$ as agent i's score assigned to alternative $l\in L$ when he announces v_i. Note that it depends only on the ranking of the alternative, not the specific ordinal type $t_i(v_i)$.

4. A motivating example

The objective of this section is to give a simple example with two agents and three alternatives, that will assist the reader to understand our motivation, main intuition and notation.

Suppose $N=\{1,2\}$, $L=\{a,b,c\}$ and $v_i^{[2]}$ is uniformly distributed in $(0,1)$ for $i=1,2$. We begin with a non-neutral environment where alternatives b and c are popular: For $i\in N$, $Pr\left[t_i\left(\hat{v}_i\right)=t\right]=\epsilon$ for $t=abc, acb, bac, cab$ and $Pr\left[t_i\left(\hat{v}_i\right)=t\right]=\frac{1-4\epsilon}{2}$ for $t=bca, cba$. We look at rules widely used in practice, the plurality and Borda count voting rules denoted by f_P

Table 1
The plurality and Borda count voting rules when $v_1 \in V_1^{abc}$ and $v_1 \in V_1^{bac}$.

	$v_1 \in V_1^{abc}$		$v_1 \in V_1^{bac}$	
	$f_P(v_1, v_2)$	$f_B(v_1, v_2)$	$f_P(v_1, v_2)$	$f_B(v_1, v_2)$
$v_2 \in V_2^{abc}$	$(1, 0, 0)$	$(1, 0, 0)$	$\left(\frac{1}{2}, \frac{1}{2}, 0\right)$	$\left(\frac{1}{2}, \frac{1}{2}, 0\right)$
$v_2 \in V_2^{acb}$	$(1, 0, 0)$	$(1, 0, 0)$	$\left(\frac{1}{2}, \frac{1}{2}, 0\right)$	$(1, 0, 0)$
$v_2 \in V_2^{bac}$	$\left(\frac{1}{2}, \frac{1}{2}, 0\right)$	$\left(\frac{1}{2}, \frac{1}{2}, 0\right)$	$(0, 1, 0)$	$(0, 1, 0)$
$v_2 \in V_2^{bca}$	$\left(\frac{1}{2}, \frac{1}{2}, 0\right)$	$(0, 1, 0)$	$(0, 1, 0)$	$(0, 1, 0)$
$v_2 \in V_2^{cab}$	$\left(\frac{1}{2}, 0, \frac{1}{2}\right)$	$(1, 0, 0)$	$\left(0, \frac{1}{2}, \frac{1}{2}\right)$	$\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$
$v_2 \in V_2^{cba}$	$\left(\frac{1}{2}, 0, \frac{1}{2}\right)$	$\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$	$\left(0, \frac{1}{2}, \frac{1}{2}\right)$	$(0, 1, 0)$

Table 2
The plurality with the standard tie-breaking rule and non-neutral tie-breaking rule. (For interpretation of the references to color in this table, the reader is referred to the web version of this article.)

	$v_1 \in V_1^{abc}$		$v_1 \in V_1^{bac}$	
	$f_P(v_1, v_2)$	$f_{P'}(v_1, v_2)$	$f_P(v_1, v_2)$	$f_{P'}(v_1, v_2)$
$v_2 \in V_2^{abc}$	$(1, 0, 0)$	$(1, 0, 0)$	$\left(\frac{1}{2}, \frac{1}{2}, 0\right)$	$(1, 0, 0)$
$v_2 \in V_2^{acb}$	$(1, 0, 0)$	$(1, 0, 0)$	$\left(\frac{1}{2}, \frac{1}{2}, 0\right)$	$(1, 0, 0)$
$v_2 \in V_2^{bac}$	$\left(\frac{1}{2}, \frac{1}{2}, 0\right)$	$(1, 0, 0)$	$(0, 1, 0)$	$(0, 1, 0)$
$v_2 \in V_2^{bca}$	$\left(\frac{1}{2}, \frac{1}{2}, 0\right)$	$(1, 0, 0)$	$(0, 1, 0)$	$(0, 1, 0)$
$v_2 \in V_2^{cab}$	$\left(\frac{1}{2}, 0, \frac{1}{2}\right)$	$(0, 0, 1)$	$\left(0, \frac{1}{2}, \frac{1}{2}\right)$	$(0, 1, 0)$
$v_2 \in V_2^{cba}$	$\left(\frac{1}{2}, 0, \frac{1}{2}\right)$	$(0, 0, 1)$	$\left(0, \frac{1}{2}, \frac{1}{2}\right)$	$(0, 1, 0)$

and f_B.[17] In our setting, f_P is a scoring rule with $s_P = (1, 0, 0)$ and f_B is a scoring rule with $s_B = \left(1, \frac{1}{2}, 0\right)$. For now, we assume the standard tie-breaking rule that ties are broken by the uniform distribution over the set of alternatives with the greatest sum of scores. Table 1 shows the comparison of the two rules and the usual trade-off between efficiency and incentive compatibility even in the class of ordinal rules.

In the array, agent 1's value announcements appear along the rows and agent 2's along the columns. Within, $f_P(v_1, v_2)$ and $f_B(v_1, v_2)$ are laid out according to agents' announcements.

The voting rule f_P is IC. Table 1 indicates that an agent's misrepresentations cannot increase the probability that the top alternative is chosen and cannot decrease the probability that the bottom one is chosen. f_P is less efficient than f_B in terms of the total expected utilities, i.e., $U_1(f_P) + U_2(f_P) < U_1(f_B) + U_2(f_B)$. In fact, as shown in the next section, f_B is the optimal ordinal rule based upon maximizing the total expected utilities. This is because the score vector of the Borda count voting rule reflects the voters' expected utilities given their ordinal types, i.e., $s_B = \mathbb{E}\left(1, \hat{v}_i^{[2]}, 0 \mid t_i(v_i)\right) = \left(1, \frac{1}{2}, 0\right)$ for $i = 1, 2$. However, f_B is not IC because an agent's misrepresentation can decrease the probability that the bottom alternative is selected (see the case when $v_1 \in V_1^{abc}$ and $v_2 \in V_2^{cba}$ and recall $Pr\left[t_2(\hat{v}_2) = cba\right] = \frac{1-4\epsilon}{2}$). When $v_1 = (1, 1-\delta, 0)$ with sufficiently small δ, agent 1 has an incentive to lie that his type is *bac*, which is an example of the usual trade-off between efficiency and incentive compatibility in the class of ordinal rules.

The first question is in what kind of environment this trade-off may not exist. We observe, from this example, that there is an incentive to lie when some alternatives are popular ex-ante. Thus, we consider a neutral environment in which no alternative is special *ex-ante*. Only one condition is changed such that $Pr\left[t_i(\hat{v}_i) = t\right] = \frac{1}{6}$ for $t \in T^{ORD}$ and $i = 1, 2$. Then, the agent expects that the probability of each alternative getting the score 1 is the same given agent 2's truthful reports. Agent 1's truthful report, thus, is better than any other report. Since the same argument can be applied to agent 2, f_B is IC.

One may think that the neutral environment is the only driving force for the incentive compatibility and thus ask if any ordinal rule is IC in a neutral environment. We can easily see that this is not the case. For example, consider a rule $g(v_1, v_2) = (0, 0, 1)$ if $v_1 \in V_1^{abc}$, otherwise $g = f_P$. The rule g is obviously ordinal, but is not IC. One may observe that g is not a scoring rule, and so ask if any scoring rule, a proper subset of the ordinal rules, is IC. However, even a scoring rule is not necessarily IC in a neutral environment (see Table 2).

The rule $f_{P'}$ is a plurality voting rule with the non-neutral tie-breaking rule that alternative a is chosen when a and b are tied, alternative b is chosen when b and c are tied, and the alternative c is chosen when c and a are tied. The cases

[17] The plurality voting rule allows each agent to vote for one alternative and chooses the alternative with the most votes. The Borda count voting rule allows each agent to give the alternatives a certain number of points corresponding to their ranking of the alternatives, and determines the alternative with the most points.

Table 3
f_B and f^* when $v_1 \in V_1^{abc^H}$ and $v_1' \in V_1^{abc^L}$. (For interpretation of the references to color in this table, the reader is referred to the web version of this article.)

	$v_1 \in V_1^{abc}$	$v_1 \in V_1^{abc^H}$	$v_1' \in V_1^{abc^L}$
	$f_B(v_1, v_2)$	$f^*(v_1, v_2)$	$f^*(v_1', v_2)$
$v_2 \in V_2^{abc^H}$	$(1, 0, 0)$	$(1, 0, 0)$	$(1, 0, 0)$
$v_2 \in V_2^{abc^L}$	$(1, 0, 0)$	$(1, 0, 0)$	$(1, 0, 0)$
$v_2 \in V_2^{acb^H}$	$(1, 0, 0)$	$(1, 0, 0)$	$(1, 0, 0)$
$v_2 \in V_2^{acb^L}$	$(1, 0, 0)$	$(1, 0, 0)$	$(1, 0, 0)$
$v_2 \in V_2^{bac^H}$	$(\frac{1}{2}, \frac{1}{2}, 0)$	$(\frac{1}{2}, \frac{1}{2}, 0)$	$(1, 0, 0)$
$v_2 \in V_2^{bac^L}$	$(\frac{1}{2}, \frac{1}{2}, 0)$	$(0, 1, 0)$	$(\frac{1}{2}, \frac{1}{2}, 0)$
$v_2 \in V_2^{bca^H}$	$(0, 1, 0)$	$(0, 1, 0)$	$(0, 1, 0)$
$v_2 \in V_2^{bca^L}$	$(0, 1, 0)$	$(0, 1, 0)$	$(0, 1, 0)$
$v_2 \in V_2^{cab^H}$	$(1, 0, 0)$	$(1, 0, 0)$	$(1, 0, 0)$
$v_2 \in V_2^{cab^L}$	$(1, 0, 0)$	$(1, 0, 0)$	$(1, 0, 0)$
$v_2 \in V_2^{cba^H}$	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	$(0, 1, 0)$	$(0, 1, 0)$
$v_2 \in V_2^{cba^L}$	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	$(0, 1, 0)$	$(\frac{1}{2}, 0, \frac{1}{2})$

where $f_{P'}(v) \neq f_P(v)$ are in red. Since we do not restrict tie-breaking rules in the definition of scoring rules, the rule $f_{P'}$ is a scoring rule. Similarly, when $v_1 = (1, 1-\delta, 0)$ with sufficiently small δ, agent 1 has an incentive to lie that his type is *bac*. Thus, $f_{P'}$ is not IC. Even in this case, we show in the next section that it is possible to construct an IC rule g such that $U_i(g) = U_i(f_{P'})$ for every i.

The second question is whether we can design an IC rule superior to f_B (recall that f_B is the optimal ordinal rule based upon maximizing the total expected utilities.) We will show the exact construction of an IC rule superior to the Borda Count rule, using preference intensity information. We consider a rule based on a finer partition than the ordinal partition. The partition divides each ordinal type set V_i^t into two sets $V_i^{t^H}$ and $V_i^{t^L}$ using a threshold $\beta \in (0, 1)$:

$$V_i^{t^H}(\beta) = \{v_i \in V^t \mid v_i^{[2]} \geq \beta\} \text{ and } V_i^{t^L}(\beta) = \{v_i \in V^t \mid v_i^{[2]} < \beta\}.$$

The H type and L type are divided according to the value of the second-ranked alternative. For each $\beta \in (0, 1)$, we find the rule f_β which maximizes the sum of expected utilities among these partition-measurable rules. In fact, f_β resembles a scoring rule, but with two score vectors, $s^H(\beta) = \left(1, \mathbb{E}\left(\hat{v}_i^{[2]} \mid v_i \in V^{t^H}\right), 0\right)$ and $s^L(\beta) = \left(1, \mathbb{E}\left(\hat{v}_i^{[2]} \mid v_i \in V^{t^L}\right), 0\right)$. Since we use the ranking information as well as the preference intensity information with this partition, each f_β is more efficient than any ordinal rules. However, we need to find a specific β for incentive compatibility. Given the environment of this example, we find the IC rule $f^* = f_{\beta^*}$ with $\beta^* = \frac{1}{\sqrt{2}}$, which implies that $s^H = \left(1, \frac{1+\sqrt{2}}{2\sqrt{2}}, 0\right)$ and $s^L = \left(1, \frac{1}{2\sqrt{2}}, 0\right)$. To help understand the intuition, we provide the following table showing f_B and f^* when agent 1 announces an H type and an L type (see Table 3).

