ORDINAL BAYESIAN INCENTIVE COMPATIBILITY WITH RESPECT TO CORRELATED BELIEFS*

Abhigyan Bose†1 and Souvik Roy‡2

[1,2]Economic Research Unit, Indian Statistical Institute, Kolkata

July, 2022

Abstract

We consider social choice functions (SCFs) that are locally robust ordinal Bayesian incentive compatible (LOBIC) with respect to correlated prior beliefs. We assume the coexistence of both positively and negatively correlated priors, as well as independent priors. We model such priors using a betweenness property. We prove the following two results in the framework: (i) an SCF is LOBIC with respect to some prior only if it satisfies the sequential ordinal non-domination property, and (ii) if an SCF satisfies the strong ordinal non-domination property, then it is LOBIC with respect to some prior that is correlated under betweenness relation. Finally, we discuss how our results can be generalized for arbitrary priors modeled using binary relations.

JEL Classification: D71, D82.

Keywords: ordinal Bayesian incentive compatibility; correlated beliefs; betweenness property; ordinal non-domination property;

*We are grateful to Arunava Sen for suggesting us the problem and for having helpful discussions on the paper.

†Contact: boseabhigyan@gmail.com

‡Corresponding Author, Contact: souvik.2004@gmail.com

1. Introduction

A social choice function (SCF) selects an alternative at every collection of preferences of the agents in a society. An SCF is called ordinal Bayesian incentive compatible (OBIC) with respect to a prior (belief) if, by misreporting his sincere preference, no agent can increase his expected utility according to his prior for any utility function representing his sincere preference. An SCF is called locally robust OBIC (LOBIC) with respect to a prior if it is OBIC with respect to all priors lying in a small neighborhood of the original prior. LOBIC ensures that agents are incentivized to reveal their sincere preferences even if the designer is *slightly* unsure about their beliefs. The objective of this paper is to explore the structure of SCFs that are LOBIC with respect to correlated priors. Relevance of correlated priors in mechanism design is well-established in the literature (see Bhargava et al. (2015), Yamashita (2018), Laffont and Martimort (2000), Albert et al. (2017a), Albert et al. (2017b) for details).

We use the notion of 'betweenness' to formulate the correlation structure of the priors. For an illustration, suppose that there are two agents 1 and 2, and three alternatives a, b and c. Suppose further that agent 1 has the preference abc.[1] Note that the preference bac differs from abc by the swap of a and b, whereas the preference bca differs from abc by the swap of both a and b, and a and c. In some sense, the preference bca is "farther away" from abc than bac is from abc. In such situations, we say that the preference bac lies between abc and bca. Now, suppose that agent 1 believes that agent 2 is of his type, that is, has a preference that is similar to his preference. Positive correlation under betweenness relation says that agent 1 will deem the preference abc more likely than the preference bac, and the preference bac more likely than the preference bca. On the other hand, if agent 1 believes that agent 2 is of his opposite type, then negative correlation under betweenness relation imposes exactly opposite structure on his belief about agent 2's preference.

[1] By abc we denote a preference where a is first-ranked, b is second-ranked, and c is third-ranked.

In this paper, we assume the coexistence of both positively correlated and negatively correlated beliefs, as well as independent beliefs. Each agent classifies other agents into three categories: the positively correlated agents, the negatively correlated agents, and the independent agents, and forms his belief accordingly.

We first consider the question whether every SCF can be LOBIC with respect to some (correlated or uncorrelated) prior. We show that the answer is "no" and consequently provide a condition on an SCF that is necessary for it to be LOBIC with respect to some prior. Next, we focus on correlated priors (under betweenness relations) and provide a sufficient condition on an SCF for it to be LOBIC with respect to such priors. Finally, we provide a discussion on how our results can be generalized for arbitrary correlated priors.

The structure of LOBIC SCFs when priors are independent is well-explored in the literature. Majumdar and Sen (2004) show that if the domain of preferences is unrestricted, then under unanimity, the notion of OBIC (in the context of incomplete information) and that of dominant strategy incentive compatibility (DSIC) (in the context of complete information) are almost surely equivalent.[2,3] Mishra (2016) extends this result for arbitrary domains under an additional assumption called elementary monotonicity. Karmokar and Roy (2018) generalize these results for RSCFs and additionally provide some more structure of OBIC RSCFs on various domains.

In contrast to the case of independent priors, to the best of our knowledge the structure of LOBIC SCFs with respect to correlated priors is relative less explored. Only paper we know in this topic is Bhargava et al. (2015). Our paper improves their result in the following ways.

(i) Bhargava et al. (2015) assume that *all* the agents are positively correlated with each other. In reality, agents might as well be negatively correlated. As we have explained earlier, we allow for the coexistence of positive correlation, negative correlation, and independence in our analysis.

[2]That is for a class of independent priors with Lebesgue measure 1.

[3]Unanimity implies that whenever all agents agree on their top-ranked alternative, that alternative is chosen by the SCF.

(ii) Bhargava et al. (2015) use the notion of top-set correlation to measure positive correlation. To the best of our understanding, it may not be an appropriate measure for positive correlation. To see this, suppose that there are two agents 1 and 2, and three alternatives a, b and c. Assume that agent 1 has the preference abc. Suppose that he believes agent 2 has the preference abc with probability slightly higher than 50%, say 50.01%, and the preference cba with probability slightly lower than 50%, say 49.99%. Note that while the opposite preference cba is believed to be highly likely (with close to 50% probability), preferences such as bac or acb that are closer to abc are believed to be impossible. According to the notion of top-set correlation, such a belief is positively correlated. However, we feel that it violates the basic intuition of positive correlation: preferences such as bac and acb should receive higher probability than cba. As discussed earlier, our notion of priors under betweenness relations rules out such possibilities.

(iii) Theorem 1 in Bhargava et al. (2015) says that a unanimous SCF is LOBIC with respect to a prior if and only if it satisfies a property called ordinal nondomination (OND). It follows from our result that the "only if" part of this theorem is not correct.[4]

The rest of the paper is organized as follows. Section 2 introduces the basic model. Section 3 and Section 4 present a necessary and a sufficient condition for LOBIC, respectively. Section 5 provides a discussion on how our results can be generalized for arbitrary priors. All proofs are collected in the appendix.

