Correlated Strategy Proofness

Parikshit De, * Abhinaba Lahiri, * Arindam Paul †

November 30, 2019

Abstract

In this paper, we introduce the notion of a correlated strategy proof social choice rule. A social choice rule is correlated strategy proof if there exists a probability distribution over the set all preference profiles satisfying two conditions. The first condition is that the choice of this probability distribution does not depend on the true preference profile. The second condition is that given this probability distribution, every agent weakly prefers to report his true preference rather than strategically manipulate the social choice rule, under the assumption that everyone else does so. We show that every unanimous rule is correlated strategy proof.

KEYWORDS Social choice, strategy proofness, correlated equilibrium.

JEL CODES D71.

*Indian Institute of Science Education and Research, Bhopal, India. Address: IISER Bhopal, Bhopal Bypass Road, Bhauri, Bhopal, India - 462038. Email: parikshitde@iiserb.ac.in, abhinaba@iiserb.ac.in.

†Centre for Studies in Social Sciences, Kolkata, India. Address: R–1, Baishnabghata Patuli Township, Kolkata, West Bengal, India – 700094. Email: arish309@gmail.com.

1 Introduction

In strategic social choice, a well researched aspect of non manipulation is strategy proofness. A rule is strategy-proof if no individual can obtain a preferred alternative by misrepresenting his preferences for any announcement of the preferences of the other individuals. Strategy-proofness ensures that for every agent, truth-telling is a weakly dominant strategy in the direct revelation game induced by the rule. It has been shown that if there are at least three alternatives, then the only unanimous and strategy proof rule is the dictatorial rule. In other words, if there are at least three alternatives, then for every profile of true preference, truthful reporting is a Nash equilibrium of the direct revelation game induced by a non constant rule if and only if the rule is dictatorial. In this paper, our notion of non manipulation is correlated strategy proofness. Essentially, if a rule is correlated strategy proof, then there exists a correlated equilibrium of the direct revelation game where for every agent reporting truthfully is weakly better than manipulating the rule strategically.

In this paper, we consider the social choice framework where individual preference of each agent is their private information. Here we assume that there are at least three but finite number of alternatives and each agent has strict preference orderings over the set of alternatives. Here for any rule, we focus on the correlated equilibria of the game induced by that rule for every profile of true preferences. Correlated equilibria of a game is a probability distribution over all possible profiles of pure strategies provided by an external agency that satisfies the following. For every agent it is weakly better to follow the probability distribution provided by the external agency under the assumption that everyone else does so. Here we are interested in those probability distributions over the set of preference profiles of the induced game which satisfies the following conditions.

Condition 1: The choice of the probability distribution does not depend on the profile of true preference.

Condition 2: For every agent it is weakly better to report the true preference rather than manipulating the rule strategically under the assumption that everyone else does so.

Note that just a rule does not induce a game. In order to induce a game from a rule, a profile of true preference is required, which is essentially the utility function of the players of the induced game. Since in strategic social choice framework, true preference is regarded as private information, the first condition ensures that the knowledge of that information is not required for designing the probability distribution. For the second condition, first we focus on the meaning of an agent manipulating the rule strategically. This means that for every possible profiles of preferences of everyone else, the agent will play one of his best responses corresponding to that profile of preference of everyone else. Next we elaborate on the assumption that everyone else does so. Note that this assumption literally implies that for everyone else it is weakly better to report the true preference rather than manipulating the rule strategically. Since this is an incomplete information setup, one agent does not know the true preferences of other agents. This is where the probability distribution over set of all preference profiles comes into play. An agent, given his true preference, can construct a conditional probability distribution over all possible preference profiles of everyone else. Then the assumption that "everyone else does so" results in a particular belief of one agent over the preferences profiles of everyone else which is the reconstructed conditional distribution.

Next we introduce the property of correlated strategy proofness in two steps as follows. For every preference profile, a signaling function chooses a probability distribution over set of all preference profiles. We define a rule to be correlated strategy proof with respect to a signaling function if for every profile of true preferences, the probability distribution chosen by the signaling function satisfies the second condition. Finally, we define a rule to be correlated strategy proof if it is correlated strategy proof with respect to a constant signaling function. Note that a constant signaling function satisfies condition one.

In this paper, first we characterize the class of all probability distribution over preference profiles that satisfies the second condition for any rule. Given a rule, an agent and his true preference, we partition the set of preference profiles of everyone else into two classes. The first class contains those preference profiles of everyone else where the agent can manipulate the rule by not revealing his true preference and be better off. The second class contains those preference profiles of everyone else where the agent cannot manipulate the rule and for him it is weakly best to report his true preference. Our characterization states that a probability distribution over preference profiles will satisfy the second condition if and only if for every agent and for every true preference of that agent, the probability distribution assigns 0 probability to the profile where the agent reports his true preference and everyone else reports some profile which belongs to the first class. Loosely speaking, we show that a probability distribution over preference profiles will satisfy the second condition if and only if it assigns 0 probabilities to manipulable profiles. One direction of the proof of this result is intuitive as, if the manipulable profiles have 0 probability, then the conditional probability of a manipulable profile given any true preference ordering is also 0. This implies that there will be no gain from strategic manipulation.