For the efficiency of f^* over f_B, note the case when $v_2 \in V_2^{bac}$ (when the top two alternatives are shared), or $v_2 \in V_2^{cba}$ (when the two agents have the opposite ordinal preferences). These cases are in blue. While f^* can distinguish these cases based on preference intensity information, the ordinal rule f_B cannot. We verify the efficiency by $U_1(f^*) + U_2(f^*) = \frac{117}{72} > U_1(f_B) + U_2(f_B) = \frac{114}{72}$.

For the incentive compatibility of f^*, note that an H type announcement by agent 1 gives him a relatively high probability of getting the middle alternative and an L type announcement has the opposite effect.[18] Consequently, most H and L type agents like to report their true types. However, to satisfy the incentive compatibility constraint, an agent with any second-ranked alternative's value should be truth-telling. We look at the agent with the second-ranked alternative's value equal to β, called the cut-off agent. We find that under the rule $f^* = f_{\frac{1}{\sqrt{2}}}$ this agent is indifferent between an H type and an L type announcement, i.e., $\left(1, \frac{1}{\sqrt{2}}, 0\right) \cdot \left(\mathbb{E}\left(f^*(v_i, \hat{v}_{-i})\right) - \mathbb{E}\left(f^*(v_i', \hat{v}_{-i})\right)\right) = 0$. Then, every agent with any second-ranked alternative's value is truth-telling. That is, f^* is IC.[19]

5. Ordinal Paretian ex-post and incentive compatibility

In this section, we restrict our attention to ordinal rules for the following reasons. First, ordinal rules are widely used in theory and practice. More importantly, deriving results within the class of ordinal rules will serve as a benchmark when we explore the relationship between efficiency and incentive compatibility.

[18] Precisely, an H type announcement also gives him a low probability of getting the top alternative (a loss), but gives him a low probability of getting the bottom alternative (a gain).
[19] The detailed arguments are in the proof of Theorem 2.

Our primary objective for this section is to show that there is no conflict between any Pareto efficiency and incentive compatibility in the class of ordinal rules. We consider Paretian ex-post as our notion of efficiency because, we believe, it is the extremely weak Pareto efficiency and the minimal normative criterion that a social planner should take into account. Let us define Paretian ex-post rules in the class of ordinal rules.

Definition 5. A rule $f \in F^{ORD}$ is *Ordinally Paretian Ex-post (OPE)* if when there exists an alternative $l \neq l'$ such that $r_l(v_i) > r_{l'}(v_i)$ for every agent $i \in N$ and for $v \in V$, then $f^{l'}(v) = 0$.

Next, we consider a *generalized scoring* rule. Simply, it has a score vector for each agent for each type profile. Let $s_{i,\tau} \equiv \left(s^1_{i,\tau}, ..., s^m_{i,\tau}\right)$ be a score vector in $\mathbb{R}^m$ for $i \in N$ and for $\tau \in \left[T^{ORD}\right]^n$ where $s^1_{i,\tau} > ... > s^m_{i,\tau}$.

Definition 6. A rule f is a *generalized scoring rule* if there exists a score vector $s_{i,\tau}$ for each $i \in N$ for each $\tau \in \left[T^{ORD}\right]^n$ such that

$$Supp\,(f(v)) \subseteq \underset{l \in L}{argmax} \sum_{i \in N} s^{r_l(v_i)}_{i,\tau},$$

where $Supp\,(f\,(v)) = \left\{l \in L : f\,(v)^l > 0\right\}$.

Proposition 1. *A rule $f \in F^{ORD}$ is OPE if and only if there exist generalized scoring rules $\{g^l\}_{l\in L}$ and numbers $\{\lambda^l_\tau\}_{l\in L}$, where $\lambda^l_\tau \geq 0$ and $\sum_{l\in L} \lambda^l_\tau = 1$ for each $\tau \in \left[T^{ORD}\right]^n$ such that $f\,(v) = \sum_{l\in L} \lambda^l_\tau g^l\,(v)$.*

Proposition 1 characterizes the OPE rules. Essentially, an OPE rule is a randomization of generalized scoring rules at each ordinal type profile. Now, we present the main result of this section.

Theorem 1. *If a rule f is OPE, then there exists an IC $g \in F^{ORD}$ such that $U_i\,(g) = U_i\,(f)$ for every agent $i \in N$.*[20]

Theorem 1 implies that the usual trade-off between ex-post Pareto efficiency and incentive compatibility within the class of ordinal rules does not exist, at least in a neutral environment. It also has an implication for the voting rules based upon maximizing different welfare functions. Apesteguia et al. (2011) described the optimal ordinal rules based on the utilitarian, maximax, and maximin welfare functions. The maximin welfare function evaluates an alternative in terms of the expected utility of the worst-off agent, disregarding the other agents' expected utilities. In contrast to the maximin welfare function, the maximax welfare function concentrates on the most well-off agent. They abstract the strategic voting, assuming voters' truthful announcements of their preferences. However, Theorem 1 indicates that we can design IC rules which deliver the same expected utilities as their rules since their rules are OPE.

Another contribution in the literature regarding incentive compatibility is the following: Majumdar and Sen (2004) showed that a voting rule that satisfies neutrality and elementary monotonicity (roughly, if a particular alternative is chosen by the voting rule given a profile of reports, it should be chosen by the rule when the alternative moves up in any report of an agent given others' reports) is ordinally Bayesian IC. Due to Proposition 1, we find that an OPE rule is a randomization of generalized scoring rules. Thus, one may ask if any OPE rules satisfy elementary monotonicity. However the following example shows that the answer is negative.

Example 2. Consider $L = \{a, b, c\}$ and $N = \{1, 2\}$. Consider a generalized scoring rule g with $s_{i,\tau} = (1, 0.75, 0)$, for $i \in N$ and $\tau \in T^{ORD} \times T^{ORD}$ but $s_{2,(abc,bca)} = (1, 0.95, 0.9)$. Suppose agent 1 announces $v_1 \in V_1^{abc}$. If agent 2 announces $v_2 \in V_2^{cba}$, alternative b is chosen. However, if agent 2 announces $v_2 \in V_2^{bca}$, then alternative a is chosen. The rule g does not satisfy elementary monotonicity, but g is OPE according to Proposition 1.

Regarding Majumdar and Sen (2004), Theorem 1 provides a method such that even for a rule that does not satisfy neutrality or elementary monotonicity, we can construct an IC rule which delivers the same expected utilities for all agents.

6. A superior incentive compatible cardinal rule

In this section, we move beyond ordinal rules by utilizing preference intensity information. It is straightforward to design a cardinal rule superior to any ordinal rule. For example, the first-best cardinal rule is a rule where a score assigned to an alternative is the realized value of the alternative and the alternative with the greatest sum of scores is chosen. However, the non-trivial question is whether we can design an *incentive compatible* cardinal rule superior to any ordinal rule.

[20] Theorem 1 holds under asymmetric value distributions across agents since they do not affect Proposition 1, Lemma 1, and the condition of incentive compatibility.

6.1. A superior IC cardinal rule

The following theorem shows the main result of this section, answering the above question.

Theorem 2. *For $n \geq 5$, there exists an IC cardinal rule that achieves higher utilitarian social welfare than any ordinal rule.*[21]

The proof consists of six steps. Step 1 shows Lemma 2 which derives a utilitarian rule based on a general partition P. In Step 2, we introduce a special family of partitions $\{P^\beta\}_{\beta\in(0,1)}$ and a utilitarian rule based on P^β, called a P^β-Utilitarian Rule. Step 3 considers three alternatives and shows a necessary and sufficient condition for incentive compatibility of a P^β-Utilitarian rule. In Step 4, we prove the existence of a rule f^* satisfying the condition. Step 5 proves that f^* achieves a higher utilitarian social welfare than any ordinal rule. Finally in step 6, we modify the P^β-Utilitarian rule and show the extension with more than three alternatives.

6.2. (A, B)-scoring rules with three alternatives

In this subsection, we focus on three alternatives, connecting our rule f^* to several well-known rules: the *plurality, negative, Borda count, and approval* voting rules. As in Myerson (2002), the general form of these voting rules for three candidates is an (A, B)-scoring rule, where each voter must choose a score vector that is a permutation of either $(1, B, 0)$ or $(1, A, 0)$. That is, the voter can give a maximum 1 point to one candidate, A or B $(0 \leq A \leq B \leq 1)$ to some other candidate, and a minimum 0 to the remaining candidate. In our environment, we define the rule in the following way.

Definition 7. A rule f is an (A, B)-scoring rule if there exist two score vectors $s \in \{(1, A, 0), (1, B, 0)\}$ such that

$$Supp\,(f(v)) \subseteq \underset{l\in L}{argmax} \sum_{i\in N} s^{r_l(v_i)}$$

The case $(A, B) = (0, 0)$ is the *plurality* voting rule, where each voter can support a single candidate. The case $(A, B) = (1, 1)$ is the *negative* voting rule, where each voter can oppose a single candidate. The case $(A, B) = (0.5, 0.5)$ is the *Borda count* voting rule, where each voter can give candidates a completely ranked score. These rules are classified as ordinal rules because information about ordinal preference is sufficient to implement the rules. However, $(A, B) = (0, 1)$, which is the *approval* voting rule where each voter can support or oppose a group of candidates, requires more than ordinal preference information, similarly to P^β-Utilitarian Rules. The following proposition clearly shows the relationship between P^β-Utilitarian Rules, f^* and (A, B)-scoring rules.

Proposition 2. *A rule is a P^β-Utilitarian Rule if and only if it is an (A, B)-scoring rule with $(A, B) = \left(s^L(\beta)^2, s^H(\beta)^2\right)$. Furthermore, f^* is an IC (A, B)-scoring rule with $(A, B) = \left(s^L(\beta^*)^2, s^H(\beta^*)^2\right)$*[22]

Proposition 2 shows that P^β-Utilitarian Rules are not purely theoretical rules because (A, B)-scoring rules are widely used in practice due to their simple structure. If we want to implement (A, B)-scoring rules, Proposition 2 helps find an IC (A, B)-scoring rule superior to any ordinal rule.

7. Second-best rule in the simplest non-trivial environment

7.1. Second-best rule: a numerical approach

In this subsection, we provide numerical findings about the second-best rule in the simplest non-trivial environment: Two agents, three alternatives, and uniform distribution of $v_i^{[2]}$. We discretize the space of $v_i^{[2]}$ into 200 equally spaced types.[23] Since we have two agents and three alternatives (6 ordinal types), the total type pairs are 1,440,000 (= 200 × 200 × 6 × 6). The numerical problem is to find $f = \left(f^a, f^b, f^c\right)$ in each type to maximize the utilitarian welfare subject to incentive constraints.[24] For a finite type space, this is a linear programming problem. To solve the problem, we use

---

[21] For $n = 2$ or 4, the result holds with a restriction of the value distribution μ that $\mathbb{E}\left(\hat{v}_i^{[2]} \mid P_i^{ORD}\right) \leq \frac{\mathbb{E}\left(\hat{v}_i^{[1]}|P_i^{ORD}\right)+\mathbb{E}\left(\hat{v}_i^{[3]}|P_i^{ORD}\right)}{2}$. The argument is in the appendix.

[22] Recall, $s^H(\beta) = \left(1, \mathbb{E}\left(\hat{v}_i^{[2]} \mid v_i \in V^{t^H}\right), 0\right)$ and $s^L(\beta) = \left(1, \mathbb{E}\left(\hat{v}_i^{[2]} \mid v_i \in V^{t^L}\right), 0\right)$.

[23] This is the maximum for the discretization with our computing facilities.

[24] As in the standard numerical approach, we consider the local incentive constraints: No type can be better off from reporting neighboring types. In our neutral environment, this implies the global incentive constraints.

