ORDINAL VERSUS CARDINAL VOTING RULES: A MECHANISM DESIGN APPROACH

SEMIN KIM

November 6, 2014

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, i.e., when alternatives are symmetric ex-ante, essentially any ex-ante Pareto efficient ordinal rule is incentive compatible. But, importantly, 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. This cardinal rule, applied in an environment with exactly three alternatives, is closely connected to the (A,B)-scoring rules in Myerson (2002).

Keywords: Ordinal rule, Pareto efficiency, Incentive compatibility, Bayesian mechanism design.

JEL Classification: C72, D01, D02, D72, D82.

I am greatly indebted to Yaron Azrieli for his thoughtful guidance. I am also grateful to Paul Healy, James Peck, Katherine Baldiga Coffman, Christian Hellwig, Dan Levin, Caleb Cox, Ian Krajbich, Tilman Börgers, Roberto Serrano, Hye Lin Choi, Jack Rouzer, Robert Munk, Dimitry Mezhvinsky, and seminar participants at The Ohio State University Theory and Experiments Group Meetings, 2013 North America Summer Meeting of Econometric Society at University of Southern California, and 6th Alexander Hamilton Center Conference on Political Economy at New York University for their helpful comments and suggestions. All errors are my own responsibility.
Department of Economics, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei, 10617 Taiwan; seminkim.@ntu.edu.tw.

1. Introduction

For the problem of aggregation of individual preferences into a group decision, the vast majority of the social choice literature uses a purely ordinal model: agents can rank the candidates but intensity of preferences is discarded. Also, the voting literature mainly studies ordinal voting rules which depend only on ranking information in agents' preferences. Furthermore, this literature usually restricts attention 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 an alternative and a slight minority strongly prefers another alternative. Which alternative is most desirable for the society? The more concrete example is a representative democracy. Representatives choose one of the alternatives. A representative knows how much his district is willing to pay for the alternatives to be chosen, which may represent his district's preference intensities over the alternatives. He not only needs to ruminate on the ordinal preference of his group, but also on the preference intensity of his group. But if a voting rule that aggregates the votes of the representatives is ordinal, then it cannot sufficiently satisfy the whole society. Second, 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 the past literature in that, in our model, agents assign cardinal valuations to the alternatives, which express their relative desirability. Rather than requiring strategy proofness, we consider the weaker concept of *incentive compatibility*, which only requires agents to be truth-telling if other agents are also truth-telling. 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 my knowledge, only few papers investigate 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

[1]These are sometimes called dominant incentive compatible rules in the mechanism design literature.
[2]See the related literature section.

environment, i.e., when alternatives are symmetric ex-ante, and show that essentially any ex-ante Pareto efficient ordinal rule is Incentive Compatible (IC). Furthermore, we prove that there exists an IC cardinal rule which outperforms any ordinal rule.

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 valuations for the alternatives. These valuations are random variables which, we assume, are drawn independently across agents, not necessarily between alternatives. We also assume that values are *neutral* between alternatives, which implies that the value distributions of each alternative is 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 consider (weak) ex-ante Pareto efficiency in the class of ordinal rules.[3] A rule f is Ordinally Pareto Efficient (OPE) if f is ordinal, and there is no ordinal rule that strictly increases the expected utilities of all agents under f. We characterize OPE rules such that f is OPE if and only if it is a scoring rule with specific scores, which are the expected values of the ranked alternatives given ordinal preferences (Proposition 1). For reference, a rule is called a scoring rule if there exists a score vector s_i for each agent i over the alternatives such that the alternative with the greatest sum of scores is chosen. Then, 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 appealing criterion of ex-ante Pareto efficiency. Note, this result implies that there is also no conflict between strong Pareto efficiency and incentive compatibility.[4]

Next, we ask whether it is possible to design a rule that is superior to ordinal rules. We show that one can design an IC cardinal rule which achieves higher utilitarian social

[3]There are two well-known notions of Pareto efficiency: weak Pareto efficiency and strong Pareto efficiency. Since we mainly consider the weak concept, we omit the term, "weak" from now, unless necessary.

[4]A strong ordinally Pareto efficient rule f can be defined as an ordinal rule such that there is no ordinal rule that weakly increases the expected utilities of all agents, at least one strictly under f.

welfare than any ordinal rule, under a symmetry assumption on agents (Theorem 2).[5] If we consider the class of cardinal rules, there exists a trade-off between Pareto efficiency and incentive compatibility. For example, Börgers and Postl (2009) show that the first best cardinal rule in their set-up is not IC.[6] However, Theorem 2 states that we can design an IC cardinal rule superior to any ordinal rule. With three alternatives, this rule is closely related to well-known voting rules: *plurality, negative, Borda count, and approval* voting rules. As in Myerson (2002), the general form of these simple voting rules is an (A, B)-scoring rule, which is similar to a scoring rule, but where there are two possible score vectors, $(1, A, 0)$ or $(1, B, 0)$ where $(0 \leq A \leq B \leq 1)$. Proposition 2 shows that our rule is an IC (A, B)-scoring rule in an environment with three alternatives. Since (A, B)-scoring rules are simple to implement, Proposition 2 could inform voting rules in a variety of institutional settings.

The 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 shows 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 concludes. The Appendix contains the more technical portions of our proofs.

2. Related literature

There is an extensive literature which examines 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 (1977), 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

[5]We focus on symmetric agents in the sense that the value distribution is identical across agents. One advantage of this assumption is that the social welfare can be normalized to the utility of one agent - no interpersonal utility comparisons are needed to compare rules.

[6]They consider two agents who have the opposite ordinal preferences and three alternatives.

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. This approach's usefulness results from the availability of information about agents' value distributions. Jackson and Sonnenschein (2007) show 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) derive 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) derive the form of the utilitarian, maximin, and maximax ordinal rules.[7] 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. A natural question, then, is whether their rules are IC. Our Proposition 1 addresses the issue of incentive compatibility of the set of efficient rules, which includes all of their rules.