2. Preliminaries

Let A be a set of alternatives with $|A| = m$. A preference P is defined as a complete, transitive, and antisymmetric binary relation on A. We denote the weak part of a preference P by R, that is, for two alternatives a and b, aRb means

[4]In the proof of Theorem 1 in Bhargava et al. (2015), the authors consider two cases. However, to our understanding, there is a third case that the authors have missed.

either aPb or $a = b$. We denote by $\mathcal{P}(A)$ the set of all preferences on A. A domain $\mathcal{D}$ (of preferences on A) is a subset of $\mathcal{P}(A)$.

Let P be a preference and $k \in \{1, \ldots, m\}$. We denote the k-th ranked alternative in P by $P(k)$. The upper contour set of top k alternatives is defined as $U_k(P) := \cup_{l \leq k} P(l)$, and the upper contour set of an alternative a is defined as $U(a, P) := \{b \in A \mid bRa\}$. Note that $U(a, P)$ contains the alternative a itself.

To minimize notations, sometimes we do not use brackets for singleton sets. Each agent $i \in N$ has a domain of (admissible) preferences $\mathcal{D}_i$. We denote by $\mathcal{D}_N$ the product set $\prod\limits_{i \in N} \mathcal{D}_i$, and for $i \in N$, we denote by $\mathcal{D}_{-i}$ the set $\prod\limits_{j \neq i} \mathcal{D}_j$. An element $P_N = (P_1, \ldots, P_n)$ of $\mathcal{D}_N$ is called a preference profile. For a preference profile P_N, we denote the restriction of P_N to $N \setminus \{i\}$ by P_{-i}.

A belief μ_i of an agent i is a probability distribution on $\mathcal{D}_N$. Note that by considering beliefs that are probability distributions on $\mathcal{D}_N$ (and not on $\mathcal{D}_{-i}$), we allow for the possibility that a belief of an agent i may depend on his own preference P_i. A collection $\mu_N = (\mu_1, \ldots, \mu_n)$ of beliefs is called a prior (profile). For a prior μ_i and a preference P_i of agent i, we denote by $\mu_i(. \mid P_i)$ the conditional belief (conditional probability distribution on $\mathcal{D}_{-i}$ given P_i) of i given the preference P_i.

A utility function is a mapping $u : A \to \mathbb{R}$. A utility function $u : A \to \mathbb{R}$ is said to represent a preference P if for all $a, b \in A$, we have aPb if and only if $u(a) > u(b)$.

A social choice function (SCF) (on $\mathcal{D}_N$) is a mapping $f : \mathcal{D}_N \to A$.

Consider an SCF f, a prior μ_N, and an agent i. Suppose that the (sincere) preference of agent i is P_i. Fix a utility function u_i of agent i that represents his preference P_i. Then, agent i's expected utility according to his conditional belief $\mu_i(. \mid P_i)$ is given by

$$\sum_{P_{-i} \in \mathcal{D}_{-i}} u_i(f(P_i, P_{-i}))\mu_i(P_{-i} \mid P_i).$$

An SCF is called ordinal Bayesian incentive compatible (OBIC) with respect to a prior if no agent can increase his expected utility (according to his belief conditional on his sincere preference) with respect to any utility function representing his preference by misreporting his sincere preference.

Definition 2.1. An SCF f is **ordinal Bayesian incentive compatible (OBIC)** with respect to a prior μ_N if for all $i \in N$, all $P_i, P_i' \in \mathcal{D}_i$, and all utility functions u_i representing P_i, we have

$$\sum_{P_{-i} \in \mathcal{D}_{-i}} u_i(f(P_i, P_{-i}))\mu_i(P_{-i} \mid P_i) \geq \sum_{P_{-i} \in \mathcal{D}_{-i}} u_i(f(P_i', P_{-i}))\mu_i(P_{-i} \mid P_i).$$

An equivalent definition of OBIC can be given by means of stochastic dominance. It says that no agent can increase the expected probability of any upper contour set of his sincere preference by misreporting his sincere preference.

Definition 2.2. An SCF f is OBIC with respect to a prior μ_N if for all $i \in N$, all $P_i, P_i' \in \mathcal{D}_i$, and all $a \in A$, we have

$$\sum_{P_{-i} \mid f(P_i, P_{-i}) \in U(a, P_i)} \mu_i(P_{-i} \mid P_i) \geq \sum_{P_{-i} \mid f(P_i', P_{-i}) \in U(a, P_i)} \mu_i(P_{-i} \mid P_i).$$

An SCF f is is locally robust with respect to a prior μ_N if it is OBIC with respect to all priors in some (small) open neighborhood of μ_N. In other words, a locally robust OBIC continues to be OBIC if the social planner makes "small" mistakes in estimating agents' priors. To model this, we need the notion distance between two priors: the distance between μ_N and $\bar{\mu}_N$ is defined as $\sum_{i \in N} \sum_{P_N \in \mathcal{D}_N} (\mu_i(P_N) - \bar{\mu}_i(P_N))^2$. We denote by $B_\epsilon(\mu_N)$ the open ball of radius ϵ centered at μ_N, that is, the set of all priors $\bar{\mu}_N$ that are at most ϵ distance from μ_N.

Definition 2.3. An SCF f is **locally robust OBIC (LOBIC)** with respect to a prior μ_N if there exists $\epsilon > 0$ such that f is OBIC with respect to all priors in $\bar{\mu}_N \in B_\epsilon(\mu_N)$.

We illustrate the notion of LOBIC by means of the following example.

Example 2.1. Suppose that there are two agents $\{1, 2\}$ and three alternatives $\{a, b, c\}$. We denote by abc the preference where a, b, and c are the top-ranked, second-ranked, and third-ranked alternatives, respectively. In Table 1, we present an SCF, say f, and in Table 2 and Table 3 we present the conditional beliefs μ_1 and μ_2 of agent 1 and agent 2, respectively. These tables are self-explanatory.