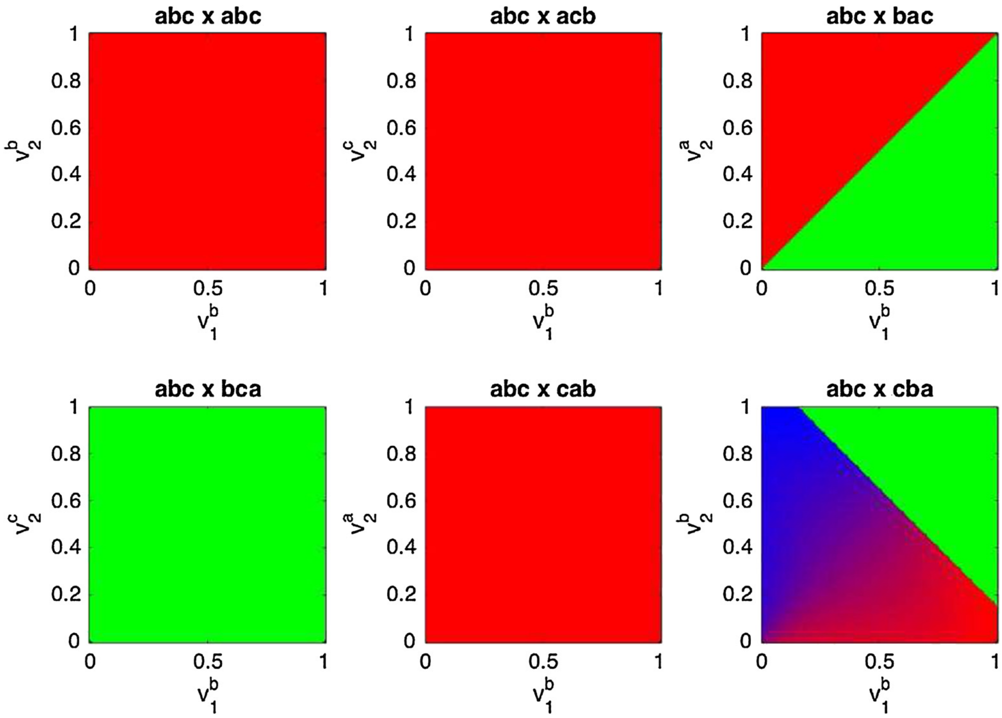

Fig. 1. Second-best rule when $v_1 \in V_1^{abc}$ and $v_2 \in V$. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

the implementation of the interior point algorithm (`'interior-point-legacy'` of `linprog` in MATLAB) available in MATLAB R2015b. Due to the size of the problem, we set the maximum number of iterations to be 200 in the following test results.

Fig. 1 shows the second-best rule when agent 1's ordinal type is abc and agent 2's ordinal types are all six types. The other ordinal types of agent 1 have the symmetric result, so we omit them. In six graphs, the horizontal axis represents the value of the second-ranked alternative for agent 1, $v_1^{[2]}$. The vertical axis represents $v_2^{[2]}$. Each graph has a grid representing the possible 200×200 type pairs. The grid points of $v_i^{[2]}$ are $\frac{1}{201}, \frac{2}{201}, \ldots, \frac{200}{201}$ and the grid lines are suppressed. Each type pair is associated with the square. The color of the square represents the probability that each alternative is chosen under the second-best rule. The red indicates alternative a, the green indicates alternative b, and the blue indicates alternative c. The mixed color represents the random choice of alternatives, which is observed only when two agents show diametrically opposite ordinal preferences.

According to Fig. 1, given $v_1 \in V_1^{abc}$ and for $v_2 \in V_2^{abc}, V_2^{acb}, V_2^{cab}$ the second-best rule chooses alternative a. For $v_2 \in V_2^{bca}$, it chooses alternative b. For $v_2 \in V_2^{bac}$, it chooses alternative a when $v_2^a > v_1^b$, and chooses alternative b otherwise. For $v_2 \in V_2^{cba}$, the rule assigns the probability 1 to alternative b when the types are above the diagonal connecting the points $\left(\frac{33}{201}, \frac{200}{201}\right)$ and $\left(\frac{200}{201}, \frac{33}{201}\right)$. Below the diagonal, it assigns mixed probability over alternatives. It assigns extremely small probability to alternative b until the type of one agent is $\frac{31}{201}$ and zero probability when the type is less than $\frac{31}{201}$.

We can compare Fig. 1 with the first-best rule. We find that the second-best rule is identical to the first-best rule except in the case when two agents show the opposite ordinal preferences (abc x cba in Fig. 1). This means that the efficiency loss from the incentive constraints exists only in that case. Fig. 2 shows the first-best rule when $v_1 \in V_1^{abc}$ and $v_2 \in V^{bac}, V^{cba}$.

The surprising aspects of the second-best rule compared with the first-best rule are two-fold. First, the second-best rule has no efficiency loss when $v_1 \in V_1^{abc}$ and $v_2 \in V^{bac}$, which is not observed in our proposed f^* and the ordinal utilitarian rule f_{OUR}. Second, the probabilities that alternative a and b are chosen, respectively, are increasing when v_1^b approaches 1 in the case of abc x cba. If we consider only this case, this pattern is inconsistent with incentive compatibility.[25] In fact, these two aspects are related. Since the probability that alternative a is chosen is decreasing as v_1^b increases when $v_2 \in V_2^{bac}$ although the probability is increasing when $v_2 \in V_2^{cba}$, incentive compatibility can hold. This is similar to the logic we find in Section 6 that the H type agent obtains the probability that alternative b is chosen, but loses the probability that alternative a is chosen relative to the L type agent under the IC rule, f^*.

[25] This aspect is fairly different from the numerical finding in Börgers and Postl (2009) since they consider only the case of abc x cba.

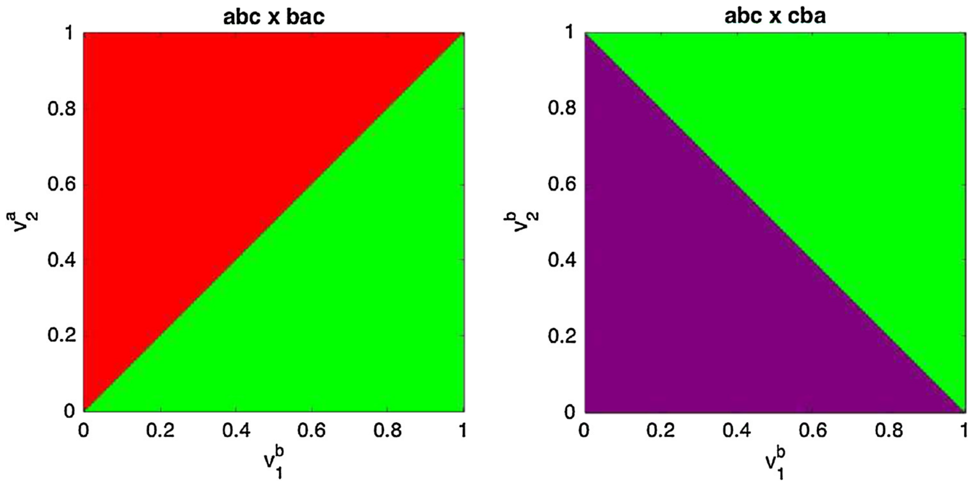

Fig. 2. First-best rule when $v_1 \in V_1^{abc}$ and $v_2 \in V^{bac}, V^{cba}$, respectively.

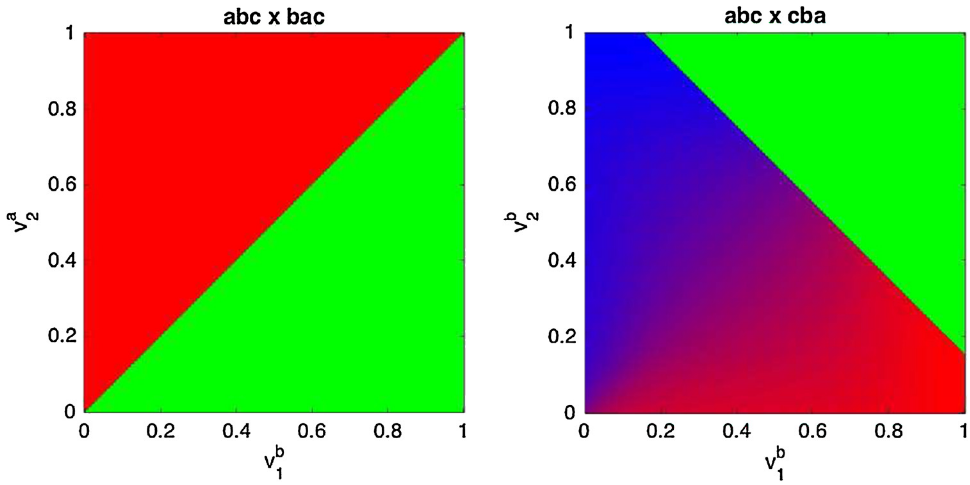

Fig. 3. f_{SB} when $v_1 \in V_1^{abc}$ and $v_2 \in V^{bac}, V^{cba}$ (500 × 500 grids).

7.2. Welfare comparison

We calculate the utilitarian welfare of the first-best rule denoted by f_{FB}, the second-best denoted by f_{SB}, and ordinal utilitarian rule f_{OUR}.

To calculate the welfare of f_{SB} more accurately, we fix four cases as f_{FB} excluding the cases of abc x bac and abc x cba to decrease the burden of computation. Then, we can discretize the space of $v_i^{[2]}$ into 500 equally spaced types.

In Fig. 3, the end points on the diagonal in abc x cba case are changed from $\left(\frac{33}{201}, \frac{200}{201}\right)$ and $\left(\frac{200}{201}, \frac{33}{201}\right)$ to $\left(\frac{80}{501}, \frac{500}{501}\right)$ and $\left(\frac{500}{501}, \frac{80}{501}\right)$, respectively. The welfare of f_{SB} is approximately 1.637. The welfare of f_{FB} is $\frac{59}{36}$, approximately 1.639.

Fig. 4 shows f_{OUR} when $v_1 \in V_1^{abc}$ and $v_2 \in V^{bac}, V^{cba}$. Given the current environment, f_{OUR} is Borda-count voting rule with the score $\left(1, \frac{1}{2}, 0\right)$. We assume the standard tie-breaking rule. Since this rule shows the difference compared to the first-best rule only when $v_2 \in V^{bac}, V^{cba}$ given $v_1 \in V_1^{abc}$, we omit other symmetric cases. The following table shows the welfare comparison of the above three rules.

Table 4 helps us measure the welfare loss for incentive constraints and restriction of ordinal rules. The loss for incentive constraints is 0.12% $\left(= \left(1 - \frac{W(f_{SB})}{W(f_{FB})}\right) \times 100\right)$ among all cardinal rules. Among incentive compatible rules, the loss for the restriction of ordinal rules is 3.27% $\left(= \left(1 - \frac{W(f_{OUR})}{W(f_{SB})}\right) \times 100\right)$. Note that for the latter loss, we can use f_{OUR} due to our Theorem 1. Thus, we observe that the loss for incentive constraints is much smaller than the loss for the restriction of ordinal rules.

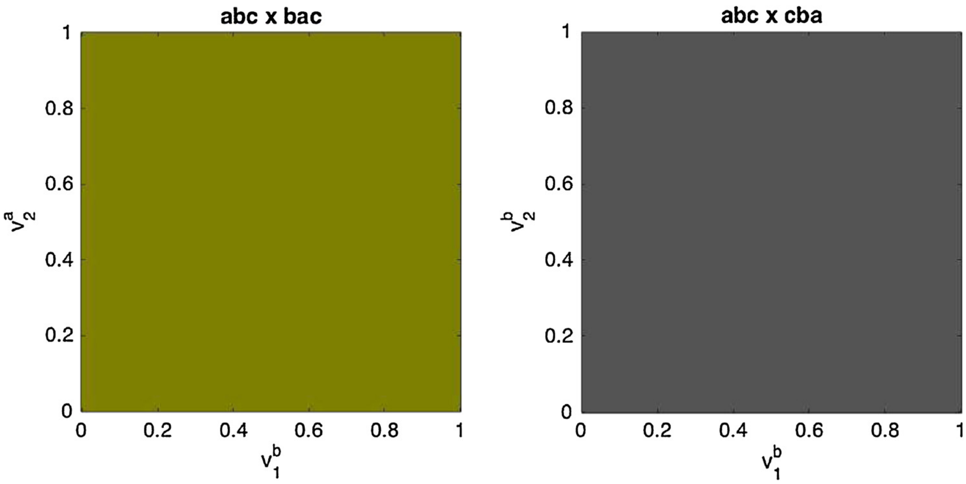

Fig. 4. f_{OUR} when $v_1 \in V_1^{abc}$ and $v_2 \in V^{bac}, V^{cba}$.

Table 4
Welfare and loss of rules relative to f_{FB} and f_{SB}.