The paper most closely related to our work is Majumdar and Sen (2004). They analyze the implications of weakening incentive constraints from strategy proofness to Ordinal Bayesian Incentive Compatibility. They show that under uniform priors a wide class of rules satisfies Ordinal Bayesian Incentive Compatibility, which is related to our Theorem 1. But, unlike their work, 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. Furthermore, we use cardinal value information and characterize the ordinally Pareto efficient rules (Proposition 1).

This paper is 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) study symmetric Bayes Nash equilibria under (A, B)-scoring rules and the welfare implication of the rules. They show that some (A, B)-scoring rules can increase social welfare relative to common voting rules such as the plurality, negative, Borda count, and approval voting rules. Their result is related to Theorem 2, but they use numerical methods for the welfare analysis.

[7]We discuss these rules in Section 5.

There are several papers that investigate voting rules only with two alternatives. Azrieli and Kim (2014) show that there is no incentive compatible rule which outperforms optimal ordinal rules, and they characterize the set of interim incentive efficient rules as the set of weighted majority rules. Schmitz and Tröger (2012) prove that the optimal rule among all anonymous and neutral rules in both symmetric and neutral environments is majority rule, which is indeed ordinal. However, we show that the investigation with three or more alternatives leads to greatly different results regarding efficiency and incentive compatibility.

Lastly, many papers including this paper do not allow monetary transfers although they apply the mechanism design approach. For example, Gershkov et al. (2014) deal with a voting mechanism, Miralles (2012) examines an allocation mechanism, and Carrasco and Fuchs (2009)[8] study a communication protocol. The main reason for the exclusion of monetary transfers is that, in many cases monetary transfers are infeasible or excluded for ethical or institutional reasons (e.g., child placement in public schools, organ transplants, collusion in markets, voting, etc).

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 $\mid L \mid = m$.[9] The typical elements of L will be denoted by l, l', l'', etc. Each agent $i \in N$ has a von-Neumann Morgenstein value (utility) function $v_i : L \to \mathbb{R}$. 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 $l, l' \in L$.[10] 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}, and let t denote an element of T^{ORD}. 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, so that$\mid \{t : t_i(v_i) = t\} = T^{ORD} \mid = m!$. As usual, a subscript $-i$ on a vector means that the ith coordinate is excluded.

[8]Their procedure to derive the optimal communication protocol has the similar flavor with our Theorem 2 in the sense that each type is divided into two types. But, unlike our paper, they cover two players, infinite action spaces and, more importantly, develop a dynamic mechanism.
[9]For every finite set X, $\mid X \mid$ denotes the number of elements in X .
[10]This assumption is not essential for results but simplifies definitions and notation.

To capture the preference intensity, it will be 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 $i \in N$.[11] 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. And the preference intensities of the other intermediate alternatives are measured between 0 and 1.

The information structure is the following. We assume that the vector of values of agent i is a random variable $\hat{v}_i$ with values in V_i. The assumption of strict preference and the normalization establish agent $i's$ value space that $V_i = [0,1]^m \diagdown \left\{\hat{v}_i : \hat{v}_i^l = \hat{v}_i^{l'} \; for \; l, l' \in L\right\}$. We consider the conditional distribution of $v_i^{[k]}$ given type $t \in T^{ORD}$ which is denoted by $\mu_i^{k,t}$. Let μ_i^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 $t, t' \in T^{ORD}$,

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

As a consequence, the structure is simplified that $\mu_i^t = \mu_i$ 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 $i \in N$.

We also assume that values are independent across agents, so the distribution of $\left(\left(\hat{v}_i^{[k]}\right)_{k=\{1,...,m\}}\right)_{i \in N}$ is the product distribution $\mu = \mu_1 \times \cdots \times \mu_n$. The above features of the environment are common knowledge among agents. But, the realized value v_i is assumed to be observed only by agent i.

A voting rule is a measurable mapping $f : V \to \triangle(L)$,[12] which means that we allow randomized rules. 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.[13] To

illustrate the features of our environment, we present the following example.

[11]Although our results except Proposition 2 hold without this normalization, we use this normalization because it simplifies many arguments and provides Proposition 2. The similar normalization appears in Börgers and Postl (2009). But, it is arguable that this normalization can capture the preference intensity and can validate the interpersonal comparisons of utility. In fact, the issue of cardinal utility and interpersonal comparisons of utility has been controversial in the literature (Strotz (1953), Harsanyi (1955), Roberts (1980), etc). It is beyond the scope of our paper.
[12]For every finite set X , $\triangle(X)$ denotes the set of probability measures on X.
[13]$x \cdot y$ denotes the inner product of the vector x and y.

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 that $t_i(v_i) = abc$, 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 environment, $Pr[t_i(\hat{v}_i) = t] = \frac{1}{6}$. For example, if $f(v) = (\frac{1}{2}, \frac{1}{2}, 0)$, it means that 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 $f(v^\sigma) = f(v)^\sigma$.

A neutral rule treats all alternatives symmetrically.

Finally in this section, we employ 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 $\mid T^{ORD} \mid$ 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: 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 $i \in N$. As usual, 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 $f(v) = f(v')$ whenever $P^{ORD}(v) = P^{ORD}(v')$. 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 $r_l : V_i \to \{1, ..., m\}$ denote the ranking of alternative l in v_i.

Next, we consider an important class of ordinal rules: *scoring* rules. Let $s_i \equiv (s_i^1, ..., s_i^m)$ be a score vector in $\mathbb{R}^m$ where $s_i^1 \geq ... \geq s_i^m$ for $i \in N$.

