Collusion-proof Mechanisms for Full Surplus Extraction

Huiyi Guo*

January 19, 2023

Abstract

The paper examines information structures that can guarantee full surplus extraction via collusion-proof mechanisms. We consider two collusion-proofness notions, the coalition incentive compatibility condition and the robust collusion-proofness condition, and both notions assume that all coalitions may be formed. When the mechanism designer is restricted to use standard Bayesian mechanisms, we show that under almost every prior distribution of agents' types, there exist payoff structures under which there is no collusion-proof full surplus extracting mechanism. However, when ambiguous mechanisms are allowed, we provide a weak necessary and sufficient condition on the prior such that collusion-proof full surplus extraction can be guaranteed. Thus, the paper sheds light on how the collusion-proofness requirement resolves the full surplus extraction paradox of Crémer and McLean (1985, 1988) and how engineering ambiguity in mechanism rules restores the paradox.

Keywords: Collusion-proofness; Multiple coalitions; Full surplus extraction; Bayesian mechanism; Ambiguous mechanism; Correlated beliefs.

JEL: D81; D82.

*Texas A&M University, Department of Economics, 4228 TAMU, College Station, TX 77843. Email address: huiyiguo@tamu.edu. The author is grateful for helpful discussions with Samson Alva, Sarah Auster, Sophie Bade, Tilman Börgers, Subir Bose, Gabriel Carroll, Simone Cerreia-Vioglio, Jimmy Chan, Federico Echenique, Hülya Eraslan, Kirill Evdokimov, Wei He, Shaowei Ke, Rich McLean, Xiaosheng Mu, Harry Pei, Mike Peters, Miklós Pintér, Phil Reny, Luca Rigotti, Mikhail Safronov, Vasiliki Skreta, Chris Shannon, Sergei Severinov, Balázs Szentes, Rodrigo Velez, Nicholas Yannelis, as well as participants at various conferences.

1 Introduction

In mechanism design theory, most works assume that agents (he) behave noncooperatively when revealing their private information to the mechanism designer (MD, she). However, there are many real-life mechanisms, such as auctions and voting, where collusion arises as a common practice, and the MD has limited power in banning it or detecting its makeup.[1] When a group of agents sees room to profit from collusion, imposing individual incentive constraints on the mechanism alone may not ensure the MD's desired outcome. In response, we wish to design mechanisms that are immune from all coalitions' joint manipulations.

This paper explores what information structures can guarantee full surplus extraction (FSE) via collusion-proof mechanisms. FSE mechanisms, giving the MD the first-best total surplus and leaving agents zero rent, are rarely seen in practice. Crémer and McLean (1985, 1988) have characterized information structures that guarantee FSE: such information structures necessarily involve correlated private information (types), yet exist broadly in any finite-dimensional type space. Hence, the theoretically permissive result on FSE is often interpreted as a paradox. In Crémer and McLean (1985, 1988)'s FSE mechanisms, each agent's monetary transfer is highly sensitive to other agents' types reported to the MD, which makes it natural for agents to contemplate colluding. Hence, it is of interest to study the extent to which the collusion-proofness requirement restricts the information structures that guarantee FSE and serves as a resolution of the paradox.

In this paper, we adopt two collusion-proofness notions. Similar to the strong Nash equilibrium and the core, both of our collusion-proofness notions assume that all coalitions can be formed. Our first notion, the coalition incentive compatibility (CIC) condition, views each coalition as a pseudo agent, whose payoff is the sum of members' payoffs and whose information is the pooling of members' information. This condition requires that no pseudo agent can benefit from misreporting his information to the MD, and is consistent with the collusion-proofness notions adopted by Green and Laffont (1979) and Safronov (2018), among others. Our second notion, the robust collusion-proofness (RCP) condition, assumes that members of a coalition can collude via any incentive compatible and individually rational

[1] See Che and Kim (2006) for a review of the literature on collusion against exogenously given institutions.

side contract. The RCP condition requires that collusion of every coalition, if occurs, should not harm the MD. This condition is the same as the one of Che and Kim (2006) except that we require the mechanism to be immune from manipulations of all coalitions, while they focus on a particular coalition.

Although FSE can be guaranteed under a broad set of information structures, we find that achieving it via collusion-proof standard Bayesian mechanisms is challenging. In particular, Proposition 1 shows that under some mild dimensional restrictions on agents' type space, for almost all prior distributions over it, there exist payoff structures under which it is neither possible to achieve FSE via a mechanism satisfying the CIC condition, nor possible via one satisfying the RCP condition. These findings contrast with the result in Che and Kim (2006): in a model with at least three agents and one known colluding coalition, they show that there is a broad class of information structures for which collusion-proof FSE can be guaranteed. Thus, the peril that more coalitions may be formed significantly reduces information structures that guarantee FSE and offers a resolution of the FSE paradox.

As it is difficult to guarantee collusion-proof FSE by adopting standard Bayesian mechanisms, the MD might be motivated to use a broader collection of tools called ambiguous mechanisms (e.g., Bose and Renou, 2014; Di Tillio et al., 2017; Guo, 2019). An ambiguous mechanism has vaguely described rules: the MD can secretly commit to a standard Bayesian mechanism, but strategically announce multiple potential mechanisms. In reality, the vague tax audit scheme can be viewed as an ambiguous mechanism. We assume that agents are ambiguity-averse towards the unknown mechanism rule and make decisions with the maxmin expected utility of Gilboa and Schmeidler (1989).

We show that collusion-proof FSE can be guaranteed, if and only if agents' prior distribution satisfies the Coalition Beliefs Determine Preferences (CBDP) property. This is true under two collusion-proofness notions, the CIC condition and the RCP* condition, where the latter is slightly stronger than the RCP condition. The CBDP property strengthens the Beliefs Determine Preferences (BDP) property of Neeman (2004) by also requiring the knowledge of any non-grand coalition's posterior belief over types of agents out of the coalition to pin down this coalition's type profile. In any fixed finite type space, the CBDP property imposes a weak restriction on the prior over the type space. Thus, there is a broad class of prior distributions under which collusion-proof FSE can be achieved via ambiguous mechanisms. In particular, ambiguity can be engineered to soften the unpermissive result on collusion-proof FSE in the two-agent environment of Laffont and Martimort (2000), and to address the peril that all coalitions can be formed in the multiple-agent model of Che and Kim (2006). As such, the use of ambiguous mechanisms can restore the FSE paradox.

Literature Review. The paper is related to three strands of the literature.

First, the paper is related to the literature on mechanism design under correlated beliefs.

Among others, Crémer and McLean (1985, 1988), McAfee and Reny (1992), and Lopomo et al. (2022) have characterized conditions on the information structures so that FSE can be guaranteed. In particular, in a finite type space, Crémer and McLean (1988) have shown that Convex Independence is the necessary and sufficient condition on the prior to guarantee FSE. A related question is what information structures with correlated beliefs can guarantee the implementability of all efficient allocations, with or without additional individual rationality and/or budget balance restrictions on the mechanism. The proper scoring rule (see Börgers et al., 2015, for a reference) and the work of d'Aspremont et al. (1990, 2004), McLean and Postlewaite (2004, 2015), and Kosenok and Severinov (2008), among others, have provided answers to these questions. The methodology adopted in the current paper is related to Crémer and McLean (1988) and Kosenok and Severinov (2008): we also focus on a finite type space and establish the existence of a desirable mechanism via the duality approach. A key difference is that we design the mechanism in a way that is immune from collusion.

The current paper is directly related to the literature on collusion-proof mechanisms.

One approach in the literature considers all possible coalitions' deviations and imposes the collusion-proofness requirement on the mechanism axiomatically without explicitly modeling strategic interactions due to information asymmetry within a coalition. Green and Laffont (1979), Chen and Micali (2012), Bierbrauer and Hellwig (2016), and Safronov (2018), among others, adopt this approach. This approach provides a benchmark to study collusion-proof mechanisms since the worst-case scenario from the MD's perspective is that all coalitions may be formed and that agents collude without encountering information frictions. In this strand of the literature, Green and Laffont (1979), Chen and Micali (2012), and Bierbrauer and Hellwig (2016) focus on ex-post collusion-proofness notions, and Safronov (2018) adopts an interim notion. Our CIC condition follows from Safronov (2018), except that our agents/coalitions use maxmin expected utility when facing an ambiguous mechanism. Safronov (2018) shows that in private-value environments with independent beliefs, every efficient allocation rule is implementable via an incentive compatible, budget balanced, and collusion-proof mechanism. Our results complement this approach by deriving contrasting implications for FSE under standard Bayesian mechanisms and ambiguous mechanisms.

Another approach to studying collusion-proof mechanisms focuses on one particular coalition that can be formed and explicitly considers within-coalition information asymmetry, which may undermine the coalition's ability to collude. Laffont and Martimort (1997, 2000), Che and Kim (2006), and Meng et al. (2017) follow this approach. Laffont and Martimort (1997, 2000) and Meng et al. (2017) characterize the optimal collusion-proof mechanisms in their specific payoff structures with two agents and two types. One observation from them is that the collusion-proofness requirement does not have an additional bite into the MD's ability to extract agents' surplus in the independent belief case, but can have an additional bite if agents have correlated beliefs. Che and Kim (2006) consider environments with more agents and general payoff structures. They extend the positive result on collusion-proof mechanism design under independent beliefs and also identify a sufficient condition on the information structure that guarantees collusion-proof FSE. Our RCP and RCP* conditions are built on the collusion-proofness notion of Che and Kim (2006): we strengthen their notion by requiring a mechanism to be immune from all coalitions' manipulations. Our work contributes to this approach by revealing the importance of the MD's knowledge of the composition of the colluding coalition in the FSE problem: when all coalitions can be formed, it may be impossible to design an RCP standard Bayesian mechanism that achieves FSE.

The paper also fits into the literature on mechanism design with ambiguity-averse agents.

Some works in this literature, including the current one, explore if it is possible to strategically engineer ambiguity in the mechanism to improve its performance. One approach involves endogenously generating ambiguity in agents' beliefs towards other agents' types: Bose and Renou (2014) do so via an ambiguous communication device and the endogenously engineered ambiguous beliefs may allow the MD to implement social choice functions that are not implementable otherwise. A second approach generates ambiguity on the payoff rule more directly. For example, Di Tillio et al. (2017), Bose and Daripa (2017), Guo (2019), and Tang and Zhang (2021) demonstrate that ambiguous mechanisms are more potent than standard Bayesian mechanisms in screening, preferences elicitation, FSE, and implementing social choice correspondences. To the best of our knowledge, the current paper is the first one that studies how ambiguous mechanisms can be introduced to address collusion.

In some other works, agents are assumed to hold ambiguous beliefs about other agents' types exogenously, and the MD designs the optimal/efficient standard Bayesian mechanisms, e.g., Bose et al. (2006), Renou (2015), Wolitzky (2016), De Castro and Yannelis (2018), Song (2018, 2022), Kocherlakota and Song (2019), and Lopomo et al. (2020). The current work differs from these papers, as we do not assume ambiguous beliefs about other agents' types.

The rest of the paper proceeds as follows. Section 2 sets up the model. Section 3 defines the two collusion-proofness notions and provides a group of theoretically unpermissive results on collusion-proof FSE via standard Bayesian mechanisms. Section 4 presents a group of possibility results by adopting ambiguous mechanisms. Section 5 concludes. All proofs are relegated to the Appendix.

2 Set-up

We study an environment with one mechanism designer (MD, she) and a finite set of agents (he) $I = \{1, 2, ..., n \geqslant 2\}$. Each $i \in I$ privately observes his type $\theta_i \in \Theta_i$, where Θ_i is i's type set satisfying $2 \leqslant |\Theta_i| < +\infty$ and $\Theta \equiv \prod_{i \in I} \Theta_i$ is the finite type space. Assume that there is a full-support common prior over Θ, i.e., a $p \in \Delta(\Theta)$ such that $p(\theta) > 0$ for each $\theta \in \Theta$. The pair (Θ, p) is called an **information structure**.

For any $S \in 2^I \backslash \{\varnothing\}$, S is an agent if $|S| = 1$, a **coalition** if $2 \leqslant |S| \leqslant n$, and the grand coalition if $|S| = n$. For S with $1 \leqslant |S| < n$, given type profile $\theta_S \equiv (\theta_i)_{i \in S} \in \Theta_S \equiv \prod_{i \in S} \Theta_i$, we let $p(\cdot|\theta_S) \equiv (p(\theta_{-S}|\theta_S))_{\theta_{-S} \in \Theta_{-S}}$ denote the posterior belief over types of agents out of S, where $p(\theta_{-S}|\theta_S) \equiv \frac{p(\theta_S, \theta_{-S})}{p(\theta_S)}$, $p(\theta_S) \equiv \sum_{\theta'_{-S} \in \Theta_{-S}} p(\theta_S, \theta'_{-S})$, and $\theta_{-S} \equiv (\theta_j)_{j \in I \backslash S}$. Agents are said to have independent beliefs, if for all $i \in I$, $p(\cdot|\theta_i)$ is constant across different $\theta_i \in \Theta_i$; otherwise, agents have correlated beliefs. For simplicity, denote θ_I by θ.

The MD's quasi-linear utility function is of the form $u_0(a) - \sum_{i \in I} \bar{t}_i$ and each agent's quasi-linear utility function is of the form $u_i(a,\theta) + \bar{t}_i$, where $a \in A$ is an element in the compact set of feasible outcomes, $u_0(a)$ and $u_i(a,\theta)$ are the MD's and agent i's payoff from outcome a, respectively, and $\bar{t}_i \in \mathbb{R}$ is the monetary transfer from the MD to agent i.[2] Assume that there is a feasible outcome in A that gives each agent and the MD zero payoffs, i.e., an outside option. We call the profile of utility functions $(u_0, (u_i)_{i\in I})$ a **payoff structure**.

For example, in the single-unit auction case, we may have $A = \{0, 1, ..., n\}$, which means that the good is not produced, allocated to agent 1,..., and allocated to agent n, respectively. View $-u_0(a)$ as the MD's cost of producing outcome a. For each $i \in I$, $u_i(a,\theta)$ may depend on all agents' private information (e.g., as in a common value auction), or in the degenerate **private-value** case, depend on θ_i only (e.g., in a private-value auction).

Let (q,t) denote a (direct) **standard Bayesian mechanism**, where $q : \Theta \to A$ is the allocation rule that assigns the outcome and $t : \Theta \to \mathbb{R}^n$ is the transfer rule that describes the monetary payment received by agents. We may simply call (q,t) a mechanism. Let $\sigma_i : \Theta_i \to \Delta(\Theta_i)$ denote agent i's reporting strategy, under which $\sigma_i[\theta_i](\theta_i')$ is the probability that type-θ_i agent reports θ_i'. Let $\bar{\sigma}_i$ denote the truthful reporting strategy, i.e., the one such that $\bar{\sigma}_i[\theta_i](\theta_i) = 1$ for all $\theta_i \in \Theta_i$.

