Robust Binary Voting*

Tangren Feng† and Qinggong Wu‡

November, 2019

Click here for the latest version

Abstract

In this paper, we study a new robustness concept in mechanism design with interdependent values: interim dominant strategy incentive compatibility (IDSIC). It requires each agent to have an interim dominant strategy, i.e., conditional on her own private information, the strategy maximizes her expected payoff for all possible strategies the other agents could use. In a simple setting with two alternatives and no transfers, we characterize IDSIC together with two other well studied concepts: dominant strategy incentive compatibility (DSIC) and ex post incentive compatibility (EPIC). While both DSIC and EPIC permit only constant mechanisms in sufficiently rich environments, non-constant IDSIC mechanisms exist in any environment. The characterization of IDSIC suggests a simple class of (indirect) binary voting rules: Each agent reports Yes/No. Moreover, if the binary voting rule is also additive, then the indirect mechanism is versatile: It admits an interim dominant strategy equilibrium on all payoff environments and all corresponding type spaces.

Keywords: Robust mechanism design, interdependent values, interim dominant strategy, binary voting.

*We are grateful to Tilman Börgers for his continuous encouragement and guidance in developing this project. We thank Aislinn Bohren, Laura Doval, Matthew Gentzkow, Faruk Gul, Kevin He, Emir Kamenica, Shaowei Ke, Elliot Lipnowski, Heng Liu, George Mailath, David Miller, Esteban Peralta, Collin Raymond, Alexander Wolitzky, Weijie Zhong, as well as seminar participants at Michigan, Michigan State, Peking, Tsinghua, and HKUST for helpful comments.

†Department of Economics, University of Michigan. Email: tangren@umich.edu.

‡Department of Economics, Hong Kong University of Science and Technology. Email: wqg@ust.hk.

1 Introduction

To vote wisely is not easy. To that end, an agent typically needs to carefully evaluate the candidates and understand how other people will vote. Such strategic consideration can get complex and exhausting. Therefore, an ideal mechanism would make it easy for agents to determine their unambiguously best vote without having to resort to intricate strategic considerations. In addition to easing agents' cognitive burdens, such a mechanism would also function in a more stable or *robust* fashion, because an agent is likely to adopt that unambiguously best strategy regardless of the many confounding factors she may encounter.

The notion of robustness is captured by the concept of *dominant strategy incentive compatibility (DSIC)* that has been heavily studied in the literature. However, the literature is based on the keynote that DSIC is too restrictive. Indeed, as the famous Gibbard–Satterthwaite Theorem states, when there are at least three alternatives and preferences are private-value only dictatorships achieve DSIC on unrestricted preference domain (Gibbard (1973) and Satterthwaite (1975)). By this logic, it is then foreseeable that DSIC becomes even more difficult to achieve in the more general interdependent value setting where private information does not pin down one's preference. This we confirm. By studying a model where agents need to collectively choose from *two* alternatives, we observe that, despite there being fewer than enough alternatives to entail the Gibbard–Satterthwaite Theorem, DSIC is nonetheless too restrictive because of preference interdependence. Typically, only constant mechanisms satisfy DSIC, that is, even dictatorships fail the DSIC scrutiny. A non-constant DSIC mechanism exists only if some voters' preferences are *de facto* private-value, and the mechanism is responsive only to these voters.

Many collective choice situations in real life fit the interdependent value setting better, particularly when information about the values of the alternatives is fragmentarily distributed among the population. Classical examples include decision making in committees, legislatures, and juries. It is thus important to understand whether, in the interdependent value setting, there are mechanisms that retain *a reasonable degree of robustness* but are not as austere as ones that satisfy DSIC. This is the main question we address in the paper.

We observe that, in the interdependent value setting, DSIC captures a strong notion of robustness. Given DSIC, an agent can unambiguously determine the best strategy free of any belief about (1) other people's information that directly affects her own preference, and (2) other people's strategies. In other words, DSIC has two orthogonal properties:

1. **Informationally belief-free**: For each agent, there is a strategy that remains a best response given any interim belief about the distribution of the other agents' types, *conditional on equilibrium strategies*.
2. **Strategically belief-free**: For each agent, there is a strategy that remains a best response given any belief about the other agents' strategies, *conditional on her interim belief about the distribution of the other agents' types*.

The informationally belief-free property implies that to come up with a best response, agents do not need to confer with (or even to be conscious of) their own possibly very complicated belief hierarchies. Therefore, indeterminacy in what beliefs the agents would have does not shake the mechanism designer's confidence that they will follow the belief-free optimal strategies, as long as everyone expects everyone else to follow those strategies. In other words, mechanisms that have the informational belief-free property are robust to misspecification, or the lack of specification, of the agents beliefs about the other agents' types ("informational beliefs").

In contrast, the strategic belief-free property implies that there is an *interim dominant strategy* for every agent that remains a best response regardless of whether she has a correct, or any at all, *strategic* belief about the other agents' strategies. Hence it is easy for an agent to find and take that interim dominant strategy regardless of how she reasons about other people's strategies. Mechanisms that have the strategic belief-free property are robust to the possibility that agents lack adequate strategic sense or understanding ("strategic beliefs").

DSIC mechanisms are exactly the mechanisms that possess both belief-free properties. Having both properties at once certainly makes DSIC mechanisms attractive to a mechanism designer who only has limited or unreliable knowledge about the agents' informational *and* strategic beliefs. The cost is, as we have pointed out, of course the restrictiveness of DSIC. To mitigate this conflict between robustness and leniency, we take the approach of exploring the middle ground by maintaining one belief-free property a time while relaxing the other.

Mechanisms that have the informational belief-free property are, as is well known in the literature (Bergemann and Morris (2005)), ones that satisfy ex post incentive compatibility (EPIC). We develop a characterization for the set of all EPIC (direct revelation) mechanisms in our simple setting with two alternatives and no transfers, and use the characterization to find necessary and sufficient conditions for there to be non-constant EPIC mechanisms. If all agents have the same ex post preference and every agent's preference changes monotonically in her and the other agents' payoff-relevant information in the same fashion, then there exist non-constant EPIC mechanisms.[1] Moreover some of them are ex post Pareto efficient.

Should we be excited about the positive results? To explore how generic environments that admit non-constant EPIC mechanisms might be, we look at continuous type spaces subject to standard regularity conditions. We find that, as long as there is enough preference interdependences and preference heterogeneities so that when indifference curves intersect they do not overlap locally beyond the point of intersection, then every EPIC mechanism must be constant.

Now let us turn to mechanisms that have the strategically belief-free property. We call a mechanism *interim dominant strategy incentive compatible (IDSIC)* if it has the strategic belief-free property, because in such a mechanism every player has an *interim* dominant strategy given her private information. The strategically belief-free property is a popular motivation for DSIC (or strategy-proofness in the literature of voting and market design) mechanism design in complete information or private value settings. However, when it comes to the interdependent value setting, the strategically belief-free property by itself has not received the due attention that we think it deserves.[2] When the strategically belief-free property is discussed in the interdependent values setting, it is discussed as an addendum to DSIC and is hence not analytically separated from the informationally belief-free property. As a result, while DSIC mechanism de-

[1]The classical Condorcet Jury model, or common value voting in general, assumes this condition.

[2]To the extent that there still lack commonly agreed terminologies for the property and the associated incentive compatibility condition. See the Literature subsection for more detailed discussion.

sign prevails, IDSIC mechanism design almost does not exist in the literature. We think ISDIC deserves more attentions than it has. First, as in a complete information or private value setting, a mechanism with the strategically belief-free property allows agents, who might not be strategically sophisticated or "correct" in anticipating an equilibrium, to have an unambiguous optimal strategy and hence behave in a predictable fashion. Second, as we show in the paper, IDSIC is more permissive than DSIC in the interdependent values setting, because non-constant IDSIC mechanisms always exist. This additional permissiveness is desirable for the design of actual mechanisms, especially in cases where there is little ambiguity about the underlying type space and hence the informational belief-free property that comes with DSIC is less important, for instance, when the type space is a common prior type space where the type-generating process is objective and straightforward. Last, as we will also show, even though an agent's interim dominant strategy depends on her type, i.e., her infinite belief hierarchies, in effect only her first-order belief matters. In other words, the agent can determine her interim dominant strategy without being aware of her higher-order beliefs.[3] Therefore, although in an IDSIC mechanism an agent needs to examine her belief to come up with a best response — whereas in a DSIC or EPIC mechanism such examination is not needed at all — this examination is relatively simple.

In the paper, we characterize the set of IDSIC (direct revelation) mechanisms, show that non-constant IDSIC mechanisms always exist, and that they have a very simple structure: For the same agent, all types that have the same strict interim preference regarding the two alternatives are treated equally. The simple structure implies that IDSIC allows a mechanism to dispense with all the bells and whistles that would otherwise be necessary to cater to the rich type space where belief hierarchies could be complicated, because interim preference can be pinned down with just first-order belief. In other words, any IDSIC choice rule can be implemented in a reduced direct mechanism where the agents only report their payoff type and first-order belief. IDSIC mechanisms have a distinct feature that a IDSIC mechanism can only elicits preference rankings (not intensities), intensity still constrains it. Further more, the characterization of IDSIC suggests a class of straightforward indirect mechanisms, which we call *binary additive voting mechanisms*. In a binary additive voting mechanism, each agent

[3]This result applies to more general settings with any finite alternatives.

casts a Yes/No vote. Unlike majority voting mechanisms in which an agent's vote is either pivotal or not, the effect of an agent's vote in binary additive voting mechanisms is independent of the other agents' votes. We show that a binary additive voting mechanism is versatile: it is IDSIC for any type space with respect to any payoff environment.

Literature

There is a massive literature that emphasizes what we mean by robustness, and DSIC is held up as the most thorough. In particular, the literature on robust social choice and voting mostly draws upon the celebrated Gibbard–Satterthwaite impossibility result (Gibbard (1973) and Satterthwaite (1975)) and its extension Hylland (1980), which states that in the private value setting, when there are at least three candidates and any preference profile is possible, then only (random) dictatorship satisfies DSIC.

Follow-up papers look for more positive results in two natural ways. The first one is restricting the preference domain (for example, Moulin (1980), Gershkov et al. (2016) and see Barberà (2011) for an excellent survey). In contrast, by allowing interdependent preferences, our setting is distinctly different than, and in some cases embeds, the private value setting with unrestricted preference domain that underlies the impossibility results. Moreover, in our setting there are only two candidates, for if there were more then our passage would also be blocked by the impossibility results as long as the private setting with unrestricted domain is embedded in our model, because DSIC, IDSIC and EPIC are equivalent in the private value setting.

The other way of exploring positive results is weakening DSIC in the private value settings. Azevedo and Budish (2019) proposes strategy-proofness in the large (SP-L) which shares the same spirit as IDSIC. SP-L weakens DSIC in two ways. First, it only requires truth-telling be *approximately* optimal *in a large market*. This part is orthogonal to IDSIC. Second, SP-L requires that in the interim stage, truth-telling is best responding to a *subset* (full support, iid distributions) of all possible other agents' reports, rather than best responding to *all* ex post realization of other agents' reports as IDSIC asks. Both SP-L and IDSIC evaluate ICs in the interim stage, but IDSIC is reduced to DSIC with private values and is stronger than SP-L.

There is a growing literature that studies EPIC mechanisms. As far as we know, ours is the first paper that studies EPIC mechanisms for general social choice or voting. Jehiel et al. (2006) shows when the payoff state spaces are continuous, agents have interdependent values and multidimensional signals, generically, any deterministic EPIC mechanism (with transfers) is constant. Jehiel et al. is a direct comparison with our impossibility result in EPIC. While we do not allow for transfers which makes it harder to find non-constant EPIC mechanisms, we allow for stochastic choice functions and do not require agents' signal to be multidimensional. There are several cases that open the door to positive results in Jehiel et al., for example, separable values and one-dimensional signals, neither of which would work in our setting.

The notion of IDSIC is not new (Crémer and McLean (1985)), although it has received far less attention than what we think it deserves. Out of a similar robustness concern, Börgers and Li (forthcoming) proposes a condition, termed as "strategic simplicity", that resembles IDSIC. The main difference is that, under strategic simplicity, an agent has a strategy that is a best response to any scenario that (1) is consistent with her interim belief and (2) where the other agents do not play weakly dominated strategies, whereas, under IDSIC the interim dominant strategy is a best response to any scenario that satisfies (1).[4] Therefore, strategic simplicity is a conceptually weaker condition than IDSIC. Börgers and Li (forthcoming) show that, in private value voting, strategically simple mechanisms are "local dictatorships" in general. There is a large literature on robust mechanism design where robustness is interpreted differently from what we mean in the paper: A mechanism is said to be robust if it is interim incentive compatible[5] with respect to many type spaces à la Harsanyi (1967/1968), or to a very rich type space. There, robustness is interpreted as versatility. In a seminal paper, Bergemann and Morris (2005) show that a mechanism is interim incentive compatible with respect to every type space (or to the universal type space à la Mertens and Zamir (1985)) if and only if it is an EPIC mechanism. Therefore EPIC mechanisms are robust in this sense, too. IDSIC also has a robustness-as-versatility flavor, since an IDSIC mechanism is expected to function well in all situations that share the same type space but differ in

[4]Although Börgers and Li (forthcoming) focus on the private value setting, the condition can be generalized to the interdependent value setting.

[5]Interim incentive compatibility generalizes Bayesian incentive compatibility to non-common prior type spaces. See Bergemann and Morris (2005).

the voters' ideas about other people's strategies.

Organization

The paper is organized as follows. Section 2 describes the environment and solution concepts. Section 3 discusses dominant strategy incentive compatibility. Section 4 presents results on ex post incentive compatibility. Section 5 considers interim dominant strategy incentive compatibility. Section 6 concludes. Appendix A presents minor results on ex post incentive compatibility and Appendix B collects the proofs.

2 Model

Environment

N agents $I = \{1, 2, \ldots, N\}$ need to make a collective choice from two alternatives $a \in A = \{S, R\}$. Every agent receives a payoff of 0 if S — the *S*tatus quo or *S*afe option — is chosen. On the other hand, if R — the *R*eform or *R*isky option — is chosen, payoffs to the agents depend on an N-dimensional payoff state of the world $\theta = (\theta_1, ..., \theta_N) \in \Theta_1 \times ... \times \Theta_N = \Theta$. In particular, the payoff to agent i from R being chosen in state θ is expressed as $u_i(\theta)$. In most part of the paper we assume that Θ is finite; however, in one subsection in section 4 we allow Θ to be infinite. We fix sets I and A through out the paper and call $\langle \Theta, \{u_i\}_{i \in I} \rangle$ the **payoff environment** of the collective choice problem. This payoff environment is common knowledge among the agents.

Information about the true state θ is dispersed among the agents. More specifically, agent i only privately observes θ_i, which we say is agent i's **payoff type**. Agent i also has a (subjective) belief about the other agents' payoff types, and this belief is said to be agent i's first-order belief. Moreover, agent i also has a belief about the other agents' payoff types *and* first-order beliefs, and this belief is said to be agent i's second-order belief. Agent i's higher-order beliefs are defined analogously *ad infinitum*. We call β_i, the agent's (infinite) belief hierarchy, her **belief type**. The agent's payoff type and belief type constitute her **type**.

Types are cumbersome objects to think about and work with because they involve infinite belief hierarchies. Harsanyi (1967/1968) and Mertens and Zamir (1985) show that any type space has a much simpler formulation. Following Bergemann and Morris (2005), the type space is defined as follows. We denote the set of all probability measures on the Borel field of a metric space X by $\Delta(X)$.

Definition 2.1. *A type space is a list $\mathcal{T} = \langle T_i, \hat{\theta}_i, \hat{\beta}_i \rangle_{i \in I}$ where for each agent i, T_i is a nonempty finite set of types, and $\hat{\theta}_i, \hat{\beta}_i$ are functions of the form:*

$$\hat{\theta}_i : T_i \to \Theta_i \ \ \textit{and} \ \ \hat{\beta}_i : T_i \to \Delta(T_{-i})$$

which respectively reflect type t_i's payoff type and belief type.

For each type $t_i \in T_i$ of agent i, $\hat{\theta}_i(t_i)$ is her payoff type and $\hat{\beta}_i(t_i)$ is her belief type. For each payoff type θ_i, there at least exists one type t_i such that $\hat{\theta}_i(t_i) = \theta_i$. We denote $\hat{\beta}_i(t_i)[E]$ the probability that type t_i of agent i assigns to other agents having types t_{-i} in a measurable set $E \subset T_{-i}$.

The infinite belief hierarchy of an agent can be recovered from the simple formulation of types given in Definition 2.1. For example, agent i's **first-order belief** function $\hat{b}_i : T_i \to \Delta(\Theta_{-i})$ (which will be of particular importance) can be computed as follows:

$$\hat{b}_i(t_i)[E'] = \sum_{\{t_{-i} | \hat{\theta}_{-i}(t_{-i}) \in E'\}} \hat{\beta}_i(t_i)[t_{-i}],$$

so that $\hat{b}_i(t_i)[E']$ is the probability that type t_i of agent i assigns to the event that the other agents' payoff type profile θ_{-i} is in $E' \subset \Theta_{-i}$.

We fix the type space through out the paper expect in subsections 5.4 and 5.5.