In what follows, we argue that f is LOBIC with respect to $\mu_N = (\mu_1, \mu_2)$. We follow Definition 2.2 for this. Suppose that the sincere preference of agent 1 is $P_1 = abc$. Suppose further that he considers a misreport as $P_1' = acb$. Note that his conditional belief $\mu_1(. \mid abc)$ at $P_1 = abc$ is given in the first row of Table 2. Further note that the (non-trivial) upper contour sets of the preference abc are $\{a\}$ and $\{a, b\}$. The believed (through $\mu_1(. \mid abc)$) probability that the outcome lies in the upper contour set $\{a\}$ (that is, the outcome is a) when 1 reports abc is $\mu_1(abc \mid abc) + \mu_1(acb \mid abc) + \mu_1(bca \mid abc) = 0.25 + 0.25 + 0.20 = 0.70$. Similarly, the believed (through $\mu_1(. \mid abc)$, and not through $\mu_1(. \mid acb)$) probability that the outcome is a when 1 misreports his preference as acb is $\mu_1(abc \mid abc) + \mu_1(acb \mid abc) + \mu_1(bac \mid abc) + \mu_1(cba \mid abc) = 0.25 + 0.25 + 0.04 + 0.06 = 0.60$. Since $\mu_1(\{P_2 | f(P_1, P_2) = a\} \mid abc) \geq \mu_1(\{P_2 | f(P_1', P_2) = a\} \mid abc)$, we have that the requirement of OBIC is satisfied for this instance. One can verify that f satisfies this requirement for other instances as well. Furthermore, one can check that all these requirements will be satisfied if we slightly perturb the prior (profile) μ_N, ensuring that the SCF remains LOBIC.

$1 \setminus 2$	abc	acb	bac	bca	cab	cba
abc	a	a	c	a	b	b
acb	a	a	a	b	c	a
bac	b	a	b	b	a	c
bca	c	b	b	b	a	c
cab	a	a	b	c	c	c
cba	a	c	b	a	c	c

Table 1: Example of an SCF that does not satisfy OND but satisfies LOBIC

$1 \setminus 2$	abc	acb	bac	bca	cab	cba
abc	0.25	0.25	0.04	0.20	0.20	0.06
acb	0.25	0.25	0.20	0.01	0.09	0.20
bac	0.20	0.15	0.25	0.25	0.14	0.01
bca	0.15	0.20	0.25	0.25	0.01	0.14
cab	0.15	0.14	0.01	0.20	0.25	0.25
cba	0.01	0.25	0.23	0.01	0.25	0.25

Table 2: Conditional belief of Agent 1

$1 \setminus 2$	abc	acb	bac	bca	cab	cba
abc	0.25	0.25	0.01	0.01	0.01	0.09
acb	0.25	0.25	0.09	0.25	0.25	0.01
bac	0.09	0.20	0.25	0.25	0.15	0.20
bca	0.01	0.01	0.25	0.25	0.09	0.20
cab	0.20	0.20	0.20	0.23	0.25	0.25
cba	0.20	0.09	0.20	0.01	0.25	0.25

Table 3: Conditional belief of Agent 2

3. A NECESSARY CONDITION FOR LOBIC WITH RESPECT TO A(NY) CORRELATED PRIOR

In this section, we provide a necessary condition for an SCF to be LOBIC. Our necessary condition uses the notion of sequential ordinal non-domination (sequential OND). Bhargava et al. (2015) introduce the notion of ordinal non-domination (OND), sequential OND is a modification of that.

First, we present the notion of OND. For an illustration, consider an agent i with (sincere) preference P_i and a preference profile P_{-i} of all agents except i. Suppose that agent i can "manipulate" by misrepresenting his preference as P'_i when others have the profile P_{-i}, that is, suppose $f(P'_i, P_{-i}) P_i f(P_i, P_{-i})$. Informally speaking, OND says that there must be another preference profile P'_{-i} of the agents in $N \setminus \{i\}$ at which such a "manipulated gain" is lost. More formally, there must be P'_{-i} such that for i (i) reporting his sincere preference P_i against P'_{-i} is weakly better than the outcome i obtained by manipulating, that is, $f(P_i, P'_{-i}) R_i f(P'_i, P_{-i})$, and (ii) the (sincere) outcome $f(P_i, P_{-i})$ is weakly better than the outcome obtained by misreporting P'_i against P'_{-i}, that is, $f(P_i, P_{-i}) R_i f(P'_i, P'_{-i})$.

Definition 3.1. An SCF $f : \mathcal{D}_N \to A$ satisfies the **ordinal nondomination (OND)** property if for all $i \in N$, all $P_i, P'_i \in \mathcal{D}_i$, and all $P_{-i} \in \mathcal{D}_{-i}$ such that $f(P'_i, P_{-i}) P_i f(P_i, P_{-i})$, there exists $P'_{-i} \in \mathcal{D}_{-i}$ with the property that $f(P_i, P'_{-i}) R_i f(P'_i, P_{-i})$ and $f(P_i, P_{-i}) R_i f(P'_i, P'_{-i})$.

In what follows, we argue by means of Example 2.1 that the OND property is not necessary for LOBIC. Consider the LOBIC SCF f in Example 2.1. Consider $P_1 = abc$, $P'_1 = acb$, and $P_2 = bac$. We have $f(P'_1, P_2) P_1 f(P_1, P_2)$. However, there is no P'_2 such that $f(P_1, P_2) R_1 f(P'_1, P'_2)$ and $f(P_1, P'_2) R_1 f(P'_1, P_2)$. Therefore, f does not satisfy the OND property.

Note that f satisfies the requirement of OND for all other situations. For instance, when $P_1 = acb$, $P'_1 = abc$ and $P_2 = bca$, the requirement of OND is satisfied by taking $P'_2 = cba$.

In view of Example 1, we modify the OND property as sequential OND. In contrast to the OND property where a gain of agent i by manipulation can be paid back at *exactly one* preference profile P'_{-i}, in case of sequential OND the same can happen through a sequence of preference profiles $(P^1_{-i}, \ldots, P^k_{-i})$ for some $k \geq 1$. Note that OND is a special case of sequential OND where the length of the sequence is 1.

Definition 3.2. For an SCF $f : \mathcal{D}_N \to A$ and a pair of distinct preferences (P_i, P'_i) in $\mathcal{D}_i$, a sequence $(P^1_{-i}, \ldots, P^k_{-i})$ of elements of $\mathcal{D}_{-i}$ is called an **OND sequence** for f with respect to (P_i, P'_i) if for all $l = 1, \ldots, k$, we have $f(P_i, P^l_{-i}) P_i f(P'_i, P^l_{-i})$ and $f(P_i, P^{l+1}_{-i}) R_i f(P'_i, P^l_{-i})$, where $P^{k+1}_{-i} = P^k_{-i}$.