An allocation rule q is said to be **efficient**, if it maximizes the ex-ante total surplus, i.e.,

$$\sum_{\theta\in\Theta}[u_0\big(q(\theta)\big) + \sum_{i\in I} u_i\big(q(\theta),\theta\big)]p(\theta) = \max_{\tilde{q}:\Theta\to A}\sum_{\theta\in\Theta}[u_0\big(\tilde{q}(\theta)\big) + \sum_{i\in I} u_i\big(\tilde{q}(\theta),\theta\big)]p(\theta) \equiv FS,$$

where FS stands for the full surplus, or equivalently, if q maximizes the ex-post total surplus pointwise, i.e., $u_0\big(q(\theta)\big) + \sum_{i\in I} u_i\big(q(\theta),\theta\big) \geqslant u_0(a) + \sum_{i\in I} u_i\big(a,\theta\big)$ for all $a \in A$ and $\theta \in \Theta$.

If type-θ_i agent i follows strategy σ_i and other agents truthfully report, his utility is

$$V_i[q,t](\theta_i,\sigma_i) \equiv \sum_{\theta_i'\in\Theta_i}\sum_{\theta_{-i}\in\Theta_{-i}}[u_i\big(q(\theta_i',\theta_{-i}),(\theta_i,\theta_{-i})\big) + t_i(\theta_i',\theta_{-i})]p(\theta_{-i}|\theta_i)\sigma_i[\theta_i](\theta_i').$$

When $\sigma_i[\theta_i](\theta_i') = 1$, we may let $V_i[q,t](\theta_i,\theta_i')$ denote $V_i[q,t](\theta_i,\sigma_i)$ for simplicity.

[2] If A contains lotteries, we assume that $u_0(\cdot)$ and $u_i(\cdot,\theta)$ defined over lotteries are consistent with the expected utility theory. This assumption is used in the proof of Claim 3. We also remark that the analysis of the paper does not change if we generalize u_0 so that it depends on both a and θ. However, we focus on the current setup, which is the same as that of Che and Kim (2006), to highlight our observation that their possibility result on collusion-proof FSE is overturned when all coalitions can be formed.

The mechanism (q, t) is said to be **feasible** if it satisfies the interim individual rationality (IR) condition and the interim incentive compatibility (IC) condition below:

$$\text{IR} \qquad V_i[q,t](\theta_i, \bar{\sigma}_i) \geqslant 0, \forall i \in I, \theta_i \in \Theta_i,$$

$$\text{IC} \qquad V_i[q,t](\theta_i, \bar{\sigma}_i) \geqslant V_i[q,t](\theta_i, \sigma_i), \forall i \in I, \theta_i \in \Theta_i, \sigma_i : \Theta_i \to \Delta(\Theta_i).$$

Given a payoff structure $(u_0, (u_i)_{i \in I})$, we say (q, t) achieves **full surplus extraction** (FSE), if (q, t) is feasible and makes the MD's ex-ante payoff equal to the FS, i.e.,

$$\sum_{\theta \in \Theta} [u_0\big(q(\theta)\big) - \sum_{i \in I} t_i(\theta)] p(\theta) = FS.$$

FSE requires that q must be efficient and agents' IR constraints must bind. We say an information structure (Θ, p) **guarantees FSE** if for any payoff structure $(u_0, (u_i)_{i \in I})$, there exists a mechanism (q, t) that achieves FSE. Crémer and McLean (1988) have shown that an information structure (Θ, p) guarantees FSE, if and only if the prior p satisfies the Convex Independence condition (Definition 2 in Appendix A.1), which necessitates correlated beliefs. For almost all priors p over a finite type space Θ, Convex Independence is satisfied.

3 Standard Bayesian Mechanisms

3.1 Definitions

When a coalition can secretly coordinate members' reports to the MD, merely focusing on individual incentives may not be sufficient for the MD to extract the full surplus. Motivated by this, we consider two collusion-proofness notions that have been introduced in the literature and will impose them on FSE mechanisms.

To begin with, we define a joint reporting strategy. For any $S \in 2^I \backslash \{\varnothing\}$, denote a **joint reporting strategy** by $\delta^S : \Theta_S \to \Delta(\Theta_S)$. When S is an agent, δ^S can be viewed as this agent's strategy. When S is a coalition, by adopting δ^S, coalition S with type profile θ_S jointly reports type profile θ'_S with probability $\delta^S[\theta_S](\theta'_S)$ to mechanism (q, t). Let $\bar{\delta}^S : \Theta_S \to \Delta(\Theta_S)$ be the truthful joint reporting strategy.

One strand of the literature, e.g., Green and Laffont (1979) and Safronov (2018), views every coalition S as a pseudo agent with type set Θ_S and requires that all types of all pseudo agents have no incentive to misreport. This pseudo agent approach implicitly assumes that members within a coalition pool/disclose their information to each other. As in Safronov (2018), the pseudo agent has a "utility function" which is the sum of members' utility functions and maximizes the interim utility. By adopting δ^S, a type-θ_S pseudo agent earns interim utility $V_S[q,t](\theta_S,\delta^S) \equiv \sum_{i\in S} V_i[q,t](\theta_S,\delta^S)$, where

$$V_i[q,t](\theta_S,\delta^S) \equiv \sum_{\theta'_S\in\Theta_S}\sum_{\theta_{-S}\in\Theta_{-S}} [u_i\big(q(\theta'_S,\theta_{-S}),(\theta_S,\theta_{-S})\big) + t_i(\theta'_S,\theta_{-S})]p(\theta_{-S}|\theta_S)\delta^S[\theta_S](\theta'_S).$$

When $\delta^S[\theta_S](\theta'_S) = 1$, let $V_S[q,t](\theta_S,\theta'_S)$ denote $V_S[q,t](\theta_S,\delta^S)$. When $S = I$, the definition should read as: $V_i[q,t](\theta,\delta^I) \equiv \sum_{\theta'\in\Theta}[u_i\big(q(\theta'),\theta\big) + t_i(\theta')]\delta^I[\theta](\theta')$.

With the above notations, we present our first collusion-proofness notion. A mechanism (q,t) is said to satisfy the **coalition incentive compatibility** (CIC) condition, if for all $S \in 2^I\backslash\{\varnothing\}$ with $|S| \geqslant 2$, $\theta_S \in \Theta_S$, and $\delta^S : \Theta_S \to \Delta(\Theta_S)$, $V_S[q,t](\theta_S,\bar{\delta}^S) \geqslant V_S[q,t](\theta_S,\delta^S)$.

Instead of assuming that members in a coalition disclose private information to each other, another strand of the literature, e.g., Laffont and Martimort (2000) and Che and Kim (2006), adopts the mechanism design approach to model information transmission within a coalition. Before types are reported to (q,t), a mediator benevolent to a coalition S can secretly approach S and offer an S**-side contract** (δ^S,ψ^S), which consists of a joint reporting strategy δ^S and a **side transfer rule**, i.e., a mapping $\psi^S : \Theta_S \to \mathbb{R}^{|S|}$ such that $\sum_{i\in S}\psi_i^S(\theta_S) = 0$ for all $\theta_S \in \Theta_S$.

Following Che and Kim (2006), we focus on the outcome implemented as a result of the S-side contract. Given (q,t) and S, we say $(\tilde{q} : \Theta \to A, \tilde{t} : \Theta \to \mathbb{R}^n)$ is an S**-reallocational manipulation**, if there exists an S**-side contract** (δ^S,ψ^S), such that for all $\theta \in \Theta$,

$$\tilde{t}_i(\theta) = \begin{cases} \sum_{\theta'_S\in\Theta_S} t_i(\theta'_S,\theta_{-S})\delta^S[\theta_S](\theta'_S) + \psi_i^S(\theta_S) & \text{if } i \in S, \quad (1)\\ \sum_{\theta'_S\in\Theta_S} t_i(\theta'_S,\theta_{-S})\delta^S[\theta_S](\theta'_S) & \text{otherwise}, \quad (2)\end{cases}$$

$$u_0(\tilde{q}(\theta)) = \sum_{\theta'_S\in\Theta_S} u_0\big(q(\theta'_S,\theta_{-S})\big)\delta^S[\theta_S](\theta'_S), \tag{3}$$

$$u_i(\tilde{q}(\theta),\theta) = \sum_{\theta'_S\in\Theta_S} u_i\big(q(\theta'_S,\theta_{-S}),\theta\big)\delta^S[\theta_S](\theta'_S), \forall i \notin S. \tag{4}$$

By (3) and (4), the reallocation can only take place within S and thus should be undetectable by the MD and the noncollusive agents (agents out of S). If equation (4) also holds for all $i \in S$, then $(\tilde{q}, \tilde{t})$ reduces to an **S-communicative manipulation**, which does not involve reallocation of the outcome. Given (q, t), we may denote $(\tilde{q}, \tilde{t})$, an S-reallocational manipulation (resp. S-communicative manipulation) induced by (δ^S, ψ^S), as $(\tilde{q}, t^{\delta^S} + \psi^S)$ (resp. $(q^{\delta^S}, t^{\delta^S} + \psi^S)$) to highlight its structure. The S-reallocational manipulation $(\tilde{q}, \tilde{t})$ is said to be **S-feasible**, if no agent in S has the incentive to decline or misreport in $(\tilde{q}, \tilde{t})$, i.e.,

$$S\text{-IR} \qquad V_i[\tilde{q}, \tilde{t}](\theta_i, \bar{\sigma}_i) \geqslant V_i[q, t](\theta_i, \bar{\sigma}_i), \forall i \in S, \theta_i \in \Theta_i,$$

$$S\text{-IC} \qquad V_i[\tilde{q}, \tilde{t}](\theta_i, \bar{\sigma}_i) \geqslant V_i[\tilde{q}, \tilde{t}](\theta_i, \sigma_i), \forall i \in S, \theta_i \in \Theta_i, \sigma_i : \Theta_i \to \Delta(\Theta_i).$$

A weighted sum of the above S-IR constraints implies that engaging in this collusion is at least not harmful to S in terms of members' aggregate ex-ante utility.

In the collusion-proofness notion adopted by Che and Kim (2006), only one coalition, denoted by S, may be formed. Agents can collude via any S-feasible reallocational manipulation, but if collusion takes place, it should neither hurt the MD nor lead to an infeasible mechanism. In particular, if manipulation of S is anticipated but violates noncollusive agents' IR or IC constraints, then the noncollusive agents will decline or misreport in the main mechanism, which may jeopardize the MD's payoff. Formally, for any $S \in 2^I \backslash \{\varnothing\}$ with $|S| \geqslant 2$, a feasible mechanism (q, t) is said to be **robustly collusion-proof with respect to S** (RCP with respect to S), if every S-feasible S-reallocational manipulation $(\tilde{q}, \tilde{t})$ is a feasible mechanism and satisfies $\sum_{\theta \in \Theta}[u_0(q(\theta)) - \sum_{i \in I} t_i(\theta)]p(\theta) = \sum_{\theta \in \Theta}[u_0(\tilde{q}(\theta)) - \sum_{i \in I} \tilde{t}_i(\theta)]p(\theta)$.

In the current paper, we assume that the MD is concerned about joint manipulations of all possible coalitions, rather than just one. Hence, the second collusion-proofness notion considered by the current paper is more demanding than requiring RCP with respect to any particular coalition. A feasible mechanism (q, t) is said to be **robustly collusion-proof** (RCP) if it is RCP with respect to every $S \in 2^I \backslash \{\varnothing\}$ with $|S| \geqslant 2$.

3.2 Results

To the best of our knowledge, the literature has not explored how the CIC condition affects FSE problems. Che and Kim (2006) have shown that when there are at least three agents, for any finite type space and almost every prior over it, the corresponding information structure can guarantee FSE via mechanisms that are RCP with respect to the grand coalition.

However, we show with an example below that when more coalitions can be formed, their possibility result may no longer hold. In fact, the additional CIC or RCP condition drastically restricts the information structures that guarantee FSE. Hence, without the knowledge of the makeup of the colluding coalition, collusion-proof FSE remains a challenging question.

Example 1. *To show that the additional CIC or RCP condition significantly restricts the information structures that guarantee FSE, let Θ be any finite type space with $n \geqslant 2$ agents, and we focus on a broad class of prior distributions over Θ. By relabeling the indices, we assume without loss that agent n has a relatively small type set in the sense that $|\Theta_n| \leqslant |\Theta_{-n}|$. For each of these priors, it suffices to show that there exists a payoff structure under which collusion-proof FSE is impossible. We demonstrate an example of such a payoff structure in a single-unit auction setup, where the good is costlessly produced by the MD (auctioneer).*

For simplicity, assume that agents (bidders) have private-value utility functions and agent n is the strongest bidder who always dominates all others. Namely, $\theta_n > \theta_j > 0$ for all $\theta_n \in \Theta_n$, $j \in I\backslash\{n\}$, and $\theta_j \in \Theta_j$, where each agent i's private valuation of the good is denoted by $\theta_i \in \Theta_i$. In this case, the unique efficient allocation rule q takes a simple form: it always assigns the good to agent n, and thus there is no uncertainty about who receives the good.

Following the rest of the paper, we assume that the MD cannot ban the formation of any coalition. However, Claims 1 through 4 merely rely on the possibility that the following two coalitions can be formed: the group of weak agents $I\backslash\{n\}$ who never receive the good and the grand coalition I where every member is fully exploited.

To extract the full surplus from agents, the most natural way is to charge agent n only and make his transfer contingent on all agents' reports – this allows the MD to take advantage of correlated beliefs across agents. However, this leaves leeway for collusion because agent n can bribe all other agents so that they report in a way in his favor. Hence, if there exists a collusion-proof FSE mechanism, the transfer rule has to be more sophisticatedly designed.

We now present Claim 1 (all proofs in this section are provided in Appendix A.2), which will be used to establish the result under our first collusion-proofness notion, the CIC condition. For convenience, we label elements in Θ_{-n} by θ_{-n}^1, θ_{-n}^2,..., $\theta_{-n}^{|\Theta_{-n}|}$ and those in Θ_n by

θ_n^1, θ_n^2,..., $\theta_n^{|\Theta_n|}$ and let $\hat{X}$ be the following $|\Theta_n| \times |\Theta_n|$ matrix.

$$\hat{X} \equiv \begin{bmatrix} p(\theta_{-n}^1, \theta_n^1) & p(\theta_{-n}^1, \theta_n^2) & \dots & p(\theta_{-n}^1, \theta_n^{|\Theta_n|}) \\ p(\theta_{-n}^2, \theta_n^1) & p(\theta_{-n}^2, \theta_n^2) & \dots & p(\theta_{-n}^2, \theta_n^{|\Theta_n|}) \\ \vdots & \vdots & \ddots & \vdots \\ p(\theta_{-n}^{|\Theta_n|}, \theta_n^1) & p(\theta_{-n}^{|\Theta_n|}, \theta_n^2) & \dots & p(\theta_{-n}^{|\Theta_n|}, \theta_n^{|\Theta_n|}) \end{bmatrix}. \tag{5}$$

Claim 1. *Assume that p is such that $\hat{X}$ has full rank. There exists a collection of private-value parameters, such that for any FSE mechanism (q,t), there exists $S \in \{I \backslash \{n\}, I\}$ and two type profiles $\theta_S \neq \theta'_S \in \Theta_S$ such that $V_S[q,t](\theta_S, \theta'_S) > V_S[q,t](\theta_S, \theta_S)$.*

We adopt a duality approach to establish the above claim. The proof adjusts the private-value parameters in a way that depends on the information structure. Assuming there exists a transfer rule t such that (q,t) satisfies multiple IC, CIC, and IR constraints, we then use these constraints to reach a contradiction.