Mechanisms

We investigate mechanisms by which the agents arrive at a collective decision without the use of side payments. We formulated a mechanism as a messaging game with a choice function as follows.

Definition 2.2. *A mechanism is a list* $\langle M_1, \cdots, M_N, q\rangle$ *such that for each* $i \in I$ *the set* M_i *is a nonempty set of messages, and* $q : M_1 \times \cdots \times M_N \to [0, 1]$ *is a choice function that indicates the probability with which alternative* R *is chosen.*

Side payments are ruled out because the theoretical objective of this paper is to explore and contribute to the theory of mechanism design with non-transferable utilities. Moreover, not using transfers is a typical constraint to which the design of actual collective choice mechanisms, specifically voting mechanisms, is subject.

On the other hand, we do not require that the mechanism is deterministic. In other words, devices like lotteries can be used in a mechanism such that even if the agents take actions deterministically, the outcome can still be uncertain.

Two classes of mechanisms are of particular interest: The **direct mechanisms** where $M_i = T_i$ for every $i \in I$, and the **fully reduced direct mechanisms** where $M_i = \Theta_i$ for every $i \in I$.[6] Under the direct mechanisms, the agents are asked to report their respective types, whereas under the fully reduced direct mechanism, they are instead asked to report their respective *payoff* types.

Social Choice Rule

A social choice rule $f : T \to [0, 1]$ is a function which maps agents' types to outcomes.

Solution Concepts

As outlined in the Introduction, we analyze mechanism design subject to three incentive compatibility (IC) conditions:

- Dominant strategy incentive compatibility (DSIC)
- Ex post incentive compatibility (EPIC)
- Interim dominant strategy incentive compatibility (IDSIC)

In the upcoming sections we will formally define these IC conditions. Roughly speaking, DSIC requires that every agent has a (weakly) dominant strategy that maximizes her

[6] We save the terminology "reduced direct mechanism" for later use.

ex post payoff given any message profile from the other agents in any payoff state θ. [7]

EPIC requires that there is an ex post equilibrium in which every agent's equilibrium strategy maximizes her ex post payoff in any payoff state θ conditional on other agents following their equilibrium strategies. IDSIC requires that every agent has a (weakly) *interim* dominant strategy that maximizes her expected payoff given any message profile from the other agents conditional on her interim belief about the payoff state θ. There is one other IC condition, Bayesian incentive compatibility (BIC), that is popular in the literature. BIC requires that there is a Bayesian Nash equilibrium such that every agent's equilibrium strategy maximizes her expected payoff conditional on her interim belief about the payoff state θ and that other agents follow their equilibrium strategies. We do not discuss BIC because it lacks the robustness properties that we focus on in the paper.

DSIC is the strongest IC condition and BIC is weakest. IDSIC and EPIC are in between, as they arise from respectively relaxing the informational and the strategic belief-free properties from DSIC.

3 Dominant Strategy Incentive Compatibility

In this section, we characterizes all DSIC mechanisms and confirm the claim that DSIC is too restrictive in settings with interdependent values.

Dominant strategy incentive compatibility is formally defined as follows:

Definition 3.1. *The strategy profile σ^* is a dominant strategy equilibrium of the mechanism $\langle M_1, \cdots, M_N, q\rangle$ if*

$$u_i(\hat{\theta}(t))q(\sigma_i^*(t_i), m_{-i}) \geq u_i(\hat{\theta}(t))q(m_i, m_{-i})$$

for all $m \in M$ and all $t \in T$, and all $i \in I$.

That is, for each agent i and type t_i, $\sigma_i^*(t_i)$ maximizes her ex post utility for all possible messages m_{-i} other agents could send and all possible realizations of other agents' types

[7] We will abuse notations when there are no confusions. In particular, we require the optimal strategies to be truth telling when we say a *direct* mechanism is DSIC. The same applies to EPIC and IDSIC.

t_{-i}. If a mechanism admits a dominant strategy equilibrium, then it satisfies dominant strategy incentive compatibility.

By the revelation principle, we can focus on truth-telling dominant strategy equilibria of fully reduced direct mechanisms. Hence, without further specifications, all mechanisms in this section refer to fully reduced direct mechanisms.

Definition 3.2. *A fully reduced direct mechanism $\langle \Theta_1, \cdots, \Theta_N, q\rangle$ is dominant strategy incentive compatible if*

$$u_i(\theta)q(\theta_i, \theta'_{-i}) \geq u_i(\theta)q(\theta'_i, \theta'_{-i})$$

for all $\theta, \theta' \in \Theta$, and all $i \in I$.

That is, truth telling is a dominant strategy in the ex post stage, i.e., it maximizes each agent's ex post utility for all possible messages θ'_{-i} other agents could send and all possible realizations of other agents' payoff types θ_{-i}.

3.1 Characterization

The following definitions will be useful of the characterization.

Definition 3.3. *A correspondence $\phi_i : \Theta_i \rightrightarrows \{-1, 0, 1\}$ is an indicator correspondence of agent i if*

$$\phi_i(\theta_i) \ni \begin{cases} 1, & \textit{if } u_i(\theta_i, \theta_{-i}) > 0 \textit{ for some } \theta_{-i} \in \Theta_{-i} \\ 0, & \textit{if } u_i(\theta_i, \theta_{-i}) = 0 \textit{ for some } \theta_{-i} \in \Theta_{-i} \\ -1, & \textit{if } u_i(\theta_i, \theta_{-i}) < 0 \textit{ for some } \theta_{-i} \in \Theta_{-i} \end{cases}. \tag{3.1}$$

$\phi_i(\theta_i)$ contains agent i's possible ex post preferences over $\{S, R\}$ when her private signal is θ_i. For example, $\phi_i(\theta_i) = \{1, -1\}$ means there exists θ_{-i} and θ'_{-i} such that $u_i(\theta_i, \theta_{-i}) > 0$ and $u_i(\theta_i, \theta'_{-i}) < 0$, i.e., given her private signal θ_i, it is possible agent i prefers R over S or prefers R over S in the ex post stage.

The following lemma establishes the link between indicator correspondences and DSIC mechanisms.

Lemma 3.1. *q is DSIC if and only if for any $i = 1, ..., N$ and $\theta_i \in \Theta_i$:*

1. *If $1 \in \phi_i(\theta_i)$ then $q(\theta_i, \theta_{-i}) = \max_{\theta_i' \in \Theta_i} q(\theta_i', \theta_{-i})$ for any $\theta_{-i} \in \Theta_{-i}$.*

2. *If $-1 \in \phi_i(\theta_i)$ then $q(\theta_i, \theta_{-i}) = \min_{\theta_i' \in \Theta_i} q(\theta_i', \theta_{-i})$ for any $\theta_{-i} \in \Theta_{-i}$.*

An immediate implication of Lemma 3.1 is that if there exist one private signal θ_i such that agent i is uncertain about her ex post preferences over $\{S, R\}$ then a DSIC mechanism cannot be responsive to her private signal. Lemma 3.2 formalizes the observation.

Definition 3.4. *A mechanism $q(\theta_1, \cdots, \theta_N)$ is **responsive** to θ_i if there exist $\theta_i, \theta_i' \in \Theta_i$ and $\theta_{-i} \in \Theta_{-i}$ such that $q(\theta_i, \theta_{-i}) \neq q(\theta_i', \theta_{-i})$.*

Definition 3.5. *Agent i has **quasi-private values** if there does not exist $\theta_i \in \Theta_i$ such that $\{-1, 1\} \subset \phi_i(\theta_i)$.*

An agent who has quasi-private value is certain of whether S is weakly superior to R or R is weakly superior to S based on her private information. In other words, a quasi-private value agent's interim preferences over $\{S, R\}$ are the same as her ex post preferences.

Lemma 3.2. *Suppose q is dominant strategy incentive compatible. Then q is responsive to θ_i only if agent i has quasi-private values.*

Lemma 3.2 allows us to focus on mechanisms that only heed to agents of essentially private values. Let PI be the set of agents who have quasi-private valuations. We then introduce the binary relation $\Rightarrow_i$ on Θ_i where $i \in PI$: $\theta_i \Rightarrow_i \theta_i'$ if $\phi_i(\theta_i) \ni 1$, $\phi_i(\theta_i') \ni -1$, or $\theta_i = \theta_i'$. The binary relation captures the dominant strategy incentive compatibilities.

Proposition 3.1. *q is DSIC if and only if*

1. *$q(\theta) = q(\theta')$ if $\theta_i = \theta_i'$ for all $i \in PI$;*

2. *$q(\theta) \geq q(\theta')$ if $\theta_i \Rightarrow_i \theta_i'$ for all $i \in PI$.*

The first part of the proposition captures the essence of Lemma 3.2: A DSIC mechanism cannot be responsive to any agents other than those have quasi-private values. Also note that if $PI = \emptyset$ then the first part holds for all s and s'. Hence, only constant mechanisms are DSIC. The second part suggests the mechanism needs and only needs to respect quasi-private value agents' ordinal preferences to elicit private information from them. In other words, the mechanism is only informed by ordinal preferences.

According to Proposition 3.1, any DSIC choice rule can be indirectly implemented by a mechanism that only collects reports from those who have quasi-private values. Excluding the agents not having quasi-private valuations essentially transforms the environment into a private value setting, and therefore the proposition implies that preference interdependence is in sharp conflict with the existence of non-trivial DSIC mechanisms. In other words, DSIC cannot survive in general interdependent value setting.

Practically, a DSIC mechanism only respects those who have a strong opinion regarding the relative values of R vis-à-vis S, i.e. those who cannot be swayed by additional input from the other agents. In other words, DSIC only allows those who are stubborn to have their voices heard, possibly against objections from all the others agents. Since the purpose of mechanism design is often exactly to rectify the unfortunate situation that would arise in the absence of such a mechanism, DSIC does not achieve much in this respect.

3.2 A Partnership Example

Here we use a simple example to illustrate our results. This example will be re-examined when we introduce the other IC conditions.

Two agents need to decide whether they form a partnership or not. Both agents have two possible payoff types, *H(igh)* or *L(ow)*. S represents the joint decision of no partnership, and R represents partnership. The value of partnership depends on the types. Partnership between two high types is very productive, in which case both agents receive a payoff of 4. Partnership between two low types is less productive, in which case both agents receive a payoff of 1. Partnership where types mismatch is counterproductive, in which case both agents receive a payoff of -2.

$u_i(\theta)$	H	L
H	4	-2
L	-2	1

For simplicity, assume that there is a unique belief type associated with each payoff type (and when discussing this example we simply use the payoff type to represent the type of an agent), and the beliefs all come from a common prior such that every state has a probability of 1/4.

Neither agent has quasi-private values, because for either type of the same agent, the sign of her payoff depends on the other agent's type, and hence she is uncertain about her preference ranking regarding S and R despite her private information. It follows from Proposition 3.1 that only constant mechanisms are DSIC in this environment. This observation is striking because there is no conflict of interest between the two agents, yet it is still impossible to robustly (in the DSIC sense) incentivize truthful revelation of private information for non-trivial collective decision making.

4 Ex Post Incentive Compatibility

In this section, we first give a characterization of EPIC and two sufficient conditions on the payoff environment for the existence of non-constant EPIC mechanisms. Later, we show that when the payoff type space is continuous and the payoff environment is sufficiently rich, any EPIC mechanism is constant.

Ex post incentive compatibility is formally defined as follows.

Definition 4.1. *The strategy profile σ^* is an ex post equilibrium of the mechanism $\langle M_1, \cdots, M_N, q\rangle$ if*

$$u_i(\hat{\theta}(t))q(\sigma^*(t)) \geq u_i(\hat{\theta}(t))q(m_i, \sigma^*_{-i}(t_{-i}))$$

for all $m_i \in M_i$ and all $t \in T$, and all $i \in I$.

That is, for each agent i and type t_i, $\sigma_i^*(t_i)$ maximizes her ex post utility all possible realizations of other agents' types t_{-i} conditional on other agents would play their equilibrium strategies. If a mechanism admits a ex post equilibrium, then it satisfies ex post incentive compatibility.

By the revelation principle, we can again focus on fully reduced direct ex post incentive compatible mechanisms. Again, without further specifications, all mechanisms refer to fully reduced mechanisms in this section.

Definition 4.2. *A fully reduced direct mechanism q is ex post incentive compatible if for any $\theta \in \Theta$, $i \in I$ and $\theta'_i \in \Theta_i$,*

$$u_i(\theta)q(\theta_i, \theta_{-i}) \geq u_i(\theta)q(\theta'_i, \theta_{-i}).$$

Therefore, under an EPIC mechanism, truth-telling is a best response for every agent even if the state is common knowledge (given agent i's strategy set is Θ_i not Θ). It equivalently means that truth-telling is a best response regardless of an agent's belief about the distribution of the other agents' signals. In contrast with DSIC, EPIC requires every agent to (correctly) believe that the other agents are truthful.

4.1 Characterization

Given a payoff environment $\langle \Theta, \{u_i\}_{i \in I} \rangle$, define a binary relation $\rightarrow$ over Θ as follows: $\theta \rightarrow \theta'$ if there exists $i \in \{1, ..., N\}$ such that: (1) $\theta_{-i} = \theta'_{-i}$ and (2) $u_i(\theta) > 0$ or $u_i(\theta') < 0$.

If a list of states $(\theta^1, ..., \theta^J)$[8] satisfies $\theta^1 \rightarrow \theta^2 ... \rightarrow \theta^J$, then this list is called a **path**. Moreover, if $\theta^1 = \theta^J$, then it is a **cycle**. The notation $\theta \rightsquigarrow \theta'$ is used to denote that there is a path from θ to θ', and $\theta \leftrightsquigarrow \theta'$ denotes that there is a cycle from θ to θ'.

It is clear that $\rightsquigarrow$ and $\leftrightsquigarrow$ are both reflexive[9] and transitive, and moreover $\leftrightsquigarrow$ is also symmetric. Since $\leftrightsquigarrow$ is reflexive, symmetric and transitive, it is an equivalence relation. Let C^* denote the equivalence class partition induced by $\leftrightsquigarrow$ on Θ, i.e. θ, θ' are in the same cell in C^* if and only if $\theta \leftrightsquigarrow \theta'$.

It will be useful to abuse notation and extend the $\rightarrow$ relation to *sets* of states: For a pair of sets of states c and c', we say $c \rightarrow c'$ if there exist $\theta \in c$ and $\theta' \in c'$ such that $\theta \rightarrow \theta'$.

[8] The same state is allowed to appear multiple times in the list.

[9] The singleton list $\{\theta\}$ is a (degenerate) path and cycle

We extend $\rightsquigarrow$ and $\leftrightsquigarrow$ to sets of states analogously.

A partition C of Θ is said to be **acyclic** if there does not exist distinct $c, c' \in \Theta$ such that $c \leftrightsquigarrow c'$. In one of the proofs we will show that C^* is the finest acyclic partition of Θ

Proposition 4.1. *q is EPIC if and only if there are probabilities $\{q_c\}_{c \in C^*}$ such that:*

1. *$q(\theta) = q_c$ if $\theta \in c$.*

2. *$q_c \geq q_{c'}$ if $c \rightsquigarrow c'$.*

Proposition 4.1 shows that EPIC mechanisms are mechanisms that respect the $\rightsquigarrow$ relation and the corresponding finest acyclic partition C^* on the state space. If two states θ, θ' are in the same cell of C^*, then EPIC requires the same probability of choosing R in them; if two states are in different cells, then relationship between $q(\theta)$ and $q(\theta')$ needs to agree with the $\rightsquigarrow$ relation. In other words, EPIC mechanisms are only responsive to ordinal information.

The proposition illustrates what EPIC entails in the binary voting environment: Since there are only two alternatives, then in any state an agent either prefers that R be chosen with higher probability or L be chosen with higher probability, modulo indifference. EPIC thus requires that, for any agent, unilaterally deviating to reporting untruthfully weakly reduces the probability of R being chosen in any state where she prefers R, or increases the probability whenever she prefers S. The states are therefore chained by such potential unilateral deviations, and the probabilities of R being chosen must cascade down along the $\rightsquigarrow$ chain to prevent any upstream traffic which represents an untruthful deviation.

4.2 Examples

The Partnership Example

Here we come back to the partnership example introduced in previous section. Recall that the payoff matrix is

$u_i(\theta)$	H	L
H	4	-2
L	-2	1

In this example, each state is a cell of the finest acyclic partition C^*, and all ex post incentive compatible mechanisms can be characterized by the following four inequalities implied by the second part of Proposition 4.1: the probabilities of choosing R when both agents prefer R, q_1 and q_4, shall be greater than their conterparts, q_2 and q_3, when both agents prefer S.

$$\begin{array}{c|ccc} q & 0 & & 1 \\ \hline 0 & q_1 & \geq & q_2 \\ & \rotatebox{-90}{$\geq$} & & \rotatebox{-90}{$\leq$} \\ 1 & q_3 & \leq & q_4 \end{array}$$

Generalized Condorcet jury

We analyze a classical example of a binary social choice problems—Condorcet jury voting as an example. We consider a generalized version of it.