Definition 3.3. An SCF $f : \mathcal{D}_N \to A$ satisfies the **sequential OND** property if for all $i \in N$, all $P_i, P'_i \in \mathcal{D}_i$, and all $P_{-i} \in \mathcal{D}_{-i}$ with $f(P'_i, P_{-i}) P_i f(P_i, P_{-i})$, there exists an OND sequence $(P^1_{-i}, \ldots, P^k_{-i})$ for f with respect to (P_i, P'_i) such that $f(P_i, P_{-i}) R_i f(P'_i, P^k_{-i})$ and $f(P_i, P^1_{-i}) R_i f(P'_i, P_{-i})$.

In what follows, we argue that the SCF f in Table 1 satisfies the sequential OND property. Recall that f satisfies the OND property (and hence, the sequential OND property) for all situations except the one where $P_1 = abc$, $P_1' = acb$, and $P_2 = bac$. Consider the sequence $(P_2^1 = bca, P_2^2 = cab)$ of preferences of agent 2. Note that (i) $f(P_1, P_2^1) R_1 f(P_1', P_2)$, (ii) $f(P_1, P_2) R_1 f(P_1', P_2^2)$, and (iii)$f(P_1, P_2^2) R_1 f(P_1', P_2^1)$. Thus, f satisfies the sequential OND property.

Now, we present the main theorem of this section. It says that sequential OND is a necessary condition for an SCF to be LOBIC with respect to any prior.

Theorem 3.1. *An SCF is LOBIC with respect to some prior only if it satisfies the sequential OND property.*

The proof of this theorem is relegated to Appendix A.

4. A SUFFICIENT CONDITION FOR LOBIC WITH RESPECT TO PRIORS SATISFYING BETWEENNESS RELATIONS

In this section, we consider correlated (both positively and negatively) beliefs and provide a sufficient condition for an SCF to be LOBIC with respect to such priors.

Bhargava et al. (2015) introduced the notion of top-set (TS) correlation as a measure of positive correlation. It says that every agent i with preference P_i conditionally believes that for every $k = 1, \ldots, m-1$, other agents are likely to have preferences with the same set of top k alternatives as him.

Definition 4.1. A belief μ_i of agent i is top-set correlated if for all P_i, all $k = 1, \ldots, m-1$, and all $B \subsetneq A$ such that $B \neq U_k(P_i)$ and $|B| = k$, we have

$$\sum_{P_{-i} | U_k(P_j) = U_k(P_i) \;\; \forall j \neq i} \mu_i(P_{-i} \mid P_i) > \sum_{P_{-i} | U_k(P_j) = B \;\; \forall j \neq i} \mu_i(P_{-i} \mid P_i).$$

In what follows, we argue that top-set correlation might not be the right measure for positive correlation. Suppose that there are three alternatives $a, b,$ and c and two agents 1 and 2. Suppose further that agent 1 has preference

$P_1 = abc$. If agent 1 believes that agent 2 is of his type (that is, positively correlated with him), then he must deem preferences that are "closer" to abc more likely for 2 than the ones that are "farther away" from abc. For instance, $\mu_1(bac \mid P_1)$ and $\mu_1(acb \mid P_1)$ should be bigger than $\mu_1(cba \mid P_1)$. However, this is not ensured by the definition of top-set correlation. For instance, beliefs such as $\mu_1(abc \mid P_1) = 0.5 + \epsilon$ and $\mu_1(cba \mid P_1) = 0.5 - \epsilon$, where $\epsilon > 0$ is arbitrarily small, satisfy top-set correlation but do not really represent positive correlation.[5] At the extreme, note that for any preference P_1 (over any number of alternatives), a belief of agent 1 that gives positive probability to only P_1 and its opposite preference, with arbitrarily small higher probability to P_1, satisfies top-set correlation. However, according to such a belief, any preference which is obtained through a small change in P_1 gets much lower probability (in fact zero probability) than the preference that is completely opposite of P_1, contradicting the intuition behind positive correlation.

In view of the above discussion, we introduce a new notion of correlated belief. Our notion is based on the notion of betweenness. A preference P is said to lie between two preferences P^1 and P^2, denoted by $P \in (P^1, P^2)$, if $P^1 \Delta P \subseteq P^1 \Delta P^2$, where $P \Delta P' = \{\{x, y\} \mid xPy \text{ and } yP'x\}$ denotes the set of (unordered) pairs of alternatives whose relative orderings are different in P and P'. In other words, P lies between P^1 and P^2 if P is "more similar" to P^1 than P^2 is to P^1.

A prior of an agent i is positively correlated with respect to the betweenness property if whenever agent i has sincere preference P_i, he believes that his opponents' preferences are more likely to be "closer" (with respect to the betweenness relation) to P_i. The notion of negatively correlated priors with respect to the betweenness property is defined in a symmetrically opposite manner: an agent i believes that his opponents' preferences are more likely to be "farther away" from P_i.

We assume that for every agent $i \in N$, there exists a *fixed* (does not depend on the preferences of i) partition $\{N_i^0, N_i^+, N_i^-\}$ of $N \setminus \{i\}$ such that for any

[5] The opposite preference P' of a preference P is the one that reverses the ordering of the alternatives in P, that is, for all alternatives a and b, aPb if and only if $bP'a$.

P_i, $\mu_i(. \mid P_i)$ is independent of the preferences of the agents in N_i^0, positively correlated with the preferences of the agents in N_i^+, and negatively correlated with preferences of the agents in N_i^-. We call such a prior (profile) a correlated prior (profile) under betweenness relation and denote the set of all such correlated priors by $\mathcal{M}$. Below, we provide a formal definition of this. For ease of presentation, we write $P'_{-i} \in \langle P_i, P_{-i} \rangle$ to mean $P'_j \in (P_i, P_j)$ for all $j \in N_i^+$ and $P_j \in (P_i, P'_j)$ for all $j \in N_i^-$.

Definition 4.2. A prior μ_i of an agent $i \in N$ is **correlated under betweenness relation (CBR)** if for all $P_i \in \mathcal{D}_i$ and all $P_{-i}, P'_{-i} \in \mathcal{D}_{-i}$ with the property that $P'_j \in (P_i, P_j)$ for all $j \in N_i^+$ and $P_j \in (P_i, P'_j)$ for all $j \in N_i^-$, we have $\mu_i(P'_{-i}|P_i) > \mu_i(P_{-i}|P_i)$. A prior (profile) μ_N is correlated if μ_i is correlated for each $i \in N$.