Finally, we restrict our attention to the class of unanimous rules. Here we show that for any unanimous rule, unanimous profiles are not manipulable. We conclude by showing that every unanimous rule is correlated strategy proof by choosing a probability distribution over the set of all preference profiles that assigns 0 probability to non unanimous profiles. Note that such a choice satisfies the first condition.

Strategic social choice begins with the impossibility result, which says that if there are at least three alternatives, then the only non constant strategy proof rule is the dictatorial rule (Gibbard (1973), Satterthwaite (1975)). In the literature, the most common process of circumventing this impossibility result is to consider restricted domain. But this process does not always leads to a possibility result. There are restricted domains where one can design non dictatorial strategy-proof rules (Moulin (1980), Barberà et al. (1991)), and there are restricted domains where one cannot (Aswal et al. (2003), Sato (2010)). In this paper, we propose another process of circumventing the impossibility result, namely restricting the belief of one agent over preferences of everyone else. At the same time, this work bears similarities with Information Design perspective (Bergemann and Morris (2019)) in strategic social choice setup with ordinal preferences.

This paper is organized as follows. Section 2 introduces the model. Section 3 contains the results. Section 4 concludes by relating this work with the existing literature on information design.

2 Model

Let $A = \{a_1, a_2, \ldots, a_m\}$ denote the set containing $m \geq 3$ finitely many alternatives and $N = \{1, 2, \ldots, n\}$, $n \geq 2$, a finite set of agents/individuals. Each individual i in N has a strict preference relation P_i over A. Let $\mathcal{P}$ be the set of these preference relations. For any $i \in N$ and $P_i \in \mathcal{P}$, let $r_k(P_i)$ denote the k^{th} ranked alternative according to P_i, where $k \in \{1, 2, \ldots, m\}$. As $m < \infty$, for any $i \in N$, every $P_i \in \mathcal{P}$ can be represented by a utility function $u_i : A \longrightarrow \mathbb{R}$ such that for all $a, b \in A$, aP_ib if and only if $u_i(a) > u_i(b)$.

A preference profile is a list $P = (P_1, P_2, \ldots, P_n) \in \mathcal{P}^n$ of individuals preferences. For any coalition $S \subseteq N$ and any profile $P \in \mathcal{P}$, P_S denotes the restriction of the profile P to the coalition S *i.e.* $P_S = (P_i)_{i \in S}$. For any $S \subseteq N$ and any $P, P' \in \mathcal{P}^n$, $P'' = (P_S, P'_{N \setminus S}) \in \mathcal{P}^n$ denote that profile where $P''_i = P_i$ for all $i \in S$ and $P''_i = P'_i$ for all $i \in N \setminus S$. A profile $P' \in \mathcal{P}^n$ is defined to be a $i-$deviation from another profile $P \in \mathcal{P}^n$ if $P_{N \setminus \{i\}} = P'_{N \setminus \{i\}}$.

Definition **1.** *A rule f is a mapping from $\mathcal{P}^n$ to A i.e. $f : \mathcal{P}^n \longrightarrow A$.*

Next, we provide an example of a very well known rule in the literature.

Example **1.** *A rule f^d is dictatorial if there exists an agent $i \in N$ such that $f(P) = r_1(P_i)$ for all $P \in \mathcal{P}^n$.*

Next, we define two well known properties of a rule in the literature called strategy-proofness and unanimity.

Definition **2.** *A profile* $P \in \mathcal{P}^n$ *is an unanimous profile if* $r_1(P_i) = r_1(P_j)$ *for all* $i \neq j \in N$. *A rule* f *is unanimous if* $f(P) = r_1(P_1)$ *for all unanimous profiles* $P \in \mathcal{P}^n$.

In words, if every agent has the same alternative at top according to their reported profile, then a unanimous rule will select that alternative. Note that the dictatorial rule is unanimous. In this setup, every agent is allowed to report any preference ordering from $\mathcal{P}$. The next property of a rule talks about the situation when agents do not report their true preference.

Definition **3.** *A rule* f *is manipulable if, for any* $i \in N$ *and for any* $P \in \mathcal{P}^n$*, there exists an* $i-$*deviation* $P' \in \mathcal{P}^n$ *of* P *such that* $f(P')P_i f(P)$.
A rule f *is strategy-proof if* f *is not manipulable for all* $i \in N$ *and for all* $P \in \mathcal{P}^n$.