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

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

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

Note that $s_i^{r_l(v_i)}$ can be seen as agent $i's$ score assigned to alternative $l \in L$ when he announces v_i. Additionally, it depends only on the ranking of the alternative, not the specific ordinal type $t_i\,(v_i)$. We allow *asymmetric* score vectors across agents. If s_i is identical for all agents then we get a symmetric scoring rule as standard in the voting literature.

4. A Motivating Example

The objective of this section is to give a simple example with two agents and three alternatives, which helps to understand our motivation, notation, and main intuition.

Suppose that $N = \{1, 2\}$, $L = \{a, b, c\}$ 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 $i \in N$, $Pr\,[t_i\,(\hat{v}_i) = t] = \epsilon$ for $t = abc, acb, bac, cab$ and $Pr\,[t_i\,(\hat{v}_i) = t] = \frac{1-4\epsilon}{2}$ for $t = bca, cba$. We look at rules widely used in practice, the plurality and Borda count voting rules.[14] Denote the plurality voting rule by f_P and denote the Borda count voting rule by f_B. In our setting, f_P is a scoring rule with $s_{i,P} = (1, 0, 0)$ and f_B with $s_{i,B} = \left(1, \frac{1}{2}, 0\right)$ for $i = 1, 2$. We assume here that ties are broken by the uniform distribution over the set of alternatives with the greatest sum of scores. The following table shows the comparison of the two rules and the usual trade off between efficiency and incentive compatibility even in the class of ordinal rules.

[14]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.

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

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

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

The voting rule f_P is IC. Table 1 informs that an agent's misrepresentations do not help.[15] 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. That is because the score vector of the Borda count voting rule reflects the voters' expected utilities given their ordinal types, i.e., $s_{i,B} = \mathbb{E}\left(1, \hat{v}_i^{[2]}, 0 \mid t_i(v_i)\right) = \left(1, \frac{1}{2}, 0\right)$ 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 $Pr[t_2(\hat{v}_2) = cba] = \frac{1-4\epsilon}{2}$). When $v_1 = (1, 1-\varepsilon, 0)$ with sufficiently small ε, agent 1 has an incentive to lie that his type is *bac*, which shows 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 does not exist. We observe, from this example, that there is an incentive to lie when some alternatives are popular ex-ante. Then, we consider a neutral environment in which no alternative is special ex-ante . Only one condition is changed such that $Pr[t_i(\hat{v}_i) = t] = \frac{1}{6}$ for $t \in T^{ORD}$ and $i = 1, 2$. He expects that the probability of each alternative getting the score 1 is the same given agent 2's truthful reports. Now, agent 1's truthful report is

[15]The misrepresentations cannot increase the probability that the top alternative is chosen and cannot decrease the probability that the bottom one is chosen.

better than any other report. Since the same argument is applied to any other ordinal type and to agent 2, f_B is IC.

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

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

H type and L type are divided according to the value of the second-ranked alternative. For each $\beta \in (0,1)$, we consider the rule f_β which maximizes the sum of expected utilities among this partition-measurable rules. In fact, f_β resembles a scoring rule but with the 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 $s^L(\beta) = \left(1, \mathbb{E}\left(\hat{v}_i^{[2]} \mid v_i \in V^{t^H}\right), 0\right)$. 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. But, 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 understanding the intuition, we provide the following table shows f_B and f^* when agent 1 announces an H type and an L type.

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

TABLE 2. f_B and f^* when $v_1 \in V_1^{abc^H}$ and $v_1' \in V_1^{abc^L}$

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 agents have the opposite ordinal preferences). 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 of an agent 1 gives him a relatively high probability of getting the middle alternative and an L type announcement has the opposite effect.[16] Consequently, the most of H and L type agents like to report their true types. But, to satisfy 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 as β, called the cut-off agent. We find that under the rule $f^* = f_{\frac{1}{\sqrt{2}}}$ the cut-off 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, i.e., f^* is IC.[17]

5. Ordinal Pareto Efficiency and Incentive compatibility

In this section, we restrict our attention to ordinal rules for three reasons. First, it is consistent with previous work in the literature. Second, ordinal rules are widely used in practice. Third and most importantly, deriving results within the class of ordinal rules will serve as a benchmark as we explore the relationship between efficiency and incentive compatibility.

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

Definition 5. A rule $f \in F^{ORD}$ is *Ordinally Pareto Efficient* (OPE) if there is no rule $g \in F^{ORD}$ such that $U_i(g) > U_i(f)$ for every agent $i \in N$.

---

[16]Precisely, an H type announcement 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).

[17]The details of the arguments are in Section 6.

The following proposition characterizes OPE rules and shows the relationship between OPE rules and scoring rules.

Proposition 1. *A rule $f \in F^{ORD}$ is OPE if and only if there are numbers $\{\lambda_i\}_{i\in N}$, where $\lambda_i \geq 0$ and $\lambda = (\lambda_1, ..., \lambda_n) \neq \mathbf{0}$ such that f is a scoring rule with $s_i = \lambda_i \mathbb{E}\left(\left(\hat{v}_i^{[k]}\right)_{k=\{1,...,m\}} \mid P_i^{ORD}\right)$ for $i \in N$.*[18]