Next, we introduce the notion of strong OND. Informally speaking, in addition to OND, it says that for every preference of an agent, as the preference profiles of other agents become more similar to his preference, the outcomes of an SCF become weakly better for him. Thus, it imposes a type of monotonicity property on an SCF.

Definition 4.3. An SCF $f : \mathcal{D}_N \to A$ satisfies the **strong OND property** if f satisfies the OND property and for all $i \in N$, all $P_i \in \mathcal{D}_i$, and all $P_{-i}, P'_{-i} \in \mathcal{D}_{-i}$ such $P_{-i} \in \langle P_i, P'_{-i} \rangle$, we have $f(P_i, P_{-i}) R_i f(P_i, P'_{-i})$.

In Table 4, we present an SCF that satisfies the strong OND property.

Example 4.1. Suppose that there are three alternatives $\{a, b, c\}$ and two agents $\{1, 2\}$. Assume that both agents are positively correlated, that is, $N_i^+ = N \setminus i$ for all $i \in N$. Consider the SCF, say f, presented in Table 4. We argue that f satisfies the strong OND property. Consider $P_1 = abc$ and pick two preferences, say $P_2 = bac$ and $P'_2 = cba$, of agent 2. Note that $P_2 \in \langle P_1, P'_2 \rangle$, and $f(P_1, P_2) = b$ and $f(P_1, P'_2) = b$. This implies $f(P_1, P_2) R_1 f(P_1, P'_2)$, and hence f satisfies the strong OND property for this instance. It can be verified that f satisfies the requirement of strong OND for all other preference profiles and for all agents.

1 \ 2	abc	acb	bac	bca	cab	cba
abc	a	a	b	b	a	b
acb	a	a	a	c	c	c
bac	a	a	b	b	a	b
bca	b	c	b	b	c	c
cab	a	a	a	c	c	c
cba	b	c	b	b	c	c

Table 4: Example of an SCF that satisfies strong OND property

Now, we present the main result of this section. It says that if an SCF satisfies the strong OND property, then there must be some correlated prior with respect to which it is LOBIC.

Theorem 4.1. *Suppose that an SCF satisfies the strong OND property. Then, it is LOBIC with respect to some CBR prior.*

The proof of this theorem is relegated to Appendix B.

Since the SCF in Table 4 satisfies the strong OND property, by Theorem 4.1, there must exist some CBR prior with respect to which it is LOBIC. In Table 5 and Table 6, we present such priors for agents 1 and 2, respectively.

1 \ 2	abc	acb	bac	bca	cab	cba
abc	0.35	0.25	0.10	0.06	0.20	0.04
acb	0.25	0.35	0.20	0.04	0.10	0.06
bac	0.10	0.06	0.35	0.25	0.04	0.20
bca	0.20	0.04	0.25	0.35	0.06	0.10
cab	0.06	0.10	0.04	0.20	0.35	0.25
cba	0.04	0.20	0.06	0.10	0.25	0.35

Table 5: Conditional prior of Agent 1 for Table 4

5. A DISCUSSION ON GENERALIZING THE RESULTS

In this section, we show how the results in the paper can be generalized beyond the specific notion of betweenness we have used in the paper. We introduce a

$1 \setminus 2$	abc	acb	bac	bca	cab	cba
abc	0.30	0.20	0.20	0.10	0.15	0.05
acb	0.20	0.30	0.15	0.05	0.20	0.10
bac	0.20	0.10	0.30	0.20	0.05	0.15
bca	0.15	0.05	0.20	0.30	0.10	0.20
cab	0.10	0.20	0.05	0.15	0.30	0.20
cba	0.05	0.15	0.10	0.20	0.20	0.30

Table 6: Conditional prior of Agent 2 for Table 4

general notion of betweenness by means of a binary relation and present the result in Section 4 for this notion. Note that our necessary condition for LOBIC in Section 3 is independent of the structure of priors, therefore it does not require any generalization.

For every agent $i \in N$ and every preference P_i of i, fix a binary relation $b(P_i)$ on $\mathcal{D}_{-i}$ satisfying transitivity, anti-symmetricity, and reflexivity (but not necessarily complete). The relation $b(P_i)$ represents i's belief about the preference of other in the following manner: for distinct $P_{-i}, P'_{-i} \in \mathcal{D}_{-i}$, $(P_{-i}, P'_{-i}) \in b(P_i)$ implies that P_{-i} is more likely than P'_{-i} according to the correlated belief of i conditional on P_i. To distinguish the current notion of correlated belief from our earlier one, we call it correlated under generalized betweeness relation belief .

Definition 5.1. A prior μ_i of an agent $i \in N$ is **correlated under generalized betweenness relation (CGBR)** if for all $P_i \in \mathcal{D}_i$ and all distinct $P_{-i}, P'_{-i} \in \mathcal{D}_{-i}$ with the property that $(P_{-i}, P'_{-i}) \in b(P_i)$, we have $\mu_i(P_{-i}|P_i) > \mu_i(P'_{-i}|P_i)$. A prior (profile) μ_N is **CGBR** if μ_i is CGBR for each $i \in N$.

Note that the notion of OND sequence does not involve any prior. So, we continue to use Definition 3.2 for an OND sequence.

Definition 5.2. An SCF $f : \mathcal{D}_N \to A$ satisfies the **generalized strong OND** property if it satisfies OND property and the property that for all $P_i \in \mathcal{D}_i$ and all $P_{-i}, P'_{-i} \in \mathcal{D}_{-i}$ with $(P_{-i}, P'_{-i}) \in b(P_i)$, we have $f(P_i, P_{-i}) R_i f(P_i, P'_{-i})$.

Theorem 5.1. *Suppose that an SCF satisfies the generalized strong OND property. Then, it is LOBIC with respect to some CGBR prior.*

The proof of Theorem 5.1 is similar to the proof of Theorem 4.1, and hence it is omitted.