Notice that for a finite Θ, for almost all $p \in \Delta(\Theta)$, $\hat{X}$ has full rank since $|\Theta_n| \leqslant |\Theta_{-n}|$. Thus, Claim 1 immediately implies the following result.

Claim 2. *For any Θ and almost any prior $p \in \Delta(\Theta)$, there exists a private-value payoff structure under which it is impossible to achieve FSE via mechanisms satisfying the CIC condition.*

The above result does not rely on the value of n. In the remainder of this example, we impose the assumption that $n \geqslant 4$ so that $|I \backslash \{n\}| \geqslant 3$ under which we can establish a result regarding our second collusion-proofness notion, the RCP condition.

Let $(\Theta_{-n}, \hat{p})$ be the information structure in the sub-environment with agents in $I \backslash \{n\}$ only, where $\hat{p}$ is the marginal distribution of p over Θ_{-n}. We have the following claim.

Claim 3. *Assume that $p \in \Delta(\Theta)$ and $\hat{p} \in \Delta(\Theta_{-n})$ satisfy both the Convex Independence condition and the Identifiability condition (Definition 3 in Appendix A.1) and that $\hat{X}$ has full rank. There exists a collection of private-value parameters, such that for any FSE mechanism (q,t), there exists $S \in \{I \backslash \{n\}, I\}$, $\hat{\delta}^S : \Theta_S \to \Delta(\Theta_S)$, and a side transfer rule ψ^S such that $(q^{\hat{\delta}^S}, t^{\hat{\delta}^S} + \psi^S)$ is an S-feasible S-communicative manipulation and*

$$\sum_{\theta_S \in \Theta_S} V_S[q,t](\theta_S, \hat{\delta}^S) p(\theta_S) > \sum_{\theta_S \in \Theta_S} V_S[q,t](\theta_S, \bar{\delta}^S) p(\theta_S). \tag{6}$$

Under the same payoff structure constructed for Claim 1, for the coalition S such that the inequality in Claim 1 holds, let $\hat{\delta}^S : \Theta_S \to \Delta(\Theta_S)$ be a joint reporting strategy that maximizes $V_S[q,t](\theta_S, \delta^S)$ for each $\theta_S \in \Theta_S$. It is clear that (6) is satisfied. The key to the proof is to establish the existence of a side transfer rule ψ^S such that the S-communicative manipulation is S-feasible. This is done by adapting the main result of Kosenok and Severinov (2008) to a sub-environment with agents in S only: there exists an ex-post budget balanced $\psi^S : \Theta_S \to \mathbb{R}^{|S|}$ such that it is part of a feasible mechanism. The feasibility is then used to establish the S-feasibility of the S-communicative manipulation in the original environment.

Since q is efficient, (6) implies that the S-communicative manipulation either decreases the ex-ante payoff of the MD, or hurts at least one agent out of S. In the latter case, the S-communicative manipulation leads to an infeasible mechanism. In either case, (q,t) is not RCP with respect to S, and thus not RCP.

To formalize the message, recall some genericity observations on the assumptions imposed on the prior in Claim 3. When $|\Theta_i| \leqslant |\Theta_{-i}|$ for all $i \in I$, almost every prior $p \in \Delta(\Theta)$ satisfies the Convex Independence condition. By Kosenok and Severinov (2008), when $n = 3$ and $|\Theta| \geqslant 12$ (i.e., $n > 3$ or $n = 3$ but there exists an agent with at least three types), almost every prior $p \in \Delta(\Theta)$ satisfies the Identifiability condition. Hence, under mild dimensional restrictions on the type space, the assumptions in Claim 3 hold for almost all priors. These observations immediately lead to the following claim and we omit its proof.

Claim 4. *Suppose Θ is such that $n \geqslant 4$ and $|\Theta| \geqslant 24$, and that there exists $j \in I$ for whom $|\Theta_j| \leqslant |\Theta_{-j}|$ and $|\Theta_i| \leqslant |\Theta_{-j-i}|$ for all $i \in I\backslash\{j\}$. Under almost every prior $p \in \Delta(\Theta)$, there exists a private-value payoff structure under which it is impossible to achieve FSE via mechanisms satisfying the RCP condition.*

Notice that the dimensional restrictions in Claim 4 are satisfied under the most commonly studied case where agents' type sets are equally large and when the number of agents is large enough: when $n \geqslant 5$ or $n \geqslant 4$ and each agent has at least three types.

This example provides sufficient conditions on the priors such that collusion-proof FSE cannot be guaranteed. These sufficient conditions may not be necessary: there are other information structures that cannot guarantee collusion-proof FSE. For example, in a two-

agent two-type public good setting, Laffont and Martimort (2000) have characterized the MD's optimal collusion-proof mechanism, and when agents' beliefs are weakly positively correlated, there does not exist an FSE mechanism that is RCP with respect to $I = \{1, 2\}$.[3] *As I is the only coalition in their two-agent setup, there does not exist an FSE mechanism that is RCP. Their argument is different from the duality approach that establishes Claims 1 through 4, and we refer readers to their Proposition 5 for the proof.*

Example 1 shows that for a broad class of information structures, we can construct private-value auction payoff structures under which collusion-proof FSE cannot be guaranteed. In particular, Claim 2 and Claim 4 can be understood as two almost-everywhere impossibility results on guaranteeing FSE via CIC mechanisms and RCP mechanisms, respectively. These intuitive private-value payoff structures used in the example may be of interest on their own. However, when we are not restricted to constructing the payoff structure with private-value utility functions, the almost-everywhere impossibility result in Claim 2 can be strengthened into an everywhere impossibility result as in Proposition 1 (*i*). Proposition 1 (*ii*) is a direct corollary of Claim 4 and thus we omit this part of the proof.

Proposition 1. *(i) There does not exist any information structure (Θ, p) that guarantees FSE via mechanisms satisfying the CIC condition.*

(ii) Suppose Θ is such that $n \geqslant 4$ and $|\Theta| \geqslant 24$, and that there exists $j \in I$ for whom $|\Theta_j| \leqslant |\Theta_{-j}|$ and $|\Theta_i| \leqslant |\Theta_{-j-i}|$ for all $i \in I \backslash \{j\}$. Under almost every prior $p \in \Delta(\Theta)$, the information structure (Θ, p) cannot guarantee FSE via mechanisms satisfying the RCP condition.

Similar to Example 1, Proposition 1 only relies on the possibility that two coalitions can be formed: a coalition of $n-1$ agents and the grand coalition I. To establish Proposition 1 (*i*), we construct an interdependent-value payoff structure that admits a unique efficient allocation rule q, assume that there exists a mechanism (q, t) satisfying multiple IC and CIC constraints, and then reach a contradiction without using any IR constraints. The

[3] In fact, the collusion-proofness notion adopted by Laffont and Martimort (2000) is called the weak collusion-proofness. Che and Kim (2006) have shown that when a mechanism extracts the full surplus and thus involves an efficient allocation rule, RCP with respect to I is stronger than weak collusion-proofness.

proof can also establish the following result on implementation: there does not exist any information structure (Θ, p) that guarantees implementability of all efficient allocation rules via mechanisms satisfying the IC and CIC conditions.

By Proposition 1, the concern that coalitions may be formed makes it difficult for the MD to guarantee FSE via standard Bayesian mechanisms. If we view the FSE result of Crémer and McLean (1985, 1988) as a paradox, our Proposition 1 shows that collusion is one potential resolution of this paradox. This observation is related to Laffont and Martimort (2000) and Che and Kim (2006). Laffont and Martimort (2000) show that in a two-agent two-type framework, the MD may fail to obtain FSE when the two agents can collude, and our Proposition 1 (*ii*) further extends this message into environments with more agents and more types. According to Che and Kim (2006), in a framework with at least three agents, it is usually easy to obtain FSE via a standard Bayesian mechanism that is RCP with respect to one known colluding coalition, but our Proposition 1 (*ii*) shows that the unknown makeup of the coalition can overturn this possibility result.

4 Ambiguous Mechanisms

4.1 Definitions

Section 3 has shown that it may be difficult to achieve FSE via collusion-proof standard Bayesian mechanisms. As such, the MD may want to consider a broader set of tools called ambiguous mechanisms (see, e.g., Bose and Renou, 2014; Di Tillio et al., 2017; Guo, 2019; Tang and Zhang, 2021). An **ambiguous mechanism** is a publicly announced nonempty compact set of mechanisms, among which agents do not know the true mechanism that the MD has committed to.[4] Other than the ambiguity engineered by the MD, we assume that there is no other exogenous ambiguity and agents still share an unambiguous common prior. This simplifying assumption helps us to compare the results in the current section with those from the previous section, and thus highlight the freedom granted by ambiguous mechanisms. As we will show in this section, focusing on direct mechanisms and ambiguity

[4]We impose the compactness assumption so that the min operator can be conveniently used.

in transfer rules alone is sufficient to turn the technically unpermissive results on collusion-proof FSE in Section 3 into possibility results.[5] Hence, for FSE, we focus on the case where all mechanisms share an efficient allocation rule q, and T is the set of potential transfer rules. Formally, given $(u_0, (u_i)_{i\in I})$, we say an ambiguous mechanism (q, T) achieves FSE, if it is feasible (remains to be redefined) and

$$\sum_{\theta\in\Theta}[u_0\big(q(\theta)\big) - \sum_{i\in I} t_i(\theta)]p(\theta) = FS, \forall t \in T. \tag{7}$$

When the set T is a singleton, the ambiguous mechanism reduces to a standard Bayesian mechanism that we have been using in previous sections.

For simplify, we assume that each agent has the maxmin expected utility (MEU) of Gilboa and Schmeidler (1989).[6] We assume that an (potentially pseudo) agent $S \in 2^I\backslash\{\varnothing\}$ with type $\theta_S \in \Theta_S$ holds a nonempty, compact, and convex set of probabilistic assessments over $T \times \Theta_{-S}$: the set of all distributions over $T \times \Theta_{-S}$ whose marginal distribution over Θ_{-S} is $p(\cdot|\theta_S)$. Among the multiple-belief set, the worst-case expected utility of following strategy $\delta^S : \Theta_S \to \Delta(\Theta_S)$ is $V_S[q, T](\theta_S, \delta^S) \equiv \min_{t\in T} V_S[q, t](\theta_S, \delta^S)$. When T is a singleton, the MEU is consistent with the subjective expected utility. By replacing $V_S[q, t](\theta_S, \delta^S)$ with $V_S[q, T](\theta_S, \delta^S)$, we can redefine the IR, IC, CIC, and the feasibility conditions.

It is natural to assume that a coalition S can also collude via an ambiguous side contract, and make it further contingent on information that will be revealed from the main mechanism ($t \in T$ and $\theta_{-S} \in \Theta_{-S}$). Following this spirit, an **ambiguous *S*-side contract** is a pair (δ^S, Ψ^S), where Ψ^S is a set of potential side transfer rules and each element in it, $\psi^S : T \times \Theta \to \mathbb{R}^{|S|}$, is required to satisfy $\sum_{i\in S} \psi_i^S(t, \theta) = 0$ for all $t \in T$ and $\theta \in \Theta$. The mediator commits to one $\psi^S \in \Psi^S$ but only announces (δ^S, Ψ^S) to S in the ex-ante stage. After $\theta_S \in \Theta_S$ is elicited from the side contract and the main mechanism reveals $t \in T$ and $\theta_{-S} \in \Theta_{-S}$, the mediator reveals $\psi^S \in \Psi^S$ and redistributes money within S accordingly.

[5] Bose and Renou (2014) have introduced an indirect mechanism to engineer ambiguous beliefs. When ambiguous beliefs are present, Renou (2015) and Lopomo et al. (2020) show that it is no longer generically possible to guarantee FSE. Whether inducing ambiguous beliefs and ambiguous mechanisms simultaneously can further help the MD in achieving FSE (or collusion-proof FSE) remains an open question.

[6] At the end of Section 4, we discuss a less extreme ambiguity aversion model.

Given (q, T) and coalition S, we say $(\tilde{q}, \tilde{T})$ is an **ambiguous *S*-reallocational manipulation**, if there exists an ambiguous S-side contract (δ^S, Ψ^S) such that (i) for each $\tilde{t} \in \tilde{T}$, there exists $t \in T$ and $\psi^S \in \Psi^S$ such that for all $\theta \in \Theta$,

$$\tilde{t}_i(\theta) = \sum_{\theta'_S \in \Theta_S} t_i(\theta'_S, \theta_{-S})\delta^S[\theta_S](\theta'_S) + \psi^S_i(t, \theta) \text{ if } i \in S, \tag{8}$$

and (2), (3), and (4) are satisfied, and (ii) for each $t \in T$ and $\psi^S \in \Psi^S$, the transfer rule $\tilde{t}$ defined by (8) and (2) is an element of $\tilde{T}$. We also denote this ambiguous S-reallocational manipulation by $(\tilde{q}, T^{\delta^S} + \Psi^S)$, from which type-$\theta_i$ agent i's MEU is $V_i[\tilde{q}, T^{\delta^S} + \Psi^S](\theta_i, \sigma_i) \equiv \min_{t \in T, \psi^S \in \Psi^S} \sum_{\hat{\theta}_i \in \Theta_i} V_i[\tilde{q}, t^{\delta^S} + \psi^S(t, \cdot)](\theta_i, \hat{\theta}_i)\sigma_i[\theta_i](\hat{\theta}_i)$, where

$$V_i[\tilde{q}, t^{\delta^S} + \psi^S(t, \cdot)](\theta_i, \hat{\theta}_i) = \begin{cases} V_i[\tilde{q}, t^{\delta^S}](\theta_i, \hat{\theta}_i) + \sum_{\theta_{-i} \in \Theta_{-i}} \psi^S_i(t, \hat{\theta}_i, \theta_{-i}) p(\theta_{-i}|\theta_i), & \text{if } i \in S, \\ V_i[\tilde{q}, t^{\delta^S}](\theta_i, \hat{\theta}_i), & \text{if } i \notin S. \end{cases}$$

By replacing $V_i[\tilde{q}, t^{\delta^S} + \psi^S](\theta_i, \sigma_i)$ in Section 3 with $V_i[\tilde{q}, T^{\delta^S} + \Psi^S](\theta_i, \sigma_i)$, we can redefine the S-feasibility condition for an ambiguous S-reallocational manipulation and the **robust collusion-proofness*** (RCP*) condition for an ambiguous mechanism. Here, we allow the side transfer rule to address additional contingencies compared to the one in Section 3, and thus, RCP* can be viewed as a strengthening of the RCP condition.

4.2 Results

Do ambiguous mechanisms help the MD to guarantee FSE via collusion-proof mechanisms? If yes, to what extent? To answer this question, we first adapt the Beliefs Determine Preferences property of Neeman (2004) and introduce a definition.