A rule is strategy-proof if no individual can obtain a preferred alternative by misrepresenting her preferences for any announcement of the preferences of the other individuals. Note that the dictatorial rule is strategy proof. Strategy-proofness ensures that for every agent truth-telling is a weakly dominant strategy in the direct revelation game induced by the rule. Next, we formally define the direct revelation game induced by any rule. Given a rule f and a profile of "true" preferences $\overline{P} \in \mathcal{P}^n$, we can define a simultaneous strategic game (game in normal form) as follows. Set of agents is N. For each agent $i \in N$, the set of pure strategies or action space A_i is $\mathcal{P}$. Each agent's utility function is $v_i(a_1, a_2, \ldots, a_n) = v_i(P_1, P_2, \ldots, P_n) = \overline{u}_i(f(P_1, P_2, \ldots, P_n))$, where $\overline{u}_i : A \longrightarrow \mathbb{R}$ is any utility function representing the preference relation $\overline{P}_i$ for all $i \in N$. We denote this game by $\mathcal{G}^{f,\overline{P}}$.

It can be shown using Gibbard (1973) and Satterthwaite (1975) that for any profile of "true" preferences, reporting truthfully becomes a Nash equilibrium only in the games induced by the dictatorial rules. In other words, the pure strategy profile $(a_1, a_2, \ldots, a_n) = (\overline{P}_1, \overline{P}_2, \ldots, \overline{P}_n)$ is a Nash equilibrium of the game $\mathcal{G}^{f,\overline{P}}$ for all $\overline{P} \in \mathcal{P}^n$ if and only if $f(P) = f^d(P)$ for all $P \in \mathcal{P}^n$.

Next we incorporate the concept of correlated equilibrium introduced in Aumann (1987) to the game $\mathcal{G}^{f,\overline{P}}$. Here, we denote a communication device as c, where $c \in \mathcal{L}(\mathcal{P}^n)$, where $\mathcal{L}(\mathcal{P}^n)$ denotes the set of all probability distribution over $\mathcal{P}^n$. Next we modify the utility functions corresponding to $\mathcal{G}^{f,\overline{P}}$, to incorporate such a communication device. Given any $c \in \mathcal{L}(\mathcal{P}^n)$, and a profile of true preference $\overline{P} \in \mathcal{P}^n$, to define the utility of agent i we consider two fields:

1. whether he follows the communication device or not. If he follows the communication device, he will report his true preference ($\overline{P}_i$). Otherwise, he reports optimally corresponding to what everyone else has reported.

2. whether he believes that everyone else will follow the communication device or not. One thing to note here is that although each agent can observe the communication device, true preference relation of any agent is still his private knowledge. So if he believes that everyone else will follow the communication device, then this generates another probability distribution from c conditioned on what his true preference is.

In this paper, the objective is to look for a communication device such that every agent $i \in N$ will follow the communication device given that he believes that everyone else will follow the same. So in presence of some $c \in \mathcal{L}(\mathcal{P}^n)$ we define the utility function of agent i in two cases:

i follows c given that he believes that everyone else will follow c

$$w_i(\overline{P}_i|c) = \sum_{P_{N\setminus\{i\}} \in \mathcal{P}^{n-1}} \frac{c(\overline{P}_i, P_{N\setminus\{i\}})}{\sum_{P_{N\setminus\{i\}} \in \mathcal{P}^{n-1}} c(\overline{P}_i, P_{N\setminus\{i\}})} \overline{u}_i(f(\overline{P}_i, P_{N\setminus\{i\}}))$$

for any $\overline{u}_i$ which represents $\overline{P}_i$.

i does not follow c given that he believes that everyone else will follow c

Here, the agent i first calculates his best response for every $P_{N\setminus\{i\}}$ given his true preference $\overline{P}_i$ as follows. $b_{\overline{P}_i} : \mathcal{P}^{n-1} \longrightarrow 2^{\mathcal{P}}$, is the best response correspondence of player i given his true preference $\overline{P}_i$ if for any $P_{N\setminus\{i\}} \in \mathcal{P}^{n-1}$, $P_i \in b_{\overline{P}_i}(P_{N\setminus\{i\}})$ if and only if

$$\overline{u}_i(f(P_i, P_{N\setminus\{i\}})) \geq \overline{u}_i(f(P_i', P_{N\setminus\{i\}}))$$

for all $P_i' \in \mathcal{P}$ and for all $\overline{u}_i$ representing $\overline{P}_i$. For any $P_{N\setminus\{i\}} \in \mathcal{P}^{n-1}$, let $P_{b_{\overline{P}_i}(P_{N\setminus\{i\}})}$ denote a generic element of $b_{\overline{P}_i}(P_{N\setminus\{i\}})$. Next, his utility from disobeying is

$$w_i(b_{\overline{P}_i}|c) = \sum_{P_{N\setminus\{i\}} \in \mathcal{P}^{n-1}} \frac{c(\overline{P}_i, P_{N\setminus\{i\}})}{\sum_{P_{N\setminus\{i\}} \in \mathcal{P}^{n-1}} c(\overline{P}_i, P_{N\setminus\{i\}})} \overline{u}_i(f(P_{b_{\overline{P}_i}(P_{N\setminus\{i\}})}, P_{N\setminus\{i\}}))$$

for any $\overline{u}_i$ which represents $\overline{P}_i$.