Welfare, $W(f)$	f_{FB}	f_{SB}	f_{OUR}
	$\frac{59}{36}$	$\frac{58.932}{36}$	$\frac{57}{36}$
Loss relative to First-best rule, $\left(1 - \frac{W(f)}{W(f_{FB})}\right) \times 100$	0%	0.12%	3.39%
Loss relative to Second-best rule, $\left(1 - \frac{W(f)}{W(f_{SB})}\right) \times 100$		0%	3.27%

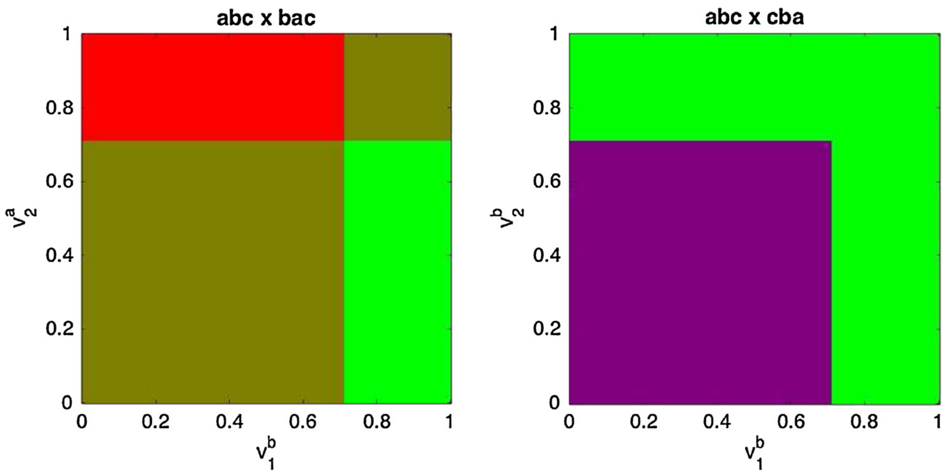

Fig. 5. f^* when $v_1 \in V_1^{abc}$ and $v_2 \in V^{bac}, V^{cba}$.

Remark. We also provide the welfare of the rule we suggest in Section 6 f^*, which is the (A, B)-scoring rule with $\left(\frac{\sqrt{2}}{2}, \frac{1+\sqrt{2}}{2}\right)$ by Proposition 2.

Fig. 5 displays f^* when $v_1 \in V_1^{abc}$ and $v_2 \in V^{bac}, V^{cba}$. As mentioned in the previous subsection, Figs. 4 and 5 inform the efficiency loss when $v_1 \in V_1^{abc}$ and $v_2 \in V^{bac}$ under f_{OUR} and f^*.

The welfare of f^* is $\frac{58.5}{36}$ $(= 1.625)$. The loss under f^* relative to the first-best rule is 0.85% and the loss relative to the second best rule is 0.73%. Relatively, the losses are significantly less than the losses under f_{OUR}.

8. Concluding comments

We investigated the efficiency and incentive compatibility of voting rules in a Bayesian environment with private values and at least three alternatives. First, we proved that in a neutral environment, essentially any ex-post Pareto efficient ordinal rule is IC, which implies that the usual conflict between efficiency and incentive compatibility does not exist in the class of ordinal rules. This conflict, however, arises if we consider cardinal rules. Further, it is not straightforward to design an IC cardinal rule which is more efficient than ordinal rules. However, we successfully constructed an IC cardinal rule superior to any ordinal rule. With three alternatives, this cardinal rule turns out to be an IC (A,B)-scoring rule as in Myerson (2002). Finally, we reported the numerical findings regarding the second-best rule in the simple environment, and the welfare loss for incentive constraints and for the restriction of ordinal rules.

We believe that our paper addresses some of the most basic questions regarding the design of IC voting rules in a Bayesian environment, and that our results are building blocks to finding the analytical characterization of the second-best rule. Further exploration of the second-best rule numerically and analytically in the general environment would be an important direction for research. In addition, the method of using a finer partition and finding a condition for incentive compatibility in the proof of Theorem 2 is novel and worthy of attention. This method may be applicable to a more general distribution of agents' values for designing a superior voting rule or even to different mechanisms such as allocation or matching rules. We hope to study these directions in the future.

Appendix

Proofs in Section 5

Proof of Proposition 1. Since the [If part] is obvious, we omit the proof. [Only if part] Assume f is OPE. Given v and $\tau(v)=\tau$, let $\bar{L}(v)=\left\{l'\in L \mid f(v)^{l'}>0\right\}$. For each $l'\in\bar{L}(v)$, we can construct $g^{l'}(v)$ with $s_{i,\tau}$ such that $s^1_{i,\tau}=1+\epsilon>\cdots>s^{r_{l'}(v_i)}_{i,\tau}=1>s^{r_{l'}(v_i)+1}_{i,\tau}=1-\delta>\cdots$, where $\epsilon,\delta>0$ for $i\in N$ with $1<r_{l'}(v_i)<m$ and ϵ or δ may not exist for some agent i with $r_{l'}(v_i)=1$ or m. Since f is OPE, for every l, there exists an agent j such that $r_l\left(v_j\right)<r_{l'}\left(v_j\right)$. Then, $\sum_{i\in N}s^{r_l(v_i)}_{i,\tau}\le(n-1)(1+\epsilon)+(1-\delta)<n=\sum_{i\in N}s^{r_{l'}(v_i)}_{i,\tau}$ with sufficiently small ϵ. Then, $g^{l'}(v)$ chooses l' with probability 1. For each $l\notin\bar{L}(v)$, we can construct $g^l(v)$ with any $s_{i,\tau}$. With the similar process for each τ, we can construct generalized scoring rules $\{g^l\}_{l\in L}$. Finally, we find $\{\lambda^l_\tau\}_{l\in L}$, where $\lambda^l_\tau=f(v)^l$ such that $f(v)=\sum_{l\in L}\lambda^l_\tau g^l(v)$. □

Proof of Theorem 1. The following lemma is useful for the proof of Theorem 1 and shows that for any rule in a neutral environment, we can construct a neutral rule that delivers the same expected utility for each agent.

Lemma 1. *For any rule $f\in F^{ORD}$, a rule $g(v)=\frac{1}{m!}\sum\limits_{\sigma\in\phi}f\left(v^\sigma\right)^{\sigma^{-1}}\in F^{ORD}$ is neutral. Moreover, $U_i(g)=U_i(f)$ for every agent i.*

Proof. Consider $g(v)=\frac{1}{m!}\sum\limits_{\sigma\in\phi}f\left(v^\sigma\right)^{\sigma^{-1}}$. First, g is neutral. For any $\sigma^*\in\phi$,

$$
\begin{aligned}
g\left(v^{\sigma^*}\right)^{\sigma^{*-1}}&=\left(\frac{1}{m!}\sum_{\sigma\in\phi}f\left(\left(v^{\sigma^*}\right)^{\sigma}\right)^{\sigma^{-1}}\right)^{\sigma^{*-1}}=\frac{1}{m!}\sum_{\sigma\in\phi}\left(f\left(\left(v^{\sigma^*}\right)^{\sigma}\right)^{\sigma^{-1}}\right)^{\sigma^{*-1}}\\
&=\frac{1}{m!}\sum_{\sigma\in\phi}f\left(v^{\sigma^*(\sigma)}\right)^{\sigma^{-1}(\sigma^{*-1})}=\frac{1}{m!}\sum_{\sigma\in\phi}f\left(v^\sigma\right)^{\sigma^{-1}}=g(v),
\end{aligned}
$$

where the second equality follows from the fact that the permutation of an aggregated vector is the same as an aggregation of the individually permutated vectors, and the fourth follows from the fact that $\{\sigma^*(\sigma):\sigma\in\phi\}=\phi$.

Moreover, $U_i(f)=\mathbb{E}\left(\hat{v}_i\cdot f\left(\hat{v}\right)\right)=\frac{1}{m!}\sum\limits_{t\in T}\mathbb{E}\left(\hat{v}_i\mid v_i\in V^t_i\right)\cdot\mathbb{E}\left(f\left(\hat{v}\right)\mid v_i\in V^t_i\right)$, where the second equality follows from the neutral environment which yields the same probability of any ordinal type $\frac{1}{m!}$. To see the connection between f and g, we change the expression of $U_i(f)$ with permutations. Fix $t\in T^{ORD}$,

$$
\begin{aligned}
U_i(f)&=\frac{1}{m!}\sum_{\sigma\in\phi}\mathbb{E}\left(\hat{v}^\sigma_i\mid v_i\in V^t_i\right)\cdot\mathbb{E}\left(f\left(\hat{v}^\sigma\right)\mid v_i\in V^t_i\right)\\
&=\frac{1}{m!}\sum_{\sigma\in\phi}\mathbb{E}\left(\hat{v}^\sigma_i\mid v_i\in V^t_i\right)^{\sigma^{-1}}\cdot\mathbb{E}\left(f\left(\hat{v}^\sigma\right)\mid v_i\in V^t_i\right)^{\sigma^{-1}}\\
&=\mathbb{E}\left(\hat{v}_i\mid v_i\in V^t_i\right)\cdot\frac{1}{m!}\sum_{\sigma\in\phi}\mathbb{E}\left(f\left(\hat{v}^\sigma\right)\mid v_i\in V^t_i\right)^{\sigma^{-1}}
\end{aligned}
$$

where the second equality follows from the fact that for $x, y \in \mathbb{R}^m$ and σ^{-1}, $x^{\sigma^{-1}} \cdot y^{\sigma^{-1}} = x \cdot y$, and the third follows from the neutral environment.

Since this formula holds for any t and any i,

$$\begin{aligned} U_i(f) &= \frac{1}{m!}\sum_{t\in T}\mathbb{E}\left(\hat{v}_i \mid v_i \in V_i^t\right)\cdot\frac{1}{m!}\sum_{\sigma\in\phi}\mathbb{E}\left(f\left(\hat{v}^{\sigma}\right) \mid v_i \in V_i^t\right)^{\sigma^{-1}} \\ &= \frac{1}{m!}\sum_{t\in T}\mathbb{E}\left(\hat{v}_i \mid v_i \in V_i^t\right)\cdot\mathbb{E}\left(\frac{1}{m!}\sum_{\sigma\in\phi} f\left(\hat{v}^{\sigma}\right)^{\sigma^{-1}} \mid v_i \in V_i^t\right) \\ &= U_i(g). \qquad \square \end{aligned}$$

The neutral rule is constructed by assigning $\frac{1}{m!}$ (the probability of an ordinal type) to every inversely permutated original rule where the value profile is permutated.[26]

Assume a rule f is OPE. By Proposition 1 f is a randomization of generalized scoring rules. The rule f may not be neutral because we allow non-neutral tie-breaking rule. By Lemma 1, however, we can construct a neutral rule g from f. This neutralization can also be seen as randomization of generalized scoring rules with scores where the value profile is permutated.

Fix $t \in T^{ORD}$. It is sufficient to consider only $v_i \in V_i^t$ instead of all ordinal types because of the neutrality of the rule and environment. Pick $v_i \in V_i^t$ and $v_i' \in V_i$. To check the incentive compatibility of g, we look at probabilities that alternatives are chosen given i's announcement under g (i.e., $\mathbb{E}\left(g(v_i, \hat{v}_{-i})\right)$ and $\mathbb{E}\left(g(v_i', \hat{v}_{-i})\right)$). First, for any τ and any sub-generalized scoring rules, consider the aggregated scores of the other agents, $\sum_{j\neq i}\left(s_{j,\tau}^{r_l(v_j)}\right)_{l\in L}$. Second, note that the probabilities that each alternative obtains the maximum aggregated scores are the same as $\frac{1}{m}$ and that $\sum_{j\neq i}\left(s_{j,\tau}^{r_l(v_j)}\right)_{l\in L}$ are symmetric for each alternative because the environment and g are neutral. Third, the true announcement adding $\left(s_{i,\tau}^{r_l(v_i)}\right)_{l\in L}$ to the previous scores in every event generates $\mathbb{E}\left(g(v_i, \hat{v}_{-i})\right)$ such that the probability that the first-ranked alternative is chosen becomes the highest and the probability that the second-ranked alternative is chosen becomes the second-highest and so on. Further, since g is neutral, the order of the probabilities resembles the order of the value announcement, i.e., $\mathbb{E}\left(g(v_i', \hat{v}_{-i})\right) = \mathbb{E}\left(g(v_i, \hat{v}_{-i})\right)^{\sigma}$.[27] Thus, the probability distribution $\mathbb{E}\left(g(v_i, \hat{v}_{-i})\right)$ first order stochastically dominates $\mathbb{E}\left(g(v_i, \hat{v}_{-i})\right)^{\sigma}$ for any $\sigma \in \phi$, so that $v_i \cdot \mathbb{E}\left(g(v_i, \hat{v}_{-i})\right) \geq v_i \cdot \mathbb{E}\left(g(v_i', \hat{v}_{-i})\right)$. Thus, g is IC. $\square$

Proofs in Section 6

Proof of Theorem 2.

Step 1) Let $W(g) = \sum_{i\in N} U_i(g)$ be the utilitarian social welfare. Consider any finite measurable partition P_i that divides V_i. Let $P = (P_1 \times \cdots \times P_n)$ be the corresponding partition product of V. Let F^P denote the set of P-measurable rules. Given a partition P, we say that a rule $f \in F^P$ is a P-Utilitarian Rule if $f \in \underset{g\in F^P}{arg\,max}\, W(g)$. The following lemma characterizes P-Utilitarian rules.