Definition 1. 1. *For $S \in 2^I \backslash \{\varnothing, I\}$, the prior p satisfies the S-**Beliefs Determine Preferences** (S-BDP) property, if $p(\cdot|\theta_S) \neq p(\cdot|\theta'_S)$ for each pair of $\theta_S \neq \theta'_S \in \Theta_S$.*

2. *The prior p satisfies the **Beliefs Determine Preferences** (BDP) property, if it satisfies the S-BDP property for all singleton $S \in 2^I$.*

3. *The prior p satisfies the **Coalition Beliefs Determine Preferences** (CBDP) property, if it satisfies the S-BDP property for all $S \in 2^I \backslash \{\varnothing, I\}$.*

The CBDP property implies that knowing the posterior belief of some $S \in 2^I \backslash \{\varnothing, I\}$ over types of agents in $I \backslash S$ can uniquely identify the type profile of S. By definition, the CBDP property implies the BDP property.[7] When agents have independent beliefs, the prior neither satisfies the BDP property nor the CBDP property. However, the BDP and CBDP properties impose a weak restriction on priors over a fixed finite-dimensional Θ: among all priors over Θ, the ones for which the CBDP property fails constitute a set of measure zero.[8]

Proposition 2 shows that the CBDP property characterizes information structures that guarantee collusion-proofness FSE. This is true for both the CIC and the RCP* conditions.

Proposition 2. *Given an information structure* (Θ, p)*, the following statements are equivalent:*

1. *The CBDP property holds for prior* p.
2. *The information structure* (Θ, p) *guarantees FSE via ambiguous mechanisms satisfying the CIC condition.*
3. *The information structure* (Θ, p) *guarantees FSE via ambiguous mechanisms satisfying the RCP* condition.*

A sketch of the proof is provided below and details are relegated to Appendix A.3. We follow the roadmap that Statement 1 $\Leftrightarrow$ Statement 2 and Statement 1 $\Leftrightarrow$ Statement 3.

In the Appendix, Lemma 4 constructs a payoff structure, assumes that (q, T) is an FSE ambiguous mechanism, and then establishes the failure of the CIC or IC condition of (q, T) when the CBDP property fails. This either leads to a contradiction with the feasibility of (q, T) or shows that the CIC condition fails, and thus proves Statement 2 $\Rightarrow$ Statement 1.

[7]To see that the CBDP property can be strictly stronger, consider an example with $I = \{1, 2, 3\}$ and $\Theta_i = \{\theta_i^1, \theta_i^2\}$ for each $i \in I$. We collapse the agent index and denote, for instance, $(\theta_1^1, \theta_2^2, \theta_3^1)$ by θ^{121}. The following prior p satisfies the BDP property, where $p(\theta^{111}) = 0.1$, $p(\theta^{112}) = 0.2$, $p(\theta^{121}) = 0.1$, $p(\theta^{122}) = 0.1$, $p(\theta^{211}) = 0.1$, $p(\theta^{212}) = 0.1$, $p(\theta^{221}) = 0.2$, and $p(\theta^{222}) = 0.1$. However, $p(\cdot|\theta_1^1, \theta_2^2) = p(\cdot|\theta_1^2, \theta_2^1) = (0.5, 0.5)$, and thus, p does not satisfy the CBDP property.

[8]Following Crémer and McLean (1988), Che and Kim (2006), and Kosenok and Severinov (2008), the current paper focuses on a fixed finite type space to study mechanism design with correlated beliefs. Without fixing the dimension of the type space a priori, there is a literature (see, e.g., Heifetz and Neeman, 2006; Chen and Xiong, 2013) discussing how generic the BDP property is.

To establish Statement 3 ⇒ Statement 1, the challenge mainly arises in the case that the BDP property holds but the CBDP property fails. For each FSE ambiguous mechanism in the same payoff structure constructed for Lemma 4, Lemma 6 identifies an ex-ante strictly profitable joint reporting strategy for a coalition. Taking advantage of the BDP property, Lemma 5 constructs an ambiguous side contract to implement the collusion. The ambiguous side contract is contingent on types elicited within the coalition, types reported by noncollusive agents to the MD, and the true mechanism adopted by the MD, but involves monetary transfers within the coalition only. The consequent communicative manipulation hurts either the MD or a noncollusive agent, and thus the FSE ambiguous mechanism cannot be RCP*.

To sketch the proof of Statement 1 ⇒ Statement 2 and Statement 1 ⇒ Statement 3, we first establish the existence of a transfer rule η, under which the MD's ex-post payoff, $u_0\big(q(\theta)\big) - \sum_{i\in I}\eta_i(\theta)$, is constant and equal to FS, and agents' IR constraints bind. This transfer rule aligns the grand coalition's incentive with the nonconstant ex-post social surplus, which establishes the grand coalition's CIC constraints. By the MD's constant ex-post payoff, no I-feasible ambiguous I-reallocational manipulation can affect the MD's payoff, which establishes the RCP* condition with respect to I. However, η neither addresses the IC constraints, nor the CIC or RCP* constraints with respect to any non-grand coalition.

To construct an FSE mechanism that is CIC, we present the following lemma and then adjust η with a group of transfer rules.

Lemma 1. *Given $S \in 2^I\backslash\{\varnothing, I\}$ and type profile $\bar{\theta}_S \in \Theta_S$, if there does not exist $\hat{\theta}_S \in \Theta_S\backslash\{\bar{\theta}_S\}$ such that $p(\cdot|\hat{\theta}_S) = p(\cdot|\bar{\theta}_S)$, then there exists a transfer rule $\phi^{\bar{\theta}_S} : \Theta \to \mathbb{R}^n$ such that*

(i) $\sum_{i\in C}\sum_{\theta_{-C}\in\Theta_{-C}} \phi_i^{\bar{\theta}_S}(\theta_C, \theta_{-C})p(\theta_{-C}|\theta_C) = 0$ *for all* $C \in 2^I\backslash\{\varnothing\}$ *and* $\theta_C \in \Theta_C$*;*

(ii) $\sum_{i\in S}\sum_{\theta_{-S}\in\Theta_{-S}} \phi_i^{\bar{\theta}_S}(\hat{\theta}_S, \theta_{-S})p(\theta_{-S}|\bar{\theta}_S) < 0$ *for all* $\hat{\theta}_S \in \Theta_S\backslash\{\bar{\theta}_S\}$*.*

As coalitions and individuals overlap with each other, (i) above imposes multiple intertwined constraints on the transfer rule $\phi^{\bar{\theta}_S}$ and complicates the problem compared to the case without coalition concerns. As a simplification, we first view the problem as a two-agent one, which contains pseudo agents S and $I\backslash S$ only. In Step 1 of the proof, we apply the transposition theorem of Motzkin (1951) to establish the existence of an ex-post budget balanced transfer rule between the two pseudo agents such that (ii) and (i) with respect to $C = S$ and $C = I\backslash S$ hold. In Step 2, we split the transfer within S and $I\backslash S$ so that the equality in (i) holds for all other C. The division has to be carefully designed rather than a simplistic equal division. In particular, when specifying $\phi_i^{\bar{\theta}_S}$ for some $i \in I$, not only the equation in (i) with respect to $C = \{i\}$ is affected, those with respect to any other coalition containing i are also affected. To tackle this challenge, we apply the alternative theorem of Fredholm (1903) to establish the existence of a division such that every agent $i \in I$ receives zero in expectation conditional on any θ_{-i}. Intuitively, this means that the pseudo agent $I\backslash\{i\}$ always believes that i receives zero in expectation. In Steps 3 to 5, we show that the division satisfies all required conditions by applying law of iterated expectations.

An ambiguous mechanism (q, T), where $T \equiv \{\eta + \lambda\phi^{\bar{\theta}_S} | S \in 2^I\backslash\{\varnothing, I\}, \bar{\theta}_S \in \Theta_S\}$ and $\lambda \in \mathbb{R}_+$ is sufficiently large, is then shown to be feasible and extract the full surplus. To see this, by (i), each $\phi^{\bar{\theta}_S}$ does not affect the MD's ex-post payoff, as setting $C = I$ implies that $\phi^{\bar{\theta}_S}$ is ex-post budget balanced; neither does $\phi^{\bar{\theta}_S}$ affect any (potentially pseudo) agent's on-path interim payoff, as $\phi^{\bar{\theta}_S}$ gives every $C \in 2^I\backslash\{\varnothing, I\}$ zero expected utility on path. However, by (ii) above, for each type-θ_S (potentially pseudo) agent $S \in 2^I\backslash\{\varnothing, I\}$, any unilateral deviation from truthful report (in potentially mixed strategy) earns him a negative expected transfer under ϕ^{θ_S}. When the multiplier λ is sufficiently large, $\eta + \lambda\phi^{\theta_S} \in T$ gives θ_S a negative expected utility, which bounds his MEU of misreporting from above and eventually establishes the IC and CIC conditions.

Although the above ambiguous mechanism (q, T) satisfies CIC, it may not satisfy RCP*. Intuitively, the above construction ensures that by misreporting, a non-grand coalition S earns a low MEU in the interim stage (after knowing θ_S), but S may find it profitable to misreport in the ex-ante stage. If the latter occurs, there might be room for an S-feasible reallocational manipulation. The inconsistency between interim and ex-ante preferences happens because randomization may be used to partially hedge against ambiguity, and ex-ante utility of misreporting as randomization of interim utilities is not necessarily low.[9]

To construct an FSE ambiguous mechanism that is RCP*, we adjust η with a different group of transfer rules. The following lemma is helpful.

[9]See Ke and Zhang (2020) for a study on randomization at different stages under ambiguity.

Lemma 2. *Fix any $S \in 2^I$ with $2 \leqslant |S| \leqslant n-1$. If the S-BDP property holds, then there exists an ex-post budget balanced transfer rule $\phi^S : \Theta \to \mathbb{R}^n$ such that*

(i) $\sum_{\theta_{-i}\in\Theta_{-i}} \phi_i^S(\theta_i, \theta_{-i})p(\theta_{-i}|\theta_i) = 0$ *for all $i \in I$ and $\theta_i \in \Theta_i$;*

(ii) $\sum_{\hat{\theta}_S\in\Theta_S} \sum_{\bar{\theta}_S\in\Theta_S} \sum_{i\in S} \sum_{\theta_{-S}\in\Theta_{-S}} \phi_i^S(\hat{\theta}_S, \theta_{-S})p(\bar{\theta}_S, \theta_{-S})\delta^S[\bar{\theta}_S](\hat{\theta}_S) < 0$ *for all $\delta^S \neq \bar{\delta}^S$.*

Different from Lemma 1, Lemma 2 only imposes the equality constraints on individual agents, and imposes the inequality constraint at the ex-ante stage. Also, Lemma 2 does not hold for $|S| = 1$, which is different from Lemma 1.

When the CBDP property holds, define $T = \{\eta + \lambda_1\phi^{\tilde{\theta}_i}|i \in I, \tilde{\theta}_i \in \Theta_i\} \cup \{\eta + \lambda_2\phi^S|S \in 2^I\backslash\{\varnothing, I\}$ with $2 \leqslant |S| \leqslant n-1\}$, where λ_1 and λ_2 are two large numbers. The set $\{\eta+\lambda_1\phi^{\tilde{\theta}_i}|i \in I, \tilde{\theta}_i \in \Theta_i\}$ makes sure that (q, T) is a feasible FSE ambiguous mechanism by the same argument as before. Suppose a mediator secretly approaches coalition S with $2 \leqslant |S| \leqslant n-1$ and proposes an ambiguous S-side contract that involves misreporting. Under transfer rule $t = \eta + \lambda_2\phi^S \in T$, the total utility of coalition S in the ex-ante stage is negative. Thus, no ambiguous S-side contract involving misreporting can lead to an S-feasible ambiguous S-reallocational manipulation. Hence, (q, T) is RCP* with respect to S.

We summarize two features of our CIC or RCP* ambiguous mechanism. First, the MD earns a constant ex-post payoff from all potential transfer rules in an ambiguous mechanism, and thus each of our ambiguous mechanisms is a full-insurance mechanism in the ex-post sense from the MD's perspective. Second, ambiguity manifests an agent/non-grand coalition differently on and off the equilibrium path. No ambiguity is perceived on path because every potential transfer rule gives him the same interim or ex-ante payoff. However, off path, his interim (in the CIC ambiguous mechanism) or ex-ante (in the RCP* ambiguous mechanism) utility levels under different transfer rules can be unequal and at least one transfer rule leads to a low interim or ex-ante utility level. The fear of encountering the worst-case mechanism discourages misreporting of the ambiguity-averse agent/non-grand coalition.

4.3 Discussions

We now discuss the connection of our Proposition 2 with the literature in detail.

If we view the results of Crémer and McLean (1985, 1988) as a paradox, our Proposition 1 shows that collusion can be one resolution of the paradox, but Proposition 2 shows that the use of ambiguous mechanisms can restore the paradox. In particular, in the two-agent setup, Proposition 2 can soften the theoretically unpermissive result of Laffont and Martimort (2000) on collusion-proof FSE. Notice that with only two agents, the CBDP property is equivalent to the BDP property, under which we can guarantee collusion-proof FSE. Our Proposition 2 is also related to Theorem 2 and Corollary 2 of Che and Kim (2006), where it is shown that if the prior satisfies Convex Independence and their Condition PI', then the corresponding information structure can guarantee FSE via standard Bayesian mechanisms that are RCP with respect to I. Our CBDP property is neither stronger nor weaker than their sufficient conditions. However, recall that our RCP* condition is stronger than RCP with respect to I, partly because we require the mechanism to be immune from all coalitions' manipulations, and partly because we allow side contracts to address more contingencies. In fact, when aiming to design an FSE ambiguous mechanism that is RCP with respect to I only, we can modify the current proofs and show that the BDP property, which is weaker than Convex Independence, is sufficient. Namely, ambiguous mechanisms are more potent in achieving collusion-proof FSE in the sense of Che and Kim (2006), even when a coalition can use ambiguous side contracts to combat the ambiguity in the main mechanism.

In an independent private-value environment, Safronov (2018) redefines the efficient, IC, and budget balanced expected externality mechanism (see, e.g., d'Aspremont and Gérard-Varet, 1979) into one that also satisfies the CIC condition. When the CBDP property holds, we can implement any efficient allocation rule q under any payoff structure with IC, CIC, and budget balanced ambiguous mechanisms.[10] To get an idea, by defining $T \equiv \{\lambda \phi^{\bar{\theta}_S} | S \in 2^I \backslash \{\varnothing, I\}, \bar{\theta}_S \in \Theta_S\}$ where λ is sufficiently large, one can show that (q, T) is IC, CIC, and budget balanced. Hence, our approach can be used to complement Safronov (2018) by going beyond independent private valuation environments. When q generates a nonnegative ex-ante surplus to agents, we can adjust this ambiguous mechanism so that it satisfies the additional IR condition. Thus, our approach can also be used to complement Kosenok and

[10] One can replace the efficient allocation rule by any allocation rule q for which $\sum_{i \in I} u_i\big(q(\theta), \theta\big) \geqslant \sum_{i \in I} u_i\big(q(\theta'), \theta\big)$ for all $\theta, \theta' \in \Theta$. The restriction is used to establish the grand coalition's CIC constraints.