Next, for any rule f and any profile of true preference $\overline{P} \in \mathcal{P}^n$, we define a communication device $c \in \mathcal{L}(\mathcal{P}^n)$ as a correlated equilibrium of the game $\mathcal{G}^{f,\overline{P}}$ as follows

Definition 4. *For any rule f and any profile of true preference $\overline{P} \in \mathcal{P}^n$, $c \in \mathcal{L}(\mathcal{P}^n)$ is a correlated equilibrium of the game $\mathcal{G}^{f,\overline{P}}$ if*

$$\begin{aligned} w_i(\overline{P}_i|c) &= \sum_{P_{N\setminus\{i\}} \in \mathcal{P}^{n-1}} \frac{c(\overline{P}_i, P_{N\setminus\{i\}})}{\sum_{P_{N\setminus\{i\}} \in \mathcal{P}^{n-1}} c(\overline{P}_i, P_{N\setminus\{i\}})} \overline{u}_i(f(\overline{P}_i, P_{N\setminus\{i\}})) \\ &\geq \sum_{P_{N\setminus\{i\}} \in \mathcal{P}^{n-1}} \frac{c(\overline{P}_i, P_{N\setminus\{i\}})}{\sum_{P_{N\setminus\{i\}} \in \mathcal{P}^{n-1}} c(\overline{P}_i, P_{N\setminus\{i\}})} \overline{u}_i(f(P_{b_{\overline{P}_i}(P_{N\setminus\{i\}})}, P_{N\setminus\{i\}})) \\ &= w_i(b_{\overline{P}_i}|c) \end{aligned}$$

for all $i \in N$ and for any $\overline{u}_i$ which represents $\overline{P}_i$.

Finally we define our notion of a correlated strategy proof rule below. Let $g : \mathcal{P}^n \longrightarrow \mathcal{L}(\mathcal{P}^n)$ denote a function that selects a communication device for any possible preference profile.

Definition **5.** *A rule f is correlated strategy proof with respect to g if $g(\overline{P})$ is a correlated equilibrium of the game $\mathcal{G}^{f,\overline{P}}$ for all $\overline{P} \in \mathcal{P}^n$.*
A rule f is correlated strategy proof, if f is correlated strategy proof with respect to g, where g is a constant function; i.e; $g(P) = c$ for all $P \in \mathcal{P}^n$ for some $c \in \mathcal{L}(\mathcal{P}^n)$.

3 Results

Let f be any correlated strategy proof rule with respect to g. In this section, we characterize the class $R(g) \subseteq \mathcal{L}(\mathcal{P}^n)$, where $R(g)$ denotes the range of g. Let $\overline{P} \in \mathcal{P}^n$ be a profile of true preferences.

Definition **6.** *A profile of $n-1$ agents $P_{N\setminus\{i\}} \in \mathcal{P}^{n-1}$ is manipulable for f at a preference ordering $\overline{P}_i$ of agent $i \in N$ if there exists a preference ordering $P'_i \in \mathcal{P}$ such that $f(P'_i, P_{N\setminus\{i\}})\overline{P}_i f(\overline{P}_i, P_{N\setminus\{i\}})$.*

Next, for any rule f, any $i \in N$, and any $\overline{P}_i \in \mathcal{P}$, we partition the set of all preference profiles of $n-1$ agents $\mathcal{P}^{n-1}$ in two sets $\mathcal{P}^M_{f,\overline{P}_i}$ and $\mathcal{P}^{NM}_{f,\overline{P}_i}$ as follows.

$$\mathcal{P}^M_{f,\overline{P}_i} = \left\{P_{N\setminus\{i\}} \in \mathcal{P}^n : P_{N\setminus\{i\}} \text{ is manipulable for } f \text{ at the preference ordering } \overline{P}_i \text{ of agent } i\right\}$$

$$\mathcal{P}^{NM}_{f,\overline{P}_i} = \mathcal{P}^{n-1} \setminus \mathcal{P}^M_{f,\overline{P}_i}$$

Note that $\mathcal{P}^M_{f,\overline{P}_i}$ is the collection of all profiles of $n-1$ agents which are manipulable for f at the preference ordering $\overline{P}_i$ by agent i. Similarly, $\mathcal{P}^{NM}_{f,\overline{P},i}$ is the collection of all profiles of $n-1$ agents which are not manipulable for f at the preference ordering $\overline{P}_i$ by agent i. Now we introduce our first lemma.