Lemma 2. *A rule $f \in F^P$ is a P-Utilitarian Rule if and only if it satisfies*

$$Supp\,(f(v)) \subseteq \underset{l\in L}{argmax}\sum_{i\in N} s^{r_l(v_i)}$$

where

$$s^{r_l(v_i)} = \mathbb{E}\left(\hat{v}_i^l \mid P\right) \; for\ i \in N\ and\ l \in L$$

Proof. For every $g \in F^P$ we have

[26] This lemma is similar to Lemma 3 in Schmitz and Tröger (2012), but it shows that this construction that preserves several properties from the original rule is possible even with more than 2 alternatives.

[27] In Examples 2 with the standard tie-breaking rule, we can calculate that $\mathbb{E}\left(f(v_1, \hat{v}_2)\right) = \left(\frac{2}{3}, \frac{1}{3}, 0\right)$. If agent 1 announces $v_1' \in V_1^{bca}$, then $\mathbb{E}\left(f(v_1', \hat{v}_2)\right) = \left(0, \frac{2}{3}, \frac{1}{3}\right)$. Note the value and the order of coordinates in $\mathbb{E}\left(f(v_1, \hat{v}_2)\right)$ and $\mathbb{E}\left(f(v_1', \hat{v}_2)\right)$.

$$\begin{aligned} W(g) &= \mathbb{E}\left(\sum_{i\in N} \hat{v}_i \cdot g(\hat{v})\right) = \mathbb{E}\left[\mathbb{E}\left(\sum_{i\in N} \hat{v}_i \cdot g(\hat{v}) \middle| P\right)\right] \\ &= \mathbb{E}\left[g(\hat{v}) \cdot \mathbb{E}\left(\sum_{i\in N} \hat{v}_i \middle| P\right)\right] \\ &= \mathbb{E}\left[g(\hat{v}) \cdot \sum_{i\in N} \mathbb{E}\left(\hat{v}_i \middle| P_i\right)\right], \end{aligned}$$

where the third equality follows from the fact that $g \in F^P$ and the fourth from our assumption of independent values across agents. Thus, a rule g is a maximizer of W in F^P if and only if it satisfies

$$Supp\,(g(v)) \subseteq \underset{l\in L}{argmax}\, \mathbb{E}\left(\hat{v}_i^l \middle| P_i(v_i)\right) = \underset{l\in L}{argmax}\, \mathbb{E}\left(\hat{v}_i^{[r_l]} \middle| P_i(v_i)\right).$$

Note $\mathbb{E}\left(\hat{v}_i^{[r_l]} \middle| P_i^{ORD}(v_i)\right)$ can function as the score assigned to alternative l when agent i announces v_i, meaning $s^{r_l(v_i)} = \mathbb{E}\left(\hat{v}_i^{[r_l]} \middle| P_i(v_i)\right)$.[28] $\square$

Step 2) We consider a special family of partitions which provides the preference intensity information as well as the ranking information. Let P_i^β be a partition which divides each ordinal type set in P_i^{ORD} into two sets. It follows that the new type set consists of $2m!$ types, $T = \{t^H, t^L : t \in T^{ORD}\}$. For every $t \in T^{ORD}$, the set V_i^t is partitioned into the two sets $V_i^{t^H}(\beta)$ and $V_i^{t^L}(\beta)$ according to the partition coefficient $\beta \in (0, 1)$.

$$\begin{aligned} V^{t^H}(\beta) &= \{v_i \in V^t \mid v_i^{[2]} \geq \beta v_i^{[1]} + (1-\beta)\, v_i^{[3]}\} = \{v_i \in V^t \mid v_i^{[2]} \geq \beta + (1-\beta)\, v_i^{[3]}\} \\ V_i^{t^L}(\beta) &= \{v_i \in V^t \mid v_i^{[2]} < \beta v_i^{[1]} + (1-\beta)\, v_i^{[3]}\} = \{v_i \in V^t \mid v_i^{[2]} < \beta + (1-\beta)\, v_i^{[3]}\} \end{aligned}$$

where $\beta \in (0, 1)$ is a partition coefficient.

Each ordinal type is divided into H and L types, and an agent i's type is determined by the relative value of the second-ranked alternative $v_i^{[2]}$. An agent i with $v_i \in V^{t^H}$, called an H type agent, values the second-ranked alternative relatively closely to the first-ranked alternative. An agent i with $v_i \in V^{t^L}$, called an L type agent, values the second-ranked alternative relatively closely to the third-ranked alternative.

Let f_β be a P^β-Utilitarian Rule and fix the standard tie-breaking rule.

We mainly consider P^β-measurable rules, so the following type-based notations are convenient. Consider a type function associated with P^β, $t_i^\beta : V_i \to T$ which maps a value vector of agent i to the corresponding type $t \in T$. For example, if $v_i \in V_i^{t^H}(\beta)$, then $t_i^\beta(v_i) = t^H$. Then, we can identify each rule f_β with g_β: $T^n \to \triangle(L)$ by $g_\beta(t_1, ..., t_n) = f_\beta\left(t_1^\beta(v_i), ..., t_n^\beta(v_n)\right)$. There are essentially two score vectors and two type probabilities, as below, because the value distributions across agents are identical and the environment is neutral.

$$\begin{aligned} & s^H(\beta) = \left(\mathbb{E}\left(\hat{v}_i^{[k]} \mid t_i^\beta(v_i) = t^H\right)\right)_{k=\{1,...,m\}}, s^L(\beta) = \left(\mathbb{E}\left(\hat{v}_i^{[k]} \mid t_i^\beta(v_i) = t^L\right)\right)_{k=\{1,...,m\}} \\ & p^H(\beta) = Pr\left(\left\{v_i : t^\beta(v_i) = t^H\right\}\right), p^L(\beta) = Pr\left(\left\{v_i : t^\beta(v_i) = t^L\right\}\right). \end{aligned}$$

Step 3) We first consider the case with three alternatives, which clearly shows how to use preference intensity information. Since g_β is neutral and we assume a neutral environment, it is sufficient to consider only one ordinal type in our examination of incentive compatibility. We fix $t \in T^{ORD}$, $t_i = t^H$, and $t_i' = t^L$. To simplify notation, let $\mathbb{E}\left(g_\beta\left(t, \hat{t}_{-i}\right)\right) = P(\beta)$ and $\mathbb{E}\left(g_\beta\left(t', \hat{t}_{-i}\right)\right) = P(\beta)'$. These are probability vectors in which the coordinates are the probabilities that the alternatives are chosen under the rule g_β given t and t' respectively. Note the following property of $P(\beta)$ and $P(\beta)'$,

$$P(\beta)^{[2]} \geq P(\beta)'^{[2]}, P(\beta)^{[1]} \leq P(\beta)'^{[1]} \text{ and} P(\beta)^{[3]} \leq P(\beta)'^{[3]} \text{ for any } \beta \in (0, 1).$$

In other words, the change of one's announcement from L to H type given the type profile of others weakly increases the probability that the second-ranked alternative is chosen and weakly decreases the probabilities that other alternatives are chosen. This results from a similar process of scoring rules, but with the two score vectors defined such that $s^H(\beta)^{[1]} = s^L(\beta)^{[1]} = 1$, $s^H(\beta)^{[3]} = s^L(\beta)^{[3]} = 0$ and $s^H(\beta)^{[2]} > s^L(\beta)^{[2]}$.

We define the function $h(\beta) = \left(P(\beta)^{[1]} - P(\beta)'^{[1]}\right) + \beta\left(P(\beta)^{[2]} - P(\beta)'^{[2]}\right)$ from the incentive constraints. The following lemma identifies the necessary and sufficient condition for incentive compatibility.

[28] Some parts of this proof are similarly found in the proof of Theorem 1 in Azrieli and Kim (2014).

Lemma 3. *$h(\beta) = 0$ if and only if g_β is IC.*

Proof. Assume $h(\beta) = 0$. First, check the incentive constraint between t and t'.

For $v_i \in V_i^t(\beta)$ and $v_i' \in V_i^{t'}(\beta)$,

$$\begin{aligned}
&v_i \cdot \left(\mathbb{E}\left(g_\beta\left(t, \hat{t}_{-i}\right)\right) - \mathbb{E}\left(g_\beta\left(t', \hat{t}_{-i}\right)\right)\right) \\
&= v_i^{[1]}\left(P(\beta)^{[1]} - P(\beta)'^{[1]}\right) + v_i^{[2]}\left(P(\beta)^{[2]} - P(\beta)'^{[2]}\right) + v_i^{[3]}\left(P(\beta)^{[3]} - P(\beta)'^{[3]}\right) \\
&= \left(1 - v_i^{[3]}\right)\left(P(\beta)^{[1]} - P(\beta)'^{[1]}\right) + \left(v_i^{[2]} - v_i^{[3]}\right)\left(P(\beta)^{[2]} - P(\beta)'^{[2]}\right) \\
&\geq \left(1 - v_i^{[3]}\right)\left(\left(P(\beta)^{[1]} - P(\beta)'^{[1]}\right) + \beta\left(P(\beta)^{[2]} - P(\beta)'^{[2]}\right)\right) \\
&= \left(1 - v_i^{[3]}\right) h(\beta) = 0.
\end{aligned}$$

The inequality comes from the fact that $v_i \in V_i^t(\beta)$ and $P(\beta)^{[2]} - P(\beta)'^{[2]} \geq 0$.

Similarly,

$$\begin{aligned}
&v_i' \cdot \left(\mathbb{E}\left(g_\beta\left(t', \hat{t}_{-i}\right)\right) - \mathbb{E}\left(g_\beta\left(t, \hat{t}_{-i}\right)\right)\right) \\
&\geq \left(1 - v_i'^{[3]}\right)\left(\left(P(\beta)'^{[1]} - P(\beta)^{[1]}\right) + \beta\left(P(\beta)'^{[2]} - P(\beta)^{[2]}\right)\right) \\
&= -\left(1 - v_i'^{[3]}\right) h(\beta) = 0.
\end{aligned}$$

The condition, $h(\beta) = 0$ can be interpreted with the cut-off agent argument. The cut-off agent is the agent with the value of the second-ranked alternative $v_i^{[2]} = \beta + (1-\beta)\, v_i^{[3]}$ which divides the H type and L type sets. If this agent is indifferent between the announcement of an H type and an L type, then every agent is willing to announce his true type.

Second, consider the remaining incentive constraints regarding other type announcements.

For $t_i'' = t^H$, $t_i''' = t^L$ where $t \in T^{ORD}$,

$$\mathbb{E}\left(g_\beta\left(t'', \hat{t}_{-i}\right)\right) = \mathbb{E}\left(g_\beta\left(t, \hat{t}_{-i}\right)\right)^\sigma \text{ for some } \sigma. \text{ Similarly,}$$

$$\mathbb{E}\left(g_\beta\left(t''', \hat{t}_{-i}\right)\right) = \mathbb{E}\left(g_\beta\left(t', \hat{t}_{-i}\right)\right)^\sigma \text{ for some } \sigma.$$

Note, that the value order of coordinates in $P(\beta)$ and $P(\beta)'$ still follows the order of value announcements from the similar argument in the proof of Theorem 1.

By the first order stochastic dominance of $\mathbb{E}\left(g_\beta\left(t, \hat{t}_{-i}\right)\right)$ over $\mathbb{E}\left(g_\beta\left(t'', \hat{t}_{-i}\right)\right)$ and $\mathbb{E}\left(g_\beta\left(t', \hat{t}_{-i}\right)\right)$ over $\mathbb{E}\left(g_\beta\left(t''', \hat{t}_{-i}\right)\right)$, and the above argument between t and t', we have

$$\begin{aligned}
&v_i \cdot \mathbb{E}\left(g_\beta\left(t, \hat{t}_{-i}\right)\right) \geq v_i \cdot \mathbb{E}\left(g_\beta\left(t'', \hat{t}_{-i}\right)\right), \\
&v_i \cdot \mathbb{E}\left(g_\beta\left(t, \hat{t}_{-i}\right)\right) \geq v_i \cdot \mathbb{E}\left(g_\beta\left(t', \hat{t}_{-i}\right)\right) \geq v_i \cdot \mathbb{E}\left(g_\beta\left(t''', \hat{t}_{-i}\right)\right), \text{ and} \\
&v_i' \cdot \mathbb{E}\left(g_\beta\left(t', \hat{t}_{-i}\right)\right) \geq v_i' \cdot \mathbb{E}\left(g_\beta\left(t''', \hat{t}_{-i}\right)\right), \\
&v_i' \cdot \mathbb{E}\left(g_\beta\left(t', \hat{t}_{-i}\right)\right) \geq v_i' \cdot \mathbb{E}\left(g_\beta\left(t, \hat{t}_{-i}\right)\right) \geq v_i' \cdot \mathbb{E}\left(g_\beta\left(t'', \hat{t}_{-i}\right)\right).
\end{aligned}$$

The other direction in the claim is obvious from the first part of this proof, so it is omitted. □

We call $h(\beta)$ the balance function for the rule g_β. That is, because $h(\beta)$ can be arranged such that $h(\beta) = \left(P(\beta)^{[1]} - P(\beta)'^{[1]}\right) + \beta\left(P(\beta)^{[2]} - P(\beta)'^{[2]}\right) = (1-\beta)\left(P(\beta)^{[1]} - P(\beta)'^{[1]}\right) + \beta\left(P(\beta)'^{[3]} - P(\beta)^{[3]}\right)$, it indicates a weighted average of a loss $\left(P(\beta)^{[1]} \leq P(\beta)'^{[1]}\right)$ and a gain $\left(P(\beta)^{[3]} \leq P(\beta)'^{[3]}\right)$ when an agent announces H type rather than L type. This lemma says that the gain and loss are balanced, $h(\beta) = 0$, if and only if $g(\beta)$ is IC.