Each agent each gets a binary signal θ_i taking value of 0 or 1. There exists a payoff function u such that $u_i(\theta) = u(\theta)$ for all i and u is permutation invariant, so the value of R depends on how many 1-signals obtain. Define $\epsilon(\theta) := \sum_{i=1}^{N} \theta_i$. Suppose $u(\theta) \geq u(\theta')$ if and only if $\epsilon(\theta) \geq \epsilon(\theta')$, that is, R is more valuable if there are more 1-signals. This situation generalizes the Condorcet Jury model with the interpretation that the agents are jurors to determine whether to convict or acquit a defendant. A 1-signal is a partial evidence that the defendant is guilty, thus the more 1's the more guilty the defendant could be. S represents the decision to acquit and R to convict. The jurors prefer to acquit if there's not enough guilty evidence or to convict otherwise.

Figure 1: Condorcet Jury with $N = 3$

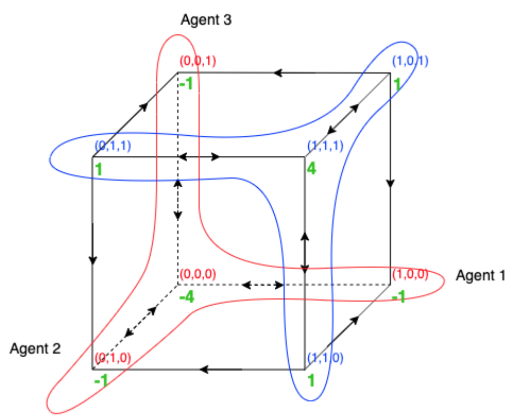
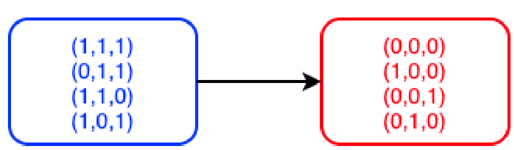


Figure 1 is an example with $N = 3$. Each vertex of the cube is a state with the corresponding payoff $u(\theta)$ underneath it. For example, the vertex $(0, 0, 0)$ is the state every agent gets signal 0, and the payoff of choosing R in this state is -4 which is below the vertex.

The $\rightsquigarrow$ relation over states and the finest acyclic partition C^* and are represented by the arrows and colored sets in the figure. The partition C^* and the $\rightsquigarrow$ relation over sets of states are represented by Figure 2.

Figure 2: Finest acyclic partition C^*



Hence, a mechanism is EPIC if and only if all the states in the blue cell share the same probability of choosing R, q_1, all the states in the blue cell share the same probability of choosing R, q_2, and $q_1 \geq q_2$.

4.3 Existence of Non-constant EPIC Mechanism

We discuss the existence and uniqueness of non-constant EPIC mechanism in this subsection.

Proposition 4.2. *The following statements are equivalent:*

1. *There exists a non-constant EPIC mechanism.*
2. *C^* is not a singleton.*
3. *Θ can be bipartitioned into c_A and c_B such that $c_B \not\rightarrow c_A$.*
4. *There exist $\theta, \theta' \in \Theta$ such that $\theta \not\rightsquigarrow \theta'$.*

How useful Proposition 4.2 is for a particular environment depends on how easy it is to determine C^*. If the state space is large and has no obvious structure, then determining C^* could be computationally exhausting. Moreover, since the proposition is formulated not directly in terms of the environment parameters, but instead indirectly in terms of the abstract structure underpinning the environment, it is difficult to extract much intuition about when and why a non-constant mechanism exists. To have a better understanding in this respect, we propose two sufficient conditions for the existence of non-constant mechanisms that are formulated directly in terms of the environment parameters.

The first condition is directly inherited from the observation in Proposition 3.1 that there are non-constant DSIC mechanisms if and only if there are agents with quasi-private values. Hence, we have the following corollary.

Corollary 4.1. *Suppose there exists at least one agent with quasi-private values. Then there exist non-constant EPIC mechanisms.*

Since a DSIC mechanism must also be an EPIC mechanism, the existence of agents with quasi-private values obviously implies that there are non-constant EPIC mechanisms.

The second condition is motivated by the Condorcet jury voting example: It is not unusual that agents might have identical ex post preferences. Formally,

Definition 4.3. *Agents have common interests if there exists a payoff function u such that sgn$(u_i(\theta)) =$ sgn$(u(\theta))$ for all $i \in I$ and $\theta \in \Theta$.*

Proposition 4.3. *Suppose agents have common interests. Then the ex post Pareto efficient mechanism q^* is EPIC where*

$$q^*(\theta) = \begin{cases} 1, & u(\theta) \geq 0 \\ 0, & u(\theta) < 0 \end{cases}$$

Proposition 4.3 indicates common interests is a sufficient condition for the existence of non-constant EPIC mechanisms.

That the Pareto efficient mechanism is EPIC is not surprising given common interests. It is easy to see that the truth-telling strategy profile that simultaneously maximizes every agent's ex post payoff is an ex post equilibrium, because it is impossible for any agent to increase the payoff by unilateral deviation as the upper bound is already reached.

These two sufficient conditions for existence of non-constant EPIC mechanisms — the existence of agents with quasi-private values and common interests — are satisfied in many common environments, yet there are many other environments in which they are violated. Indeed, any general environment in which there is sufficient preference interdependence and preference heterogeneity would violate both conditions. Do non-constant EPIC mechanisms exist in those environments? Yes, there *might* exist non-constant mechanisms. [10] However, as we will show in the extension to continuous payoff state spaces in the next subsection, these two sufficient conditions are "almost necessary" for the existence of non-constant EPIC mechanisms, in the sense that any continuous type space that violates mild generalizations of both conditions only admits constant EPIC mechanisms.

[10]The example illustrated in Figure 3 in the appendix is a non-constant EPIC mechanism in such an environment.

4.4 Continuous Type Spaces: A Negative Result

In this subsection, we consider continuous payoff state spaces and show that, apart from mild generalizations of quasi-private values and common interests, all other environments admit only constant EPIC mechanism.

To simplify exposition, we assume $I = \{1, 2\}$, $\Theta = [0, 1]^2$, and u_i is continuously differentiable on Θ.[11]

Let us introduce some useful notation: For agent i, let IC_i denote the set of all payoff states in which i is indifferent between S and R, and let BD_i denote the set of all payoff states θ such that for any $\epsilon > 0$, there is a payoff state where i strictly prefers S to R and there is also a payoff state where i strictly prefers R to S in the $\epsilon-$neighborhood of θ under the Euclidean norm. Clearly $BD_i \subset IC_i$ because u_i is continuously differentiable.

Definition 4.4. *Given Θ, the agents' preferences are **generically interdependent** if for any $i \in I$, $\theta \in BD_i$ and $j \neq i$,*

$$\frac{\partial u_i(\theta)}{\partial \theta_j} \neq 0.$$

If preferences are generically interdependent, then when agent i is in a state where she is indifferent between S and R, a slight change in the payoff type of agent j breaks the indifference. In other words, a slight change in j's payoff type matters to i when i is indifferent between S and R. This, of course, cannot hold if i has quasi-private values.

Definition 4.5. *Given Θ, the agents' preferences are **generically heterogeneous** if for any $\theta \in BD_1 \cap BD_2$ and any $\epsilon > 0$, there exists θ' in the $\epsilon-$neighborhood of θ such that $u_1(\theta')u_2(\theta') < 0$.*

If preferences are generically heterogeneous, then for any payoff state where both agents are indifferent between S and R, there is a close enough payoff state where the agents have opposite preferences. Generic preference heterogeneity says that the two agents cannot agree everywhere: In the least, when both of them agree that S and R are equally good, a slight change in the state can cause them to disagree.

We denote $\Theta \setminus (IC_1 \cup IC_2)$ by $\bar{\Theta}$ which is the set of payoff states in which both agents

[11] If non-constant EPIC mechanisms do not exist in a two agent collective choice problem, it does not exist in other collective choice problems.

have strict ex post preferences. Since u is continuously differentiable, the set $IC_1 \cup IC_2$ is of Lebesgue measure zero in $\mathbb{R}^2$.

Proposition 4.4. *If the agents' preferences are generically interdependent and generically heterogeneous, then any ex post incentive compatible mechanism is constant over* $\bar{\Theta}$.

Proposition 4.4 shows that EPIC can be restrictive at times, and we would like to suggest that this is often the case, as environments where preferences are not generically interdependent and heterogeneous are exceptions rather than the norm. To illustrate the point, suppose in addition that u_1 and u_2 are both strictly increasing in θ. Therefore BD_1 and BD_2 both are curves cutting through the $[0,1]^2$ square. The proposition implies that all EPIC mechanisms are constant as long as there is not a vertical section on the BD_1 curve, there is not a horizontal section on the BD_2 curve, and the two curves overlap only on finitely many points.

Why do non-constant mechanisms fare better on finite payoff sate spaces? One way to understand the reason is thinking about the finite state space as a coarse, low-resolution discretization of the continuous space. For example, instead of being fully conscious of her exact payoff type which can be any real number between 0 and 1, agent $i = 1,2$ only roughly rounds the her payoff type to the first decimal place, and consequently she in effect has only 10 payoff types: $< 0 \sim 1 >, ..., < 0.9 \sim 1 >$. On the continuous payoff space, such rough rounding implies that the BD_i curve traces along the $t_1 = 0, 0.1, ..., 1$ and $t_2 = 0, 0.1, ..., 1$ grid lines. Obviously this irons the otherwise swerving BD_i curve into horizontal and vertical sections, and moreover the two curves, otherwise not overlapping, are more likely to be squeezed onto the same section of a grid line. Therefore discretization makes the preferences less generically interdependent and less generically heterogeneous, thus admitting non-constant EPIC mechanisms.

5 Interim Dominant Strategy Incentive Compatibility

It is useful to first define what interim dominant strategies are.

Definition 5.1. *For agent i, strategy σ_i is an interim dominant strategy in mechanism $\langle M_1, \cdots, M_N, q\rangle$ if $U_i(\sigma_i(t_i), \sigma_{-i}|t_i) \geq U_i(m_i, \sigma_{-i}|t_i)$ for any $t_i \in T_i$, $m_i \in M_i$ and $\sigma_{-i} : T_{-i} \to M_{-i}$, where $U_i(\sigma_i, \sigma_{-i}|t_i)$ denotes agent i's interim expected payoff if she follows strategy σ_i, other agents follow strategies given by σ_{-i}, and her type is t_i.*

In plain words, σ_i prescribes for every type t_i of agent i a strategy that maximizes her interim expected payoff regardless of what strategies other agents follow. It is worth noting that agent i's *subjective* belief $\hat{\beta}_i(t_i)$ is used in computing her interim expected payoff.

We can then define interim dominant strategy equilibrium and interim dominant strategy incentive compatibility as follows.

Definition 5.2. *Strategy profile σ^* is an interim dominant strategy equilibrium of mechanism $\langle M_1, \cdots, M_N, q\rangle$ if σ_i^* is an interim dominant strategy for every agent $i = 1, ..., N$.*

If a mechanism admits an interim dominant strategy equilibrium, then it satisfies interim dominant strategy incentive compatibility.

As usual, there is a general revelation principle that can simplify analysis by allowing us to focus on direct mechanisms. Later on will discuss a variation of the direct mechanism and respectively develop another revelation principle.

Lemma 5.1. *(Revelation Principle 1) Let σ^* be an interim dominant strategy equilibrium of any mechanism $\langle M_1, \cdots, M_N, q\rangle$. Construct a direct mechanism $\langle \overline{M}_1, \overline{M}_2, \ldots, \overline{M}_N, \overline{q}\rangle$ as follows:*

1. *$\overline{M}_i = T_i$ for all $i \in I$.*

2. *$\overline{q}(t) = q(\sigma^*(t))$ for any $t \in T$.*

Then in this direct mechanism, the strategies given by $\overline{\sigma}_i^*(t_i) = t_i$ for all $i \in I$ and $t_i \in T_i$ form an interim dominant strategy equilibrium. Moreover, $\overline{\sigma}^*$ and σ^* are outcome equivalent.

Definition 5.3. *A direct mechanism $\langle T_1, ..., T_N, q\rangle$ is interim dominant strategy incentive compatible if*

$$U_i(t_i, \sigma_{-i}|, t_i) \geq U_i(t_i', \sigma_{-i}|, t_i)$$

all $\sigma_{-i} : T_{-i} \to T_{-i}$, for all $t_i, t_i' \in T_i$ and all $i \in I$, where

$$U_i(t_i', \sigma_{-i}|, t_i) = \sum_{t_{-i} \in T_i} \hat{\beta}_i(t_i)[t_{-i}] q(t_i', \sigma_{-i}(t_{-i})) u_i(\hat{\theta}(t_i, t_{-i}))$$

is agent i's expected payoff of reporting t_i' given her type is t_i and other agents' strategy is σ_{-i}.

5.1 IDSIC mechanisms are higher-order belief-free

An IDSIC mechanism is by definition strategically belief-free, but it is not informationally belief-free, because whether a strategy is interim dominant depends on an agent's belief. In this section, though, we show that an agent's *first-order belief* is sufficient to determine whether a strategy is interim dominant, whereas higher-order beliefs do not matter.

Recall that agent i's first-order belief, when her type is t_i, assigns probability $\hat{b}_i(t_i)[\theta_{-i}]$ to the event that the type profile of the other agents is θ_{-i}.

Definition 5.4. *An strategy profile σ is **higher-order belief-independent** if $\sigma_i(t_i) = \sigma_i(t_i')$ for any t_i, t_i' such that $\hat{\theta}_i(t_i) = \hat{\theta}_i(t_i')$ and $\hat{b}_i(t_i) = \hat{b}_i(t_i')$.*

In words, a higher-order belief-independent strategy profile prescribes the same strategy for all types of agent i that have the same payoff type and first-order belief.

We are ready to start showing the connection between IDSIC and the higher-order belief-free property, To set the stage, let us introduce one more notation: Let ***IDM***$_i(t_i)$ [12] denote the set of all messages m_i such that there is some interim dominant strategy σ_i where $\sigma_i(t_i) = m_i$. It is obvious that σ_i is an interim dominant strategy if and only if $\sigma_i(t_i) \in$ ***IDM***$_i(t_i)$ for every $t_i \in T_i$.

Lemma 5.2. ***IDM***$_i(t_i) =$ ***IDM***$_i(t_i')$ *for any $i \in I$ and any $t_i, t_i' \in T_i$ such that $\hat{\theta}_i(t_i) = \hat{\theta}_i(t_i')$ and $\hat{b}_i(t_i) = \hat{b}_i(t_i')$.*

The Lemma essentially shows that if two types of agent i has the same payoff type and first-order beliefs, then they are strategically "identical" if we focus on interim dominant

[12] *IDM* represents "interim dominant messages"

strategy equilibria, because the two types are presented with the same set of messages that dominate other messages at the respective interim stages. This observation leads to the following Proposition.

Proposition 5.1. *Mechanism $\langle M_1, ..., M_N, q\rangle$ has an interim dominant strategy equilibrium if and only if it has an interim dominant strategy equilibrium that is higher-order belief-independent.*

The Proposition shows that, in addition to the built-in strategically belief-free property, any IDSIC mechanism also has what we may call the "informationally higher-order belief-free" property, as there is always an interim dominant strategy equilibrium that does not depend on higher-order beliefs. In comparison, the informationally belief-free property introduced in the Introduction is, in this sense, informationally higher-order *and* first-order belief-free.

Proposition 5.1 and Lemma 5.3 suggest that, to find an interim dominant strategy equilibrium, or to find an IDSIC mechanism, it is without loss of generality to focus only on higher-order belief-independent strategy profiles where only payoff types and first-order beliefs matter. Since the payoff types and first-order beliefs type are of a particular importance with respect to IDSIC, it is useful to refer to them with special terminology:

Definition 5.5. *Agent i's **reduced type** $h_i := (\theta_i, b_i)$ consist of her payoff type and her first-order belief. Agent i's reduced type space is $H_i := \Theta_i \times B_i$.*

Naturally, there is a Revelation Principle with respect to reduced types.

Lemma 5.3. *(Revelation Principle 2) Let σ^* be a higher-order belief-independent interim dominant strategy equilibrium of a mechanism $\langle M_1, M_2, \ldots, M_N, q\rangle$. Construct a reduced direct mechanism $\langle \overline{M}_1, \overline{M}_2, \ldots, \overline{M}_N, \overline{q}\rangle$ as follows:*

1. *$\overline{M}_i = H_i$ for all $i \in I$.*
2. *$\overline{q}(h) = q\big(\sigma^*(h)\big)$ for all $h \in H = H_1 \times ... \times H_N$.*

Then in this reduced direct mechanism the strategies given by $\overline{\sigma}^*_i(t_i) = (\hat{\theta}_i(t_i), \hat{b}_i(t_i))$ for all $i \in I$ and $t_i \in T_i$ form an interim dominant strategy equilibrium that is higher-order belief-independent. Moreover, $\overline{\sigma}^*$ and σ^* are outcome equivalent.

In the reduced direct mechanism as constructed in the Lemma, an agent reports her reduced type, instead of her original type. It may seem that reporting the reduced type $h_i = (\theta_i, b_i)$ seems to be a more tedious job than reporting the original type t_i, as agent i might need to derive her payoff type and first-order belief from her type t_i, which introduces additional labor. This appearance, though, is merely an artifact of the Harsanyian formulation of the type space. Indeed, the type t_i is an abstraction of an agent's infinite belief hierarchies, which the agent is likely to be conscious of, whereas his payoff type and first-order belief are more concrete and salient objects and hence are more likely to be in her awareness and more easily to be explicitly reported.