Lemma **1.** *Any rule f is correlated strategy proof with respect to g if and only if $g(\overline{P})(\overline{P}_i, P_{N\setminus\{i\}}) = 0$ for all $\overline{P} \in \mathcal{P}^n$, for all $i \in N$, and for all $P_{N\setminus\{i\}} \in \mathcal{P}^M_{f,\overline{P}_i}$.*

Proof. Let f be any rule. Suppose that there exists a function $h : \mathcal{P}^n \longrightarrow \mathcal{L}(\mathcal{P}^n)$ such that $h(\overline{P})(\overline{P}_i, P_{N\setminus\{i\}}) = 0$ for all $\overline{P} \in \mathcal{P}^n$, for all $i \in N$, and for all $P_{N\setminus\{i\}} \in \mathcal{P}^M_{f,\overline{P}_i}$. We are going to show that f is correlated strategy proof with respect to h; i.e; we are going to show that

$h(\overline{P})$ is a correlated equilibrium of the game $\mathcal{G}^{f,\overline{P}}$ for all $\overline{P} \in \mathcal{P}^n$. In other words, we have to show that for all $\overline{P} \in \mathcal{P}^n$

$$\begin{aligned} w_i(\overline{P}_i|h(\overline{P})) &= \sum_{P_{N\setminus\{i\}}\in\mathcal{P}^{n-1}} \frac{h(\overline{P})(\overline{P}_i,P_{N\setminus\{i\}})}{\sum_{P_{N\setminus\{i\}}\in\mathcal{P}^{n-1}} h(\overline{P})(\overline{P}_i,P_{N\setminus\{i\}})}\overline{u}_i(f(\overline{P}_i,P_{N\setminus\{i\}})) \\ &\geq \sum_{P_{N\setminus\{i\}}\in\mathcal{P}^{n-1}} \frac{h(\overline{P})(\overline{P}_i,P_{N\setminus\{i\}})}{\sum_{P_{N\setminus\{i\}}\in\mathcal{P}^{n-1}} h(\overline{P})(\overline{P}_i,P_{N\setminus\{i\}})}\overline{u}_i(f(P_{b_{\overline{P}_i}},P_{N\setminus\{i\}})) \\ &= w_i(b_{\overline{P}_i}|h(\overline{P})) \end{aligned}$$

for all $i \in N$. Now we fix a $\overline{P} \in \mathcal{P}^n$ and consider the following cases based on those agents $i \in N$ who can manipulate f at their true preference $\overline{P}_i$; and those who cannot.

Case 1 Consider those $i \in N$ such that $P_{N\setminus\{i\}} \in \mathcal{P}^{NM}_{f,\overline{P}_i}$ for all $P_{N\setminus\{i\}} \in \mathcal{P}^{N-1}$:
Suppose that in this case for some $i \in N$, for some $P_i \in \mathcal{P}$ and for some $P_{N\setminus\{i\}} \in \mathcal{P}^{N-1}$, we have $f(P_i, P_{N\setminus\{i\}})\overline{P}_i f(\overline{P}_i, P_{N\setminus\{i\}})$. Then this implies that $P_{N\setminus\{i\}} \in \mathcal{P}^M_{f,\overline{P},i}$, which contradicts this case. So in this case, we have either $f(\overline{P}_i, P_{N\setminus\{i\}}) = f(P_i, P_{N\setminus\{i\}})$ or $f(\overline{P}_i, P_{N\setminus\{i\}})\overline{P}_i f(P_i, P_{N\setminus\{i\}})$ for all $i \in N$, for all $P_i \in \mathcal{P}$ and for all $P_{N\setminus\{i\}} \in \mathcal{P}^{N-1}$. Then it follows that $\overline{u}_i(f(\overline{P}_i, P_{N\setminus\{i\}})) \geq \overline{u}_i(f(P_i, P_{N\setminus\{i\}}))$ for all $i \in N$, for all $P_i \in \mathcal{P}$ and for all $P_{N\setminus\{i\}} \in \mathcal{P}^{N-1}$ in this case. Then it follows from the definition that $w_i(\overline{P}_i|h(\overline{P})) \geq w_i(b_{\overline{P}_i}|h(\overline{P}))$ for all $i \in N$ in this case.