Step 4) We proceed with three claims regarding the balance function $h(\beta)$ to show the existence of an IC cardinal rule.

Claim 1. $\lim_{\beta \to 0} h(\beta) < 0$ *and* $\lim_{\beta \to 1} h(\beta) > 0$.

Proof. For $\lim_{\beta \to 0} h(\beta) = \lim_{\beta \to 0} P(\beta)^{[1]} - P(\beta)'^{[1]} < 0$, we know that for every case, the change of announcement from H type to L type weakly increases the probability that the first-ranked alternative is chosen. Thus, it is sufficient to find the case where the change strictly increases the probability as β is close to 0. We can always find the case where $g_\beta(t_i, t_{-i}) = (\frac{1}{2}, \frac{1}{2}, 0)$ when every one's type is H type and $g_\beta\left(t_i', t_{-i}\right) = (1, 0, 0)$. For example, for $n = 2$, $t_1 = abc^H$ and $t_2 = bac^H$ and for $n = 3$, $t_1 = abc^H$, $t_2 = bca^H$, and $t_3 = cab^H$. Generally, for any $n = 2$ or $n \geq 4$, we can combine above two profiles to find the case.[29] Since all other types are H types ($\lim_{\beta \to 0} p^H(\beta) = \frac{1}{m!}$), $\lim_{\beta \to 0} P(\beta)^{[1]} - P(\beta)'^{[1]} < 0$.

[29] For example, for $n = 4$, $t_1 = abc^H$, $t_2 = bac^H$, $t_3 = abc^H$ and $t_4 = bac^H$, and for $n = 5$, $t_1 = abc^H$, $t_2 = bca^H$ $t_3 = cab^H$, $t_4 = abc^H$ and $t_5 = bac^H$, and so on.

Next, for $\lim_{\beta\to 1} h(\beta) = \lim_{\beta\to 1} P(\beta)'^{[3]} - P(\beta)^{[3]} > 0$, it is sufficient to find the case where the change of an agent's announcement from L type to H type strictly decreases the probability that the third-ranked alternative is chosen as β is close to 1.[30] We can always find the case where $g_\beta\left(t_i', t_{-i}\right) = (0, \frac{1}{2}, \frac{1}{2})$ or $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ when everyone's type is L type and $g_\beta\left(t_i, t_{-i}\right) = (0, 1, 0)$. For example, for $n = 3$, $t_1' = abc^L$, $t_2 = bca^L$ and $t_3 = cab^L$ and for $n = 5$, the same t_1',t_2,t_3, and $t_4 = bca^L$, $t_5 = cba^L$. Generally, for any $n = 3$ or $n \geq 5$, we can use the above two profiles to find the case. However, we may not find the case when $n = 2$ and $n = 4$. For $n = 2$, the possible case where $g_\beta\left(t_i', t_{-i}\right)^c > 0$ is that $t_1' = abc^L$ and $t_2 = cba^L$. However, in some distributions such that $s^L(\beta)^2 > \frac{1+s^L(\beta)}{2}$ for any $\beta \in (0, 1)$, $g_\beta\left(t_i', t_{-i}\right) = (0, 1, 0) = g_\beta\left(t_i, t_{-i}\right)$. The argument for $n = 4$ is analogous. Since all other types are L types ($\lim_{\beta\to 1} p^L(\beta) = \frac{1}{m!}$), $\lim_{\beta\to 1} P(\beta)'^{[3]} - P(\beta)^{[3]} > 0$. □

With Claim 1, if $h(\beta)$ is continuous on $(0, 1)$, then we can easily find a β^* such that $h(\beta^*) = 0$ by the intermediate value theorem. Then, the new rule f^* is an IC cardinal rule where $f^*(v) = g_{\beta^*}(t_i, t_{-i})$.

However, $h(\beta)$ may be discontinuous in some environments. We provide Examples 3 to explain the potential discontinuity of $h(\beta)$. For ease of notation, we use the vector of scores assigned to the alternatives which depends on β and the announced types, $s(\beta, t_i) = \mathbb{E}\left(\hat{v}_i \mid t_i^\beta(v_i) = t_i\right)$.

Example 3. Given $L = \{a, b, c\}$ and $t_i = abc^H$, we have $g_\beta\left(t_i, t_{-i}\right)$ is determined by the aggregated scores of the alternatives which depends on β and t_{-i}. Let $S^H(\beta, t_{-i}) = \left(\left(s^H(\beta)\right) + \sum_{j\neq i} s_j\left(\beta, t_j\right)\right)$ be the vector of aggregated scores assigned to the alternatives; then, we can express $P(\beta)^{[1]}$ with the function

$$\begin{aligned} P(\beta)^{[1]} = Pr&\left(\left\{t_{-i} : S^H(\beta, t_{-i})^a > S^H(\beta, t_{-i})^b \textit{ and } S^H(\beta, t_{-i})^c\right\}\right) + \\ &\frac{1}{2} Pr\left(\left\{t_{-i} :\ S^H(\beta, t_{-i})^a = S^H(\beta, t_{-i})^b > S^H(\beta, t_{-i})^c\right\}\right) + \\ &\frac{1}{2} Pr\left(\left\{t_{-i} :\ S^H(\beta, t_{-i})^a = S^H(\beta, t_{-i})^c > S^H(\beta, t_{-i})^b\right\}\right) + \\ &\frac{1}{3} Pr\left(\left\{t_{-i} :\ S^H(\beta, t_{-i})^a = S^H(\beta, t_{-i})^b = S^H(\beta, t_{-i})^c\right\}\right). \end{aligned}$$

$P(\beta)^{[2]}$ and $P(\beta)^{[3]}$ can be similarly expressed. $P(\beta)'^{[1]}$, $P(\beta)'^{[2]}$, and $P(\beta)'^{[3]}$ can be expressed using the corresponding $S^L(\beta, t_{-i})$.

Note that g_β is determined by the value order of $S^H(\beta, t_{-i})^a$, $S^H(\beta, t_{-i})^b$, and $S^H(\beta, t_{-i})^c$ at each t_{-i}. If g_β does not change as β changes given any t_{-i}, then $P(\beta)^{[1]}$ is continuous in β because $p^H(\beta)$, $p^L(\beta)$, and $s_i(\beta, t_i)$ are continuous in β. However, g_β could change with β because the value order of the coordinates in $S^H(\beta, t_{-i})$ could change as well. This feature could cause a jump in $P(\beta)^{[1]}$ at some $\beta's$, which means that $P(\beta)^{[1]}$ is possibly discontinuous. The analogous argument can be applied to $P(\beta)$ and $P(\beta)'$. Therefore, $h(\beta)$ may also be discontinuous.

Even in this case, we can construct a slightly different IC rule.

Define the set $D = \{\beta \in (0, 1) : h(\beta)\, \textit{is discontinuous at}\, \beta\}$.

Claim 2. *If D is not empty, then there exits a $\hat{\beta} \in D$ such that* $\underset{\beta\to\hat{\beta}-}{h(\beta)} \cdot \underset{\beta\to\hat{\beta}+}{h(\beta)} \leq 0$.

Proof. Suppose that for all $\hat{\beta} \in D$, $\underset{\beta\to\hat{\beta}-}{h(\beta)} \cdot \underset{\beta\to\hat{\beta}+}{h(\beta)} > 0$. We have two cases.

Case I: There exists $\hat{\beta}_1 < \hat{\beta}_2 \in D$ such $\underset{\beta\to\hat{\beta}_1+}{h(\beta)} \cdot \underset{\beta\to\hat{\beta}_2-}{h(\beta)} < 0$ and $h(\beta)$ is continuous on $\left(\hat{\beta}_1, \hat{\beta}_2\right)$. By the intermediate value theorem, there is a $\beta \in (\hat{\beta}_1, \hat{\beta}_2)$ such that $h(\beta) = 0$, which contradicts that $\underset{\beta\to\hat{\beta}-}{h(\beta)} \cdot \underset{\beta\to\hat{\beta}+}{h(\beta)} > 0$.

Case II: For any $\hat{\beta}_1 < \hat{\beta}_2 \in D$, we have $\underset{\beta\to\hat{\beta}_1+}{h(\beta)} \cdot \underset{\beta\to\hat{\beta}_2-}{h(\beta)} > 0$. By Claim 1, $h(\beta)$ is continuous on some $\left(0, \hat{\beta}_1\right)$ with $\underset{\beta\to\hat{\beta}_1-}{h(\beta)} > 0$ or on some $\left(\hat{\beta}_2, 1\right)$ with $\underset{\beta\to\hat{\beta}_2+}{h(\beta)} < 0$. By Claim 1 and the intermediate value theorem, there is a $\beta \in (0, \hat{\beta}_1)$ or $(\hat{\beta}_2, 1)$ such that $h(\beta) = 0$, which is also a contradiction. □

Fig. 6 is helpful in understanding the remaining argument.

---

[30] Recall $h(\beta) = \left(P(\beta)^{[1]} - P(\beta)'^{[1]}\right) + \beta\left(P(\beta)^{[2]} - P(\beta)'^{[2]}\right) = (1-\beta)\left(P(\beta)^{[1]} - P(\beta)'^{[1]}\right) + \beta\left(P(\beta)'^{[3]} - P(\beta)'^{[3]}\right)$.

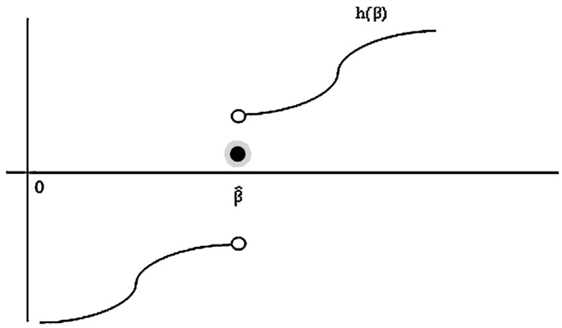

Fig. 6. An example of discontinuous $h(\beta)$ at $\hat{\beta}$.

By Claim 2, we can design an IC cardinal rule with such a $\hat{\beta}$ by using an appropriate convex combination of two rules. One rule is g^+ with the balance function at $\hat{\beta}$, $h^+\left(\hat{\beta}\right) = \underset{\beta\to\hat{\beta}+}{h(\beta)}$. The other rule is g^- with $h^-\left(\hat{\beta}\right) = \underset{\beta\to\hat{\beta}-}{h(\beta)}$. The next claim concludes Step 4 by showing what the appropriate combination is.

Claim 3. *There exists an IC cardinal rule $f^*(v)$ such that $f^*(v) = \alpha g^+ + (1-\alpha)g^-$ where $\alpha = \frac{-h^-\left(\hat{\beta}\right)}{h^+\left(\hat{\beta}\right)-h^-\left(\hat{\beta}\right)} \in [0,1]$.*

Proof. From Claim 2, fix $\hat{\beta}$ and consider a rule, g^+ based on the fixed partition $P^{\hat{\beta}}$ but with a different score vector, $s_i\left(\hat{\beta}+\epsilon, t_i\right)$. With sufficiently small $\epsilon > 0$ such that $\hat{\beta}+\epsilon \notin D$, g^+ are different from $g_{\hat{\beta}}$ only in some of the tie cases of $g_{\hat{\beta}}$. This is because any $\hat{\beta} \in D$ involves tie cases where $S^H\left(\hat{\beta}, t_{-i}\right)^l = S^H\left(\hat{\beta}, t_{-i}\right)^{l'}$ or $S^L\left(\hat{\beta}, t_{-i}\right)^l = S^L\left(\hat{\beta}, t_{-i}\right)^{l'}$ $for\ l \neq l' \in L$ as seen in Examples 3. With $P^{\hat{\beta}}$, we obtain the balance function of g^+, $h^+\left(\hat{\beta}\right) = \underset{\beta\to\hat{\beta}+}{h(\beta)}$. Due to the same partition $P^{\hat{\beta}}$ and the different decisions only in tie cases of $g_{\hat{\beta}}$, g^+ is still a maximizer of the utilitarian social welfare in $F^{P^{\hat{\beta}}}$. Similarly, we design a rule, g^- based on $P^{\hat{\beta}}$ but with $s_i\left(\hat{\beta}-\epsilon, t_i\right)$ and the balance function $h^-\left(\hat{\beta}\right) = \underset{\beta\to\hat{\beta}-}{h(\beta)}$. From this perspective, we can say that g^+ and g^- are scoring rules with the same score vectors of $g_{\hat{\beta}}$, but with different tie-breaking rules.