Remark 5.1. *All results in this subsection are not restricted to the special two-alternative setting of the paper, as they also apply to settings with any finite number of alternatives.*

5.2 Characterization

In this subsection we characterize the set of all IDSIC reduced direct mechanisms.

For any agent i and $h_i = (\theta_i, b_i) \in H_i$ define

$$\overline{\alpha}_i(\theta_i, b_i) := \sum_{\{\theta_{-i} | u_i(\theta_i, \theta_{-i}) > 0\}} b_i(\theta_{-i}) u_i(\theta_i, \theta_{-i}),$$

$$\underline{\alpha}_i(\theta_i, b_i) := \sum_{\{\theta_{-i} | u_i(\theta_i, \theta_{-i}) < 0\}} b_i(\theta_{-i}) u_i(\theta_i, \theta_{-i}),$$

and

$$\alpha_i(\theta_i, b_i) := \sum_{\theta_{-i} \in \Theta_{-i}} b_i(\theta_i) u_i(\theta_i, \theta_{-i}).$$

Clearly we have $\underline{\alpha}_i(h_i) \leq 0 \leq \overline{\alpha}_i(h_i)$ and $\underline{\alpha}_i(h_i) + \overline{\alpha}_i(h_i) = \alpha_i(h_i)$. The sign of $\alpha_i(h_i)$ is positive/negative when agent i at the interim stage prefers R/S. $\overline{\alpha}_i(h_i)$ is the expected payoff to agent i of reduced type h_i from the choice rule that chooses R whenever R is ex post preferred to S. Similarly $\underline{\alpha}_i(h_i)$ is the expected payoff from the choice rule that chooses R whenever S is ex post preferred to R.

Define $H_i^+ := \{h_i \in H_i | \alpha_i(h_i) > 0\}$, $H_i^- := \{h_i \in H_i | \alpha_i(h_i) < 0\}$, and $H_i^0 := \{h_i \in H_i | \alpha_i(h_i) = 0\}$ for each $i \in I$. H_i^+ is the set of reduced types in which agent i strictly prefers R over S in the interim stage, H_i^- is the set of reduced types in which agent i strictly prefers S over R in the interim stage, and H_i^0 is the set of reduced types in which agent i is indifferent.

The following lemma will be useful for us to develop a characterization of IDSIC reduced revelation mechanisms.

Lemma 5.4. *q is IDSIC if and only if for any $i = 1, ..., N$, $h_i, h_i' \in H_i$, and $h_{-i}, h_{-i}' \in H_{-i}$:*

$$\underline{\alpha}_i(h_i)\big(q(h_i, h_{-i}) - q(h_i', h_{-i})\big) + \overline{\alpha}_i(h_i)\big(q(h_i, h_{-i}') - q(h_i', h_{-i}')\big) \geq 0. \tag{5.1}$$

The Lemma is based on the observation that, among many possible incentive constraints, the only one that is binding for agent i of reduced type h_i corresponds to the case that the other agents coordinate on on the same message profile (h_{-i}) whenever agent i of type h_i ex post prefers S, or they coordinate on another message profile (h_{-i}') whenever agent i of type h_i ex post prefers R.

The characterization of IDSIC reduced direct mechanisms will depend on crucial parameters defined as follows: For agent i and reduced type h_i,

$$\rho_i(h_i) := \begin{cases} \dfrac{\overline{\alpha}_i(h_i)}{-\underline{\alpha}_i(h_i)} & \text{if } h_i \in H_i^+ \\ 1 & \text{if } h_i \in H_i^0 \\ \dfrac{-\underline{\alpha}_i(h_i)}{\overline{\alpha}_i(h_i)} & \text{if } h_i \in H_i^- \end{cases}$$

And for agent i,

$$\rho_i := \min_{h_i \in H_i} \rho_i(h_i).$$

It is easy to verify that $\rho_i \geq 1$.

We first characterize IDSIC reduced direct mechanisms where there are no types indifferent between S and R at the interim stage. The general characterization is given later on at Proposition 5.3.

Proposition 5.2. *Suppose $H_i^0 = \emptyset$ for all i. A reduced direct mechanism q is IDSIC if and only if for any agent i and $h_i \in H_{-i}$, there are two numbers $\overline{q}_i(h_{-i})$ and $\underline{q}_i(h_{-i})$,* *where* $\overline{q}_i(h_{-i}) \geq \underline{q}_i(h_{-i})$*, such that:*

1.

$$q(h_i, h_{-i}) = \begin{cases} \overline{q}_i(h_{-i}) & \textit{if } h_i \in H_i^+ \\ \underline{q}_i(h_{-i}) & \textit{if } h_i \in H_i^- \end{cases}$$

2.

$$\max_{h_{-i} \in H_{-i}} \Big(\overline{q}_i(h_{-i}) - \underline{q}_i(h_{-i})\Big) \leq \rho_i \min_{h_{-i} \in H_{-i}} \Big(\overline{q}_i(h_{-i}) - \underline{q}_i(h_{-i})\Big).$$

Condition 1 is noteworthy in that all types with the same interim ordinal preference are treated equally by the mechanism. In other words, an IDSIC reduced mechanism elicits preference rankings but has to ignore preference intensities. When H_i^0 is empty, the interim preference ranking that agent i can have is at most two, and hence conditional on the other agents' reports h_{-i}, agent i's marginal influence on the choice probability is binary: high (leading to $\overline{q}_i(h_{-i})$) or low (leading to $\underline{q}_i(h_{-i})$). It is then easy to see that the mechanism can be simulated by a much simpler mechanism, which we call a **binary voting mechanism**. In a binary voting mechanism, every agent is given two messages: R and S. The R message is interpreted as a vote supporting the alternative R, whereas the S message is a vote supporting the alternative S. A unilateral switch from sending message S to sending R raises the chance that alternative R is chosen. Binary voting is clearly the most common and important voting format used for bi-candidate public choice.

Condition 2 shows, despite that an IDSIC reduced direct mechanism does not elicit an agent's interim preference intensity, that intensity still constrains the mechanism. To see this, observe that when the other agents report h_{-i}, agent i's marginal voting power can be represented by $\overline{q}_i(h_{-i}) - \underline{q}_i(h_{-i})$, which is the marginal increase in the probability of R being chosen if i unilaterally switches from reporting to be a type in H_i^- to reporting to be a type in H_i^+. Condition 2 says that this marginal voting power cannot fluctuate too widely with h_{-i}, i.e, the ratio of agent i's maximal marginal voting power $\max_{h_{-i}}(\overline{q}_i(h_{-i}) - \underline{q}_i(h_{-i}))$ to his minimal marginal voting power $\min_{h_{-i}} \overline{q}_i(h_{-i}) - \underline{q}_i(h_{-i})$ cannot be higher than the parameter ρ_i which depends on the agent's interim preference intensities.

There is a special class of mechanisms such that every agent's marginal voting power is constant, and hence Condition 2 is immediately satisfied. This class of mechanism, which we will formally define as "additive mechanisms" in Definition 5.6 and discuss with more details, will turn out to be IDSIC with respect to any type space based on any payoff state space. In other words, additive mechanisms are not only robust, but also versatile.

On the other hand, majority voting mechanisms, where R is chosen if and only if there are more than k votes for R, are not IDSIC exactly because an agent's marginal voting power fluctuates too much. Indeed, when the agent is pivotal, his marginal voting power is the maximal 1, whereas when he is not pivotal, his marginal power is the minimal 0.

Now we drop the assumption that H_i^0 is empty for every i and present the characterization for IDSIC reduced direct mechanisms in this general case.

Proposition 5.3. *A reduced direct mechanism q is IDSIC if and only if for any agent i and $h_{-i} \in H_{-i}$ there are two numbers, $\overline{q}_i(h_{-i})$ and $\underline{q}_i(h_{-i})$ where $\overline{q}_i(h_{-i}) \geq \underline{q}_i(h_{-i})$, such that:*

1.

$$q(h_i, h_{-i}) \begin{cases} = \overline{q}_i(h_{-i}) & \textit{if } h_i \in H_i^+ \\ \in [\underline{q}_i(h_{-i}), \overline{q}_i(h_{-i})] & \textit{if } h_i \in H_i^0 \\ = \underline{q}_i(h_{-i}) & \textit{if } h_i \in H_i^- \end{cases}$$

2. *If $H_i^0 = \emptyset$, then*

$$\max_{h_{-i} \in H_{-i}} \Big(\overline{q}_i(h_{-i}) - \underline{q}_i(h_{-i})\Big) \leq \rho_i \min_{h_{-i} \in H_{-i}} \Big(\overline{q}_i(h_{-i}) - \underline{q}_i(h_{-i})\Big).$$

If $H_i^0 \neq \emptyset$, then $q(h_i, h_{-i}) - q(h_i', h_{-i})$ is independent of h_{-i}.

The biggest difference between Proposition 5.2 and Proposition 5.3 is that the existence of an reduced type that is interim indifferent between S and R immediately pushes the parameter ρ_i to the lower bound 1, which implies that agent i's marginal voting power has to be constant.

5.3 Example

Here we come back to the partnership example again. Recall that we assume there is a unique belief type associated with each payoff type and the beliefs all come from a common prior such that every state has a probability of 1/4. The payoff matrix is

$u_i(\theta)$	H	L
H	4	-2
L	-2	1

In order to find all IDSIC mechanisms in this example, we first calculate agents' interim preferences. For example, $\alpha_1(H) = \frac{1}{4}4 + \frac{1}{4}(-2) = 1/2$. Then we have $\alpha_2(H) = 1/2 > 0$ and $\alpha_1(L) = \alpha_2(L) = -1/4 < 0$. Monotonicity conditions require $q(H,H) \geq q(L,H), q(H,L) \geq q(L,L), q(H,H) \geq q(H,L), q(L,H) \geq q(L,L)$.

Second, we compute each agent's intensity ratio. When agent 1 is the H type, $\overline{\alpha}_1(H) = \frac{1}{4}4 = 1$ and $\underline{\alpha}_1(H) = \frac{1}{4}(-2) = -1/2$. Then $\rho_1(H) = \overline{\alpha}_1(H)/|\underline{\alpha}_1(H)| = 2$. Similarly, we have $\rho_1(H) = \rho_2(H) = \rho_2(L) = 2$ which leads $\rho_1 = \rho_2 = 2$. Smoothness conditions are the following, $\frac{1}{2} \leq \frac{q(H,H) - q(L,H)}{q(H,L) - q(L,L)} \leq 2$, $\frac{1}{2} \leq \frac{q(H,H) - q(H,L)}{q(L,L) - q(L,H)} \leq 2$.

Now suppose $q(H,H) = 1$, $q(H,L) = 1/2$, and $q(L,L) = 0$. Then q is IDSIC if and only if $1/4 \leq q(L,H) \leq 3/4$.

5.4 Universal existence of non-constant IDSIC mechanisms

In this section, we show that non-constant IDSIC mechanisms universally exist over all type spaces. In fact, they are closely related to additive mechanisms that we already informally mentioned in Section 5.2. Here we formally define them:

Definition 5.6. *An indirect mechanism $\langle M_1, ..., M_N, q\rangle$ where $|M_i| \geq 2$ for all i is an **additive mechanism** if there exist functions $\pi_i^q : M_i \to [0,1]$ such that $q(m_1, ..., m_N) = \sum_{i\in I} \pi_i^q(m_i)$.*

In an additive mechanism, every message m_i has a "score" $\pi_i^q(m_i)$ attached to it, and the eventual probability that R is chosen is the sum of the scores of the chosen messages.

Proposition 5.4. *Fix a payoff environment. A mechanism* $(M_1, ..., M_N, q)$ *is IDSIC in all type spaces if and only if it is additive.*[13]

The Proposition establishes the universal existence of non-constant IDSIC mechanisms, because additive mechanisms are always IDSIC, and because most additive mechanisms are non-constant. Moreover, it also shows that additive mechanisms are versatile, in the sense that they are IDSIC with respect too many (in this case, all) type spaces, and hence are expected to function well even if the underlying type space is uncertain. We will discuss this versatility aspect in the next subsection.

Additive mechanisms also have a close relation with a class of well-known social choice rules: random dictatorships.

Definition 5.7. *A social choice rule* $q : T \to [0,1]$ *is a* ***random dictatorship*** *if there exist numbers* $(\lambda_i^q)_{i\in I}$*, where* $\lambda_i^q \in [0,1]$ *and* $\sum_{i\in I} \lambda_i \le 1$*, and functions* $\mu_i^q : T_i \to [0,1]$ *where* $\mu_i^q(t_i) = 1$ *if* $\alpha_i(t_i) > 0$ *and* $\mu_i^q(t_i) = 0$ *if* $\alpha_i(t_i) < 0$*, and a constant* $\tilde{\lambda}^q \in [0, 1 - \sum_{i\in I} \lambda_i]$*, such that*

$$q(t) = \sum_{i\in I} \lambda_i^q \mu_i^q(t_i) + \tilde{\lambda}^q.$$

Under a random dictatorship rule, agent i has chance (with a probability of λ_i^q) to be the "dictator" in the event of which her interim preferred alternative is chosen by the rule. There is also a chance (with a probability of $\tilde{\lambda}^1$) that no agent is chosen to be the random dictator and R is in this case chosen with certainty.

The following two result establish the close connection between additive mechanisms and random dictatorship rules.

Proposition 5.5. *Suppose mechanism* $(M_1, ..., M_N, q)$ *is an additive mechanism, and* σ^* *is an interim dominant strategy equilibrium of it. Then the social choice function* $q \circ \sigma^* : T \to [0,1]$ *is a random dictatorship.*

Proposition 5.6. *If* q *is a random dictatorship, then the direct mechanism* (T, q) *is an additive voting mechanisms.*

[13]We rule out mechanisms with trivial message sets $|M_i| = 1$ for some i in the proposition.

How should we interpret the results? Do they mean that IDSIC mechanisms in general are not ideal, because they are "dictatorial" by nature? We recommend that "dictatorship" not be taken at the face value, because random dictatorships are social choice rules, not mechanisms. In other words, the actual mechanism that induces a random dictatorship need not entail the seemingly undemocratic element of someone dictating a decision, possibly against the prevailing public opinion. Proposition 5.5 shows that additive mechanisms, which can be perfectly democratic and fair *formally and actually* if every agent has the same marginal voting power π_i^q, induces a random dictatorship rule nonetheless. We observe that the random "dictator" under a random dictatorship is as dictatorial as the median voter in the Median Voter Theorem who *de facto* decides the voting outcome, which is to say not very much dictatorial at all.

5.5 Versatile IDSIC Mechanisms

In this sub-section we show that there are IDSIC mechanisms that have a nice property: versatility. A mechanism is versatile if it is able to handle a wide variety of situations and is and less reliant on the configuration of the environment. "Aye-Nay" voting used in many legislative procedures, for instance, is versatile, as it handles a budgeting bill as well as an impeachment motion, despite the great difference between these two issues. It is no wonder why such a mechanism has been consistently used over history and across the globe.

Now we set the stage to formally discuss versatility. We identify three building blocks of any collective choice environment: The underlying payoff state space Θ, the payoff functions $\{u_i\}_{i \in I}$ defined with respect to Θ, and the type space T. We jointly call these three elements the **environment** concerning the collective choice problem, and denote it as E. Let $\mathscr{E}$ denote the set of all environments, where, specifically, the underlying payoff state space may vary.

Definition 5.8. *A mechanism is **versatile** with respect to a given incentive compatibility if the mechanism satisfies that incentive compatibility given any $E \in \mathscr{E}$.*

Our discussion on DSIC and EPIC shows that no non-constant mechanisms are versatile with respect to those two incentive compatibility conditions. However, non-constant mechanisms that are versatile with respect to IDSIC do exist. In fact, as the following result shows, they have a very simple characterization.

Proposition 5.7. *Any binary additive voting rule is versatile with respect to IDSIC.*

Proposition 5.7 shows that even if the designer is agnostic not only to the belief environment (type spaces) but also to the payoff environment, she is able to find a mechanism that is IDSIC.

Moreover, when we know more about the environment, we can find more mechanisms other than binary additive voting mechanisms that are IDSIC.

Fix any type space, and an agent i and her type t_i. Let $v_i(t_i) := \dfrac{\overline{\alpha}_i(\hat{\theta}_i(t_i), \hat{b}_i(t_i))}{|\underline{\alpha}_i(\hat{\theta}_i(t_i), \hat{b}_i(t_i))|}$ be called this type's **virtual type**. For any environment E let $V := V_1 \times ... V_N$ where $V_i = \{v_i(t_i) : t_i \in T_i\}$ denotes the **virtual type space** induced by E. Note that the virtual type space of any original type space is a subspace of $(\mathbb{R} \cup \{-\infty, \infty\})^N$. Given any environment, an agent can compute her virtual value (which is a real number) by considering her payoff type and first-order belief.

Proposition 5.8. *Given a binary voting mechanism q. If an environment E induces a virtual type space V such that $v_i \notin (1/\eta_i, \eta_i)$ for all $v_i \in V_i$ and $i \in I$, then q is IDSIC on E, where*

$$\eta_i = \frac{\max_{m_{-i}} q(1, m_{-i}) - q(0, m_{-i})}{\min_{m_{-i}} q(1, m_{-i}) - q(0, m_{-i})}$$

for all $i \in I$.