Proof. We start with the convexity of the utility possibility set of F^{ORD}. The set F^{ORD}, viewed as a subset of linear space $\mathbb{R}^{m\times n}$, is convex because we allow randomized rules. In addition, the mapping from rules to the utility vectors is affine: $U_i(\alpha f + (1-\alpha) g) = \alpha U_i(f) + (1-\alpha) U_i(g)$ for any $f, g \in F^{ORD}$ and any $\alpha \in [0,1]$. Hence, $\left\{\left(U_i(f)_{i\in N}\right) : f \in F^{ORD}\right\}$ is convex. For convex utility possibility sets, it is well known that Pareto efficiency can equivalently be represented by the maximization of linear combinations of utilities with $\{\lambda_i\}_{i\in N}$, $\lambda_i \geq 0$ and $\lambda = (\lambda_1, ..., \lambda_n) \neq 0$ [Holmström and Myerson (1983)].[19] Thus, a rule f is OPE if and only if there are numbers $\{\lambda_i\}_{i\in N}$, where $\lambda_i \geq 0$ and $\lambda = (\lambda_1, ..., \lambda_n) \neq 0$ such that f maximizes the social welfare $\sum_{i\in N} \lambda_i U_i(g)$ among all functions $g \in F^{ORD}$. Fix $\{\lambda_i\}_{i\in N}$ and consider the generalized utilitarian social welfare, $W_\lambda(g) = \sum_{i\in N} \lambda_i U_i(g)$. It is the weighted sum of the expected utilities of all the agents. Given λ and P^{ORD}, then for every $g \in F^{ORD}$ we have

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

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

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

[18]$\mathbb{E}\left(\left(\hat{v}_i^{[k]}\right)_{k=\{1,...,m\}} \mid P_i^{ORD}\right)$ is the conditional expectation of $\left(\hat{v}_i^{[k]}\right)_{k=\{1,...,m\}}$ given the ordinal partition of agent i, i.e., $\mathbb{E}\left(\left(\hat{v}_i^{[k]}\right)_{k=\{1,...,m\}} \mid P_i^{ORD}\right)(v_i) = \mathbb{E}\left(\left(\hat{v}_i^{[k]}\right)_{k=\{1,...,m\}} \mid t_i(v_i)\right)$

[19]For the detail of the argument, see page 560 in Mas-colell et al. (1995) .

In a neutral environment, $\mathbb{E}\left(\hat{v}_i^l \middle| P_i^{ORD}(v_i)\right)$ equals $\mathbb{E}\left(\hat{v}_i^{[r_l]} \middle| P_i^{ORD}(v_i)\right)$, depending only on the ranking of alternative l. It does not depend on the specific ordinal type $t_i(v_i)$. Hence, $\lambda_i \mathbb{E}\left(\hat{v}_i^{[r_l]} \middle| P_i^{ORD}(v_i)\right)$ can function as the score assigned to alternative l when agent i announces v_i, meaning $s_i^{r_l(v_i)} = \lambda_i \mathbb{E}\left(\hat{v}_i^{[r_l]} \middle| P_i^{ORD}(v_i)\right)$. By the definition of a scoring rule, the maximizer f is a scoring rule with $s_i = \lambda_i \mathbb{E}\left(\left(\hat{v}_i^{[k]}\right)_{k=\{1,\ldots,m\}} \mid P_i^{ORD}\right)$ $for\ i \in N$.[20]□

Proposition 1 shows that OPE rules are scoring rules with a particular structure of scores. Notice that the optimal score s_i depends on the welfare weight λ_i and the value distribution conditional on the ordinal partition μ_i. If $\lambda_i = 1$ for every agent i, $W_\lambda(g)$ is the utilitarian social welfare. The following corollary characterizes the optimal ordinal rules in terms of the utilitarian social welfare.

Corollary 1. *A rule $f \in F^{ORD}$ is an Ordinal Utilitarian Rule if and only if f is a scoring rule with $s_i = \mathbb{E}\left(\left(\hat{v}_i^{[k]}\right)_{k=\{1,\ldots,m\}} \mid P_i^{ORD}\right)$ for $i \in N$.*

The Ordinal Utilitarian Rule is denoted by f_{OUR}. The following example is helpful to understand the rule and the relationship between the scores and agents' value announcements.

Example 2. Consider $L = \{a, b, c\}$, $N = \{1, 2\}$, and f_{OUR}. Suppose that $\hat{v}_1^{[2]}$ has a uniform distribution on $[0, 1]$, and $\hat{v}_2^{[2]}$ has a normal distribution on $[0, 1]$ with $\mathbb{E}\left(\left(\hat{v}_i^{[2]}\right) \mid P_i^{ORD}\right) = \frac{2}{3}$. From the definition of f_{OUR}, we find that $s_1 = \left(1, \frac{1}{2}, 0\right)$ and $s_2 = \left(1, \frac{2}{3}, 0\right)$. If agents 1 and 2 announce $v_1 \in V_1^{abc}$ and $v_2 \in V_2^{bca}$, then $\left(s_1^{r_a(v_1)}, s_1^{r_b(v_1)}, s_1^{r_c(v_1)}\right) = \left(1, \frac{1}{2}, 0\right)$ and $\left(s_2^{r_a(v_2)}, s_2^{r_b(v_2)}, s_2^{r_c(v_2)}\right) = \left(0, 1, \frac{2}{3}\right)$. This implies that $f_{OUR}(v_1, v_2) = (0, 1, 0)$ because $\left(s_1^{r_a(v_1)}, s_1^{r_b(v_1)}, s_1^{r_c(v_1)}\right) + \left(s_2^{r_a(v_2)}, s_2^{r_b(v_2)}, s_2^{r_c(v_2)}\right) = (1, \frac{3}{2}, \frac{2}{3})$.

This example also shows that in an environment with asymmetric value distributions across agents, the ordinal utilitarian rule may not be a symmetric scoring rule even with equal welfare weights (i.e., $\lambda_i = 1$ for all i).

We now discuss 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.*

[20] Some parts of this proof are similarly found in the proof of Theorem 1 and 2 in Azrieli and Kim (2014).

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* $g(v) = \frac{1}{m!}\sum_{\sigma\in\phi} f(v^{\sigma})^{\sigma^{-1}} \in F^{ORD}$ *is neutral. Moreover,* $U_i(g) = U_i(f)$ *for every agent* i.

The proof of Lemma 1 is in the Appendix. 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.[21]

Proof of Theorem 1:

Assume a rule f is OPE, by Proposition 1 f is a scoring rule. 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.