A. Proof of Theorem 3.1

Proof. Suppose an SCF $f : \mathcal{D}_N \to A$ is LOBIC with respect to some prior μ_N. We show that f satisfies the sequential OND property, that is, for all $i \in N$, all $P_i, P_i' \in \mathcal{D}_i$, and all $P_{-i} \in \mathcal{D}_{-i}$ with $f(P_i', P_{-i}) P_i f(P_i, P_{-i})$, there exists an OND sequence $(P^1_{-i}, \ldots, P^k_{-i})$ for f with respect to (P_i, P_i') such that $f(P_i, P_{-i}) R_i f(P_i', P^k_{-i})$ and $f(P_i, P^1_{-i}) R_i f(P_i', P_{-i})$.

Since f is LOBIC, for all agents $i \in N$, all preferences P_i of agent i, and all $k = 1, \ldots, m$, we have

$$\sum_{P_{-i} | f(P_i, P_{-i}) \in U_k(P_i)} \mu(P_{-i} \mid P_i) \geq \sum_{P_{-i} | f(P_i', P_{-i}) \in U_k(P_i)} \mu(P_{-i} \mid P_i). \tag{1}$$

Consider an agent $i \in N$, two preferences $\bar{P}_i, \bar{P}_i' \in \mathcal{D}_i$, and a preference profile $\bar{P}_{-i} \in \mathcal{D}_{-i}$ of the other agents such that $f(\bar{P}_i', \bar{P}_{-i}) \bar{P}_i f(\bar{P}_i, \bar{P}_{-i})$. If there does not exist any such instance, then f satisfies sequential OND vacuously. We proceed to show that there is an OND sequence $(P^1_{-i}, \ldots, P^k_{-i})$ for f with respect to $(\bar{P}_i, \bar{P}_i')$ such that $f(\bar{P}_i, \bar{P}_{-i}) \bar{R}_i f(\bar{P}_i', P^k_{-i})$ and $f(\bar{P}_i, P^1_{-i}) \bar{R}_i f(\bar{P}_i', \bar{P}_{-i})$. Let $f(\bar{P}_i, \bar{P}_{-i}) = a$ and $f(\bar{P}_i', \bar{P}_{-i}) = b$.

Consider the upper contour set $U(b, \bar{P}_i)$ of b at $\bar{P}_i$. Because $b \bar{P}_i a$, we have $a \notin U(b, \bar{P}_i)$. Applying (1) to the upper contour set $U(b, \bar{P}_i)$, we have

$$\sum_{P_{-i} | f(\bar{P}_i, P_{-i}) \in U(b, \bar{P}_i)} \mu(P_{-i} | \bar{P}_i) \geq \sum_{P_{-i} | f(\bar{P}_i', P_{-i}) \in U(b, \bar{P}_i)} \mu(P_{-i} | \bar{P}_i). \tag{2}$$

Because $f(\bar{P}_i, \bar{P}_{-i}) = a$ and $a \notin U(b, \bar{P}_i)$, by (2), there must exist $\hat{P}_{-i}$ such that $f(\bar{P}_i, \hat{P}_{-i}) \in U(b, \bar{P}_i)$ and $f(\bar{P}_i', \hat{P}_{-i}) \notin U(b, \bar{P}_i)$. Let $\hat{\mathcal{P}}_{-i}$ be the set of all such preferences $\hat{P}_{-i}$. Let P^1_{-i} be such that $f(\bar{P}_i{}', \hat{P}_{-i}) \bar{R}_i f(\bar{P}_i{}', P^1_{-i})$ for all $\hat{P}_{-i} \in \hat{\mathcal{P}}_{-i}$. In other words, P^1_{-i} gives agent i the worst outcome in the set $\hat{\mathcal{P}}_{-i}$ when his true preference is $\bar{P}_i$ and he reports the preference $\bar{P}_i'$. If $f(\bar{P}_i, \bar{P}_{-i}) \bar{R}_i f(\bar{P}_i{}', P^1_{-i})$, then the sequence (P^1_{-i}) is an OND sequence for f with respect to $(\bar{P}_i, \bar{P}'_i)$. Suppose instead $f(\bar{P}_i{}', P^1_{-i})\bar{P}_i f(\bar{P}_i, \bar{P}_{-i})$. Let $f(\bar{P}_i{}', P^1_{-i}) = c$.

Consider the upper contour set $U(c, \bar{P}_i)$. Applying (1) to $U(c, \bar{P}_i)$, we have

$$\sum_{P_{-i} \mid f(\bar{P}_i, P_{-i}) \in U(c, \bar{P}_i)} \mu(P_{-i} \mid \bar{P}_i) \geq \sum_{P_{-i} \mid f(\bar{P}'_i, P_{-i}) \in U(c, \bar{P}_i)} \mu(P_{-i} \mid \bar{P}_i). \tag{3}$$

Because $f(\bar{P}_i{}', P^1_{-i})\bar{P}_i f(\bar{P}_i, \bar{P}_{-i})$, we have that $f(\bar{P}_i, \bar{P}_{-i}) \notin U(c, \bar{P}_i)$. Hence, by (3) there must exist P^*_{-i} such that $f(\bar{P}_i, P^*_{-i}) \in U(c, \bar{P}_i)$ and $f(\bar{P}'_i, P^*_{-i}) \notin U(c, \bar{P}_i)$. As before, let $\mathcal{P}^*_{-i}$ be the set of all such preference profiles P^*_{-i} and let P^2_{-i} be such that $f(\bar{P}_i{}', P^*_{-i})\bar{R}_i f(\bar{P}_i{}', P^2_{-i})$ for all $P^*_{-i} \in \mathcal{P}^*_{-i}$. By the definition of P^2_{-i}, we have $f(\bar{P}'_i, P^2_{-i}) \notin U(c, \bar{P}_i)$. This, together with the fact that $f(\bar{P}'_i, P^1_{-i}) = c$, implies $f(\bar{P}'_i, P^1_{-i})\bar{P}_i f(\bar{P}'_i, P^2_{-i})$, and hence $P^1_{-i} \neq P^2_{-i}$. If $f(\bar{P}_i, \bar{P}_{-i})\bar{R}_i f(\bar{P}_i{}', P^2_{-i})$, then (P^1_{-i}, P^2_{-i}) is an OND sequence for f with respect to $(\bar{P}_i, \bar{P}'_i)$ such that $f(\bar{P}_i, \bar{P}_{-i})\bar{R}_i f(\bar{P}'_i, P^2_{-i})$ and $f(\bar{P}_i, P^1_{-i})\bar{R}_i f(\bar{P}'_i, \bar{P}_{-i})$. If not, then we proceed to the next step.