If the designer knows more about the environment, the mechanism may deviate more from binary additive voting without violating IDSIC.

5.6 *DSIC* $=$ *EPIC* $\cap$ *IDSIC*

In this subsection, we show that the joint of the two "qualified" robustness conditions — IDSIC and EPIC — is exactly DSIC.

Proposition 5.9. *A mechanism is DSIC if and only if it is EPIC and IDSIC.*

The "only if" direction is immediate. For better exposition, we assume the type space is the payoff type space. In order to understand the intuition behind the "if" direction, we shall first recall the defining property of DSIC is that it can not be responsive to any not privately informed agents. Consider a not privately informed agent i. Suppose agent i's interim preferences is R, $\alpha_i(\theta_i) > 0$, upon observing her own signal θ_i, IDSIC requires $\theta_i \in \text{argmax}_{\theta_i'} q(\theta_i', \theta_{-i})$ for all θ_{-i}. Since agent i doesn't have quasi private value, there exists θ_{-i} such that $u_i(\theta_i, \theta_{-i}) < 0$ in which her ex post preferences is L. EPIC demands $\theta_i \in \text{argmin}_{\theta_i'} q(\theta_i', \theta_{-i})$ for that θ_{-i}. As a result, $q(\theta_i, \theta_{-i})$ is a constant function over θ_i.

In other words, EPIC respects agents' ex post preferences while IDSIC follows agents' interim preferences. Disagreements between these two arise whenever the agent does not have quasi private values. The tension is so severe that complying with one leads to a violation of the other. Consequently, any EPIC and IDSIC mechanism cannot be responsive to any agent who does not have quasi private value.

6 Conclusion

We study robust mechanisms without transfers in a setting where there are two alternatives, the agents' preferences are interdependent, and the underlying type space is rich. Three notions of robustness (incentive compatibilities) are examined. The first two are widely used: dominant strategy incentive compatibility and ex post incentive compatibility. The former permits nothing but constant mechanisms and the latter is more lenient, but only so in finite type spaces — when the type space is continuous, non-constant mechanisms are again ruled out under weak conditions.

The interdependent values setting enables us to study another natural notion of robustness— *interim dominant strategy incentive compatibility*. It requires that each agent has a weakly interim dominant strategy— that is, conditional on each agent's own private information, her strategy must maximize her expected payoff for all possible strategies the other agents could use. We establish a revelation principle that allows the mechanisms to simply ask the agents to reveal their payoff type and first-order beliefs. The characterization suggests a simple binary voting rule: Each agent reports Yes/No to the mechanism. Moreover, if the binary voting rule is also additive (each agent's influence is independent with other agents reports), then the indirect mechanism is versatile:

It admits interim dominant strategy equilibrium on all payoff environments and all corresponding type spaces.

References

Azevedo, Eduardo M and Eric Budish (2019), "Strategy-proofness in the large." *The Review of Economic Studies*, 86, 81–116.

Barberà, Salvador (2011), "Strategy-proof social choice." In *Handbook of Social Choice and Welfare* (Kenneth J Arrow, Amartya Sen, and Kotaro Suzumura, eds.), volume 2, chapter 25, 731–831, Elsevier, Netherlands: North Holland.

Bergemann, Dirk and Stephen Morris (2005), "Robust mechanism design." *Econometrica*, 73, 1771–1813.

Börgers, Tilman and Jiangtao Li (forthcoming), "Strategically simple mechanisms." *Econometrica*.

Crémer, Jacques and Richard P McLean (1985), "Optimal selling strategies under uncertainty for a discriminating monopolist when demands are interdependent." *Econometrica*, 53, 345–361.

Gershkov, Alex, Benny Moldovanu, and Xianwen Shi (2016), "Optimal voting rules." *The Review of Economic Studies*, 84, 688–717.

Gibbard, Allan (1973), "Manipulation of voting schemes: a general result." *Econometrica*, 41, 587–601.

Harsanyi, J (1967/1968), "Games with incomplete information played by bayesian agents i-iii." *Management Science*, 14.

Hylland, Aanund (1980), "Strategy proofness of voting procedures with lotteries as outcomes and infinite sets of strategies." *Unpublished paper, University of Oslo*.

Jehiel, Philippe, Moritz Meyer-ter Vehn, Benny Moldovanu, and William R Zame (2006), "The limits of ex post implementation." *Econometrica*, 74, 585–610.

Mertens, Jean-François and Shmuel Zamir (1985), "Formulation of bayesian analysis for games with incomplete information." *International Journal of Game Theory*, 14, 1–29.

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

Satterthwaite, Mark Allen (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.

A Properties of EPIC Mechanisms

A.1 Monotonicity and efficiency

It turns out Monotonicity plays a major role on the efficiency property of EPIC mechanisms. We define $\Theta_+ := \{\theta \in \Theta | u_i(\theta) > 0$ for all $i \in I\}$ and $\Theta_- := \{\theta \in \Theta | u_i(\theta) < 0$ for all $i \in I\}$. Θ_+ is the set of states in which all agents strictly prefer R over S, and Θ_- is the set of states in which all agents strictly prefer S over R.

Definition A.1. *A mechanism q is ex post Pareto efficient if $q(\theta) = 1$ for all $\theta \in \Theta_+$ and $q(\theta) = 0$ for all $\theta \in \Theta_-$.*

A mechanism q is ex post efficient if it respects unanimity of strict preference on the part of the agents.

Definition A.2. *A mechanism q is ex post efficient in the range if $\min_{\theta \in \Theta_+} q(\theta) \geq \max_{\theta \in \Theta_-} q(\theta)$.*

A mechanism q is ex post efficient in the range if the smallest probability of choosing R when all agents prefer R is higher than the greatest probability of choosing R when when all agents prefer S.

Definition A.3. *A mechanism q is grossly Pareto inefficient if $\max_{\theta \in \Theta_+} q(\theta) < \min_{\theta \in \Theta_-} q(\theta)$.*

A mechanism q is grossly efficient in the range if the largest probability of choosing R when all agents prefer R is lower than the smallest probability of choosing R when when all agents prefer S.

In any state θ, agent i may be of the following three types in terms of her preference over R vs. S: he may prefer R to S (sgn$(u_i(\theta)) > 0$), he may be indifferent (sgn$(u_i(\theta)) = 0$), or he may prefer S to R (sgn$(u_i(\theta)) < 0$).

Definition A.4. *$u_i(\theta)$ is monotone with respect to θ_j if there is a linear order $\succ_j^{(i)}$ on Θ_j such that $\theta_j \succ_j^{(i)} \theta'_j$ implies*

$$\textit{sgn}(u_i((\theta_j, \theta_{-j})) \geq \textit{sgn}(u_i((\theta'_j, \theta_{-j}))$$

for all θ_{-j}.

Monotonicity means that agent i has a clear interpretation of agent j's private signal: higher ranked θ_j implies higher payoff of choosing R, regardless of θ_{-j}. In the case that u_i is not monotone w.r.t. θ_j, agent i's interpretation of θ_j depends on the realization of other private signals.

Also observe that, according to our definition, every private value utility function is monotone.

We say that the environment or $\{u_i\}_{i\in I}$ is monotone if every agent has monotone preference. Formally,

Definition A.5. *$\{u_i\}_{i\in I}$ is monotone if for each agent i there is a collection of linear orders $\{\succ_j^{(i)}\}_{j\in I}$ on $\{\Theta_j\}_{j\in I}$ such that if $\theta_j \succ_j^{(i)} \theta'_j$, then*

$$\textit{sgn}(u_i((\theta_j, \theta_{-j})) \geq \textit{sgn}(u_i((\theta'_j, \theta_{-j}))$$

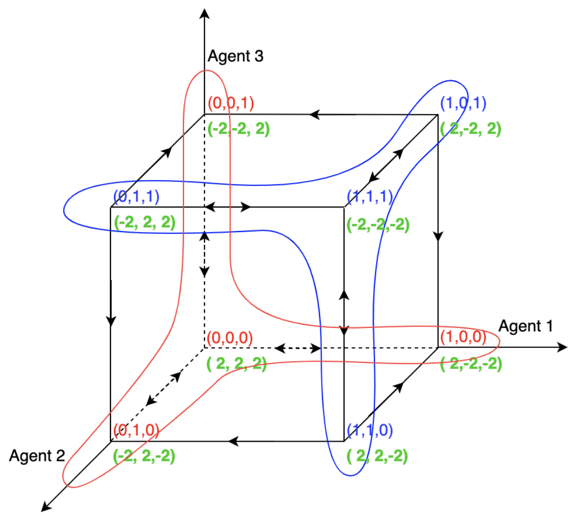
for all θ_{-j} and all i.

Though monotonicity seems to be strong, it is not strong enough to insure every EPIC mechanism is Pareto efficient. Figure 3 shows an extreme counterexample. Similar to Figure 1, each agent has two possible signals $\{0, 1\}$, each vertex of the cube represents a state, and the green vector under each state is the corresponding payoff vector. For example, vertex $(0, 1, 1)$ represents the state agent 1 gets signal 0 and both agent 2 and 3 get signal 1. The payoff vector $(-2, 2, 2)$ means the ex post payoff of choosing R in state $(0, 1, 1)$ is -2 for agent 1, and 2 for agent 2 and agent 3. It can be verified that $\{u_i\}_{i\in I}$ is monotone, but no EPIC mechanism is Pareto efficient. Furthermore, since $\Theta_- = \{(1, 1, 1)\}$, $\Theta_+ = \{(0, 0, 0)\}$ and $(1, 1, 1) \rightsquigarrow (0, 0, 0)$, any non-constant EPIC mechanism is grossly Pareto inefficient, i.e. $\max_{s\in\Theta_+} q(s) \leq \min_{s\in\Theta_-} q(s)$.

Hence, in this extreme example, any EPIC mechanism "cannot" response to any private information in the sense that responding to private information leads to efficiency loss.

A key feature of the environment is that though each agent i has an interpretation of each signal θ_j which is captured by the linear order $\succ_j^{(i)}$, agents disagree with each other regarding the interpretations of signals, for example, agent 1 thinks $\theta_1 = 1$ is a "good" signal for R, while agents 2 and 3 think $\theta_1 = 1$ is a "bad" signal for R. This observation

Figure 3: An example



leads to the following strengthening of monotonicity.

Definition A.6. *$\{u_i\}_{i\in I}$ is uniformly monotone if it is monotone and the linear order $\{\succ_j^{(i)}\}_{j\in I}$ is independent of i. That is, there is a collection of linear orders $\{\succ_j\}_{j\in I}$ on $\{\Theta_j\}_{j\in I}$ such that such that if $\theta_j \succ_i \theta_j'$, then*

$$sgn(u_i((\theta_i, \theta_{-i})) \geq sgn(u_i((\theta_i', \theta_{-i}))$$

for all θ_{-j} and all i.

Monotonicity requires that each agent i has clear (and her own) interpretation of θ_j. On top of that, uniform monotonicity makes sure every agent's interpretation is the same which is much stronger than what monotonicity asks for. However, both cases are common in reality. For example, most college admission officers would agree higher SAT scores could be the signal of competitive applicants; while different voters may have different tastes over political candidates' ideologies and policies.

Proposition A.1. *Suppose $\{u_i(s)\}_{i\in I}$ is uniformly monotone, then any EPIC mechanism is Pareto efficient in the range.*

A.2 Uniqueness

We focus on the existence of non-constant EPIC mechanisms in the main text. Here we provide a uniqueness result.

Definition A.7. *A mechanism q is a monotone transformation of q' if there is a weakly increasing function $f : [0,1] \to [0,1]$ such that*

$$q(\theta) = f(q'(\theta))$$

for all $\theta \in \Theta$.

Proposition A.2. *Suppose agents have common interests and $\{u_i\}_{i \in I}$ is uniformly monotone, then any non-constant EPIC mechanism q is a monotone transformation of q^*.*

Proposition A.2 states that, under uniform monotonicity and common interests, there is a unique class of non-constant EPIC mechanisms, and the Pareto efficient mechanism is one of them.

A.3 Optimal Design

Suppose there is a mechanism designer whose preference over mechanisms is represented by the utility function $d(q)$ that is linear in q:

$$d(q) = \sum_{s \in S} \delta(\theta) q(\theta).$$

Note that $\delta(\theta)$ can reflect the designer's own preference over R vs S in θ, or it can reflect the designer's concern for the agents' welfare. For instance, if $\delta(\theta) = \sum_{i=1}^{N} u_i(\theta)$ then d is utilitarian.

Fix a payoff environment, we denote EP the set of all EPIC mechanisms.

Definition A.8. *A mechanism $\bar{q}$ is optimal among all EPIC mechanisms if*

$$d(\bar{q}) = \sup_{q \in EP} d(q)$$

Lemma A.1. *There is an optimal EPIC mechanism that is deterministic.*

Lemma A.1 and Proposition 4.1 jointly imply that optimal design is simplified to assigning $q_c \in \{0, 1\}$ to every $c \in C^*$ respecting that $q_c \geq q_{c'}$ if $c \rightsquigarrow c'$.

Below we explicitly describe a procedure that returns an optimal mechanism.

Step 1. Construct the set $\mathcal{C}^B$ of all acyclic bipartitions of C^*.[14] Thus, any $\gamma \in C^*$ takes the form of $\gamma = \{C_1(\gamma), C_2(\gamma)\}$ where $c' \not\rightsquigarrow c$ for any $c \in C_1(\gamma)$ and $c' \in C_2(\gamma)$.

Step 2. For every $\gamma \in \mathcal{C}^B$:

- If there exist $c \in C_1(\gamma)$ and $c' \in C_2(\gamma)$ such that $c \to c'$, then define $\hat{c}(\gamma) = \cup_{c \in C_1(\gamma)} c$.
- Otherwise, define $\hat{c}(\gamma) = \text{argmax}_{C \in \{C_1(\gamma), C_2(\gamma)\}} \sum_{c \in C} \sum_{s \in c} \delta(\theta)$.

Let $v(\gamma) := \sum_{s \in \hat{c}(\gamma)} \delta(\theta)$.

Step 3. Pick any $\gamma^* \in \text{argmax}_{\gamma \in \mathcal{C}^B} v(\gamma)$.

Step 4. Compare $v(\gamma^*)$ and $\max\{0, \sum_{\theta \in \Theta} \delta(\theta)\}$:

- If $v(\gamma^*) > \max\{0, \sum_{\theta \in \Theta} \delta(\theta)\}$, then return $\bar{q}$ where $\bar{q}(\theta) = 1$ if $\theta \in \hat{c}(\gamma^*)$ and $\bar{q}(\theta) = 0$ if $\theta \notin \hat{c}(\gamma^*)$.
- If $v(\gamma^*) \leq \max\{0, \sum_{\theta \in \Theta} \delta(\theta)\}$ and $\max\{0, \sum_{\theta \in \Theta} \delta(\theta)\} \geq 0$, then return $\bar{q}$ where $\bar{q}(\theta) = 1$ for every $\theta \in \Theta$.
- If $v(\gamma^*) \leq \max\{0, \sum_{\theta \in \Theta} \delta(\theta)\}$ and $\max\{0, \sum_{\theta \in \Theta} \delta(\theta)\} < 0$, then return $\bar{q}$ where $\bar{q}(\theta) = 0$ for every $\theta \in \Theta$.

Proposition A.3. *$\bar{q}$ is an optimal EPIC mechanism.*

B Proofs

This section collects proofs. We omit some proofs which are similar to others or straightforward.

[14]Lemma B.4 implies that $(C^*, \rightsquigarrow)$ corresponds to a directed acyclic graph (DAG). The construction of $\mathcal{C}^B$ is equivalent to finding all acyclic bipartitions of the DAG induced by $(C^*, \to)$.

B.1 DSIC

B.1.1 Proof of lemma 3.1

Proof. "If": Suppose conditions 1 and 2 hold. Pick any $s \in \Theta_p$ and i. If $u_i(s) > 0$, then $\phi_i(\theta_i) \ni 1$, and hence for any θ'_{-i} and θ'_i:

$$u_i(s)q(\theta_i, \theta'_{-i}) = u_i(s) \max_{\theta''_i \in \Theta_i} q(\theta''_i, \theta'_{-i}) \geq u_i(s)q(\theta'_i, \theta'_{-i}).$$

The same equality for $u_i(s) \leq 0$ holds analogously. Therefore q is DSIC.