Fix $t \in T$. 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 $v_i' \in V_i$, note that $\left(s_i^{r_l(v_i')}\right)_{l\in L} = \left(s_i^{r_l(v_i)}\right)_{l\in L}^{\sigma}$ for some σ. To check 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, consider the aggregated scores of the other agents, $\sum_{j\neq i}\left(s_j^{r_l(v_j)}\right)_{l\in L}$. The probabilities that each alternative obtains the maximum aggregated scores are the same as $\frac{1}{m}$ because the environment is neutral. Second, the true announcement adds $\left(s_i^{r_l(v_i)}\right)_{l\in L}$ to the previous scores in every event, which 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 the second-highest and so on. Also, the order of the probabilities resembles the order of the value announcement, i.e., $\mathbb{E}\left(g(v_i',\hat{v}_{-i})\right) = \mathbb{E}\left(g(v_i,\hat{v}_{-i})\right)^{\sigma}$.[22] 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 $v_i \cdot \mathbb{E}\left(g(v_i,\hat{v}_{-i})\right) \geq v_i \cdot \mathbb{E}\left(g(v_i',\hat{v}_{-i})\right)$. It concludes that g is IC.□

---

[21] 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.

[22] In Example 2 with the standard tie breaking rule such that ties are broken by uniform distribution over the set of maximizers, 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 $\mathbb{E}\left(f(v_1',\hat{v}_2)\right)$.

Theorem 1 implies that there is non-existent the usual trade off between ex-ante Pareto efficiency and incentive compatibility within the class of ordinal rules, 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) describe the optimal ordinal rules based on the utilitarian, maximax, and maximin welfare functions. They use similar but stronger assumptions than the current paper.[23] 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. Obviously, all of their rules are ordinally Pareto efficient. But, they abstract the strategic voting, assuming voters' truthful announcements of their preferences. However, Theorem 1 shows that all of their rules are IC even without the assumption.

In addition, we can find an important feature of the set of OPE rules in Proposition 1. The set of OPE rules is in the class of scoring rules that allows *asymmetric* scores (even zero score vectors) across agents. Apesteguia et al. (2011) show that an optimal rule based on a maximax or maximin welfare function in the symmetric environment may not be a *symmetric* scoring rule, but the approximation of the rule is a symmetric scoring rule. Proposition 1 complements their analysis because the optimal rule can be a scoring rule with $s_i = \mathbf{0}$ for some $i \in N$.

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, [24] but the non-trivial question is whether we can design an *incentive compatible* cardinal rule superior to any ordinal rule.

6.1. **A Superior IC Cardinal Rule.** The following theorem shows the main result of this section, answering the above question.

Theorem 2. *Assume $n \geq 5$ and that μ_i is an identical distribution for all $i \in N$, then there exists an IC cardinal rule that achieves higher utilitarian social welfare than any ordinal rule.*

[23]Specifically, they assume identical value distributions across agents.

[24]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.

Proof of Theorem 2:

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 $\left\{P^{\beta}\right\}_{\beta\in(0,1)}$ 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.

Step 1) 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 $f \in \underset{g\in F^P}{argmax}\, W_{\lambda}(g)$ where $\lambda_i = 1$ for every $i \in N$. The following lemma generalizes Proposition 1 with a general partition, restricting λ to $\mathbf{1}$.[25]

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

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

where

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

The proof of this lemma is omitted because it is nearly identical to the proof of Proposition 1.

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, $T = \left\{t^H, t^L : t \in T^{ORD}\right\}$. 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 $\beta \in (0,1)$.

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

[25]The restriction of λ is not necessary for the generalization of Proposition 1, but is useful for Theorem 1.

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. Roughly speaking, the H type agents hate the third-ranked alternative and L type agents love the first-ranked alternative.

Let f_β be a P^β-Utilitarian Rule and fix the standard tie breaking rule such that ties are broken by uniform distribution over the set of the maximizers.

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 because the value distributions across agents are identical and the environment is neutral.

$$s^H(\beta) = \left(\mathbb{E}\left(\hat{v}_i^{[k]} \mid t_i^\beta(v_i) = t^H\right)\right)_{k=\{1,...,m\}}, \; s^L(\beta) = \left(\mathbb{E}\left(\hat{v}_i^{[k]} \mid t_i^\beta(v_i) = t^L\right)\right)_{k=\{1,...,m\}}$$

$$p^H(\beta) = Pr\left(\left\{v_i : t^\beta(v_i) = t^H\right\}\right), \; p^L(\beta) = Pr\left(\left\{v_i : t^\beta(v_i) = t^L\right\}\right).$$

Step 3) We first consider the three alternative case, 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 to examine 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 (see the proof of Lemma 3 in Appendix). The following lemma identifies the necessary and sufficient condition for incentive compatibility.

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

The proof of the lemma is in Appendix. 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 shows 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, and the proofs are in Appendix.

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

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 provides Example 3 in Appendix that explains the potential discontinuity of $h(\beta)$. Even in this case, we construct a slightly different IC rule. Define the set $D = \{\beta \in (0,1) : h(\beta) \textit{ is discontinuous at } \beta\}$.

Claim 2. *If* $D \neq \phi$, *then there exits a* $\hat{\beta} \in D$ *such that* $\underset{\beta\to\hat{\beta}-}{h(\beta)} \cdot \underset{\beta\to\hat{\beta}+}{h(\beta)} \leq 0$.

The following figure is helpful in understanding the remaining argument.