Severinov (2008) by fulfilling the CIC requirement on top of their feasibility and budget balance requirements.

The current paper also extends the approach of Guo (2019), where it is shown that FSE can be guaranteed under ambiguous mechanisms if and only if the prior satisfies the BDP property. The leading distinction is that the current paper focuses on collusion concerns, which bring new challenges in splitting a pseudo agent's transfer to individuals' in Lemma 1, preventing an ex-ante profitable deviation in Lemma 2, and designing the ambiguous side contract in Lemma 5, etc. In addition, we remark that the earlier paper, as well as many works on mechanism design under ambiguity, only focuses on pure strategy deviations, which may be with loss of generality because mixed strategies sometimes can be played to hedge against ambiguity. However, the current paper explicitly addresses misreporting in mixed strategies and thus the concern of hedging does not apply.

For simplicity, the paper has been assuming that all agents, including pseudo ones, adopt the MEU model. This assumption is a bit extreme, but there are less extreme models of decision making under ambiguity in the literature under which our results at least hold partially. As a brief illustration, we adjust the CIC ambiguous mechanism constructed for Proposition 2 into a "symmetric" one so that it works under α-MEU of Ghirardato and Marinacci (2002), where $\alpha > 0.5$ and $1 - \alpha$ are the weights an (potentially pseudo) agent assigns to the worst-case and best-case expected utility. In particular, we can do so by defining $T \equiv \{\eta + \lambda\phi^{\theta_C} | C \in 2^I \backslash \{\varnothing, I\}, \theta_C \in \Theta_C\} \cup \{\eta - \lambda\phi^{\theta_C} | C \in 2^I \backslash \{\varnothing, I\}, \theta_C \in \Theta_C\}$, where λ is sufficiently large so that for all $C \in 2^I \backslash \{\varnothing, I\}$ and $\theta_C \in \Theta_C$, type-θ_C pseudo agent earns a negative expected utility under $\eta + \lambda(\alpha - (1 - \alpha))\phi^{\theta_C}$ by misreporting. It is worth noting that for the ambiguity neutral case ($\alpha = 0.5$), ambiguous mechanisms are perceived as standard Bayesian mechanisms, and are not helpful to achieve collusion-proof FSE.

5 Concluding Remarks

We end the paper with two open questions.

First, there are alternative collusion-proofness notions of interest. For example, one may consider a variant of the RCP condition where the mediator can only coordinate joint deviations with side contracts that are immune from further deviations of coalitions. This requirement restricts the class of side contracts that a mediator can use and hence weakens the RCP condition. Whether such an alternative collusion-proofness notion can soften the theoretically unpermissive result in Poposition 1 or relax the necessity of the CBDP property in Proposition 2 remains an open question.

Second, as in Crémer and McLean (1985, 1988), Che and Kim (2006), and Kosenok and Severinov (2008), we focus on properties of the information structure such that collusion-proof FSE can be guaranteed. To establish Proposition 1 and the necessity direction of Proposition 2, we adopt particular payoff structures for which FSE cannot be achieved. This approach does not exclude the possibility that there are payoff structures for which collusion-proof FSE can be achieved. Identifying those payoff structures remains an open question.

References

Bierbrauer, F. J. and Hellwig, M. F. (2016). Robustly coalition-proof incentive mechanisms for public good provision are voting mechanisms and vice versa. *The Review of Economic Studies*, 83(4):1440–1464.

Börgers, T., Krähmer, D., and Strausz, R. (2015). *An introduction to the theory of mechanism design*. Oxford University Press, USA.

Bose, S. and Daripa, A. (2017). Eliciting ambiguous beliefs under alpha-maxmin preference. Working Paper.

Bose, S., Ozdenoren, E., and Pape, A. (2006). Optimal auctions with ambiguity. *Theoretical Economics*, 1(4):411–438.

Bose, S. and Renou, L. (2014). Mechanism design with ambiguous communication devices. *Econometrica*, 82(5):1853–1872.

Che, Y.-K. and Kim, J. (2006). Robustly collusion-proof implementation. *Econometrica*, 74(4):1063–1107.

Chen, J. and Micali, S. (2012). Collusive dominant-strategy truthfulness. *Journal of Economic Theory*, 147(3):1300–1312.

Chen, Y.-C. and Xiong, S. (2013). Genericity and robustness of full surplus extraction. *Econometrica*, 81(2):825–847.

Crémer, J. and McLean, R. P. (1985). Optimal selling strategies under uncertainty for a discriminating monopolist when demands are interdependent. *Econometrica*, 53(2):345–361.

Crémer, J. and McLean, R. P. (1988). Full extraction of the surplus in Bayesian and dominant strategy auctions. *Econometrica*, 56(6):1247–1257.

d'Aspremont, C., Cremer, J., and Gérard-Varet, L.-A. (1990). Incentives and the existence of Pareto-optimal revelation mechanisms. *Journal of Economic Theory*, 51(2):233–254.

d'Aspremont, C., Crémer, J., and Gérard-Varet, L.-A. (2004). Balanced Bayesian mechanisms. *Journal of Economic Theory*, 115(2):385–396.

d'Aspremont, C. and Gérard-Varet, L.-A. (1979). Incentives and incomplete information. *Journal of Public Economics*, 11(1):25–45.

De Castro, L. and Yannelis, N. C. (2018). Uncertainty, efficiency and incentive compatibility: Ambiguity solves the conflict between efficiency and incentive compatibility. *Journal of Economic Theory*, 177:678–707.

Di Tillio, A., Kos, N., and Messner, M. (2017). The design of ambiguous mechanisms. *The Review of Economic Studies*, 84(1):237–274.

Fredholm, I. (1903). Sur une classe d'équations fonctionnelles. *Acta Mathematica*, 27:365–390.

Ghirardato, P. and Marinacci, M. (2002). Ambiguity made precise: A comparative foundation. *Journal of Economic Theory*, 102(2):251–289.

Gilboa, I. and Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18(2):141–153.

Green, J. and Laffont, J.-J. (1979). On coalition incentive compatibility. *The Review of Economic Studies*, 46(2):243–254.

Guo, H. (2019). Mechanism design with ambiguous transfers: An analysis in finite dimensional naive type spaces. *Journal of Economic Theory*, 183:76–105.

Heifetz, A. and Neeman, Z. (2006). On the generic (im) possibility of full surplus extraction in mechanism design. *Econometrica*, 74(1):213–233.

Ke, S. and Zhang, Q. (2020). Randomization and ambiguity aversion. *Econometrica*, 88(3):1159–1195.

Kocherlakota, N. R. and Song, Y. (2019). Public goods with ambiguity in large economies. *Journal of Economic Theory*, 182:218–246.

Kosenok, G. and Severinov, S. (2008). Individually rational, budget-balanced mechanisms and allocation of surplus. *Journal of Economic Theory*, 140(1):126–161.

Laffont, J.-J. and Martimort, D. (1997). Collusion under asymmetric information. *Econometrica*, 65(4):875–911.

Laffont, J.-J. and Martimort, D. (2000). Mechanism design with collusion and correlation. *Econometrica*, 68(2):309–342.

Lopomo, G., Rigotti, L., and Shannon, C. (2020). Uncertainty and robustness of surplus extraction. *Journal of Economic Theory*, page 105088.

Lopomo, G., Rigotti, L., and Shannon, C. (2022). Detectability, duality, and surplus extraction. *Journal of Economic Theory*, page 105465.

McAfee, R. P. and Reny, P. J. (1992). Correlated information and mechanism design. *Econometrica*, 60(2):395–421.

McLean, R. and Postlewaite, A. (2004). Informational size and efficient auctions. *The Review of Economic Studies*, 71(3):809–827.

McLean, R. P. and Postlewaite, A. (2015). Implementation with interdependent valuations. *Theoretical Economics*, 10(3):923–952.

Meng, D., Tian, G., and Yang, Z. (2017). Two-agent collusion-proof implementation with correlation and arbitrage. *Review of Economic Design*, 21(3):177–229.

Motzkin, T. S. (1951). Two consequences of the transposition theorem on linear inequalities. *Econometrica*, 19(2):184.

Neeman, Z. (2004). The relevance of private information in mechanism design. *Journal of Economic Theory*, 117(1):55–77.

Renou, L. (2015). Rent extraction and uncertainty. *Working Paper*.

Safronov, M. (2018). Coalition-proof full efficient implementation. *Journal of Economic Theory*, 177:659–677.

Song, Y. (2018). Efficient implementation with interdependent valuations and maxmin agents. *Journal of Economic Theory*, 176:693–726.

Song, Y. (2022). Approximate bayesian implementation and exact maxmin implementation: An equivalence. Working Paper.

Tang, R. and Zhang, M. (2021). Maxmin implementation. *Journal of Economic Theory*, 194:105250.

Wolitzky, A. (2016). Mechanism design with maxmin agents: theory and an application to bilateral trade. *Theoretical Economics*, 11:971–1004.

A Appendix

A.1 Preparatory Notations, Definitions, and Results

This paper establishes the existence of a transfer rule satisfying certain constraints by applying (a corollary of) the transposition theorem of Motzkin (1951) or the alternative theorem of Fredholm (1903). We present them and establish a preparatory lemma for constructing an ambiguous mechanism and an ambiguous side contract in Lemmas 1 and 5.

Theorem 1 (Motzkin, 1951)**.** *Let $B \in \mathbb{R}^{m\times l}$ and $D \in \mathbb{R}^{k\times l}$ be matrices. Exactly one of the following holds: either the system $Bx < \boldsymbol{0}_{m\times 1}$, $Dx = \boldsymbol{0}_{k\times 1}$ has a column vector solution $x \in \mathbb{R}^l$, or there exist column vectors $y_1 \in \mathbb{R}^m_+\backslash\{\boldsymbol{0}\}$ and $y_2 \in \mathbb{R}^k$ such that $B'y_1 + D'y_2 = \boldsymbol{0}_{l\times 1}$.*

We remark that $Bx < \mathbf{0}_{m\times 1}$ means that all m strict inequalities must hold.

Theorem 2 (Fredholm, 1903)**.** *Let $B \in \mathbb{R}^{m\times l}$ be a matrix and b be a column vector in $\mathbb{R}^m$. Exactly one of the following holds: either the system $Bx = b$ has a column vector solution $x \in \mathbb{R}^l$, or $B'y = \boldsymbol{0}_{l\times 1}$ has a column vector solution $y \in \mathbb{R}^m$ with $y'b \neq 0$.*

To apply these theorems, it is important to construct matrices B and D. As a preparation, we fix any order of the elements in Θ and define some row vectors in $\mathbb{R}^{n|\Theta|}$. For each $x \in \mathbb{R}^{n|\Theta|}$, divide its dimensions into n blocks of $|\Theta|$ dimensions. Let the first block of $|\Theta|$ dimensions corresponds to agent 1, ..., and the last block corresponds to agent n. Within each block, the dimensions correspond to elements of Θ. Hence, each dimension of $x \in \mathbb{R}^{n|\Theta|}$ corresponds to an agent and a type profile.

For each $S \in 2^I \backslash \{\varnothing\}$, $C \in 2^I \backslash \{\varnothing, I\}$ and type profiles $\theta_C, \theta'_C \in \Theta_C$ (may be identical), we define a row vector $p^S_{\theta_C \theta'_C} \in \mathbb{R}^{n|\Theta|}_+ \backslash \{\mathbf{0}\}$ as follows. For each $i \in S$ and $\theta_{-C} \in \Theta_{-C}$, let the dimension of $p^S_{\theta_C \theta'_C}$ corresponding to agent i and type profile (θ'_C, θ_{-C}) be equal to $p(\theta_C, \theta_{-C})$, where p is the prior. Thus, we have defined $|S||\Theta_{-C}|$ dimensions of $p^S_{\theta_C \theta'_C}$. Let all other dimensions of $p^S_{\theta_C \theta'_C}$ be 0.

For each $\theta \in \Theta$ and $S \in 2^I \backslash \{\varnothing\}$, define a row vector $e^S_\theta \in \mathbb{R}^{n|\Theta|}_+ \backslash \{\mathbf{0}\}$ as follows. For each $i \in S$, let the dimension of e^S_θ corresponding to i and θ be equal to 1. Thus, we have defined $|S|$ dimensions of e^S_θ. Let all other dimensions of e^S_θ be 0.

For example, let $I = \{1, 2, 3\}$ and $\Theta_i = \{\theta^1_i, \theta^2_i\}$ for each $i \in I$. We order the eight elements of Θ by: θ^{111}, θ^{112}, θ^{121}, θ^{122}, θ^{211}, θ^{212}, θ^{221}, θ^{222}, where for instance, $\theta^{121} \equiv (\theta^1_1, \theta^2_2, \theta^1_3)$. For each vector in $\mathbb{R}^{24}$, its first, second, and third blocks of eight dimensions correspond to agents 1, 2, and 3, respectively. Let $\mathbf{0}_{1\times k}$ denote a zero row vector in $\mathbb{R}^k$. We illustrate with two vectors below and use a box to group every block of eight dimensions:

$$p^{\{3\}}_{(\theta^1_1,\theta^1_2)(\theta^1_1,\theta^2_2)} = (\boxed{\mathbf{0}_{1\times 8}}, \boxed{\mathbf{0}_{1\times 8}}, \boxed{\mathbf{0}_{1\times 2}, p(\theta^{111}), p(\theta^{112}), \mathbf{0}_{1\times 4}});$$
$$e^{\{1,2\}}_{\theta^{112}} = (\boxed{0, 1, \mathbf{0}_{1\times 6}}, \boxed{0, 1, \mathbf{0}_{1\times 6}}, \boxed{\mathbf{0}_{1\times 8}}).$$

Lemma 3 below provides a unified approach to establish two technical observations in the proof of Lemmas 1 and 5. It establishes the existence of a budget balanced monetary transfer within a group of agents $S^1 \cup ... \cup S^K$, where the group may be I or a proper subset of I. Each S^k may represent an agent (when $|S^k| = 1$) or a coalition (when $|S^k| > 1$), and these K sets are nonempty and disjoint. The monetary transfer is contingent on all agents' reported types, rather than those in the group $S^1 \cup ... \cup S^K$ only. The transfer rule gives each (potentially pseudo) agent S^k zero on-path interim transfer, but gives type-$\bar{\theta}_{S^1}$ (potentially pseudo) agent S^1 a negative interim transfer when he unilaterally misreports.