"Only if": Assume q is DSIC. Pick any i and $\theta_i \in \Theta_i$ where $\phi_i(\theta_i) \ni 1$. There is some $\theta_{-i} \in \Theta_{-i}$ where $(\theta_i, \theta_{-i}) \in \Theta$. Therefore $e_i(\theta_i, \theta_{-i})q(\theta_i, \theta'_{-i}) \geq e_i(\theta_i, \theta_{-i})q(\theta'_i, \theta'_{-i})$ for any θ'_i and θ'_{-i}, which implies that $q(\theta_i, \theta'_{-i}) \geq q(\theta'_i, \theta'_{-i})$ for any θ'_i and θ'_{-i}. Observe that condition 1 is satisfied in this case. Similarly, if $\phi_i(\theta_i) \ni -1$, then we have condition 2. □

B.1.2 Proof of Lemma 3.2

Proof. Suppose q is DSIC and agent i is not privately informed. There exists θ_i such that $\{1, -1\} \subset \phi_i(\theta_i)$. From Lemma 3.1 we have $\max_{\theta'_i \in \Theta_i} q(\theta'_i, \theta_{-i}) = q(\theta_i, \theta_{-i}) = \min_{\theta'_i \in \Theta_i} q(\theta'_i, \theta_{-i})$ for any θ_{-i}, which implies that $q(\cdot, \theta_{-i})$ is constant over Θ_i. □

B.1.3 Proof of Proposition 3.1

Proof. Let $\Theta^{PI} := \times_{i \in PI} \Theta_i$. Therefore q is DSIC only if there is some $\hat{q} : \Theta^{PI} \to [0, 1]$ such that $q(\theta) = \hat{q}(k(\theta))$ where $k : \Theta \to \Theta^{PI}$ orthogonally projects θ from Θ to Θ^{PI}.

We can then derive binary relation $\Rightarrow$ on S^{PI}: $\theta \Rightarrow \theta'$ if $\theta_i \Rightarrow_i \theta'_i$ for every $i \in PI$. Note that $\Rightarrow$ is a partial weak order.

"If": Suppose $q(\theta) = \hat{q}(k(\theta))$ for some $\hat{q} : \Theta^{PI} \to [0, 1]$ where $\hat{q}(\theta) \geq \hat{q}(\theta')$ if $\theta \Rightarrow \theta'$. Pick any $\theta \in \Theta$, agent i, and $\theta'_i \in \Theta_i$ and $\theta'_{-i} \in \Theta_{-i}$. If $i \notin PI$ then $q(\theta_i, \theta'_{-i}) = \hat{q}(k(\theta_i, \theta'_{-i})) = \hat{q}(k(\theta'_i, \theta'_{-i})) = q(\theta'_i, \theta'_{-i})$, implying that i has no incentive to misreport θ'_i. Now consider $i \in PI$. If $\phi_i(\theta_i) = \{0\}$ then i is indifferent between S and R and hence has no incentive to misreport θ_i'. If $\phi_i(\theta_i) \ni 1$ then $u_i(\theta) \geq 0$ and also $k(\theta_i, \theta_{-i}') \Rightarrow k(\theta_i', \theta_{-i}')$. Therefore $u_i(\theta)q(\theta_i, \theta_{-i}') = u_i(\theta)\hat{q}(k(\theta_i, \theta_{-i}')) \geq u_i(\theta)\hat{q}(k(\theta_i', \theta_{-i}')) = u_i(\theta)q(\theta_i', \theta_{-i}')$, that is, there is no incentive to misreport θ_i'. The same can be established analogously if $\phi_i(\theta_i) \ni -1$. Hence q is DSIC.

"Only if": Suppose q is DSIC. Lemma 3.2 implies there is some $\hat{q} : \Theta^{PI} \rightarrow [0,1]$ such that $q(\theta) = \hat{q}(k(\theta))$. Fix any $\theta_{PI}, \theta_{PI}' \in \Theta^{PI}$ such that $\theta_{PI} \Rightarrow \theta_{PI}'$. We want to show that $\hat{q}(\theta_{PI}) \geq \hat{q}(\theta_{PI}')$, which is equivalent to showing that $q(\theta) \geq q(\theta')$ for any $\theta \in k^{-1}(\theta_{PI})$ and $\theta' \in k^{-1}(\theta_{PI}')$. For such θ, θ' consider the sequence $(\theta^0, ..., \theta^N)$ where θ^k agrees with θ' in the first k entries and with θ in the last $N-k$ entries. Therefore $\theta^0 = \theta$, $\theta^N = \theta'$, and θ^k and θ^{k-1} differ only in the k th entry. If $k \notin PI$ then $q(\theta^k) = q(\theta^{k-1})$ by Lemma 3.2. If $k \in PI$ then we have $\theta_k \Rightarrow_k \theta_k'$, which implies that either $\phi_k(\theta_k^{k-1}) = \phi_k(\theta_k) \ni 1$ or $\phi_k(\theta_k^k) = \phi^k(\theta_k') \ni -1$. In both cases we have $q(\theta^{k-1}) = q(\theta_k^{k-1}, \theta_{-k}^k) \geq q(\theta_k^k, \theta_{-k}^k) = q(\theta^k)$ by Lemma 3.1. It follows that $q(\theta) = q(\theta^0) \geq q(\theta^1) \geq ... \geq q(\theta^N) = q(\theta')$. □

B.2 EPIC

B.2.1 Some Useful Lemmas

Lemma B.1. *q is EPIC if and only if $q(\theta) \geq q(\theta')$ for any θ, θ' where $\theta \rightarrow \theta'$.*

Proof of Lemma B.1

Proof. "If": Suppose $\theta \rightarrow \theta'$ implies $q(\theta) \geq q(\theta')$. Pick any state θ where $p(\theta) > 0$ and any agent i. If $u_i(\theta) = 0$, then i is indifferent between R and L, and hence he has no incentive to misreport in θ under any mechanism. If $u_i(\theta) > 0$, then we have $\theta \rightarrow (\theta_i', \theta_{-i})$ for any $\theta_i' \neq \theta_i$. It follows that $q(\theta) \geq q(\theta_i', \theta_{-i})$ for any $\theta_i' \in \Theta_i$, which implies that it is not profitable for i to misreport in θ, because i seeks to maximize the probability R being chosen in θ. Similarly it is not profitable for i to misreport if $u_i(\theta) < 0$. Therefore q is EPIC.

"Only if": Suppose q is EPIC. Pick any θ, θ' where $\theta \rightarrow \theta'$. By definition, there exists $i \in \{1, ..., N\}$ such that: (1) $\theta_{-i} = \theta_{-i}'$ and (2) $u_i(\theta) > 0$ or $u_i(\theta') < 0$. If $u_i(\theta) > 0$, then i prefers a higher probability of R being chosen in θ, and hence EPIC implies $\theta_i \in$

$\mathrm{argmax}_{\hat{\theta}_i \in \Theta_i} q(\hat{\theta}_i, \theta_{-i})$, which in turn implies that $q(\theta) = q(\theta_i, \theta_{-i}) \geq q(\theta_i', \theta_{-i}) = q(\theta')$. The case where $u_i(\theta') < 0$ is analogous. $\square$

Lemma B.2 is a corollary of Lemma B.1.

Lemma B.2. *If q is EPIC if and only if $q(\theta) \geq q(\theta')$ for any θ, θ' where $\theta \rightsquigarrow \theta'$.*

The following lemma will be useful.

Lemma B.3. *For any partition C of S, $c, c' \in C$, $\theta \in c$ and $\theta' \in c'$, $c \rightsquigarrow c'$ if $\theta \rightsquigarrow \theta'$. Moreover, if $C = C^*$ then $c \rightsquigarrow c'$ only if $\theta \rightsquigarrow \theta'$*

Proof. Suppose $\theta \rightsquigarrow \theta'$, then there exists a list $(\theta^0, ..., \theta^J)$ such that $\theta^0 = \theta \to \theta^1 ... \to \theta^J = \theta'$. Let c^j denote the cell in C that contains θ^j, and hence $c = c^0 \to c^1 ... \to c^J = c'$, which implies that $c \rightsquigarrow c'$.

Suppose $C = C^*$ and $c \to c'$. By definition there exist $\hat{\theta} \in c$ and $\hat{\theta}' \in c'$ such that $\hat{\theta} \to \hat{\theta}'$, which implies that $\hat{\theta} \rightsquigarrow \hat{\theta}'$. By construction of C^*, $\theta \rightsquigarrow \hat{\theta}$ and $\hat{\theta}' \rightsquigarrow \theta'$, hence $\theta \rightsquigarrow \theta'$ because $\rightsquigarrow$ is transitive. With a straightforward inductive argument it is easy to generalize this observation as long as $c \rightsquigarrow c'$. $\square$

We say that a partition C of S is **acyclic** if there does not exist distinct $c, c' \in S$ such that $c \leftrightsquigarrow c'$.

Lemma B.4. *C^* is the finest acyclic partition of Θ.*

Proof. For any $c, c' \in C^*$, if $c \leftrightsquigarrow c'$ then $\theta \leftrightsquigarrow \theta'$ for any $\theta \in c$ and $\theta' \in c'$ by Lemma B.3, which implies that $c = c'$ by the construction of C^*. Thus there does not exist distinct $c, c' \in \Theta$ such that $c \leftrightsquigarrow c'$, implying that C^* is acyclic.

Now we show that C^* is the finest acyclic partition of Θ. Consider another acyclic partition $\overline{C}$ of Θ. Pick any $c \in C^*$. Suppose, in order to lead to a contradiction, that there are distinct $\overline{c}, \overline{c}' \in \overline{C}$ both of which intersect with c. Pick $\theta \in c \cap \overline{c}$ and $\theta' \in c \cap \overline{c}'$. It follows that $\theta \leftrightsquigarrow \theta'$ by construction of C^*, and hence $\overline{c} \leftrightsquigarrow \overline{c}'$ by Lemma B.3, a contradiction as $\overline{C}$ is assumed to be acyclic. Therefore $c \subset \overline{c}$ for some $\overline{c} \in \overline{C}$, implying that C^* is finer than C. $\square$

B.2.2 Proof of Proposition 4.1

Proof. "If": If there are such probabilities then $\theta \rightsquigarrow \theta'$ implies $C^*(\theta) \rightsquigarrow C^*(\theta')$ (where $C^*(\theta)$ denotes the cell in C^* that contains θ) and hence $q(\theta) = q_{C^*(\theta)} \geq q_{C^*(\theta')} = q(\theta')$, which by Lemma B.2 implies q is EPIC.

"Only if": If q is EPIC then Lemma B.2 implies that $q(\theta) = q(\theta')$ if $C^*(\theta) = C^*(\theta')$, because $\theta \rightsquigarrow \theta' \rightsquigarrow \theta$. Therefore q is constant on any $c \in C^*$, and we can denote this value as q_c. Moreover if $c \rightsquigarrow c'$ for $c, c' \in C^*$ then Lemma B.3 implies that $\theta \rightsquigarrow \theta'$ for any $\theta \in c$ and $\theta' \in c'$ and hence $q(\theta) \geq q(\theta')$ by Lemma B.2, implying $q_c \geq q_{c'}$. □

B.2.3 Proof of Proposition 4.2

Proof. That part 1 and part 2 are equivalent follows immediately from Lemma 4.1.

Suppose part 2 is true. That C^* is acyclic implies that there exists $c \in C^*$ such that $c' \not\rightsquigarrow c$ for every $c' \in C^*$. Let $\hat{c}$ denote the set in C^* containing c. Construct $c_A = \{c' : c' \in C \text{ and } c' \rightsquigarrow \hat{c}\}$ and $c_B := \Theta \backslash c_A$. By construction $c_B \not\rightsquigarrow c_A$, proving part 3.

That part 3 implies part 4 is obvious.

Suppose part 4 is true. There exists $\theta \in \Theta$ such that $\theta' \not\rightsquigarrow \theta$ for some $\theta' \in \Theta_p$. Construct $c_A := \{\theta' : \theta' \in \Theta \text{ and } \theta' \rightsquigarrow \theta\}$, and $c_B := \Theta \backslash c_A$. Note that by construction $c_B \not\rightsquigarrow c_A$, hence $C = \{c_A, c_B\}$ is an acyclic partition of S where both elements intersect with Θ_p. Since C^* is a refinement of C by Lemma B.4, part 2 follows. □

B.2.4 Proof of Proposition A.1

Definition B.1. *A vertex θ is called a source if there is a walk from θ to θ' for any $\theta' \in \Theta$; A vertex θ is called a sink if there is a walk from θ' to θ for any $\theta' \in \Theta$.*

When $\{u_i\}_{i \in I}$ is weakly monotone, we can find $\bar{\theta} = (\bar{\theta}_1, \bar{\theta}_2, \ldots, \bar{\theta}_N)$ such that $\bar{\theta}_i \succ_i^{(i)} \theta_i$ for all $\theta_i \neq \bar{\theta}_i$ and $\underline{\theta} = (\underline{\theta}_1, \ldots, \underline{\theta}_N)$ such that $\theta_i \succ_i^{(i)} \underline{\theta}_i$ for all $\theta_i \neq \bar{\theta}_i$. Alternatively, we can rearrange Θ_i by $\succ_i^{(i)}$ such that $\theta_i^{(1)} \succ_i^{(i)} \theta_i^{(2)} \succ_i^{(i)} \cdots$, then $\bar{\theta} = (\theta_1^{(1)}, \theta_2^{(1)}, \ldots, \theta_N^{(1)})$.

Lemma B.5. *Suppose $\{u_i(\theta)\}_{i \in I}$ is monotone, then $\bar{\theta}$ is a source and $\underline{\theta}$ is a sink.*

Lemma B.6. *Suppose $\{u_i(\theta)\}_{i\in I}$ is uniformly monotone, then*

1. *if there is θ such that $u_i(\theta) > 0$ for all $i \in I$ then there is a $\bar{\theta} - \theta$ walk;*
2. *if there is θ such that $u_i(\theta) < 0$ for all $i \in I$ then there is a $\theta - \underline{\theta}$ walk;*

Proof of Proposition 4.3

Proof. We want to show that for any $\theta \in \Theta$, $i \in I$ and $\theta_i' \in \Theta_i$,

$$u(\theta)q(\theta_i, \theta_{-i}) \geq u(\theta)q(\theta_i', \theta_{-i}).$$

Suppose $u(\theta) > 0$, then above equation becomes $q(\theta) = 1 \geq q(\theta_i', \theta_{-i})$. Suppose $u(\theta) < 0$, then above equation becomes $q(\theta) = 0 \leq q(\theta_i', \theta_{-i})$. Suppose $u(\theta) = 0$, then above equation becomes $0 \geq 0$. $\square$

Proof of Lemma A.1

Proof. Given Lemma B.1, it is straightforward to verify that the set of EPIC mechanisms is a polytope in $\mathbb{R}^{|S|}$ with extreme points consisting of either 0s or 1s. Since the objective function d is linear in q, there is a maximizer that coincides with an extreme point, which is then a deterministic mechanism. $\square$

B.2.5 Proof of Proposition A.3

Proof. For any $\gamma \in \mathcal{C}^B$, define $q_\gamma : \Theta \to \{0, 1\}$ such that $q_\gamma(\theta) = 1$ if $\theta \in \hat{c}(\gamma)$ and $q_\gamma(\theta) = 0$ if $\theta \notin \hat{c}(\gamma)$.

Lemma 4.1 imply that any non-constant deterministic EPIC mechanism must be measurable with respect to some $\gamma \in \mathcal{C}^B$, i.e. $q^{-1}(x) \subset \gamma$ for $x = 0, 1$. Moreover, if there exist $c \in C_1(\gamma)$ and $c' \in C_2(\gamma)$ such that $c \to c'$, then Lemma 4.1 implies that q_γ is the only non-constant deterministic EPIC mechanism measurable with respect to γ. Otherwise, there are two non-constant deterministic EPIC mechanisms measurable with respect to γ: q_γ^1 that assigns 1 to states in every $c \in C_1(\gamma)$ and 0 to states in every $c \in C_2(\gamma)$, and q_γ^2 that assigns 0 to states in every $c \in C_1(\gamma)$ and 1 to states in every $c \in C_2(\gamma)$. Observe that, in this case q_γ is set to be the mechanism that the designer prefers between q^1_γ or q^2_γ.

It follows that every optimal non-constant deterministic EPIC mechanism is in $\text{argmax}_{\gamma\in\mathcal{C}^B} d(q_\gamma)$. It is straightforward to verify that $v(\gamma) = d(q_\gamma)$. Thus γ^* is an optimal non-constant deterministic EPIC mechanism.

Step 4 simply compares an optimal non-constant deterministic EPIC mechanism to a constant deterministic mechanism (which must be EPIC), and yields the optimal deterministic EPIC mechanism as $\bar{q}$. It then follows from Lemma A.1 that $\bar{q}$ is an optimal EPIC mechanism. □

B.2.6 Proof of Proposition 4.4

Proof. Suppose a fully reduced mechanism q is EPIC. We want to show q is constant over $\bar{\Theta} := \Theta \setminus (IC_1 \cup IC_2)$ where $\bar{\Theta}$ is the set of payoff states in which both agents have strict ex post preferences.

We prove the Proposition in four steps.

Step 1: We introduce a binary relation $\sim$ on $\bar{\Theta}$.

Definition B.2. *We Two payoff states $\theta \sim \theta'$ if*

1. $u_i(\theta)u_i(\theta') > 0$ *for* $i = 1, 2$*;*
2. *There exists a path,* $(\theta = \theta^{(0)}, \theta^{(1)}, \cdots, \theta^{(n)} = \theta')$*, such that* $\theta^{(k)}$ *and* $\theta^{(k+1)}$ *differ in one entry, and* $\theta'' \in \{\theta | \theta = t\theta^{(k)} + (1-t)(\theta^{(k+1)}), \forall k = 1, 2, ..., n, \forall t \in [0, 1]\}$ *implies* $u_i(\theta'')u_i(\theta'') > 0$ *for* $i = 1, 2$*.*

Condition 1 means each agent has the same ex post preferences on θ and θ'. Condition 2 means thee exist a continuous "manhattan path" links θ and θ' and all the payoff states along the path give each agent the same ex post preferences as θ and θ' give. It is easy to verify that $\sim$ is reflexive, symmetric, and transitive. Hence, $\sim$ is an equivalence relation which induces in partition $P := \{P_k\}$ on $\bar{\Theta}$. If P is a singleton, then Proposition 4.4 holds trivially. From now on, we assume P contains at least two elements/blocks.