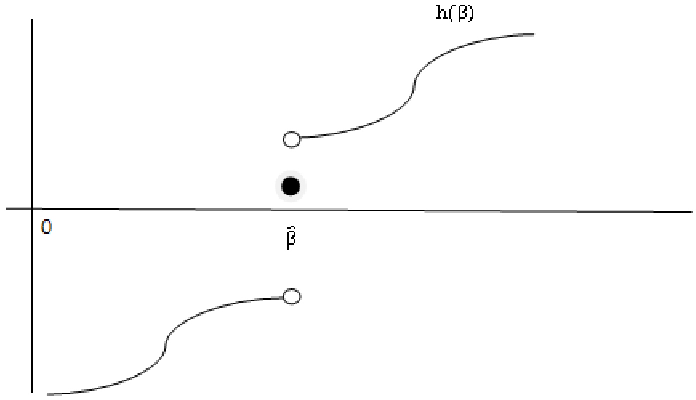

Figure 1. An example of discontinuous $h(\beta)$ at $\hat{\beta}$

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, 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 $\alpha = \frac{-h^{-}(\hat{\beta})}{h^{+}(\hat{\beta}) - h^{-}(\hat{\beta})} \in [0,1]$.*

Step 5) We prove that f^{*} has a higher utilitarian social welfare than the ordinal utilitarian rule, 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, $f^{*}(v) \neq f^{*}(v_i', v_{-i})$ while $f_{OUR}(v) = f_{OUR}(v_i', v_{-i})$ because v and (v_i', v_{-i}) 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 only assign the third-ranked alternative the same score as the g_{OUR}. That is,

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

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

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

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

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, $g_\beta\left(t_i,t_{-i}\right)$ equals $\hat{g}\left(t_i,t_{-i}\right)$. 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.

We should 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 $\mathbb{E}\left(\hat{g}_\beta\left(t,\hat{t}_{-i}\right)\right)=P\left(\beta\right)$ and $\mathbb{E}\left(\hat{g}_\beta\left(t',\hat{t}_{-i}\right)\right)=P\left(\beta\right)'$,

$P\left(\beta\right)^{[k]}=P\left(\beta\right)'^{[k]}$ for $k\geq 4$.

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

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

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

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

The next corollary covers the $n=2$ or $n\geq 4$ case with a restriction of the value distribution μ_i. The proof of the corollary is in Appendix.

Corollary 2. *Assume $n=2$ or $n\geq 4$ and that for all $i\in N$ μ_i is an identical distribution such that $\mathbb{E}\left(\hat{v}_i^{[2]}\mid P_i^{ORD}\right)\leq\frac{\mathbb{E}\left(\hat{v}_i^{[1]}|P_i^{ORD}\right)+\mathbb{E}\left(\hat{v}_i^{[3]}|P_i^{ORD}\right)}{2}$, then there exists an IC cardinal rule which achieves a higher utilitarian social welfare than any ordinal rule.*

When agents are symmetric ex-ante, it is often interesting to focus on anonymous rules. An anonymous rule treats every agent equally.

Definition 6. A rule f is *anonymous* if, for all profiles v and all permutations π over the set of agents,
$f\left(v_{\pi(1)}, ..., v_{\pi(n)}\right) = f(v)$

The following corollary demonstrates a valuable implication of Theorem 1 within the set of anonymous rules.

Corollary 3. *Under the same assumptions as Theorem 2 or Corollary 2, there exists an anonymous and IC cardinal rule that strictly Pareto dominates any anonymous ordinal rule.*

Proof. The restriction of anonymous rules and the identical value distribution across agents implies that $W(f) = nU_1(f)$ and $U_1(f) = U_i(f)$ for all $i \in N$. By Theorem 2, there exists a rule f^* such that $U_i(f^*) > U_i(f_{OUR})$ for all $i \in N$. Additionally, f^* and f_{OUR} are anonymous. □

Corollary 3 may be more interesting for people who dislike the utilitarian social welfare because it depends on the concept of Pareto dominance.

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 which is a permutation of either $(1, B, 0)$ or $(1, A, 0)$. That is, the voter can give a maximum of 1 point to one candidate, A or B ($0 \leq A \leq B \leq 1$) to some other candidate, and a minimum of 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 exists two score vectors $s_i \in \{(1, A, 0), (1, B, 0)\}$ for each $i \in N$ such that

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

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)=$(0,1)$ - the *approval* voting rule where each voter can support or oppose a group of candidates - requires more than ordinal preferences information, which is similar to P^{β}-Utilitarian Rules. The following proposition shows the relationship between P^{β}-Utilitarian Rules, f^* and (A,B)-scoring rules.

Proposition 2. *Assume that μ_i is an identical distribution for all $i \in N$. Then, 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 $(A,B) = \left(s^L(\beta^*)^2, s^H(\beta^*)^2\right)$*

The proof is in Appendix. Proposition 2 shows that P^{β}-Utilitarian Rules are not purely theoretical rules because (A,B) scoring rules are widely used in practice for the simple structure. When we implement (A,B)-scoring rules, Proposition 2 helps finding an IC (A,B)-scoring rule superior to any ordinal rule.

7. Concluding Comments

We have investigated the efficiency and incentive compatibility of voting rules in a Bayesian environment with independent private values and at least three alternatives. First, we characterize the ex-ante Pareto frontier of the set of ordinal rules. Furthermore, we prove that in a neutral environment, essentially any ex-ante Pareto efficient ordinal rule is IC, which implies that the traditional conflict between efficiency and incentive compatibility does not exist in the class of ordinal rules. This conflict, however, arises if we consider cardinal rules. Also, it is not straightforward to design a cardinal rule which is more efficient than ordinal rules as well as IC. However, we successfully construct an IC cardinal rule superior to any ordinal rule. With three alternatives, this rule turns out to be an IC (A,B)-scoring rule in Myerson (2002).

We do not attempt to derive an optimal cardinal rule subject to incentive compatibility because an analytical characterization of this second best rule is difficult to obtain. But 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 steps to find this second-best rule. In addition, the method of using a finer partition and finding the 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 fields such as allocation or matching rules. We hope to study these directions in future.

Appendix: Proofs

Proof in Section 5

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

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

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 from the fact that $\{\sigma^*(\sigma) : \sigma \in \phi\} = \phi$.