Case 2 Consider those $i \in N$ such that $P_{N\setminus\{i\}} \in \mathcal{P}^M_{f,\overline{P}_i}$ for some $P_{N\setminus\{i\}} \in \mathcal{P}^{N-1}$:
Note that in this case, whenever $f(P_i, P_{N\setminus\{i\}})\overline{P}_i f(\overline{P}_i, P_{N\setminus\{i\}})$, we have $h(\overline{P})(\overline{P}_i, P_{N\setminus\{i\}}) = 0$. So these $P_{N\setminus\{i\}}$ profiles of $n-1$ agents will not contribute anything to the sum. In all other profiles of $n-1$ agents we have either $f(P_i, P_{N\setminus\{i\}}) = f(\overline{P}_i, P_{N\setminus\{i\}})$ or $f(\overline{P}_i, P_{N\setminus\{i\}})\overline{P}_i f(P_i, P_{N\setminus\{i\}})$. This is exactly similar to case 1.

Combining these two cases, we can conclude that if $g : \mathcal{P}^n \longrightarrow \mathcal{L}(\mathcal{P}^n)$ be such a function that $g(\overline{P})(\overline{P}_i, P_{N\setminus\{i\}}) = 0$ for all $\overline{P} \in \mathcal{P}^n$, for all $i \in N$, and for all $P_{N\setminus\{i\}} \in \mathcal{P}^M_{f,\overline{P}_i}$, then the rule f is correlated strategy proof with respect to g. Next, we prove the other direction.

Suppose the rule f is correlated strategy proof with respect to some function $g : \mathcal{P}^n \longrightarrow \mathcal{L}(\mathcal{P}^n)$. We have to show that $g(\overline{P})(\overline{P}_i, P_{N\setminus\{i\}}) = 0$ for all $\overline{P} \in \mathcal{P}^n$, for all $i \in N$, and for all $P_{N\setminus\{i\}} \in \mathcal{P}^M_{f,\overline{P}_i}$. We show this by means of contradiction. So suppose that there exists a $\overline{P} \in \mathcal{P}^n$, an agent $i \in N$ and a $P'_{N\setminus\{i\}} \in \mathcal{P}^M_{f,\overline{P},i}$ such that $g(\overline{P})(\overline{P}_i, P'_{N\setminus\{i\}}) = \epsilon > 0$. As $P'_{N\setminus\{i\}} \in \mathcal{P}^M_{f,\overline{P}_i}$, so there exists a preference ordering $P'_i \in \mathcal{P}$ such that $f(P', P'_{N\setminus\{i\}})\overline{P}_i f(\overline{P}_i, P'_{N\setminus\{i\}})$. So

$\overline{u}_i(f(P', P'_{N\setminus\{i\}})) > \overline{u}_i(f(\overline{P}_i, P'_{N\setminus\{i\}}))$ for any $\overline{u}_i$ representing $\overline{P}_i$. Now consider the following.

$$\begin{aligned}
& w_i(b_{\overline{P}_i}|g(\overline{P})) - w_i(\overline{P}_i|g(\overline{P})) \\
= \;& \sum_{P^\star_{N\setminus\{i\}} \in \mathcal{P}^{n-1}} \frac{g(\overline{P})(\overline{P}_i, P^\star_{N\setminus\{i\}})}{\sum_{P^\star_{N\setminus\{i\}} \in \mathcal{P}^{n-1}} g(\overline{P})(\overline{P}_i, P^\star_{N\setminus\{i\}})} \left\{\overline{u}_i(f(P_{b_{\overline{P}_i}}, P^\star_{N\setminus\{i\}})) - \overline{u}_i(f(\overline{P}_i, P^\star_{N\setminus\{i\}}))\right\} \\
= \;& \sum_{P^\star_{N\setminus\{i\}} \in \mathcal{P}^{M}_{f,\overline{P}_i}} \frac{g(\overline{P})(\overline{P}_i, P^\star_{N\setminus\{i\}})}{\sum_{P^\star_{N\setminus\{i\}} \in \mathcal{P}^{n-1}} g(\overline{P})(\overline{P}_i, P^\star_{N\setminus\{i\}})} \left\{\overline{u}_i(f(P_{b_{\overline{P}_i}}, P^\star_{N\setminus\{i\}})) - \overline{u}_i(f(\overline{P}_i, P^\star_{N\setminus\{i\}}))\right\} + \\
& \sum_{P^\star_{N\setminus\{i\}} \in \mathcal{P}^{NM}_{f,\overline{P}_i}} \frac{g(\overline{P})(\overline{P}_i, P^\star_{N\setminus\{i\}})}{\sum_{P^\star_{N\setminus\{i\}} \in \mathcal{P}^{n-1}} g(\overline{P})(\overline{P}_i, P^\star_{N\setminus\{i\}})} \left\{\overline{u}_i(f(P_{b_{\overline{P}_i}}, P^\star_{N\setminus\{i\}})) - \overline{u}_i(f(\overline{P}_i, P^\star_{N\setminus\{i\}}))\right\}
\end{aligned}$$