Step 2: We show that q is constant in each P_i.

For any $\theta, \theta' \in P_i$, there exists a path $(\theta = \theta^{(0)}, \theta^{(1)}, \cdots, \theta^{(n)} = \theta')$ links θ and θ'. Ex post incentive compatibility requires $q(\theta^{(k)}) = q(\theta^{(k+1)})$. Hence, $q(\theta) = q(\theta')$.

Step 3: We prove that q is constant between adjacent P_i and P_j.

We first give the definition of adjacency. Note that u_1 and u_2 are continuous, each P_i is an open set in Θ. We denote the closure of P_i by $\bar{P}_i$.

Definition B.3. *Two blocks of the partition P, P_i, P_j, are adjacent if $m(\bar{P}_i \cap \bar{P}_j) > 0$ where m is the Lebesgue measure of $\mathbb{R}$.*

That is, $P_i, P_j \in P$ are adjacent if their closures intersect with each other in a non-degenerated fashion. Hence, we can find $\theta \in \bar{P}_i \cap \bar{P}_j$ and $\delta > 0$ such that $B_\delta(\theta) \subset \bar{P}_i \cup \bar{P}_j$ where $B_\delta(\theta)$ is the $\delta-$neighborhood of θ under the Euclidean norm. We discuss three cases.

Case 1: P_i and P_j share the same signs of (u_1, u_2).

Since $m(\bar{P}_i \cap \bar{P}_j) > 0$, there exist $\theta' \in P_i$ and $\theta'' \in P_j$ such that sgn$u_1(\theta') =$ sgn$u_1(\theta'')$, sgn$u_2(\theta') =$ sgn$u_2(\theta'')$, and θ' and θ'' differ in one entry. Without loss of generality, we assume $\theta'_1 = \theta''_1$. Then, by agent 1's incentive compatibilities, $q(\theta') = q(\theta'')$.

Case 2: P_i and P_j differ in one sign of (u_1, u_2).

Without loss of generality, we assume they differ in u_1. Then $\bar{P}_i \cap \bar{P}_j$ is a subset of $BD_1 \setminus BD_2$. For the θ we found right after the Definition B.3, we know $\theta \in \bar{P}_i \cap \bar{P}_j \subset BD_1$, we have $\dfrac{\partial u_1(\theta)}{\partial \theta_2} \neq 0$ by generic interdependence. Therefore, $u_1(\theta_1, \theta_2 + \delta/2) u_1(\theta_1, \theta_2 - \delta/2) < 0$. Then we know $u_2(\theta_1, \theta_2 + \delta/2) u_2(\theta_1, \theta_2 - \delta/2) > 0$. Agent 2's incentive compatibilities require that $q(\theta_1, \theta_2 + \delta/2) = q(\theta_1, \theta_2 - \delta/2)$. Since $(\theta_1, \theta_2 + \delta/2)$ and $(\theta_1, \theta_2 - \delta/2)$ belongs to different blocks, q is constant between P_i and P_j follows immediately.

Case 3: P_i and P_j differ in both signs of (u_1, u_2).

Suppose $u_1(\theta') > 0$, $u_2(\theta') > 0$ for all $\theta' \in P_i$. Then $u_1(\theta'') < 0$, $u_2(\theta'') < 0$ for all $\theta'' \in P_j$. Thus $\theta \in (\bar{P}_i \cap \bar{P}_j) \subset (BD_1 \cap BD_2)$. It is easy to show that all for all $\theta' \in B_\delta(\theta)$, we have $u_1(\theta') u_2(\theta') > 0$. Generic heterogeneity condition is violated. Similarly, the subcase $u_1(\theta') < 0$, $u_2(\theta') < 0$ violates generic heterogeneity condition.

Now suppose $u_1(\theta') < 0$, $u_2(\theta') > 0$ for all $\theta' \in P_i$. Then $u_1(\theta'') > 0$, $u_2(\theta'') < 0$ for all $\theta'' \in P_j$. Then $(\bar{P}_i \cap \bar{P}_j) \subset (BD_1 \cap BD_2)$. We can pick two payoff states within the δ-neighborhood of θ: $(\theta_1 - \epsilon, \theta_2)$, $(\theta_1 + \epsilon, \theta_2)$. We know $u_1(\theta_1 - \epsilon, \theta_2)u_1(\theta_1 + \epsilon, \theta_2) < 0$. Without loss of generality, we assume $u_1(\theta_1 - \epsilon, \theta_2) < 0$ and $u_1(\theta_1 + \epsilon, \theta_2) > 0$. Then, $(\theta_1 - \epsilon, \theta_2) \in P_i$, $(\theta_1 + \epsilon, \theta_2) \in P_j$, and $q(\theta_1 - \epsilon, \theta_2) \leq q(\theta_1 + \epsilon, \theta_2)$.

We can also pick another two payoff states within the δ-neighborhood of $(\theta_1, \theta_2 - \epsilon')$, $(\theta_1, \theta_2 + \epsilon')$. Similarly, we know $u_2(\theta_1, \theta_2 - \epsilon')u_2(\theta_1, \theta_2 - \epsilon') < 0$. Without loss of generality, we assume $u_2(\theta_1, \theta_2 - \epsilon') > 0$ and $u_\epsilon(\theta_1, \theta_2 + \epsilon') < 0$. Then, $(\theta_1, \theta_2 - \epsilon') \in P_i$, $(\theta_1, \theta_2 + \epsilon') \in P_j$, and $q(\theta_1, \theta_2 - \epsilon') \geq q(\theta_1, \theta_2 + \epsilon')$. By step 2, we know $q(\theta_1 - \epsilon, \theta_2) = q(\theta_1, \theta_2 - \epsilon')$, $q(\theta_1 + \epsilon, \theta_2) = q(\theta_1, \theta_2 + \epsilon')$. Therefore, $q(\theta_1 - \epsilon, \theta_2) = q(\theta_1, \theta_2 - \epsilon') = q(\theta_1 + \epsilon, \theta_2) = q(\theta_1, \theta_2 + \epsilon')$.

The case $u_1(\theta') > 0$, $u_2(\theta') < 0$ for all $\theta' \in P_i$ can be proved in the same way.

Step 4: Since every block P_i has at least one adjacent block, any two blocks P_i and P_j are linked by a sequence of adjacent blocks. Hence, q is constant over $\bar{\Theta}$. □

B.3 IDSIC

B.3.1 Proof of Lemma 5.2

Proof. For player i of type t_i, given message profile $m_{-i} \in M_{-i}$, payoff-type profile $\theta_{-i} \in \Theta_{-i}$ and joint strategy $\sigma_{-i} : T_{-i} \to \Delta(M_{-i})$ of the other players, define

$$q_i(m_{-i}|\sigma_{-i}, \theta_{-i}, t_i) := \sum_{\{t_{-i} \in T_{-i} : \hat{\theta}_{-i}(t_{-i}) = \theta_{-i}\}} \Pr\left[t_{-i} \middle| \theta_{-i}, t_i\right] \sigma_{-i}(t_{-i})[m_{-i}]$$

where $\Pr\left[t_{-i} \middle| \hat{\theta}_{-i}(t_{-i}) = \theta_{-i}, t_i\right] = \dfrac{\hat{\beta}_i(t_i)[t_{-i}]}{\hat{b}_i(t_i)[\theta_{-i}]}$ is agent i's belief over t_{-i} conditional on her own signal is t_i and other agents' payoff types are θ_{-i}.

Thus $q_i(m_{-i}|\sigma_{-i}, \theta_{-i}, t_i)$ is what player i of type t_i evaluates as the probability that the message profile from the other players will be m_{-i} conditional on their payoff-type profile being θ_{-i} and them following σ_{-i}.

For player i of type t_i, her expected payoff from message m_i conditional on other players following joint strategy σ_{-i} can be written as:

$$U_i(m_i|\sigma_{-i}, t_i)$$
$$:= \sum_{\theta_{-i}\in\Theta_{-i}} \hat{b}_i(t_i)[\theta_{-i}] \sum_{m_{-i}\in M_{-i}} q_i(m_{-i}|\sigma_{-i}, \theta_{-i}, t_i) \sum_{a\in A} u_i(a, \hat{\theta}_i(t_i), \theta_{-i}) q(m_i, m_{-i})[a].$$

Suppose m_i is an interim dominant action for player i of type t_i. Pick any $t_i' \in T_i$ where $\hat{\theta}_i(t_i') = \hat{\theta}_i(t_i)$ and $\hat{b}_i(t_i) = \hat{b}_i(t_i')$. For any $\sigma_{-i} : T_{-i} \to \Delta(M_{-i})$ there always exists $\chi_{\sigma_{-i}} : T_{-i} \to \Delta(M_{-i})$ such that $q_i(m_{-i}|\sigma_{-i}, \theta_{-i}, t_i') = q_i(m_{-i}|\chi_{\sigma_{-i}}, \theta_{-i}, t_i)$ for any $m_{-i} \in M_{-i}$ and $\theta_{-i} \in \Theta_{-i}$.[15]

It is straightforward to verify that $U_i(\cdot|\sigma_{-i}, t_i') = U_i(\cdot|\chi_{\sigma_{-i}}, t_i)$. It follows that for any σ_{-i} and m_i',

$$U_i(m_i|\sigma_{-i}, t_i') = U_i(m_i|\chi_{\sigma_{-i}}, t_i) \geq U_i(m_i'|\chi_{\sigma_{-i}}, t_i) = U_i(m_i'|\sigma_{-i}, t_i').$$

Thus m_i is also an interim dominant action for player i when her type is t_i'. □

B.3.2 Proof of Proposition 5.1

Proof. **"Only if'.'** Suppose σ is an interim dominant-strategy equilibrium of the mechanism. For each $i \in I$, $\theta_i \in \Theta_i$ and $b_i \in \Delta(\Theta_{-i})$ pick some $\tau_i(\theta_i, b_i) \in T_i$ where $\hat{\theta}_i(\tau_i(\theta_i, b_i)) = \theta_i$ and $\hat{b}_i(\tau_i(\theta_i, b_i)) = b_i$ whenever possible. Consider strategy profile σ' where $\sigma_i'(t_i) = \sigma_i'(\tau_i(\theta_i, b_i))$ for any t_i where $\hat{\theta}_i(t_i) = \theta_i$ and $\hat{b}_i(t_i) = b_i$. Observe that σ' is higher-order belief-independent. Moreover Lemma 5.2 implies that σ' is also an interim dominant strategy equilibrium. □

B.3.3 Proof of Lemma 5.4

Proof. **"If":** Fix agent i. Suppose the other agents report $\sigma_{-i}(\theta_{-i}) \in H_{-i}$ when their true payoff types is θ_{-i}. The difference in payoff to i between truthfully reporting h_i and

[15] For instance, set $\chi_{\sigma_{-i}}(m_{-i}|t_{-i}) = q_i(m_{-i}|\sigma_{-i}, \theta_{-i}, t_i')$ for every t_{-i} where $\hat{\theta}_{-i}(t_{-i}) = \theta_{-i}$.

misreporting as h_i' is

$$
\begin{aligned}
D :=& \sum_{\theta_{-i}\in\Theta_{-i}} b_i(\theta_{-i})u_i(\theta_i,\theta_{-i})\big(q(h_i,\sigma_{-i}(\theta_{-i})) - q(h_i',\sigma_{-i}(\theta_{-i}))\big) \\
=& \Big[\sum_{\{\theta_{-i}|u_i(\theta)>0\}} b_i(\theta_{-i})u_i(\theta_i,\theta_{-i})\big(q(h_i,\sigma_{-i}(\theta_{-i})) - q(h_i',\sigma_{-i}(\theta_{-i}))\big) \\
&+ \sum_{\{\theta_{-i}|u_i(\theta)<0\}} b_i(\theta_{-i})u_i(\theta_i,\theta_{-i})\big(q(h_i,\sigma_{-i}(\theta_{-i})) - q(h_i',\sigma_{-i}(\theta_{-i}))\big)\Big].
\end{aligned}
$$

Define $\overline{d} := \max_{h_{-i}\in H_{-i}} \big(q(h_i,h_{-i}) - q(h_i',h_{-i})\big)$ and $\underline{d} := \min_{h_{-i}\in H_{-i}} \big(q(h_i,h_{-i}) - q(h_i',h_{-i})\big)$. It is straightforward to verify that

$$
\begin{aligned}
D &\geq \overline{d} \sum_{\{\theta_{-i}|u_i<0\}} b_i(\theta_{-i})u_i(\theta_i,\theta_{-i}) + \underline{d} \sum_{\{\theta_{-i}|u_i>0\}} b_i(\theta_{-i})u_i(\theta_i,\theta_{-i}) \\
&= \underline{\alpha}_i(h_i)\overline{d} + \overline{\alpha}_i(h_i)\underline{d} \\
&\geq 0,
\end{aligned}
$$

where the last inequality is due to the assumption that for any $h_{-i}, h_{-i}' \in H_{-i}$:

$$
\underline{\alpha}_i(h_i)\Big(q(h_i,h_{-i}) - q(h_i',h_{-i})\Big) + \overline{\alpha}_i(h_i)\Big(q(h_i,h_{-i}') - q(h_i',h_{-i}')\Big) \geq 0.
$$

Therefore misreporting is not profitable, and q is IDSIC.

"Only if": Suppose q is IDSIC. Fix agent i, $h_i, h_i' \in H_i$ and $h_{-i}, h_{-i}' \in H_{-i}$. Suppose agents other than i jointly report h_{-i} whenever their true payoff types $\hat{\theta}_{-i}$ satisfy $u_i(\theta_i,\hat{\theta}_{-i}) < 0$, and they jointly report h_{-i}' otherwise. Thus IDSIC requires that

$$
\begin{aligned}
&\sum_{\{\hat{\theta}_{-i}|u_(\theta_i,\hat{\theta}_{-i})<0\}} b_i(\theta_{-i})u_i(\theta_i,\hat{\theta}_{-i})\big(q(h_i,h_{-i}) - q(h_i',h_{-i})\big) \\
&\quad + \sum_{\{\hat{\theta}_{-i}|u_i(\theta_i,\hat{\theta}_{-i})>0\}} b_i(\theta_{-i})u_i(\theta_i,\hat{\theta}_{-i})\big(q(h_i,h_{-i}') - q(h_i',h_{-i}')\big) \geq 0,
\end{aligned}
$$

implying that $\underline{\alpha}_i(h_i)\big(q(h_i,h_{-i}) - q(h_i',h_{-i})\big) + \overline{\alpha}_i(h_i)\big(q(h_i,h_{-i}') - q(h_i',h_{-i}')\big) \geq 0.$ $\square$

B.3.4 Proof of Proposition 5.3

Proof. **"Only if":** Suppose q is IDSIC. Then by Lemma 5.4, Equation 5.1 holds for all $h_i, h_i', h_{-i}, h_{-i}'$. Let $h_{-i} = h_{-i}'$, we have

$$\begin{aligned}\underline{\alpha}_i(h_i)\big(q(h_i,h_{-i}) - q(h_i',h_{-i})\big) \geq & -\overline{\alpha}_i(h_i)\big(q(h_i,h_{-i}) - q(h_i',h_{-i})\big) \\ \Leftrightarrow \alpha_i(h_i)q(h_i,h_{-i}) \geq & \alpha_i(h_i)q(h_i',h_{-i})\end{aligned} \tag{B.1}$$

for all $h_i, h_i' \in H_i$ and all $h_{-i} \in H_{-i}$.

If $\alpha_i(h_i) > 0$, then $q(h_i, h_{-i}) \geq q(h_i', h_{-i})$ for all $h_i' \in H_i$; if $\alpha_i(h_i) < 0$, then $q(h_i, h_{-i}) \leq q(h_i', h_{-i})$ for all $h_i' \in H_i$; if $\alpha_i(h_i) = 0$, then $q(h_i', h_{-i}) \leq q(h_i, h_{-i}) \leq q(h_i'', h_{-i})$ for all $h_i' \in H_i^-, h_i' \in H_i^+$. Hence, the monotonicity condition is proved.

For each agent i, we have two cases for the variation condition (condition 2).

1. $H_i^0 = \emptyset$. First, let $h_i \in H_i^+, h_i' \in H_i^-$, $h_{-i} \in$ argmax $[q(h_i, h_{-i}) - q(h_i', h_{-i})]$, $h_{-i}' \in$ argmin $[q(h_i, h_{-i}) - q(h_i', h_{-i})]$. Equation 5.1 gives

   $$-\underline{\alpha}_i(h_i) \max_{h_{-i} \in H_{-i}} \Big(\overline{q}_i(h_{-i}) - \underline{q}_i(h_{-i})\Big) \leq \overline{\alpha}_i(h_i) \min_{h_{-i} \in H_{-i}} \Big(\overline{q}_i(h_{-i}) - \underline{q}_i(h_{-i})\Big),$$

   for all $h_i \in H_i^+$. Therefore,

   $$\max_{h_{-i} \in H_{-i}} \Big(\overline{q}_i(h_{-i}) - \underline{q}_i(h_{-i})\Big) \leq \rho_i(h_i) \min_{h_{-i} \in H_{-i}} \Big(\overline{q}_i(h_{-i}) - \underline{q}_i(h_{-i})\Big),$$

   for all $h_i \in H_i^+$.