Lemma 3. *Let S^1, ..., S^K be $K \geqslant 2$ nonempty disjoint subsets of I and $\bar{\theta}_{S^1}$ be an element of Θ_{S^1}. If there does not exist $\hat{\theta}_{S^1} \in \Theta_{S^1} \backslash \{\bar{\theta}_{S^1}\}$ such that $p(\cdot|\hat{\theta}_{S^1}) = p(\cdot|\bar{\theta}_{S^1})$, then there exists a transfer rule $\xi^{\bar{\theta}_{S^1}} \equiv (\xi^{\bar{\theta}_{S^1}}_i : \Theta \to \mathbb{R})_{i \in S^1 \cup ... \cup S^K}$ such that*

(i) $\sum_{i \in S^1 \cup ... \cup S^K} \xi^{\bar{\theta}_{S^1}}_i(\theta) = 0$ *for all* $\theta \in \Theta$,

(ii) $\sum_{\theta_{-S^k} \in \Theta_{-S^k}} \sum_{i \in S^k} \xi^{\bar{\theta}_{S^1}}_i(\theta_{S^k}, \theta_{-S^k}) p(\theta_{-S^k}|\theta_{S^k}) = 0$ *for all* $k \in \{1, ..., K\}$ *and* $\theta_{S^k} \in \Theta_{S^k}$,

(iii) $\sum_{\theta_{-S^1}\in\Theta_{-S^1}}\sum_{i\in S^1}\xi_i^{\bar{\theta}_{S^1}}(\hat{\theta}_{S^1},\theta_{-S^1})p(\theta_{-S^1}|\bar{\theta}_{S^1})<0$ *for all* $\hat{\theta}_{S^1}\in\Theta_{S^1}\backslash\{\bar{\theta}_{S^1}\}$.

Proof. With the vectors defined in this section, we construct matrices B and D of dimensions $m\times l$ and $k\times l$, respectively, where $m=|\Theta_{S^1}|-1$, $k=\sum_{k=1,...,K}|\Theta_{S^k}|+|\Theta|$, and $l=n|\Theta|$. Matrix B is obtained by vertically stacking up m row vectors $p^{S^1}_{\bar{\theta}_{S^1}\hat{\theta}_{S^1}}\in\mathbb{R}^l_+$ for all $\hat{\theta}_{S^1}\in\Theta_{S^1}\backslash\{\bar{\theta}_{S^1}\}$. Construct matrix D by stacking up $\sum_{k=1,...,K}|\Theta_{S^k}|$ row vectors $p^{S^k}_{\theta_{S^k}\theta_{S^k}}\in\mathbb{R}^l_+$ for all $k\in\{1,...,K\}$ and $\theta_{S^k}\in\Theta_{S^k}$ as well as $|\Theta|$ row vectors $e^{S^1\cup...\cup S^K}_{\theta}\in\mathbb{R}^l_+$ for all $\theta\in\Theta$.

Suppose by way of contradiction that there is no transfer rule $\xi^{\bar{\theta}_{S^1}}$ satisfying the three conditions stated in the lemma. Notice that within each row of B and D, the dimensions that correspond to an agent out of $S^1\cup...\cup S^K$, if any, are equal to zero. Then we can claim that $Bx<\mathbf{0}_{m\times1}$, $Dx=\mathbf{0}_{k\times1}$ has no column vector solution $x\in\mathbb{R}^l$. By Theorem 1, there are column vectors $y_1\in\mathbb{R}^m_+\backslash\{\mathbf{0}\}$ and $y_2\in\mathbb{R}^k$, such that $B'y_1+D'y_2=\mathbf{0}_{l\times1}$, or equivalently $-y_2'D=y_1'B$ where both sides are row vectors in $\mathbb{R}^l$. As a result, there exists a profile of nonnegative numbers $(c_{\hat{\theta}_{S^1}}\in\mathbb{R}_+)_{\hat{\theta}_{S^1}\in\Theta_{S^1}\backslash\{\bar{\theta}_{S^1}\}}$ with $c_{\hat{\theta}_{S^1}}\in\mathbb{R}_{++}$ for some $\hat{\theta}_{S^1}\in\Theta_{S^1}\backslash\{\bar{\theta}_{S^1}\}$ and two profiles of numbers $(a_{\theta_{S^k}}\in\mathbb{R})_{\theta_{S^k}\in\Theta_{S^k},k\in\{1,...,K\}}$ and $(b_\theta\in\mathbb{R})_{\theta\in\Theta}$, such that

$$\sum_{k\in\{1,...,K\}}\sum_{\theta_{S^k}\in\Theta_{S^k}}a_{\theta_{S^k}}p^{S^k}_{\theta_{S^k}\theta_{S^k}}+\sum_{\theta\in\Theta}b_\theta e^{S^1\cup...\cup S^K}_\theta=\sum_{\hat{\theta}_{S^1}\in\Theta_{S^1}\backslash\{\bar{\theta}_{S^1}\}}c_{\hat{\theta}_{S^1}}p^{S^1}_{\bar{\theta}_{S^1}\hat{\theta}_{S^1}},\tag{9}$$

where both sides are row vectors in $\mathbb{R}^l$.

Fix any $\hat{\theta}_{S^1}$ with $c_{\hat{\theta}_{S^1}}\neq0$, $k\in\{2,...,K\}$, $\theta_{-S^1}\in\Theta_{-S^1}$, agent $i\in S^1$, and agent $j\in S^k$.

Recall that on each side of expression (9), each dimension in the row vector corresponds to an agent and a type profile. From the dimensions corresponding to i and $(\bar{\theta}_{S^1},\theta_{-S^1})$ on both sides of (9), we have $a_{\bar{\theta}_{S^1}}p(\bar{\theta}_{S^1},\theta_{-S^1})+b_{(\bar{\theta}_{S^1},\theta_{-S^1})}=0$; from the dimensions corresponding to j and $(\bar{\theta}_{S^1},\theta_{-S^1})$, we have $a_{\theta_{S^k}}p(\bar{\theta}_{S^1},\theta_{-S^1})+b_{(\bar{\theta}_{S^1},\theta_{-S^1})}=0$. As p has full support, $a_{\bar{\theta}_{S^1}}=a_{\theta_{S^k}}$.

Similarly, by focusing on the dimensions corresponding to i and $(\hat{\theta}_{S^1},\theta_{-S^1})$ and corresponding to j and $(\hat{\theta}_{S^1},\theta_{-S^1})$ on both sides of expression (9), we have $a_{\hat{\theta}_{S^1}}p(\hat{\theta}_{S^1},\theta_{-S^1})+b_{(\hat{\theta}_{S^1},\theta_{-S^1})}=c_{\hat{\theta}_{S^1}}p(\bar{\theta}_{S^1},\theta_{-S^1})$ and $a_{\theta_{S^k}}p(\hat{\theta}_{S^1},\theta_{-S^1})+b_{(\hat{\theta}_{S^1},\theta_{-S^1})}=0$. By the observation from the previous paragraph, we have $(a_{\hat{\theta}_{S^1}}-a_{\bar{\theta}_{S^1}})p(\hat{\theta}_{S^1},\theta_{-S^1})=c_{\hat{\theta}_{S^1}}p(\bar{\theta}_{S^1},\theta_{-S^1})$.

As $c_{\hat{\theta}_{S^1}}\neq0$ and the above equation holds for all $\theta_{-S^1}\in\Theta_{-S^1}$, it must be the case that $p(\cdot|\hat{\theta}_{S^1})=p(\cdot|\bar{\theta}_{S^1})$, a contradiction with the supposition of the lemma. □

At last, we review two conditions on the prior introduced by Crémer and McLean (1988) and Kosenok and Severinov (2008) to guarantee the existence of first-best mechanisms.

Definition 2. *The prior p is said to satisfy the* ***Convex Independence*** *condition if for all $i \in I$ and $\theta_i \in \Theta_i$, $p(\cdot|\theta_i) \notin \text{con}\{p(\cdot|\hat{\theta}_i) : \hat{\theta}_i \in \Theta_i \backslash \{\theta_i\}\}$.*

Definition 3. *The prior p is said to satisfy the* ***Identifiability*** *condition if for any full-support $p' \in \Delta(\Theta)$ with $p' \neq p$, there exists $i \in I$ and $\theta_i \in \Theta_i$, for which $p'(\cdot|\theta_i) \notin \text{con}\{p(\cdot|\hat{\theta}_i) : \hat{\theta}_i \in \Theta_i\}$.*

A.2 Omitted Details in Section 3

Proof of Claim 1. Suppose by way of contradiction that there exists an FSE mechanism (q, t) satisfying Properties (i) to (iv) below.

Property (i), binding IR of agent n. Namely, for all $\theta_n \in \Theta_n$,

$$\text{IR}(\theta_n) : \sum_{\theta_{-n} \in \Theta_{-n}} t_n(\theta_{-n}, \theta_n) p(\theta_{-n}|\theta_n) = -\theta_n.$$

Property (ii), IC of agent n. Since q is constant, it is required that for all θ_n, $\theta'_n \in \Theta_n$,

$$\text{IC}(\theta_n; \theta'_n) : \sum_{\theta_{-n} \in \Theta_{-n}} t_n(\theta_{-n}, \theta'_n) p(\theta_{-n}|\theta_n) - \sum_{\theta_{-n} \in \Theta_{-n}} t_n(\theta_{-n}, \theta_n) p(\theta_{-n}|\theta_n) \leqslant 0.$$

Property (iii), ex-post constant revenue. By CIC constraints of I and the fact that the efficient allocation rule q is constant, $\sum_{i \in I} t_i(\theta)$ must be constant in $\theta \in \Theta$. Moreover, since u_0 is equal to zero and (q, t) is an FSE mechanism, $\sum_{i \in I} t_i(\theta) = -FS$ for all $\theta \in \Theta$.

Property (iv), CIC of coalition $I \backslash \{n\}$ (or IC if $|n| = 2$). Namely, for all $\theta_{-n}, \theta'_{-n} \in \Theta_{-n}$,

$$\text{CIC}(\theta_{-n}; \theta'_{-n}) : \sum_{\theta_n \in \Theta_n} t_n(\theta_{-n}, \theta_n) p(\theta_n|\theta_{-n}) - \sum_{\theta_n \in \Theta_n} t_n(\theta'_{-n}, \theta_n) p(\theta_n|\theta_{-n}) \leqslant 0,$$

which follows from $\sum_{j \in I \backslash \{n\}} \sum_{\theta_n \in \Theta_n} t_j(\theta'_{-n}, \theta_n) p(\theta_n|\theta_{-n}) - \sum_{j \in I \backslash \{n\}} \sum_{\theta_n \in \Theta_n} t_j(\theta_{-n}, \theta_n) p(\theta_n|\theta_{-n}) \leqslant 0$ and Property (iii).

Recall the definition of row vectors $p^{\{n\}}_{\theta_{-n}\theta'_{-n}}$ and $p^{\{n\}}_{\theta_n\theta'_n}$ in Section A.1, only their last $|\Theta|$ columns, i.e., those corresponding to agent n, may be nonzero. Let $\bar{p}^{\{n\}}_{\theta_{-n}\theta'_{-n}}$ and $\bar{p}^{\{n\}}_{\theta_n\theta'_n}$ denote the truncated vectors that only keep the last $|\Theta|$ dimensions of $p^{\{n\}}_{\theta_{-n}\theta'_{-n}}$ and $p^{\{n\}}_{\theta_n\theta'_n}$, respectively. Each dimension in these vectors corresponds to a type profile in Θ. Define a set:

$$L \equiv \left\{ \bar{p}^{\{n\}}_{\theta_{-n}\theta_{-n}} - \bar{p}^{\{n\}}_{\theta_{-n}\theta'_{-n}} \;\middle|\; \begin{array}{l} \theta_{-n} \neq \theta'_{-n}, \\ \theta_{-n}, \theta'_{-n} \in \Theta_{-n} \end{array} \right\} \cup \{\bar{p}^{\{n\}}_{\theta^2_n\theta^2_n}\} \cup \left\{ \bar{p}^{\{n\}}_{\theta_n\theta_n} - \bar{p}^{\{n\}}_{\theta_n\theta'_n} \;\middle|\; \begin{array}{l} \theta_n \neq \theta'_n, \\ \theta_n, \theta'_n \in \Theta_n \end{array} \right\}. \quad (10)$$

We establish the following claim after the current proof.

Claim 5. $\text{span}(L) = \mathbb{R}^{|\Theta|}$.

By Claim 5 , as $\bar{p}^{\{n\}}_{\theta^1_n\theta^1_n} \in \mathbb{R}^{|\Theta|}$, there exists a profile of numbers $(b_{\theta_{-n}\theta'_{-n}} \in \mathbb{R})_{\theta_{-n},\theta'_{-n}\in\Theta_{-n},\theta_{-n}\neq\theta'_{-n}}$, $a_{\theta^2_n} \in \mathbb{R}$, and a profile $(b_{\theta_n\theta'_n} \in \mathbb{R})_{\theta_n,\theta'_n\in\Theta_n,\theta_n\neq\theta'_n}$ such that

$$-\bar{p}^{\{n\}}_{\theta^1_n\theta^1_n} = \sum_{\substack{\theta_{-n}\neq\theta'_{-n},\\ \theta_{-n},\theta'_{-n}\in\Theta_{-n}}} b_{\theta_{-n}\theta'_{-n}}[\bar{p}^{\{n\}}_{\theta_{-n}\theta_{-n}} - \bar{p}^{\{n\}}_{\theta_{-n}\theta'_{-n}}] + a_{\theta^2_n}\bar{p}^{\{n\}}_{\theta^2_n\theta^2_n} + \sum_{\substack{\theta_n\neq\theta'_n,\\ \theta_n,\theta'_n\in\Theta_n}} b_{\theta_n\theta'_n}[-\bar{p}^{\{n\}}_{\theta_n\theta_n} + \bar{p}^{\{n\}}_{\theta_n\theta'_n}]. \quad (11)$$

Also, by definition of these vectors, it is easy to prove that

$$\mathbf{0}_{1\times|\Theta|} = \sum_{\substack{\theta_{-n}\neq\theta'_{-n},\\ \theta_{-n},\theta'_{-n}\in\Theta_{-n}}} p(\theta'_{-n})[\bar{p}^{\{n\}}_{\theta_{-n}\theta_{-n}} - \bar{p}^{\{n\}}_{\theta_{-n}\theta'_{-n}}] + \sum_{\substack{\theta_n\neq\theta'_n,\\ \theta_n,\theta'_n\in\Theta_n}} p(\theta'_n)[-\bar{p}^{\{n\}}_{\theta_n\theta_n} + \bar{p}^{\{n\}}_{\theta_n\theta'_n}], \quad (12)$$

since in the right-hand-side vector, the dimension corresponding to $\theta \equiv (\theta_{-n}, \theta_n)$ is equal to

$$\sum_{\theta'_{-n}\in\Theta_{-n}\setminus\{\theta_{-n}\}} p(\theta'_{-n})p(\theta) - \sum_{\theta'_{-n}\in\Theta_{-n}\setminus\{\theta_{-n}\}} p(\theta_{-n})p(\theta'_{-n}, \theta_n) - \sum_{\theta'_n\in\Theta_n\setminus\{\theta_n\}} p(\theta'_n)p(\theta) + \sum_{\theta'_n\in\Theta_n\setminus\{\theta_n\}} p(\theta_n)p(\theta_{-n}, \theta'_n)$$
$$= [1 - p(\theta_{-n})]p(\theta) - p(\theta_{-n})[p(\theta_n) - p(\theta)] - [1 - p(\theta_n)]p(\theta) + p(\theta_n)[p(\theta_{-n}) - p(\theta)] = 0.$$