Now consider the second term of the sum, where $P^\star_{N\setminus\{i\}} \in \mathcal{P}^{NM}_{f,\overline{P}_i}$. Note that for any such $P^\star_{N\setminus\{i\}}$, we have $b_{\overline{P}_i}(P^\star_{N\setminus\{i\}}) = \{\overline{P}_i\}$. This implies that the second term of the sum is 0. So, we have

$$\begin{aligned}
& w_i(b_{\overline{P}_i}|g(\overline{P})) - w_i(\overline{P}_i|g(\overline{P})) \\
= \;& \sum_{P^\star_{N\setminus\{i\}} \in \mathcal{P}^{M}_{f,\overline{P}_i}} \frac{g(\overline{P})(\overline{P}_i, P^\star_{N\setminus\{i\}})}{\sum_{P^\star_{N\setminus\{i\}} \in \mathcal{P}^{n-1}} g(\overline{P})(\overline{P}_i, P^\star_{N\setminus\{i\}})} \left\{\overline{u}_i(f(P_{b_{\overline{P}_i}}, P^\star_{N\setminus\{i\}})) - \overline{u}_i(f(\overline{P}_i, P^\star_{N\setminus\{i\}}))\right\} \\
= \;& \frac{g(\overline{P})(\overline{P}_i, P'_{N\setminus\{i\}})}{\sum_{P^\star_{N\setminus\{i\}} \in \mathcal{P}^{n-1}} g(\overline{P})(\overline{P}_i, P^\star_{N\setminus\{i\}})} \left\{\overline{u}_i(f(P_{b_{\overline{P}_i}}, P'_{N\setminus\{i\}})) - \overline{u}_i(f(\overline{P}_i, P'_{N\setminus\{i\}}))\right\} + \\
& \sum_{P^\star_{N\setminus\{i\}} \in \mathcal{P}^{M}_{f,\overline{P}_i} \setminus \{P'_{N\setminus\{i\}}\}} \frac{g(\overline{P})(\overline{P}_i, P^\star_{N\setminus\{i\}})}{\sum_{P^\star_{N\setminus\{i\}} \in \mathcal{P}^{n-1}} g(\overline{P})(\overline{P}_i, P^\star_{N\setminus\{i\}})} \left\{\overline{u}_i(f(P_{b_{\overline{P}_i}}, P^\star_{N\setminus\{i\}})) - \overline{u}_i(f(\overline{P}_i, P^\star_{N\setminus\{i\}}))\right\}
\end{aligned}$$

Now consider the first term. We have assumed that $g(\overline{P})(\overline{P}_i, P'_{N\setminus\{i\}}) = \epsilon > 0$. Also $P'_i \in b_{\overline{P}_i}(P'_{N\setminus\{i\}})$. So it follows that the first term is positive. In the second term, note that we consider all $P^\star_{N\setminus\{i\}} \in \mathcal{P}^M_{f,\overline{P}_i} \setminus \{P'_{N\setminus\{i\}}\}$. Then it follows that for any such $P^\star_{N\setminus\{i\}}$, there exists a $P^\star_i \in \mathcal{P}$ such that $\overline{u}_i(f(P^\star_i, P^\star_{N\setminus\{i\}})) > \overline{u}_i(f(\overline{P}_i, P^\star_{N\setminus\{i\}}))$. As $g(\overline{P})(\overline{P}_i, P^\star_{N\setminus\{i\}}) \geq 0$ for all such $P^\star_{N\setminus\{i\}}$, the second term is non negative. So combining we have

$$w_i(b_{\overline{P}_i}|g(\overline{P})) - w_i(\overline{P}_i|g(\overline{P})) > 0$$

This contradicts the fact that $g(\overline{P})$ is a correlated equilibrium of the game $\mathcal{G}^{f,\overline{P}}$ and concludes the proof of Lemma 1. □

For any rule f, let $\mathcal{C}_f \subseteq \mathcal{L}(\mathcal{P}^n)$ be defined as follows.

$$\mathcal{C}_f = \left\{c \in \mathcal{L}(\mathcal{P}^n) : c(\overline{P}_i, P_{N\setminus\{i\}}) = 0 \text{ for all } i \in N, \text{ for all } \overline{P}_i \in \mathcal{P}^n, \text{ and for all } P_{N\setminus\{i\}} \in \mathcal{P}^M_{f,\overline{P}_i}\right\}$$

Then Lemma 1 implies that rule f is correlated strategy proof with respect to some function $g : \mathcal{P}^n \longrightarrow \mathcal{L}(\mathcal{P}^n)$ if and only if $R(g) \subseteq \mathcal{C}_f$.
Now we define a subset $\mathcal{P}^n_a$ of $\mathcal{P}^n$ as follows.