Moreover, $U_i(f) = \mathbb{E}(\hat{v}_i \cdot f(\hat{v})) = \frac{1}{m!}\sum_{t\in T}\mathbb{E}(\hat{v}_i \mid v_i \in V_i^t)\cdot\mathbb{E}(f(\hat{v}) \mid v_i \in V_i^t)$, where the second equality follows from the neutral environment. To see the connection between f and g, we change the expression of $U_i(f)$ with permutations. Fix $t \in T^{ORD}$,

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

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 from the neutral environment.

Since this formula holds for any t and any i,

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

Proofs and Example 3 in Section 6

Proof of Lemma 3. 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 by 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 or her true type.

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

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

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

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',

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

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

Proof of Lemma 4. 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 $0 < \beta_2 < \beta_1 < 1$,

$$s^H\left(\beta_1\right)^3 = \int_0^1 \int_{v^{[m]}}^1 \dots \int_{v_i^{[4]}}^1 \int_{v_i^{[3]}}^1 \int_{v_i^{[4]}}^{\frac{v_i^{[2]}-\beta_1}{1-\beta_1}} v_i^{[3]} \mu_i\left(v_i^{[1]}, ..., v_i^{[m]}\right) dv_i^{[3]} dv_i^{[2]} ... dv_i^{[m]} / p^H\left(\beta_1\right),$$
$$s^H\left(\beta_2\right)^3 = \int_0^1 \int_{v^{[m]}}^1 \dots \int_{v_i^{[4]}}^1 \int_{v_i^{[3]}}^1 \int_{v_i^{[4]}}^{\frac{v_i^{[2]}-\beta_2}{1-\beta_2}} v_i^{[3]} \mu_i\left(v_i^{[1]}, ..., v_i^{[m]}\right) dv_i^{[3]} dv_i^{[2]} ... dv_i^{[m]} / p^H\left(\beta_2\right).$$

Since $\frac{v_i^{[2]}-\beta_1}{1-\beta_1} \leq \frac{v_i^{[2]}-\beta_2}{1-\beta_2}$ and $p^H\left(\beta_2\right) > p^H\left(\beta_1\right)$, for any $\alpha \geq 0$,
$Pr\left[v_i^{[3]} \leq \alpha \mid v_i \in V_i^{t^H}\left(\beta_2\right)\right] \leq Pr\left[v_i^{[3]} \leq \alpha \mid v_i \in V_i^{t^H}\left(\beta_1\right)\right]$

By the first order stochastic dominance, $s^H\left(\beta_2\right)^3 \geq s^H\left(\beta_1\right)^3$. It means that $s^H\left(\beta\right)^3$ are (weakly) decreasing in β.

Similarly,

$$s^L\left(\beta_1\right)^3 = \int_0^1 \int_{v^{[m]}}^1 \dots \int_{v_i^{[4]}}^1 \int_{v_i^{[3]}}^1 \int_{\frac{v_i^{[2]}-\beta_1}{1-\beta_1}}^{v_i^{[2]}} v_i^{[3]} \mu_i\left(v_i^{[1]}, ..., v_i^{[m]}\right) dv_i^{[3]} dv_i^{[2]} ... dv_i^{[m]} / p^L\left(\beta_1\right)$$
$$s^L\left(\beta_2\right)^3 = \int_0^1 \int_{v^{[m]}}^1 \dots \int_{v_i^{[4]}}^1 \int_{v_i^{[3]}}^1 \int_{\frac{v_i^{[2]}-\beta_2}{1-\beta_2}}^{v_i^{[2]}} v_i^{[3]} \mu_i\left(v_i^{[1]}, ..., v_i^{[m]}\right) dv_i^{[3]} dv_i^{[2]} ... dv_i^{[m]} / p^L\left(\beta_2\right).$$

Since $\frac{v_i^{[2]}-\beta_1}{1-\beta_1} \leq \frac{v_i^{[2]}-\beta_2}{1-\beta_2}$ and $P^L\left(\beta_2\right) < P^L\left(\beta_1\right)$, for any $\alpha \geq 0$,
$Pr\left[v_i^{[3]} \geq \alpha \mid v_i \in V_i^{t^L}\left(\beta_2\right)\right] \geq Pr\left[v_i^{[3]} \geq \alpha \mid v_i \in V_i^{t^L}\left(\beta_1\right)\right]$.

By the first order stochastic dominance, $s^L\left(\beta_2\right)^3 \geq s^L\left(\beta_1\right)^3$.□

Proof of Claim 1. For $\lim_{\beta\to 0} h(\beta) = \lim_{\beta\to 0} P(\beta)^{[1]} - P(\beta)'^{[1]} < 0$, we know that at 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 $g_\beta(t_i', t_{-i}) = (1, 0, 0)$. 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.[26] Since all other types are H types ($\lim_{\beta\to 0} p^H(\beta) = \frac{1}{m!}$), $\lim_{\beta\to 0} P(\beta)^{[1]} - P(\beta)'^{[1]} < 0$

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.[27] We can always find the case where $g_\beta(t_i', t_{-i}) = (0, \frac{1}{2}, \frac{1}{2})$ or $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ when everyone's type is L type and $g_\beta(t_i, t_{-i}) = (0, 1, 0)$. For example, for $n = 3$, $t_1' = abc^L$, $t_2 = bca^L$ and $t_3 = cab^L$ and for $n = 5$, the above t_1',t_2,t_3 with $t_4 = bca^L$ and $t_3 = cba^L$. Generally, for any $n = 3$ or $n \geq 5$, we can use the above two profiles to find the case.[28] Since all other types are L types ($\lim_{\beta\to 1} p^L(\beta) = \frac{1}{m!}$), $\lim_{\beta\to 1} P(\beta)'^{[3]} - P(\beta)^{[3]} > 0$.