Continuing in this manner we can construct an OND sequence $(P^1_{-i}, P^2_{-i}, \ldots, P^k_{-i})$ for f with respect to $(\bar{P}_i, \bar{P}'_i)$ such that $f(\bar{P}_i, \bar{P}_{-i})\bar{R}_i f(\bar{P}'_i, P^k_{-i})$ and $f(\bar{P}_i, P^1_{-i})\bar{R}_i f(\bar{P}'_i, \bar{P}_{-i})$. The termination of the process is guaranteed by the fact that $P^1_{-i}, P^2_{-i}, \ldots, P^k_{-i}$ are all distinct. To see why they are distinct, note that, in a similar way as we have shown $f(\bar{P}'_i, P^1_{-i})\bar{P}_i f(\bar{P}'_i, P^2_{-i})$ in the preceding paragraph, we can show $f(\bar{P}'_i, P^1_{-i})\bar{P}_i f(\bar{P}'_i, P^2_{-i})\bar{P}_i \ldots \bar{P}_i f(\bar{P}'_i, P^k_{-i})$. This in particular means $P^1_{-i}, P^2_{-i}, \ldots, P^k_{-i}$ are all distinct. ■

B. Proof of Theorem 4.1

Proof. Suppose an SCF $f : \mathcal{D}_N \to A$ satisfies strong OND. We show that there exists a CBR prior μ_N such that f is LOBIC with respect to μ_N. We proceed to construct the mentioned CBR prior μ_N. Let μ_N be such that

(i) $\mu_i(P_{-i}|P_i) > 0$ for all $i \in N$, all $P_i \in \mathcal{D}_i$, and all $P_{-i} \in \mathcal{D}_{-i}$

(ii) $\mu_i(P_{-i}|P_i) > \mu_i(P'_{-i}|P_i)$ for all $i \in N$, all $P_{-i} \in \mathcal{D}_i$ such that $P_{-i} \in \langle P_i, P'_{-i}\rangle$,

(iii) $\mu_i(P_{-i}|P_i) > \sum_{P'_{-i}|f(P_i,P_{-i})P_i f(P_i,P'_{-i})} \mu_i(P'_{-i}|P_i)$ for all $i \in N$ and all $P_i \in \mathcal{D}_i$.

Here, (i) is a requirement for locally robustness, (ii) ensures that μ_N is a CBR prior, and (iii) is a technical condition that we need to ensure that f is LOBIC with respect to μ_N.

Since f satisfies the strong OND property, for all $i \in N$, all $P_i \in \mathcal{D}_i$, and all $P_{-i}, P'_{-i} \in \mathcal{D}_{-i}$ such $P_{-i} \in \langle P_i, P'_{-i} \rangle$, we have $f(P_i, P_{-i}) R_i f(P_i, P'_{-i})$. This implies that for any given P_{-i}, there is no P''_{-i} in the set $\{P'_{-i} \mid f(P_i, P_{-i}) P_i f(P_i, P'_{-i})\}$ such that $P''_{-i} \in \langle P_i, P_{-i} \rangle$. Hence, the prior μ_N is indeed a CBR prior. We proceed to show that f is LOBIC with respect μ_N.

Consider (arbitrary) $i \in N, P_i, P'_i \in \mathcal{D}_i$ and a utility representation u_i of P_i. Let $\overline{M}(P_i, P'_i) = \{\bar{P}_{-i} \in \mathcal{D}_{-i} \mid f(P'_i, \bar{P}_{-i}) P_i f(P_i, \bar{P}_{-i})\}$ be the set of all preference profiles of the agents in $N \setminus i$ such that agent i can manipulate by misreporting his sincere preference P_i as P'_i. Furthermore, let $M(P_i, P'_i) = \{P_{-i} \in \mathcal{D}_{-i} \mid f(P_i, P_{-i}) P_i f(P'_i, P_{-i})\}$ be the set of all preference profiles of the agents in $N \setminus i$ such that agent i cannot manipulate by misreporting his sincere preference P_i as P'_i. Clearly, $M(P_i, P'_i) \cap \overline{M}(P_i, P'_i) = \emptyset$. To show that f is LOBIC, we need to show that

$$\sum_{P_{-i} \in M(P_i, P'_i)} \mu_i(P_{-i}|P_i)[u_i(f(P_i, P_{-i})) - u_i(f(P'_i, P_{-i}))] \geq \sum_{\bar{P}_{-i} \in \overline{M}(P_i, P'_i)} \mu_i(\bar{P}_{-i}|P_i)[u_i(f(P'_i, \bar{P}_{-i})) - u_i(f(P_i, \bar{P}_{-i}))]. \tag{4}$$

For each $P_{-i} \in M(P_i, P'_i)$, we define

$$S(P_{-i}) = \{\bar{P}_{-i} \in \overline{M}(P_i, P'_i) \mid f(P_i, P_{-i}) R_i f(P'_i, \bar{P}_{-i}) \text{ and } f(P_i, \bar{P}_{-i}) R_i f(P'_i, P_{-i})\}.$$

Since f satisfies the strong OND property (and hence, the OND property in particular), we obtain the following two facts.

(i) For all $P_{-i} \in M(P_i, P'_i)$ with $S(P_{-i}) \neq \emptyset$ and all $\bar{P}_{-i} \in S(P_{-i})$, we have

$$u_i(f(P_i, P_{-i})) - u_i(f(P'_i, P_{-i})) \geq u_i(f(P'_i, \bar{P}_{-i})) - u_i(f(P_i, \bar{P}_{-i})),$$

which means

$$u_i\big(f(P_i, P_{-i})\big) - u_i\big(f(P_i', P_{-i})\big) \geq \max_{\bar{P}_{-i} \in S(P_{-i})} u_i\big(f(P_i', \bar{P}_{-i})\big) - u_i\big(f(P_i, \bar{P}_{-i})\big). \tag{5}$$