Fix $\lambda \in \mathbb{R}_+$ sufficiently large such that $\tilde{b}_{\theta_{-n}\theta'_{-n}} \equiv b_{\theta_{-n}\theta'_{-n}} + \lambda p(\theta'_{-n}) \in \mathbb{R}_+$, $\tilde{b}_{\theta_n\theta'_n} \equiv b_{\theta_n\theta'_n} + \lambda p(\theta'_n) \in \mathbb{R}_+$, for all $\theta_{-n}, \theta'_{-n} \in \Theta_{-n}$ with $\theta_{-n} \neq \theta'_{-n}$, and $\theta_n, \theta'_n \in \Theta_n$ with $\theta_n \neq \theta'_n$. When multiplying both sides of expression (12) by λ and adding up with (11), we have

$$-\bar{p}^{\{n\}}_{\theta^1_n\theta^1_n} = \sum_{\substack{\theta_{-n}\neq\theta'_{-n},\\ \theta_{-n},\theta'_{-n}\in\Theta_{-n}}} \tilde{b}_{\theta_{-n}\theta'_{-n}}[\bar{p}^{\{n\}}_{\theta_{-n}\theta_{-n}} - \bar{p}^{\{n\}}_{\theta_{-n}\theta'_{-n}}] + a_{\theta^2_n}\bar{p}^{\{n\}}_{\theta^2_n\theta^2_n} + \sum_{\substack{\theta_n\neq\theta'_n,\\ \theta_n,\theta'_n\in\Theta_n}} \tilde{b}_{\theta_n\theta'_n}[-\bar{p}^{\{n\}}_{\theta_n\theta_n} + \bar{p}^{\{n\}}_{\theta_n\theta'_n}]. \quad (13)$$

Scale constraint IR(θ^1_n) by the factor 1, each CIC($\theta_{-n}; \theta'_{-n}$) by $\tilde{b}_{\theta_{-n}\theta'_{-n}} \in \mathbb{R}_+$, IR($\theta^2_n$) by $a_{\theta^2_n}$, and each IC($\theta_n; \theta'_n$) by $\tilde{b}_{\theta_n\theta'_n} \in \mathbb{R}_+$, and then aggregate the scaled (in)equalities. By expression (13), all coefficients of $t_n(\cdot)$ are equal to zero, and thus the left-hand side of the aggregate inequality is zero. The right-hand side of the aggregate inequality is $-\theta_n^1 - a_{\theta_n^2}\theta_n^2$. When we begin with choosing θ_n^1 sufficiently large, we have $0 \leqslant -\theta_n^1 - a_{\theta_n^2}\theta_n^2 < 0$, a contradiction.

Hence, whenever (q,t) is an FSE mechanism (which requires Properties (i) and (ii)), Properties (iii) and (iv) cannot hold simultaneously, which establishes the claim. □

Proof of Claim 5. Reorder the elements in Θ as follows: $(\theta_{-n}^1, \theta_n^1)$, $(\theta_{-n}^1, \theta_n^2)$, ..., $(\theta_{-n}^1, \theta_n^{|\Theta_n|})$, $(\theta_{-n}^2, \theta_n^1)$, $(\theta_{-n}^2, \theta_n^2)$, ..., $(\theta_{-n}^2, \theta_n^{|\Theta_n|})$, ..., $(\theta_{-n}^{|\Theta_{-n}|}, \theta_n^1)$, $(\theta_{-n}^{|\Theta_{-n}|}, \theta_n^2)$, ..., $(\theta_{-n}^{|\Theta_{-n}|}, \theta_n^{|\Theta_n|})$. Recall the definition of $\hat{X}$ in (5). Define the following $|\Theta| \times |\Theta|$ matrix X, where I is the $|\Theta_n| \times |\Theta_n|$ identity matrix:

$$X \equiv \begin{bmatrix} \hat{X} & -\hat{X} & \mathbf{0}_{|\Theta_n|\times|\Theta_n|} & \cdots & \mathbf{0}_{|\Theta_n|\times|\Theta_n|} \\ \hat{X} & \mathbf{0}_{|\Theta_n|\times|\Theta_n|} & -\hat{X} & \cdots & \mathbf{0}_{|\Theta_n|\times|\Theta_n|} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \hat{X} & \mathbf{0}_{|\Theta_n|\times|\Theta_n|} & \mathbf{0}_{|\Theta_n|\times|\Theta_n|} & \cdots & -\hat{X} \\ p(\theta_{-n}^1, \theta_n^2)I & p(\theta_{-n}^2, \theta_n^2)I & p(\theta_{-n}^3, \theta_n^2)I & \cdots & p(\theta_{-n}^{|\Theta_{-n}|}, \theta_n^2)I \end{bmatrix}. \tag{14}$$

Notice that each of the first $|\Theta_n|(|\Theta_{-n}|-1)$ rows is in $\text{span}(\{\bar{p}_{\theta_{-n}\theta_{-n}}^{\{n\}} - \bar{p}_{\theta_{-n}\theta'_{-n}}^{\{n\}} | \theta_{-n}, \theta'_{-n} \in \Theta_{-n}, \theta_{-n} \neq \theta'_{-n}\})$. Each of the last $|\Theta_n|$ rows is in $\text{span}(\{\bar{p}_{\theta_n^2\theta_n^2}^{\{n\}}\} \cup \{\bar{p}_{\theta_n^2\theta_n^2}^{\{n\}} - \bar{p}_{\theta_n^2\theta_n}^{\{n\}} | \theta_n \in \Theta_n \backslash \{\theta_n^2\}\}) \subseteq \text{span}(\{\bar{p}_{\theta_n^2\theta_n^2}^{\{n\}}\} \cup \{\bar{p}_{\theta_n\theta_n}^{\{n\}} - \bar{p}_{\theta_n\theta'_n}^{\{n\}} | \theta_n, \theta'_n \in \Theta_n, \theta_n \neq \theta'_n\})$. By (10), $X \subseteq \text{span}(L)$.

We now show that $X\beta = \mathbf{0}_{|\Theta|\times 1}$, where $\beta \in \mathbb{R}^{|\Theta|}$ is a column vector, only has zero solution.

Let β_1 be the first block of $|\Theta_n|$ elements of β,..., and $\beta_{|\Theta_{-n}|}$ be the $|\Theta_{-n}|$-th block of $|\Theta_n|$ elements. The first $|\Theta_n|$ equations in $X\beta = \mathbf{0}_{|\Theta|\times 1}$ require that $\hat{X}\beta_1 - \hat{X}\beta_2 = \hat{X}(\beta_1 - \beta_2) = \mathbf{0}_{|\Theta_n|\times 1}$. The full rank of $\hat{X}$ implies $\beta_1 = \beta_2$. Repeating this argument, one can see that the first $|\Theta_n|(|\Theta_{-n}|-1)$ equations in $X\beta = \mathbf{0}_{|\Theta|\times 1}$ require that $\beta_1 = \beta_2 = ... = \beta_{|\Theta_{-n}|}$. The last $|\Theta_n|$ equations in $X\beta = \mathbf{0}_{|\Theta|\times 1}$ require that $p(\theta_{-n}^1, \theta_n^2)I\beta_1 + ... + p(\theta_{-n}^{|\Theta_{-n}|}, \theta_n^2)I\beta_{|\Theta_{-n}|} = p(\theta_n^2)\beta_1 = \mathbf{0}_{|\Theta_n|\times 1}$. Thus, β_1 is a zero vector, and so is β.

We have shown that $X\beta = \mathbf{0}_{|\Theta|\times 1}$ only has zero solution, and thus X has full rank with $\text{span}(X) = \mathbb{R}^{|\Theta|}$. As $\mathbb{R}^{|\Theta|} = \text{span}(X) \subseteq \text{span}(L) \subseteq \mathbb{R}^{|\Theta|}$, we have $\text{span}(L) = \mathbb{R}^{|\Theta|}$. □

Proof of Claim 3. We focus on the case that $S = I\backslash\{n\}$ for the remainder of the proof, since the argument for the case that $S = I$ is a more direct application of Kosenok and Severinov (2008). Consider a sub-environment with agents in $S = I\backslash\{n\}$ and information structure $(\Theta_{-n}, \hat{p})$, where $\hat{p}$ is the marginal distribution of p on Θ_{-n}. For each $i \in S$, define a utility function $\hat{u}_i$ over type space Θ_S and an extended set of feasible outcomes $\tilde{A} \equiv \Delta(A \times \mathbb{R}^n)$ as follows: for each $\theta_S \in \Theta_S$ and deterministic pair $(\bar{a}, \bar{t}) \in A \times \mathbb{R}^n$, $\hat{u}_i((\bar{a}, \bar{t}), \theta_S) \equiv \sum_{\theta_{-S} \in \Theta_{-S}} [u_i(\bar{a}, (\theta_S, \theta_{-S})) + \bar{t}_i] p(\theta_{-S}|\theta_S)$, and then extend the definition of $\hat{u}_i$ to $\Delta(A \times \mathbb{R}^n)$ by the standard expected utility.

Define an allocation rule $\hat{q} : \Theta_S \to \tilde{A}$ such that $\hat{q}(\theta_S) \equiv (q(\theta'_S, \theta_{-S}), t(\theta'_S, \theta_{-S}))$ with probability $\hat{\delta}^S[\theta_S](\theta'_S) p(\theta_{-S}|\theta_S)$ for each $\theta_S, \theta'_S \in \Theta_S$. FSE requires that the right-hand side of expression (6) is equal to 0. Hence, (6) implies that $\sum_{i \in S} \sum_{\theta_S \in \Theta_S} \hat{u}_i(\hat{q}(\theta_S), \theta_S) \hat{p}(\theta_S) = \sum_{\theta_S \in \Theta_S} V_S[q, t](\theta_S, \hat{\delta}^S) p(\theta_S) > 0$. Then by Kosenok and Severinov (2008), in the sub-environment, there exists an ex-post budget balanced transfer rule, $\psi^S : \Theta_S \to \mathbb{R}^{|S|}$, such that $(\hat{q}, \psi^S)$ is interim IR and IC. Since for each $i \in S$, θ_i, $\hat{\theta}_i \in \Theta_i$,

$$\sum_{\theta_{S\backslash\{i\}} \in \Theta_{S\backslash\{i\}}} [\hat{u}_i(\hat{q}(\hat{\theta}_i, \theta_{S\backslash\{i\}}), (\theta_i, \theta_{S\backslash\{i\}})) + \psi_i^S(\hat{\theta}_i, \theta_{S\backslash\{i\}})] \hat{p}(\theta_{S\backslash\{i\}}|\theta_i) = V_i[q^{\hat{\delta}_S}, t^{\hat{\delta}_S} + \psi^S](\theta_i, \hat{\theta}_i).$$

The interim IR and IC of $(\hat{q}, \psi^S)$ in the sub-environment imply that the S-communicative manipulation $(q^{\hat{\delta}_S}, t^{\hat{\delta}_S} + \psi^S)$ in the original environment satisfies S-IR and S-IC, and thus is S-feasible. □

Proof of Proposition 1 (i). Fix any $p \in \Delta(\Theta)$, $i \in I$, $\bar{\theta} \in \Theta$, and $\epsilon \in (0, \frac{2p(\bar{\theta}_i)(1-p(\bar{\theta}_i))}{2p(\bar{\theta}_i)(1-p(\bar{\theta}_i))+3|\Theta|})$.

Step 1. Construct a payoff structure and an efficient allocation rule q.

The set of feasible outcomes is $A = \{x_0, x_1, x_2\}$. Agents' payoffs only depend on the type of agent i and are given in the following table. In each parenthesis, the first component is the payoff of agent i and the second one is that of each $j \in I\backslash\{i\}$. Let $u_0(a) = 0$ for all $a \in A$.

(u_i, u_j)	x_0	x_1	x_2
$\theta_i = \bar{\theta}_i$	$(0, 0)$	$(1, 1)$	$(2 - \epsilon, \frac{n-2}{n-1})$
$\theta_i \neq \bar{\theta}_i$	$(0, 0)$	$(2 - \epsilon, \frac{n-2}{n-1})$	$(1, 1)$

Table 1: Agents' utility functions

Let q be the unique efficient allocation rule: $q(\theta) = x_1$ if $\theta_i = \bar{\theta}_i$, and $q(\theta) = x_2$ elsewhere. The outcome assigned by q changes only when agent i misreports $\bar{\theta}_i$ when he has another type or the other way around.

Step 2. Show that there does not exist a feasible mechanism (q, t) satisfying CIC.

Suppose there exists a feasible mechanism (q, t) satisfying the CIC condition. Similar to the proof of Claim 1, at least the following properties have to be satisfied by (q, t).

Property (i), IC of agent i. Namely, for types $\theta_i \neq \theta_i'$ where either θ_i or θ_i' is equal to $\bar{\theta}_i$,

$$\text{IC}(\theta_i; \theta_i') : \sum_{\theta_{-i} \in \Theta_{-i}} t_i(\theta_i, \theta_{-i}) p(\theta_{-i}|\theta_i) - \sum_{\theta_{-i} \in \Theta_{-i}} t_i(\theta_i', \theta_{-i}) p(\theta_{-i}|\theta_i) \geqslant 1 - \epsilon,$$

where the right-hand-side expression is the change in agent i's payoff due to the changed outcome; for types $\theta_i \neq \theta_i'$ where $\theta_i, \theta_i' \in \Theta_i \backslash \{\bar{\theta}_i\}$,

$$\text{IC}(\theta_i; \theta_i') : \qquad \sum_{\theta_{-i} \in \Theta_{-i}} t_i(\theta_i, \theta_{-i}) p(\theta_{-i}|\theta_i) - \sum_{\theta_{-i} \in \Theta_{-i}} t_i(\theta_i', \theta_{-i}) p(\theta_{-i}|\theta_i) \geqslant 0.$$

Property (ii), CIC of coalition $I \backslash \{i\}$ (or IC if $|n| = 2$). Namely, for each pair of $\theta_{-i} \neq \theta_{-i}'$,

$$\text{CIC}(\theta_{-i}; \theta_{-i}') : \qquad \sum_{j \in I \backslash \{i\}} \sum_{\theta_i \in \Theta_i} t_j(\theta_i, \theta_{-i}) p(\theta_i|\theta_{-i}) - \sum_{j \in I \backslash \{i\}} \sum_{\theta_i \in \Theta_i} t_j(\theta_i, \theta_{-i}') p(\theta_i|\theta_{-i}) \geqslant 0.$$

Property (iii), CIC of coalition I. Namely, for type profiles $\theta \neq \theta'$, where either θ_i or θ_i' is equal to $\bar{\theta}_i$ but not both,

$$\text{CIC}(\theta; \theta') : \qquad \sum_{j \in I} t_j(\theta) - \sum_{j \in I} t_j(\theta') \geqslant 2 - \epsilon + (n-1)\frac{n-2}{n-1} - n = -\epsilon;$$

for type profiles $\theta \neq \theta'$, where $\theta_i, \theta_i' \in \Theta_i \backslash \{\bar{\theta}_i\}$ or $\theta_i = \theta_i' = \bar{\theta}_i$,

$$\text{CIC}(\theta; \theta') : \qquad \sum_{j \in I} t_j(\theta) - \sum_{j \in I} t_j(\theta') \geqslant 0.$$

We scale each constraint $\text{IC}(\theta_i; \theta_i')$ by the factor $p(\theta_i)p(\theta_i')$, each constraint $\text{CIC}(\theta_{-i}; \theta_{-i}')$ by $p(\theta_{-i})p(\theta_{-i}')$, each $\text{CIC}(\bar{\theta}; \theta)$ where $\theta = (\theta_i, \theta_{-i}) \in \Theta \backslash \{\bar{\theta}\}$ by $|p(\theta) - p(\theta_i)p(\theta_{-i})| + p(\theta) - p(\theta_i)p(\theta_{-i})$, and each $\text{CIC}(\theta; \bar{\theta})$ where $\theta \in \Theta \backslash \{\bar{\theta}\}$ by $|p(\theta) - p(\theta_i)p(\theta_{-i})|$. Then aggregate these scaled constraints. Following a similar argument to derive expression (12), we can cancel all terms containing t on the left-hand side, and eventually have $0 \geqslant 2(1-\epsilon)p(\bar{\theta}_i)(1 - p(\bar{\theta}_i)) - \epsilon \sum_{\theta_{-i} \in \Theta_{-i}} \sum_{\theta_i \in \Theta_i \backslash \{\bar{\theta}_i\}} [2|p(\theta) - p(\theta_i)p(\theta_{-i})| + p(\theta) - p(\theta_i)p(\theta_{-i})] > 2(1-\epsilon)p(\bar{\theta}_i)(1 - p(\bar{\theta}_i)) - 3\epsilon|\Theta| > 0$, a contradiction. □

A.3 Proof of Proposition 2

Proof of Lemma 1. **Step 1.** Show that there exists a transfer rule $\phi : \Theta \to \mathbb{R}^n$ such that

(a) $\sum_{i\in C}\sum_{\theta_{-C}\in\Theta_{-C}}\phi_i(\theta_C,\theta_{-C})p(\theta_{-C}|\theta_C)=0$ for all $C\in\{S,I\backslash S,I\}$ and $\theta_C\in\Theta_C$;

(b) $\sum_{i\in S}\sum_{\theta_{-S}\in\Theta_{-S}}\phi_i(\hat{\theta}_S,\theta_{-S})p(\theta_{-S}|\bar{\theta}_S)<0$ for all $\hat{\theta}_S\in\Theta_S\backslash\{\bar{\theta}_S\}$.