Finally, consider a rule $\tilde{g}_\alpha = \alpha g^+ + (1-\alpha)g^-$, where[31] $\alpha = \frac{-h^-\left(\hat{\beta}\right)}{h^+\left(\hat{\beta}\right)-h^-\left(\hat{\beta}\right)} \in [0,1]$ such that the balance function of $\tilde{g}_\alpha$, $\tilde{h}_\alpha\left(\hat{\beta}\right)$ is equal to 0. Hence, $\tilde{g}_\alpha$ is an IC cardinal rule. For comparison with ordinal rules, we use the identical function, $f^*(v) = \tilde{g}_\alpha(t_1, ..., t_n)$ in Step 5. □

Step 5) We prove that f^* has a higher utilitarian social welfare than the ordinal utilitarian rule, denoted by f_{OUR}. The construction of f^* is based on a *finer* partition than f_{OUR}, which implies that $W(f^*) \geq W(f_{OUR})$ (the equality holds when $f^*(v) = f_{OUR}(v)$ for almost every $v \in V$). However, we can always find a set of v with non-zero measure such that $f^*(v) \neq f_{OUR}(v)$, as shown in the proof of Claim 1. In these cases, where $v_i \in V_i^{t^H}$ and $v_i' \in V_i^{t^L}$ given the others' types, we have $f^*(v) \neq f^*\left(v_i', v_{-i}\right)$ while $f_{OUR}(v) = f_{OUR}(v_i', v_{-i})$ because v and $\left(v_i', v_{-i}\right)$ are in the same ordinal type set. These cases guarantee a difference in social welfare between f^* and f_{OUR}, implying that $W(f^*) > W(f_{OUR})$.

Step 6) The above steps prove the theorem with three alternatives. For the extension with more than three alternatives, we need to change Step 3 by constructing a rule with g_β and two other rules. First, recall the ordinal utilitarian rule g_{OUR} which has one score vector $s_{OUR} = \left(\mathbb{E}\left(\hat{v}_i^{[k]} \mid t_i(v_i) = t\right)\right)$ for $t \in T^{ORD}$. Second, we consider a modified rule $\hat{g} \in F^{P^\beta}$ with slightly different score vectors than $s^H(\beta), s^L(\beta)$. The score vectors assign the third-ranked alternative the same score as the g_{OUR}. That is,

$$\hat{s}^H(\beta) = \left(s^H(\beta)^1, s^H(\beta)^2, s^3_{OUR}, s^H(\beta)^4, ..., s^H(\beta)^m\right),$$
$$\hat{s}^L(\beta) = \left(s^L(\beta)^1, s^L(\beta)^2, s^3_{OUR}, s^L(\beta)^4, ..., s^L(\beta)^m\right).$$

Now, we construct a new rule $\hat{g}_\beta$ in the following manner.

31 Recall $h^+\left(\hat{\beta}\right) \cdot h^-\left(\hat{\beta}\right) \leq 0$ from Claim 2.

$$\hat{g}_\beta(t_i, t_{-i}) = \begin{cases} g_\beta(t_i, t_{-i}) & \text{if } \{l \in L : r_l(v_i) \le 3\} = \{l \in L : r_l(v_j) \le 3\} \\ & \quad \text{for all } i, j \in N \text{ and } g_\beta = \hat{g} \\ g_{OUR}(t_i, t_{-i}) & \text{otherwise} \end{cases}$$

In words, $\hat{g}_\beta$ is the same as g_β when the two conditions are satisfied. First, every agent has the same top three alternatives. Second, $g_\beta(t_i, t_{-i})$ equals $\hat{g}(t_i, t_{-i})$. Otherwise, $\hat{g}_\beta$ is the same as g_{OUR}.

We construct $\hat{g}_\beta$ to apply the main intuition in previous steps with three alternatives and directly compare $\hat{g}_\beta$ with g_{OUR}. This rule may be more meaningful when we are concerned with the preference intensities and competition of the top three alternatives. After the construction of $\hat{g}_\beta$, the following arguments are almost the same, replacing g_β with $\hat{g}_\beta$ in the remaining steps. Note that Claim 1 still holds with more than 3 alternatives, considering that the shared top 3 alternatives are replaced with the 3 alternatives in the cases from Claim 1 and g_β is replaced with $\hat{g}_\beta$.

We must be careful with the replacement. First, $\hat{g}_\beta$ is neutral because g_β, $\hat{g}$, and g_{OUR} are neutral. Second, we need to examine another property of $\mathbb{E}\left(\hat{g}_\beta\left(t, \hat{t}_{-i}\right)\right) = P(\beta)$ and $\mathbb{E}\left(\hat{g}_\beta\left(t', \hat{t}_{-i}\right)\right) = P(\beta)'$,

$$P(\beta)^{[k]} = P(\beta)'^{[k]} \text{ for } k \ge 4.$$

This property results from the fact that $\hat{g}_\beta$ chooses the alternatives among the top three alternatives shared for all agents regardless of agent i's announcement of H or L type.

Unlike the case with three alternatives, $s^H(\beta)^3$ and $s^L(\beta)^3$ are non-zero.

Lemma 4. *For any $\beta \in (0, 1)$, $s^H(\beta)^3 \le s^L(\beta)^3$.*

Proof. According to the definition of the score vectors, $\lim_{\beta\to 0} s^H(\beta)^3 = \lim_{\beta\to 1} s^L(\beta)^3$. Thus, it is sufficient to show that $s^H(\beta)^3$ and $s^L(\beta)^3$ are (weakly) decreasing in β. Given $0 < \beta_2 < \beta_1 < 1$,

$$s^H(\beta_1)^3 = \int_0^1 \ldots \int_{v_i^{[3]}}^1 \int_{v_i^{[4]}}^{\frac{v_i^{[2]}-\beta_1}{1-\beta_1}} v_i^{[3]} \mu\left(v_i^{[2]}, \ldots, v_i^{[m]-1}\right) dv_i^{[3]} dv_i^{[2]} \ldots dv_i^{[m]-1} / p^H(\beta_1),$$

$$s^H(\beta_2)^3 = \int_0^1 \ldots \int_{v_i^{[3]}}^1 \int_{v_i^{[4]}}^{\frac{v_i^{[2]}-\beta_2}{1-\beta_2}} v_i^{[3]} \mu\left(v_i^{[2]}, \ldots, v_i^{[m]-1}\right) dv_i^{[3]} dv_i^{[2]} \ldots dv_i^{[m]-1} / p^H(\beta_2).$$

Since $\frac{v_i^{[2]}-\beta_1}{1-\beta_1} \le \frac{v_i^{[2]}-\beta_2}{1-\beta_2}$ and $p^H(\beta_2) > p^H(\beta_1)$, for any $\alpha \ge 0$,

$$Pr\left[v_i^{[3]} \le \alpha \mid v_i \in V_i^{t^H}(\beta_2)\right] \le Pr\left[v_i^{[3]} \le \alpha \mid v_i \in V_i^{t^H}(\beta_1)\right]$$

By the first order stochastic dominance, $s^H(\beta_2)^3 \ge s^H(\beta_1)^3$. This means that $s^H(\beta)^3$ are (weakly) decreasing in β. Similarly,

$$s^L(\beta_1)^3 = \int_0^1 \ldots \int_{v_i^{[3]}}^1 \int_{\frac{v_i^{[2]}-\beta_1}{1-\beta_1}}^{v_i^{[2]}} v_i^{[3]} \mu\left(v_i^{[2]}, \ldots, v_i^{[m]-1}\right) dv_i^{[3]} dv_i^{[2]} \ldots dv_i^{[m]-1} / p^L(\beta_1)$$

$$s^L(\beta_2)^3 = \int_0^1 \ldots \int_{v_i^{[3]}}^1 \int_{\frac{v_i^{[2]}-\beta_2}{1-\beta_2}}^{v_i^{[2]}} v_i^{[3]} \mu\left(v_i^{[2]}, \ldots, v_i^{[m]-1}\right) dv_i^{[3]} dv_i^{[2]} \ldots dv_i^{[m]-1} / p^L(\beta_2).$$

Since $\frac{v_i^{[2]}-\beta_1}{1-\beta_1} \le \frac{v_i^{[2]}-\beta_2}{1-\beta_2}$ and $P^L(\beta_2) < P^L(\beta_1)$, for any $\alpha \ge 0$,

$$Pr\left[v_i^{[3]} \ge \alpha \mid v_i \in V_i^{t^L}(\beta_2)\right] \ge Pr\left[v_i^{[3]} \ge \alpha \mid v_i \in V_i^{t^L}(\beta_1)\right].$$

By the first order stochastic dominance, $s^L(\beta_2)^3 \ge s^L(\beta_1)^3$. $\square$

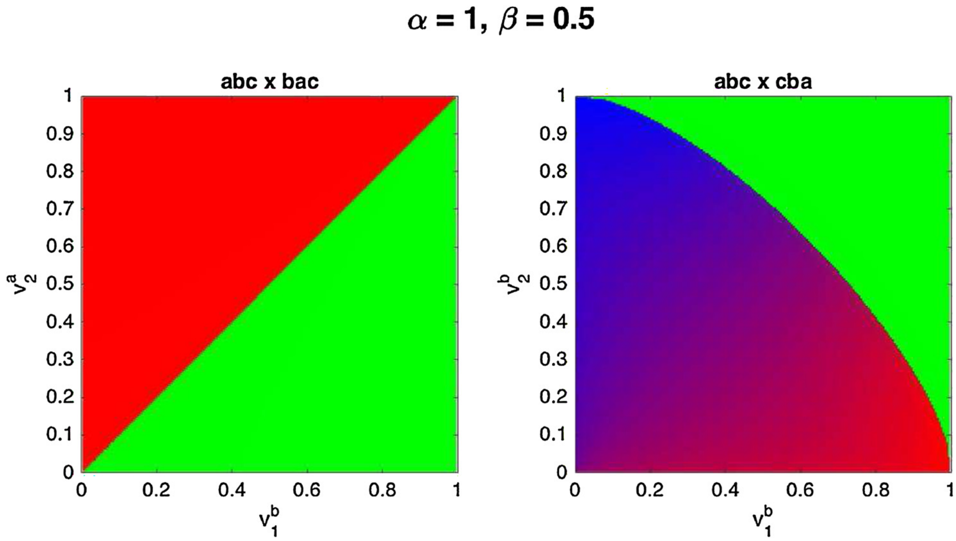

Fig. 7. Second-best rule under the beta distribution with $\alpha = 1$ and $\beta = 0.5$.

Lemma 4 implies that we may lose the property $P(\beta)^{[2]} \geq P(\beta)'^{[2]}$, $P(\beta)^{[1]} \leq P(\beta)'^{[1]}$ and $P(\beta)^{[3]} \leq P(\beta)'^{[3]}$ for any $\beta \in (0,1)$ without the replacement. However, we can recover this property by mixing g_β and $\hat{g}$ since $\hat{s}^H(\beta)^3 = \hat{s}^L(\beta)^3 = s^3_{OUR}$. The other steps hold with the replacement.

Argument for $n = 2$ **or 4.** Every step is the same as in the proof of Theorem 2 except the second part in the proof of Claim 1. For $n = 2$, consider $t_1' = abc^L$ and $t_2 = cba^L$. Since $\mathbb{E}\left(\hat{v}_i^{[2]} \mid P_i^{ORD}\right) = s^L(1)^2 \leq \frac{1+s^L(1)^3}{2} = \frac{\mathbb{E}\left(\hat{v}_i^{[1]} | P_i^{ORD}\right) + \mathbb{E}\left(\hat{v}_i^{[3]} | P_i^{ORD}\right)}{2}$, $s^L(\beta)^2$ increases and $s^L(\beta)^3$ (weakly) decreases in β, we have $g_\beta\left(t_i', t_{-i}\right) = (\frac{1}{2}, 0, \frac{1}{2})$ and $g_\beta(t_i, t_{-i}) = (0, 1, 0)$ as β approaches 1. For $n = 4$, we have the case where $t_1' = abc^L$, $t_2 = cba^L$, $t_3 = abc^L$ and $t_4 = cba^L$.