$$\mathcal{P}^n_a = \{P \in \mathcal{P}^n : r_1(P_i) = r_1(P_j) \text{ for all } i, j \in N\}$$

Note that $\mathcal{P}^n_a$ is the collection of all unanimous profile. Next, we define a class of communication devices $\mathcal{C}_a \subseteq \mathcal{L}(\mathcal{P}^n)$ as follows.

$$\mathcal{C}_a = \left\{ c \in \mathcal{L}(\mathcal{P}^n) : \sum_{P \in \mathcal{P}^n_a} c(P) = 1 \right\}$$

Note that, for any $c \in \mathcal{C}_a$, $c(P) = 0$ for any $P \notin \mathcal{P}^n_a$. This brings us to our main theorem.

Theorem **1.** *Every unanimous rule f is correlated strategy proof.*

We prove Theorem 1 with the help of the following lemma. First, for any $c \in \mathcal{C}_a$, we define a function $h_c : \mathcal{P}^n \longrightarrow \mathcal{L}(\mathcal{P}^n)$ as $h_c(P) = c$ for all $P \in \mathcal{P}^n$.

Lemma **2.** *Let f be an unanimous rule. Then f is correlated strategy proof with respect to h_c for any $c \in \mathcal{C}_a$.*

Proof. In order to prove Lemma 2, in view of Lemma 1, we have to show that $h_c(\overline{P})(\overline{P}_i, P_{N\setminus\{i\}}) = 0$ for all $\overline{P} \in \mathcal{P}^n$, for all $i \in N$, and for all $P_{N\setminus\{i\}} \in \mathcal{P}^M_{f,\overline{P}_i}$. Note that for any $\overline{P} \in \mathcal{P}^n$, for any $i \in N$, and for any $P_{N\setminus\{i\}} \in \mathcal{P}^M_{f,\overline{P}_i}$, if $(\overline{P}_i, P_{N\setminus\{i\}}) \notin \mathcal{P}^n_a$, then from the definition, it follows that $h_c(\overline{P})(\overline{P}_i, P_{N\setminus\{i\}}) = 0$. Now consider a $\overline{P} \in \mathcal{P}^n$, a $i \in N$, and a $P_{N\setminus\{i\}} \in \mathcal{P}^{n-1}$ such that $(\overline{P}_i, P_{N\setminus\{i\}}) \in \mathcal{P}^n_a$. Then unanimity of f implies that $f(\overline{P}_i, P_{N\setminus\{i\}})\overline{P}_i((P'_i, P_{N\setminus\{i\}})$ for all $P'_i \in \mathcal{P}$. This implies that $P_{N\setminus\{i\}} \notin \mathcal{P}^M_{f,\overline{P}_i}$. This concludes the proof of Lemma 2. □

Proof of Theorem 1: From Lemma 2, it follows that every unanimous rule is correlated strategy proof with respect to h_c. From the definition of h_c, it follows that h_c is a constant rule. This concludes the proof of Theorem 1. □

4 Conclusion

In this section, we briefly discuss how this work fits into the theory of information design (Bergemann and Morris (2019)). Given a game with uncertain payoffs, information design analyzes the extent to which the provision of information alone can influence the behavior of the agents. In information design setup, each agent's utility not only depends on his action but also on the realization of a payoff relevant state. To convert it to the context of this paper, one can assume that an agent's utility depends on his reported preference ordering along with the payoff relevant states, which are the announcements of everyone else. Here the goal of the information designer or the external agency is to elicit the true preference orderings of every agent. Here we assume that the information designer has no informational advantage over the agents. In such a case, under ordinal preferences, we show that the information designer, through the choice of the information provided, can influence the individually optimal behavior of the players to achieve his objective.

References

Aswal, N., S. Chatterji, and A. Sen (2003): "Dictatorial domains," *Economic Theory*, 22, 45–62.

Aumann, R. J. (1987): "Correlated Equilibrium as an Expression of Bayesian Rationality," *Econometrica*, 55, 1–18.

Barberà, S., H. Sonnenschein, and L. Zhou (1991): "Voting by committees," *Econometrica: Journal of the Econometric Society*, 595–609.

Bergemann, D. and S. Morris (2019): "Information Design: A Unified Perspective," *Journal of Economic Literature*, 57, 44–95.

Gibbard, A. (1973): "Manipulation of Voting Schemes: A General Result," *Econometrica*, 41, 587–601.

Moulin, H. (1980): "On strategy-proofness and single peakedness," *Public Choice*, 35, 437–455.

Sato, S. (2010): "Circular domains," *Review of Economic Design*, 14, 331–342.

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.