(ii)

$$\cup_{P_{-i} \in M(P_i, P_i')} S(P_{-i}) = \overline{M}(P_i, P_i'). \tag{6}$$

Consider $P_{-i} \in M(P_i, P_i')$. For every $\bar{P}_{-i} \in S(P_{-i})$, we have $f(P_i, P_{-i}) P_i f(P_i, \bar{P}_{-i})$. This, together with Part (iii) of the definition of μ_N, gives us

$$\mu_i(P_{-i}|P_i) > \sum_{\bar{P}_{-i} \in S(P_{-i})} \mu_i(\bar{P}_{-i}|P_i). \tag{7}$$

Using standard algebra, we have

$$\begin{aligned}
&\sum_{\bar{P}_{-i} \in S(P_i)} \mu_i(\bar{P}_{-i} \mid P_i) \Big[u_i\big(f(P_i', \bar{P}_{-i})\big) - u_i\big(f(P_i, \bar{P}_{-i})\big)\Big] \\
&\leq \sum_{\bar{P}_{-i} \in S(P_i)} \mu_i(\bar{P}_{-i} \mid P_i) \max_{\bar{P}_{-i} \in S(P_i)} \Big[u_i\big(f(P_i', \bar{P}_{-i})\big) - u_i\big(f(P_i, \bar{P}_{-i})\big)\Big] \\
&= \max_{\bar{P}_{-i} \in S(P_i)} \Big[u_i\big(f(P_i', \bar{P}_{-i})\big) - u_i\big(f(P_i, \bar{P}_{-i})\big)\Big] \sum_{\bar{P}_{-i} \in S(P_i)} \mu_i(\bar{P}_{-i} \mid P_i).
\end{aligned} \tag{8}$$

By (5), we have (9), and by (7) we have (10).

$$\begin{aligned}
&\max_{\bar{P}_{-i} \in S(P_i)} \Big[u_i\big(f(P_i', \bar{P}_{-i})\big) - u_i\big(f(P_i, \bar{P}_{-i})\big)\Big] \sum_{\bar{P}_{-i} \in S(P_i)} \mu_i(\bar{P}_{-i} \mid P_i) \\
&\leq \Big[u_i\big(f(P_i, P_{-i})\big) - u_i\big(f(P_i', P_{-i})\big)\Big] \sum_{\bar{P}_{-i} \in S(P_i)} \mu_i(\bar{P}_{-i} \mid P_i).
\end{aligned} \tag{9}$$

$$\begin{aligned}
&\Big[u_i\big(f(P_i, P_{-i})\big) - u_i\big(f(P_i', P_{-i})\big)\Big] \sum_{\bar{P}_{-i} \in S(P_i)} \mu_i(\bar{P}_{-i} \mid P_i) \\
&\leq \Big[u_i\big(f(P_i, P_{-i})\big) - u_i\big(f(P_i', P_{-i})\big)\Big] \mu_i(P_{-i} \mid P_i).
\end{aligned} \tag{10}$$

Combining (8), (9), and (10), we obtain (11).

$$\Big[u_i\big(f(P_i, P_{-i})\big) - u_i\big(f(P_i', P_{-i})\big)\Big]\mu_i(P_{-i} \mid P_i) \geq \sum_{\bar{P}_{-i} \in S(P_i)} \mu_i(\bar{P}_{-i} \mid P_i)\Big[u_i\big(f(P_i', \bar{P}_{-i})\big) - u_i\big(f(P_i, \bar{P}_{-i})\big)\Big] \tag{11}$$

By summing both sides of (11) over the elements P_{-i} of $M(P_i, P_i')$, we have

$$\begin{aligned} &\sum_{P_{-i} \in M(P_i, P_i')} \Big[u_i\big(f(P_i, P_{-i})\big) - u_i\big(f(P_i', P_{-i})\big)\Big]\mu_i(P_{-i} \mid P_i) \\ \geq &\sum_{P_{-i} \in M(P_i, P_i')} \sum_{\bar{P}_{-i} \in S(P_i)} \mu_i(\bar{P}_{-i} \mid P_i)\Big[u_i\big(f(P_i', \bar{P}_{-i})\big) - u_i\big(f(P_i, \bar{P}_{-i})\big)\Big]. \end{aligned} \tag{12}$$

Since by (6) $\cup_{P_{-i} \in M(P_i, P_i')} S(P_{-i}) = \overline{M}(P_i, P_i')$, all the elements of $\overline{M}(P_i, P_i')$ appear in the summand in the R.H.S. of (12). Therefore, it follows that

$$\begin{aligned} &\sum_{P_{-i} \in M(P_i, P_i')} \Big[u_i\big(f(P_i, P_{-i})\big) - u_i\big(f(P_i', P_{-i})\big)\Big]\mu_i(P_{-i} \mid P_i) \\ \geq &\sum_{\bar{P}_{-i} \in \overline{M}(P_i, P_i')} \mu_i(\bar{P}_{-i} \mid P_i)\Big[u_i\big(f(P_i', \bar{P}_{-i})\big) - u_i\big(f(P_i, \bar{P}_{-i})\big)\Big]. \end{aligned}$$

as required by (4). This completes the proof of the theorem.

■

References

ALBERT, M., V. CONITZER, G. LOPOMO, AND P. STONE (2017a): "Mechanism design with correlated valuations: Efficient methods for revenue maximization," *Under Submission at Operations Research*.

ALBERT, M., V. CONITZER, AND P. STONE (2017b): "Mechanism Design with Unknown Correlated Distributions: Can We Learn Optimal Mechanisms?" in *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems*, 69–77.

BHARGAVA, M., D. MAJUMDAR, AND A. SEN (2015): "Incentive-compatible voting rules with positively correlated beliefs," *Theoretical Economics*, 10, 867–885.

KARMOKAR, M. AND S. ROY (2018): "ON THE EQUIVALENCE OF ORDINAL BAYESIAN INCENTIVE COMPATIBILITY AND DOMINANT STRATEGY INCENTIVE COMPATIBILITY FOR RANDOM RULES," .

LAFFONT, J.-J. AND D. MARTIMORT (2000): "Mechanism design with collusion and correlation," *Econometrica*, 68, 309–342.

MAJUMDAR, D. AND A. SEN (2004): "Ordinally Bayesian incentive compatible voting rules," *Econometrica*, 72, 523–540.

MISHRA, D. (2016): "Ordinal Bayesian incentive compatibility in restricted domains," *Journal of Economic Theory*, 163, 925–954.

YAMASHITA, T. (2018): "Revenue guarantees in auctions with a (correlated) common prior and additional information," .