Proof of Proposition 2. The first part follows from Theorem 2. Note that with three alternatives $s^L(\beta) = \left(1, s^L(\beta)^2, 0\right)$ and $s^H(\beta) = \left(1, s^H(\beta)^2, 0\right)$. For the second part, consider Step 4 in the proof of Theorem 2. When $h(\beta)$ is continuous, f^* is obviously an IC (A, B)-scoring rule with $(A, B) = \left(s^L(\beta^*)^2, s^H(\beta^*)^2\right)$ and the standard tie-breaking rule. However, in the other case, f^* does not look like an (A, B)-scoring rule because of the convex combination of the two rules. Nevertheless, it is shown in the proof of Claim 3 that f^* is the same as $g_{\hat{\beta}}$ except for some ties. By the first part of Proposition 2, $g_{\hat{\beta}}$ is an IC (A, B)-scoring rule with $\left(s^L\left(\hat{\beta}\right)^2, s^H\left(\hat{\beta}\right)^2\right)$ and the standard tie-breaking rule. Since the definition of an (A, B)-scoring rule makes no restriction on tie-breaking rules, the new rule f^* is an IC (A, B)-scoring rule with $\left(s^L\left(\hat{\beta}\right)^2, s^H\left(\hat{\beta}\right)^2\right)$, but with a different tie-breaking rule from $g_{\hat{\beta}}$. Then, setting $\beta^* = \hat{\beta}$ completes the proof. □

Beta distribution in Section 7

We consider a beta distribution for discretized types of $v_i^{[2]}$. Both shape-parameters (α, β) vary from 0.5 to 5 in increments of 0.5, so we have 100 different type distributions. Then, we computed the second-best rule numerically in each type distribution to check how it changes.

Figs. 7 and 8 display the second-best rule when $v_1 \in V_1^{abc}$ and $v_2 \in V^{bac}, V^{cba}$ under the beta distribution with $\alpha = 1$ and $\beta = 0.5$ and 1.5, respectively. For other cases of v_2, the second-best rule is almost identical to the first-best rule,[32] so we present the cases when $v_1 \in V_1^{abc}$ and $v_2 \in V^{bac}, V^{cba}$. It is helpful to compare the above figures with Fig. 3 which shows the second-best rule under the beta distribution with $\alpha = 1$ and $\beta = 1$. First, we find that abc x bac case loses no efficiency relative to the first-best rule when $\beta \leq 1$, but loses some efficiency when $\beta > 1$.[33] In Fig. 8, the right upper

---

[32] We find mixed color on the very small area around the origin and the $(1, 1)$ under some distributions. Since we find a similar pattern for the first-best rule with the numerical approach, we conjecture that this pattern comes from the numerical approach and that the second-best rule is the same as the first-best rule for those cases.

[33] The incentive constraints cannot be satisfied in the numerical program when we allow only the case where agents show the opposite ordinal preferences, fixing other cases as the first-best rule. Thus, we observe the necessary efficiency loss when $\beta > 1$ for incentive compatibility.

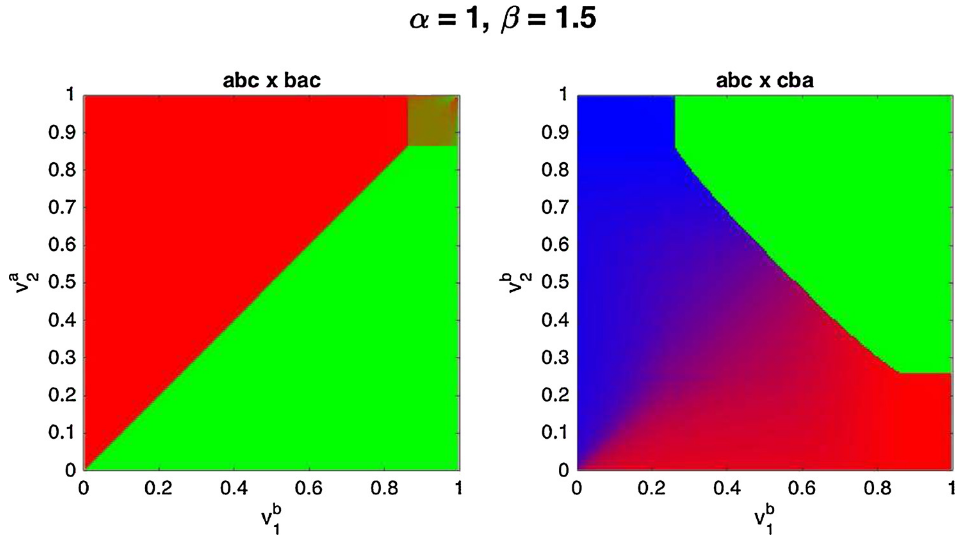

Fig. 8. Second-best rule under the beta distribution with $\alpha = 1$ and $\beta = 1.5$.

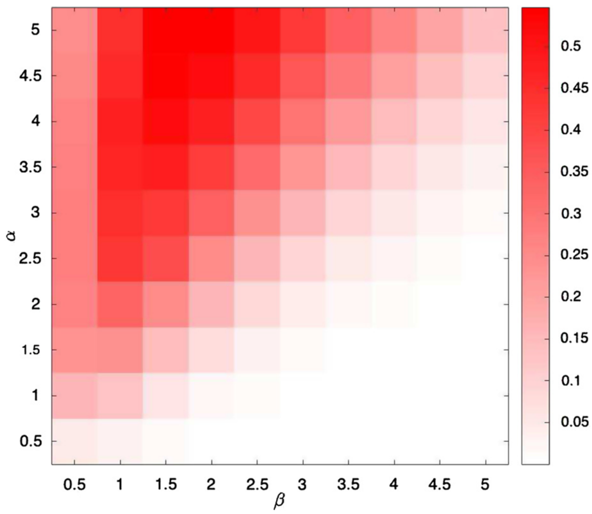

Fig. 9. Welfare Loss of f_{SB} relative to f_{FB} for 100 type distributions. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

square in abc x bac case shows mixed probability over alternative a and b, approximately 0.5 for each. The square shape is maintained but the size varies when $\beta > 1$ and disappears when $\beta \leq 1$. Second, in the case abc x cba, the line that divides whether alternative b is chosen or not is concave from the origin when $\beta < 1$, linear when $\beta = 1$, and convex when $\beta > 1$. For all $\beta > 1$, the lines are convex, but the locations are changed so the area where alternative b is chosen varies. The above pattern similarly appears for $\alpha \neq 1$, so we omit those cases.

Fig. 9 shows the welfare loss of f_{SB} relative to f_{FB} for 100 type distributions. The vertical axis represents the values of α and the horizontal axis represents the values of β. At each grid point, the welfare loss is colored with a level of red, where the darker red implies a bigger loss. The maximum loss is approximately 0.55%. The number of cases where the loss is smaller than 0.05% is 35 (right bottom areas). Note that the loss distribution is asymmetrically skewed to the top left.

To compare the welfare loss for incentive constraints with the welfare loss for the restriction of ordinal rules, we calculated the welfare loss of f_{OUR} relative to f_{SB} for each type distribution.

Fig. 10 shows the welfare loss of f_{OUR} relative to f_{SB} for 100 type distributions.[34] The maximum loss is approximately 3.9% and the loss for every case is greater than 0.22%. Thus, in most cases, the losses for incentive constraints are much

[34] For the welfare calculations on the beta distribution, we fixed 200 x 200 grids.

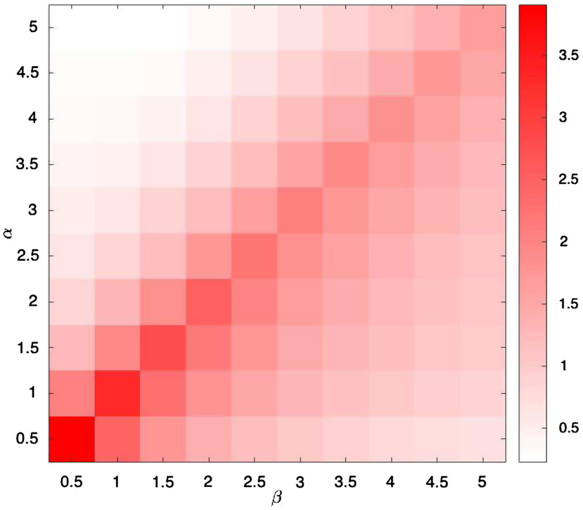

Fig. 10. Welfare Loss of f_{OUR} relative to f_{SB} for 100 type distributions.

smaller than the losses for the restriction of ordinal rules. However, this loss distribution in Fig. 10 is symmetric, so we find 10 cases that show the opposite (top left areas).[35]

References

Abdulkadiroğlu, A., Che, Y., Yasuda, Y., 2011. Resolving conflicting preferences in school choice: the "Boston Mechanism" reconsidered. Amer. Econ. Rev. 101, 1–14.
Apesteguia, J., Ballester, M.A., Ferrer, R., 2011. On the justice of decision rules. Rev. Econ. Stud. 78, 1–16.
Azrieli, Y., Kim, S., 2014. Pareto efficiency and weighted majority rules. Int. Econ. Rev. 55, 1067–1088.
Barberà, S., Jackson, M.O., 2006. On the weights of nations: assigning voting weights in a heterogeneous union. J. Polit. Economy 114, 317–339.
Börgers, T., Postl, P., 2009. Efficient compromising. J. Econ. Theory 144, 2057–2076.
Carrasco, V., Fuchs, W., 2009. A Procedure for Taking Decisions with Non-transferable Utility. Working paper.
Casella, A., 2005. Storable votes. Games Econ. Behav. 51, 391–419.
Ehlers, L., Majumdar, D., Mishra, D., Sen, A., 2016. Continuity and Incentive Compatibility in Cardinal Voting Mechanisms. Working paper.
Freixas, X., 1984. A cardinal approach to straightforward probabilistic mechanisms. J. Econ. Theory 34, 227–251.
Gershkov, A., Moldovanu, B., Shi, X., 2017. Optimal voting rules. Rev. Econ. Stud 84, 688–717.
Gibbard, A., 1973. Manipulation of voting schemes: a general result. Econometrica 41, 587–601.
Gibbard, A., 1978. Straightforwardness of game forms with lotteries as outcomes. Econometrica 46, 595–613.
Giles, A., Postl, P., 2014. Equilibrium and effectiveness of two-parameter scoring rules. Math. Soc. Sci. 68, 31–52.
Harsanyi, J.C., 1955. Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. J. Polit. Economy 63, 309–321.
Holmström, B., Myerson, R.B., 1983. Efficient and durable decision rules with incomplete information. Econometrica 51, 1799–1819.
Hortala-Vallve, R., 2012. Qualitative voting. J. Theor. Politics 24, 526–554.
Jackson, M.O., Sonnenschein, H.F., 2007. Overcoming incentive constraints by linking decisions. Econometrica 75, 241–257.
Majumdar, D., Sen, A., 2004. Ordinally Bayesian incentive compatible voting rules. Econometrica 72, 523–540.
McLennan, A., 1980. Randomized preference aggregation: additivity of power and strategy proofness. J. Econ. Theory 22, 1–11.
Miralles, A., 2012. Cardinal Bayesian allocation mechanisms without transfers. J. Econ. Theory 147, 179–206.
Moulin, H., 1980. On strategy-proofness and single peakedness. Public Choice 35, 437–455.
Myerson, R.B., 2002. Comparison of scoring rules in Poisson voting games. J. Econ. Theory 103, 219–251.
Pattanaik, P.K., Peleg, B., 1986. Distribution of power under stochastic social choice rules. Econometrica 54, 909–921.
Roberts, K.W.S., 1980. Interpersonal compatibility and social choice theory. Rev. Econ. Stud. 47, 421–439.
Satterthwaite, M.A., 1975. Strategy-proofness and arrow's conditions: existence and correspondence theorems for voting procedures and social welfare functions. J. Econ. Theory 10, 187–217.
Schmitz, P.W., Tröger, T., 2012. The (sub-)optimality of the majority rule. Games Econ. Behav. 74, 651–665.
Strotz, R.H., 1953. Cardinal utility. Amer. Econ. Rev. 43, 384–397.

[35] It could be interesting if we can characterize the set of distributions that show the opposite in the general environment. We leave this direction for future research.