Note that Claim 1 still holds with more than 3 alternatives, considering that the shared top 3 alternatives are replaced with the 3 alternates in the above cases and g_β is replaced with $\hat{g}_\beta$ constructed in Step 6.□

Example 3. For ease of notation, we use the vector of scores assigned to the alternatives which depends on β and the announced types, $s_i(\beta, t_i) = \mathbb{E}\left(\hat{v}_i \mid t_i^\beta(v_i) = t_i\right)$. Given $L = \{a, b, c\}$ and $t_i = abc^H$, then $g_\beta(t_i, t_{-i})$ is determined by the aggregated scores of the alternatives which depends on β and t_{-i}. Let $S^H(\beta, t_{-i}) = \left(\left(s^H(\beta)\right) + \sum_{j\neq i} s_j(\beta, t_j)\right)$ be the vector of aggregated scores assigned to the alternatives, then we can express $P(\beta)^{[1]}$ with this function

[26]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.

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

[28]We, however, may not find the case when $n = 2$ and $n = 4$. For $n = 2$, the possible case where $g_\beta(t_i', t_{-i})^c > 0$ is that $t_1' = abc^L$ and $t_2 = cba^L$. But , in some distributions such that $s^L(\beta)^2 > \frac{1+s^L(\beta)}{2}$ for any $\beta \in (0, 1)$, $g_\beta(t_i', t_{-i}) = (0, 1, 0)$= $g_\beta(t_i, t_{-i})$. The argument for $n = 4$ is analogous.

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

$P(\beta)^{[2]}$ 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.

Proof of Claim 2. Suppose not (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 $h(\beta)=0$, contradiction again.□

Proof of Claim 3. 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 $g_{\hat\beta}$. That 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 Example 2. With $P^{\hat\beta}$, we obtain the balance function of g^+, $h^+\left(\hat{\beta}\right)=\underset{\beta\to\hat{\beta}+}{h(\beta)}$. Because of 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[29] $\alpha=\frac{-h^-\left(\hat{\beta}\right)}{h^+\left(\hat{\beta}\right)-h^-\left(\hat{\beta}\right)}\in[0,1]$ 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. In order to compare with ordinal rules, we will use the identical function, $f^*(v)=\tilde{g}_\alpha(t_1,...,t_n)$ in the next step.□

Proof of Corollary 2. Every step is the same as 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 $g_\beta(t_i',t_{-i})=(\frac{1}{2},0,\frac{1}{2})$ and $g_\beta(t_i,t_{-i})=(0,1,0)$ as β closes to 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 1. 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. But, in the other case f^* does not look like an (A,B)-scoring rule because of the convex combination of the two rules. However, 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. □

References

[1] Apesteguia, J., M.A. Ballester, and R. Ferrer (2011), On the justice of decision rules, *Review of Economic Studies,* **78**, 1-16.

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

[2] Azrieli, Y. and S. Kim (2014), Pareto efficiency and weighted majority rules, *International Economic Review*, **55**, 1067-1088.

[3] Barberà, S. and M. O. Jackson (2006), On the weights of nations: Assigning voting weights in a heterogeneous union, *The Journal of Political Economy,* **114**, 317-339.

[4] Börgers, T. and P. Postl (2009), Efficient compromising, *Journal of Economic Theory,* **144**, 2057-2076.

[5] Carrasco, V., W. Fuchs (2009), A Procedure for Taking Decisions with Non-transferable Utility, Working paper.

[6] Freixas, X. (1984), A Cardinal Approach to Straightforward Probabilistic Mechanisms, *Journal of Economic Theory*, **34**, 227-251.

[7] Gershkov, A., B. Moldovanu, X. Shi (2014), Optimal voting rules, Working paper.

[8] Gibbard, A. (1973), Manipulation of Voting Schemes: A General Result, *Econometrica,* **41**, 587-601.

[9] Gibbard, A. (1978), Straightforwardness of Game forms with Lotteries as Outcomes, *Econometrica,* **46**, 595-613.

[10] Giles, A. and P. Postl (2014), Equilibrium and Effectiveness of Two-Parameter Scoring Rules, *Mathematical Social Science*, **68**, 31-52.

[11] Harsanyi, J. C. (1955), Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility, *The Journal of Political Economy*, **63**, 309-321

[12] Holmström, B., and R. B. Myerson (1983), Efficient and durable decision rules with incomplete information, *Econometrica,* **51**, 1799-1819.

[13] Jackson, M. O. and H. F. Sonnenschein (2007), Overcoming incentive constraints by linking decisions, *Econometrica,* **75**, 241-257.

[14] Majumdar, D. and A. Sen (2004), Ordinally Bayesian incentive compatible voting rules. *Econometrica*, **72**, 523-540.

[15] Mas-colell, A., M. D. Whinston, and J. R. Green (1995), Microeconomic Theory, Oxford University Press.

[16] Miralles, A. (2012), Cardinal Bayesian allocation mechanisms without transfers, *Journal of Economic Theory*, **147**, 179-206.

[17] Moulin, H. (1980), On strategy-proofness and single peakedness, *Public Choice*, **35**, 437-455.

[18] Myerson, R. B. (2002), Comparison of Scoring Rules in Poisson Voting Games, *Journal of Economic Theory*, **103**, 219-251.

[19] Roberts, K. W. S. (1980), Interpersonal Compatibility and Social Choice Theory, *Review of Economic Studies*, **47**, 421-439

[20] Satterthwaite, M. A. (1975), Strategy-proofness and arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions, *Journal of Economic Theory*, **10**, 187-217.

[21] Schmitz, P. W. and T. Tröger (2012), The (sub-)optimality of the majority rule, *Games and Economic Behavior*, **74**, 651-665.

[22] Strotz, R. H. (1953), Cardinal Utility, *American Economic Review*, **43**, 384-397