   Second, let $h_i \in H_i^-, h_i' \in H_i^+$, $h_{-i} \in$ argmin $[q(h_i, h_{-i}) - q(h_i', h_{-i})]$, $h_{-i}' \in$ argmax $[q(h_i, h_{-i}) - q(h_i', h_{-i})]$. Equation 5.1 gives

   $$-\underline{\alpha}_i(h_i) \min_{h_{-i} \in H_{-i}} \Big(\overline{q}_i(h_{-i}) - \underline{q}_i(h_{-i})\Big) \geq \overline{\alpha}_i(h_i) \max_{h_{-i} \in H_{-i}} \Big(\overline{q}_i(h_{-i}) - \underline{q}_i(h_{-i})\Big)$$

   Therefore,

   $$\max_{h_{-i} \in H_{-i}} \Big(\overline{q}_i(h_{-i}) - \underline{q}_i(h_{-i})\Big) \leq \rho_i(h_i) \min_{h_{-i} \in H_{-i}} \Big(\overline{q}_i(h_{-i}) - \underline{q}_i(h_{-i})\Big),$$

for all $h_i \in H_i^-$. Recall that $\rho_i = \min_{h_i} \rho_i(h_i)$. Hence,

$$\max_{h_{-i} \in H_{-i}} \Big(\overline{q}_i(h_{-i}) - \underline{q}_i(h_{-i})\Big) \leq \rho_i \min_{h_{-i} \in H_{-i}} \Big(\overline{q}_i(h_{-i}) - \underline{q}_i(h_{-i})\Big).$$

2. $H_i^0 \neq \emptyset$. Now consider $h_i^* \in H_i^0$. Since $-\overline{\alpha}_i(h_i^*) = \overline{\alpha}_i(h_i^*)$, Equation 5.1 gives

$$q(h_i^*, h_{-i}) - q(h_i', h_{-i}) \leq q(h_i^*, h_{-i}') - q(h_i', h{-}i')$$

for all $h_i' \in H_i$, and all $h_{-i}, h_{-i}' \in H_{-i}$. Then $q(h_i^*, h_{-i}) - q(h_i', h_{-i}) = q(h_i^*, h_{-i}') - q(h_i', h_{-i}')$ for all $h_i' \in H_i$ and all $h_{-i}, h_{-i}' \in H_{-i}$, which mean $q(h_i^*, h_{-i}) - q(h_i', h_{-i})$ is constant over h_{-i}, for all $h_i' \in H_i$. We then have $q(h_i, h_{-i}) - q(h_i', h_{-i}) = [q(h_i, h_{-i}) - q(h_i^*, h_{-i})] + [q(h_i^*, h_{-i}) - q(h_i', h_{-i})]$ is a constant function with respect to h_{-i}, for all $h_i, h_i' \in H_i$.

"If": We want to show that if both monotonicity condition and variation condition are satisfied, then Equation 5.1 holds for all $i \in I$, $h_i, h_i' \in H_i$, and $h_{-i}, h_{-i}' \in H_{-i}$. For each $i \in I$, we discuss two situations.

1. $H_i^0 = \emptyset$. If both $h_i, h_i' \in H_i^+$ or both $h_i, h_i' \in H_i^-$, then it is obvious that Equation 5.1 holds. If $h_i \in H_i^+, h_i' \in H_i^-$, then

$$\begin{aligned}
&\underline{\alpha}_i(h_i)\big(q(h_i, h_{-i}) - q(h_i', h_{-i})\big) + \overline{\alpha}_i(h_i)\big(q(h_i, h_{-i}') - q(h_i', h_{-i}')\big) \\
=& \underline{\alpha}_i(h_i)\big(\overline{q}(h_{-i}) - \underline{q}(h_{-i})\big) + \overline{\alpha}_i(h_i)\big(\overline{q}(h_{-i}') - \underline{q}(h_{-i}')\big) \\
\geq& \underline{\alpha}_i(h_i) \max_{h_{-i}} \big(\overline{q}(h_{-i}) - \underline{q}(h_{-i})\big) + \overline{\alpha}_i(h_i) \min_{h_{-i}'} \big(\overline{q}(h_{-i}') - \underline{q}(h_{-i}')\big) \\
\geq& 0
\end{aligned}$$

for all $h_{-i}, h_{-i}' \in H_{-i}$. The first equality follows the monotonicity condition, the second equality is due to the fact $\underline{\alpha}_i(h_i) \leq 0$ and $\overline{\alpha}_i(h_i) > 0$, the last inequality follows the variation condition. The case in which $h_i \in H_i^-, h_i' \in H_i^+$ can be prove in a similar way.

2. $H_i^0 \neq \emptyset$. Then variation condition says $q(h_i, h_{-i}) - q(h_i', h_{-i})$ is a constant function

with respect to h_{-i}, for all $h_i, h_i' \in H_i$. Equation 5.1 becomes

$$\alpha_i(h_i)[q(h_i, h_{-i}) - q(h_i', h_{-i})] \geq 0$$

which is true for all $h_i, h_i' \in H_i$ and $h_{-i} \in H_{-i}$ by the monotonicity condition.

$\square$

B.3.5 Proof of Proposition 5.4

Proof. **"If":** Fix any type space T and additive mechanism $(M_1, ..., M_N, q)$. It suffices to show that every agent has an interim dominant strategy. For agent i, consider the strategy σ_i that prescribes her to send a message $\overline{m}_i$ that maximizes $\pi_i^q(\cdot)$ if her type t_i satisfies $\alpha_i(\hat{\theta}_i(t_i), \hat{b}_i(t_i)) \geq 0$, or to send a message $\underline{m}_i$ that minimizes $\pi_i^q(\cdot)$ if her type satisfies $\alpha_i(\hat{\theta}_i(t_i), \hat{b}_i(t_i)) < 0$. For any strategy profile σ_{-i} from the other agents, if $\alpha_i(\hat{\theta}_i(t_i), \hat{b}_i(t_i)) \geq 0$, then we have

$$\begin{aligned}
& U_i(\overline{m}_i|\sigma_{-i}, t_i) \geq U_i(m_i|\sigma_{-i}, t_i) \\
\Leftrightarrow \sum_{t_{-i} \in T_i} \hat{\beta}_i(t_i)[t_{-i}] q(\overline{m}_i, \sigma_{-i}(t_{-i})) u_i(\hat{\theta}(t_i, t_{-i})) &\geq \sum_{t_{-i} \in T_i} \hat{\beta}_i(t_i)[t_{-i}] q(m_i, \sigma_{-i}(t_{-i})) u_i(\hat{\theta}(t_i, t_{-i})) \\
\Leftrightarrow \sum_{t_{-i} \in T_i} \hat{\beta}_i(t_i)[t_{-i}] \pi_i^q(\overline{m}_i) u_i(\hat{\theta}(t_i, t_{-i})) &\geq \sum_{t_{-i} \in T_i} \hat{\beta}_i(t_i)[t_{-i}] \pi_i^q(m_i) u_i(\hat{\theta}(t_i, t_{-i})) \\
\Leftrightarrow \pi_i^q(\overline{m}_i) \alpha_i(\hat{\theta}_i(t_i), \hat{b}_i(t_i)) &\geq \pi_i^q(m_i) \alpha_i(\hat{\theta}_i(t_i), \hat{b}_i(t_i)) \\
\Leftarrow \pi_i^q(\overline{m}_i) &\geq \pi_i^q(m_i),
\end{aligned}$$

from which we conclude that $\overline{m}_i$ is a best response against σ_{-i}. That $\underline{m}_i$ is a best response against σ_{-i} when $\alpha_i(\hat{\theta}_i(t_i), \hat{b}_i(t_i)) < 0$ is established analogously. Therefore σ_i is indeed an interim dominant strategy for agent i.

"Only if": Suppose $(M_1, ..., M_N, q)$ is IDSIC in all type spaces. We want to show that there exist functions $\pi_i^q : M_i \to [0, 1]$ such that $q(m_1, ..., m_N) = \sum_{i \in I} \pi_i^q(m_i)$.

Fix a type space where for every $i \in I$ there is a type $\tilde{t}_i$ such that $\overline{\alpha}_i(\hat{\theta}_i(\tilde{t}_i), \hat{b}_i(\tilde{t}_i)) = -\underline{\alpha}_i(\hat{\theta}_i(\tilde{t}_i), \hat{b}_i(\tilde{t}_i)) > 0$. Let m_i^* denote the message that agent i of type $\tilde{t}_i$ sends in a given interim dominant strategy equilibrium. It follows from an argument analogous to how

Lemma 5.4 is proved that

$$\underline{\alpha}_i(\hat{\theta}_i(\tilde{t}_i),\hat{b}_i(\tilde{t}_i))\big(q(m_i^*,m_{-i})-q(m_i,m_{-i})\big)+\overline{\alpha}_i(\hat{\theta}_i(t_i),\hat{b}_i(t_i))\big(q(m_i^*,m'_{-i})-q(m_i,m'_{-i})\big)\geq 0$$

for any $m_i \in M_i, m_{-i}, m'_{-i} \in M_{-i}$. That $\underline{\alpha}_i(\hat{\theta}_i(t_i),\hat{b}_i(t_i))+\overline{\alpha}_i(\hat{\theta}_i(\tilde{t}_i),\hat{b}_i(\tilde{t}_i))=0$ implies $q(m_i^*,m_{-i})-q(m_i,m_{-i})$ in invariant to m_{-i}. Thus the expression

$$q(m_i,m_{-i})-q(m_i^*,m_{-i})+\frac{1}{N}q(m_1^*,...,m_N^*).$$

does not depend on m_{-i}, and hence we can denote this expression as $\pi_i^q(m_i)$. Observe that for any $m_1,...,m_N$,

$$\begin{aligned}
q(m_1,...,m_N) = \; & q(m_1,m_2,...,m_N)-q(m_1^*,m_2,...,m_N)+\frac{1}{N}q(m_1^*,...,m_N^*)\\
& +q(m_1^*,m_2,...,m_N)-q(m_1^*,m_2^*,...,m_N)+\frac{1}{N}q(m_1^*,...,m_N^*)\\
& \quad \cdots\\
& +q(m_1^*,...,m_{N_1}^*,m_N)-q(m_1^*,...,m_{N-1}^*,m_N^*)+\frac{1}{N}q(m_1^*,...,m_N^*)\\
= \; & \sum_{i\in I}\pi_i^q(m_i).
\end{aligned}$$

Therefore $(M_1,...,M_N,q)$ is an additive mechanism. □

B.3.6 Proof of Proposition 5.5

Proof. Let σ^* be any interim dominant strategy equilibrium of the mechanism. It follows that for player i, any type t_i where $\alpha_i(\hat{\theta}_i(t_i),\hat{b}_i(t_i)) > 0$ only sends messages that maximize $\pi_i^g(\cdot)$ with positive probability (and denote that maximized value as $\overline{\pi}_i$), and any type t_i where $\alpha_i(\hat{\theta}_i(t_i),\hat{b}_i(t_i)) < 0$ only sends messages that minimize $\pi_i^g(\cdot)$ with positive probability (and denote that minimized value as $\underline{\pi}_i$). For any t_i where $\alpha_i(\hat{\theta}_i(t_i),\hat{b}_i(t_i))=0$ let $\hat{\pi}_i(t_i)$ denote the expected value of $\pi_i(\cdot)$ conditional on t_i's (mixed) strategy under σ. Define $\lambda_i := \overline{\pi}_i - \underline{\pi}_i$ and

$$\mu_i(t_i) = \begin{cases} 1 & \text{if } \alpha_i(\hat{\theta}_i(t_i), \hat{b}_i(t_i)) > 0 \\ 1 & \text{if } \alpha_i(\hat{\theta}_i(t_i), \hat{b}_i(t_i)) = 0 \text{ and } \lambda_i = 0 \\ (\hat{\pi}_i(t_i) - \underline{\pi}_i)/\lambda_i & \text{if } \alpha_i(\hat{\theta}_i(t_i), \hat{b}_i(t_i)) = 0 \text{ and } \lambda_i \neq 0 \\ 0 & \text{if } \alpha_i(\hat{\theta}_i(t_i), \hat{b}_i(t_i)) < 0. \end{cases}$$

It is straightforward to verify that the induced social choice function q_σ satisfies $q_\sigma(t) = \sum_{i\in I} \lambda_i \mu_i(t_i) + \sum_{i\in I} \underline{\pi}_i$. Thus q_σ is a random dictatorship. □

B.3.7 Proof of Proposition 5.8

Proof. Without loss, we assume $q(1, m_{-i}) \geq q(0, m_{-i})$. It is easy to check that the following strategy profile σ^* is an interim dominant strategy equilibrium,

$$\sigma_i^*(t_i) = \begin{cases} 1 & \text{if } \alpha_i(\hat{\theta}_i(t_i), \hat{b}_i(t_i)) \geq 0 \\ 0 & \text{if } \alpha_i(\hat{\theta}_i(t_i), \hat{b}_i(t_i)) < 0. \end{cases}$$

□

B.3.8 Proof of Proposition 5.9

Proof. The "only if" direction is obvious. For the "if" direction we prove its contrapositive: If a mechanism is not DSIC, then it cannot be both EPIC and IDSIC. Suppose, in order to lead to a contradiction, that there exists q that is not DSIC, yet it is EPIC and IDSIC. That q is not DSIC implies that

$$u_i(\theta)q(\theta_i, \theta'_{-i}) < u_i(\theta)q(\theta'_i, \theta'_{-i})$$

for some agent i, $\theta_i, \theta'_i \in \Theta_i$ and $\theta'_{-i} \in \Theta_{-i}$. Suppose $u_i(\theta) > 0$, then we have $q(\theta_i, \theta'_{-i}) < q(\theta'_i, \theta'_{-i})$, which implies that $\alpha_i(\theta_i) \leq 0$ and $\alpha_i(\theta'_i) \geq 0$ by Corollary 5.3.

Suppose $\alpha_i(\theta_i) < 0$. It then follows from Proposition 5.3 that $q(\theta_i, \theta_{-i}) \leq q(\theta'_i, \theta_{-i})$. If $q(\theta_i, \theta_{-i}) < q(\theta'_i, \theta_{-i})$ then $u_i(\theta)q(\theta_i, \theta_{-i}) < u_i(\theta)q(\theta'_i, \theta_{-i})$, contradicting EPIC. If $q(\theta_i, \theta_{-i}) =$

$q(\theta_i', \theta_{-i})$ then $q(\theta_i', \theta_i) - q(\theta_i, \theta_{-i}) = 0$, which implies, by condition 1 of Proposition 5.3, that $\min_{\hat{\theta}_{-i} \in \Theta_{-i}} \Big(\overline{q}(\theta_i, \hat{\theta}_{-i}) - \underline{q}(\theta_i, \hat{\theta}_{-i})\Big) = q(\theta_i', \theta_{-i}) - q(\theta_i, \theta_{-i}) = 0$. Also by condition 1 of Proposition 5.3 we have $\max_{\hat{\theta}_{-i} \in \Theta_{-i}} \Big(\overline{q}(\theta_i', \hat{\theta}_{-i}) - \underline{q}(\theta_i, \hat{\theta}_{-i})\Big) \geq q(\theta_i', \theta_{-i}') - q(\theta_i, \theta_{-i}') > 0$, therefore

$$\underline{\alpha}_i(\theta_i) \max_{\hat{\theta}_{-i} \in \Theta_{-i}} \Big(\overline{q}_i(\hat{\theta}_{-i}) - \underline{q}_i(\hat{\theta}_{-i})\Big) + \overline{\alpha}_i(\theta_i) \min_{\hat{\theta}_{-i} \in \Theta_{-i}} \Big(\overline{q}_i(\hat{\theta}_{-i}) - \underline{q}_i(\hat{\theta}_{-i})\Big) < 0$$

because $\alpha_i(\theta_i) < 0$ implies $\underline{\alpha}_i < 0$. This, however, contradicts condition 2 of Proposition 5.3.

We can thus deduce that $\alpha_i(\theta_i) = 0$. That $u_i(\theta) > 0$ implies $\overline{\alpha}_i(\theta_i) > 0$, and hence $\underline{\alpha}_i(\theta_i) = -\overline{\alpha}_i(\theta_i) < 0$. Substituting this into inequality 5.1, we have $\overline{\alpha}_i(\theta_i)\Big((q(\theta_i, \theta_{-i}') - q(\theta_i', \theta_{-i}')) - (q(\theta_i, \theta_{-i}) - q(\theta_i', \theta_{-i}))\Big) = 0$, which implies that $q(\theta_i, \theta_{-i}) - q(\theta_i', \theta_{-i}) = q(\theta_i, \theta_{-i}') - q(\theta_i', \theta_{-i}') < 0$, contradicting EPIC because $u_i(\theta) > 0$.

If $u_i(\theta) < 0$ then similar contradictions arise analogously. $\square$