The above result follows from Lemma 3 by defining $S^1=S$ and $S^2=I\backslash S$, and $K=2$.

Step 2. Prove the existence of a transfer rule $\hat{\phi}:\Theta\to\mathbb{R}^n$ such that for all $C\in\{S,I\backslash S\}$,

(a) $\sum_{\theta_i\in\Theta_i}\hat{\phi}_i(\theta_i,\theta_{-i})p(\theta_i|\theta_{-i})=0$ for all $i\in C$ and $\theta_{-i}\in\Theta_{-i}$;

(b) $\sum_{i\in C}\hat{\phi}_i(\theta)=\sum_{i\in C}\phi_i(\theta)$ for all $\theta\in\Theta$.

Suppose by way of contradiction that there does not exist $\hat{\phi}$ such that for $C=S$, (a) and (b) are satisfied (the case that $C=I\backslash S$ can be proved in an analogous way). It must be the case that $|S|\geqslant 2$, because (a) and (b) with respect to $C=S$ are satisfied by setting $\hat{\phi}=\phi$ when $|S|=1$.

To apply Theorem 2, with the vectors defined in Section A.1, we construct matrices B and b of dimensions $m\times l$ and $m\times 1$, respectively, where $m=\sum_{i\in S}|\Theta_{-i}|+|\Theta|$ and $l=n|\Theta|$. Matrix B is obtained by vertically stacking up $\sum_{i\in S}|\Theta_{-i}|$ row vectors $p^{\{i\}}_{\theta_{-i}\theta_{-i}}\in\mathbb{R}^l_+$ for all $i\in S$ and $\theta_{-i}\in\Theta_{-i}$ (the order does not matter), and $|\Theta|$ row vectors e^S_θ for all $\theta\in\Theta$ (following the order of elements in Θ). Construct matrix b by vertically stacking up $\sum_{i\in S}|\Theta_{-i}|$ zeros and $|\Theta|$ numbers $\sum_{i\in S}\phi_i(\theta)$ for all $\theta\in\Theta$ (following the order of elements in Θ). The supposition above implies that $Bx=b$ has no column vector solution $x\in\mathbb{R}^l$.

By Theorem 2, $B'y=\mathbf{0}_{l\times 1}$ has a column vector solution $y\in\mathbb{R}^m$ with $y'b\neq 0$, i.e., there exists a profile of numbers $(a_{\theta_{-i}}\in\mathbb{R})_{\theta_{-i}\in\Theta_{-i},i\in S}$ and a profile $(b_\theta\in\mathbb{R})_{\theta\in\Theta}$ such that

$$\sum_{i\in S}\sum_{\theta_{-i}\in\Theta_{-i}}a_{\theta_{-i}}p^{\{i\}}_{\theta_{-i}\theta_{-i}}+\sum_{\theta\in\Theta}b_\theta e^S_\theta=\mathbf{0}_{1\times l},\tag{15}$$

$$\sum_{i\in S}\sum_{\theta\in\Theta}b_\theta\phi_i(\theta)\neq 0.\tag{16}$$

Recall from Section A.1, each side of (15) is a vector in $\mathbb{R}^l$ and each dimension corresponds to an agent and a type profile. By the dimensions corresponding to $i\in S$ and $\theta\in\Theta$,

$$a_{\theta_{-i}}p(\theta)+b_\theta=0,\tag{17}$$

where θ_{-i} is part of θ. By varying $i \in S$ under a fixed θ, and then varying $\theta \in \Theta$, we have

$$a_{\theta_{-i}} = a_{\theta_{-j}}, \forall \theta \in \Theta, i, j \in S. \tag{18}$$

Fix any $\theta_{-S} \in \Theta_{-S}$ for now.

For any $\theta_S \neq \theta'_S$ that differ from each other at agent $i \in S$ only, we have $a_{(\theta_{S\backslash\{i\}}, \theta_{-S})} = a_{(\theta'_{S\backslash\{i\}}, \theta_{-S})}$. By (18), $a_{(\theta_{S\backslash\{j\}}, \theta_{-S})} = a_{(\theta_{S\backslash\{i\}}, \theta_{-S})} = a_{(\theta'_{S\backslash\{i\}}, \theta_{-S})} = a_{(\theta'_{S\backslash\{j\}}, \theta_{-S})}$ for any $j \in S\backslash\{i\}$.

For any pair of type profiles $\theta_S \neq \theta'_S$, there exists a finite sequence $(\theta^h_S)_{h=1,\ldots,H}$ where $\theta^1_S = \theta_S$, $\theta^H_S = \theta'_S$, and every pair of adjacent type profiles differ at exactly one agent. By applying the previous argument to each pair of adjacent type profiles recursively, we have that $a_{(\theta_{S\backslash\{j\}}, \theta_{-S})} = a_{(\theta_{S\backslash\{i\}}, \theta_{-S})} = a_{(\theta'_{S\backslash\{i\}}, \theta_{-S})} = a_{(\theta'_{S\backslash\{j\}}, \theta_{-S})}$ for any $i, j \in S$. Namely, there exists a constant number $\kappa_{\theta_{-S}}$ such that $a_{(\theta_{S\backslash\{i\}}, \theta_{-S})} = \kappa_{\theta_{-S}}$ for all $i \in S$ and $\theta_{S\backslash\{i\}} \in \Theta_{S\backslash\{i\}}$.

Hence, by (17), for each $\theta \in \Theta$, we have $\kappa_{\theta_{-S}} p(\theta) + b_\theta = 0$, i.e., $b_\theta = -\kappa_{\theta_{-S}} p(\theta)$.

By setting $C = I$ and $C = I\backslash S$ in (a) of Step 1, we have budget balance of ϕ and $\sum_{i \in I\backslash S} \sum_{\theta_S \in \Theta_S} p(\theta_S|\theta_{-S}) \phi_i(\theta_S, \theta_{-S}) = 0$ for all $\theta_{-S} \in \Theta_{-S}$, respectively. They jointly imply that $\sum_{i \in S} \sum_{\theta_S \in \Theta_S} p(\theta_S|\theta_{-S}) \phi_i(\theta_S, \theta_{-S}) = 0$ for all $\theta_{-S} \in \Theta_{-S}$. As a result,

$$\begin{aligned}\sum_{i \in S} \sum_{\theta \in \Theta} b_\theta \phi_i(\theta) &= \sum_{\theta_{-S} \in \Theta_{-S}} \sum_{i \in S} \sum_{\theta_S \in \Theta_S} b_{(\theta_S, \theta_{-S})} \phi_i(\theta_S, \theta_{-S}) \\ &= \sum_{\theta_{-S} \in \Theta_{-S}} -\kappa_{\theta_{-S}} p(\theta_{-S}) \underbrace{\sum_{i \in S} \sum_{\theta_S \in \Theta_S} p(\theta_S|\theta_{-S}) \phi_i(\theta_S, \theta_{-S})}_{=0} = 0,\end{aligned}$$

a contradiction with expression (16).

Step 3. By budget balance of ϕ and (b) from Step 2, $\hat{\phi}$ also satisfies budget balance. As a result, we can rewrite our result from Step 2 into the following one: there exists a budget balanced transfer rule $\hat{\phi} : \Theta \to \mathbb{R}^n$ such that,

(a) $\sum_{\theta_i \in \Theta_i} \hat{\phi}_i(\theta_i, \theta_{-i}) p(\theta_i|\theta_{-i}) = 0$ for all $i \in I$ and $\theta_{-i} \in \Theta_{-i}$;

(b) $\sum_{i \in C} \hat{\phi}_i(\theta) = \sum_{i \in C} \phi_i(\theta)$ for all $C \in \{S, I\backslash S\}$ and $\theta \in \Theta$.

Step 4. Show that

$$\sum_{i \in C} \sum_{\theta_{-C} \in \Theta_{-C}} \hat{\phi}_i(\theta_C, \theta_{-C}) p(\theta_{-C}|\theta_C) = 0, \forall C \in 2^I\backslash\{\varnothing, I\}, \theta_C \in \Theta_C. \tag{19}$$

By (a) and budget balance of $\hat{\phi}$ derived from Step 3,

$$\sum_{j \in I \backslash \{i\}} \sum_{\theta_i \in \Theta_i} \hat{\phi}_j(\theta_i, \theta_{-i}) p(\theta_i | \theta_{-i}) = 0, \forall i \in I, \theta_{-i} \in \Theta_{-i}. \tag{20}$$

Hence, for any $C \in 2^I \backslash \{\varnothing, I\}$ with $|C| = n - 1$, the equation in (19) holds. It remains to show that this equation holds for any $C \in 2^I \backslash \{\varnothing, I\}$ with $|C| < n - 1$. Fix such a C and $K \in 2^I \backslash \{\varnothing, I\}$ such that $C \cap K = \varnothing$ and $|C \cup K| = n - 1$. By (20),

$$\sum_{i \in C \cup K} \sum_{\theta_{-C \cup K} \in \Theta_{-C \cup K}} \hat{\phi}_i(\theta_{C \cup K}, \theta_{-C \cup K}) p(\theta_{-C \cup K} | \theta_{C \cup K}) = 0, \forall \theta_{C \cup K} \in \Theta_{C \cup K}.$$

Since $C \subseteq C \cup K$, the law of iterated expectations implies that

$$\sum_{i \in C \cup K} \sum_{\theta_{-C} \in \Theta_{-C}} \hat{\phi}_i(\theta_C, \theta_{-C}) p(\theta_{-C} | \theta_C) = 0, \forall \theta_C \in \Theta_C. \tag{21}$$

For any $i \in K$, since $C \subseteq I \backslash \{i\}$, the law of iterated expectations and (a) from Step 3 imply

$$\sum_{\theta_{-C} \in \Theta_{-C}} \hat{\phi}_i(\theta_C, \theta_{-C}) p(\theta_{-C} | \theta_C) = 0, \ \forall \theta_C \in \Theta_C. \tag{22}$$

From (21) and the fact that (22) holds for all $i \in K$, we know that (19) holds.

Step 5. Rename $\hat{\phi}$ as $\phi^{\bar{\theta}_S}$, which satisfies the conditions required by the lemma. □

Proof of Lemma 2. Suppose the S-BDP property holds and there exists $\delta^S \neq \bar{\delta}^S$ such that whenever an ex-post budget balanced ϕ satisfies (i), it violates (ii). We take two steps to establish a contradiction.

Step 1. Prove that for all $\theta \in \Theta$, $p(\theta) = \sum_{\bar{\theta}_S \in \Theta_S} p(\bar{\theta}_S, \theta_{-S}) \delta^S[\bar{\theta}_S](\theta_S)$.

By applying Theorem 1 in a way similar to Lemma 3, the supposition implies that there exists a profile of numbers $(a_{\theta_i} \in \mathbb{R})_{\theta_i \in \Theta_i, i \in I}$ and a profile $(b_\theta \in \mathbb{R})_{\theta \in \Theta}$ such that

$$\sum_{i \in I} \sum_{\theta_i \in \Theta_i} a_{\theta_i} p^{\{i\}}_{\theta_i \theta_i} + \sum_{\theta \in \Theta} b_\theta e^I_\theta = \sum_{\bar{\theta}_S, \hat{\theta}_S \in \Theta_S} \delta^S[\bar{\theta}_S](\hat{\theta}_S) p^S_{\bar{\theta}_S \hat{\theta}_S}. \tag{23}$$

By definitions of vectors in Section A.1, $\sum_{\theta \in \Theta} p(\theta) e^I_\theta = \sum_{i \in I} \sum_{\theta_i \in \Theta_i} p^{\{i\}}_{\theta_i \theta_i}$. Multiply this equation by a large $\lambda \in \mathbb{R}_+$, such that $\tilde{a}_{\theta_i} \equiv \lambda - a_{\theta_i} \in \mathbb{R}_{++}$ for all $i \in I$ and $\theta_i \in \Theta_i$ and $\tilde{b}_\theta \equiv \lambda p(\theta) + b_\theta \in \mathbb{R}_+$ for all $\theta \in \Theta$, and add the scaled equation with (23). We have

$$\sum_{\theta \in \Theta} \tilde{b}_\theta e^I_\theta = \sum_{i \in I} \sum_{\theta_i \in \Theta_i} \tilde{a}_{\theta_i} p^{\{i\}}_{\theta_i \theta_i} + \sum_{\bar{\theta}_S, \hat{\theta}_S \in \Theta_S} \delta^S[\bar{\theta}_S](\hat{\theta}_S) p^S_{\bar{\theta}_S \hat{\theta}_S}. \tag{24}$$

Fix any $\theta = (\theta_S, \theta_{-S}) \in \Theta$ for now. Each side of (24) is a row vector in $\mathbb{R}^{n|\Theta|}$, and each dimension corresponds to an agent and a type profile. By focusing on the dimensions corresponding to $i \in S$ and θ and then those corresponding to $j \notin S$ and θ, we have

$$\tilde{b}_\theta = \tilde{a}_{\theta_i} p(\theta) + \sum_{\bar{\theta}_S \in \Theta_S} \delta^S[\bar{\theta}_S](\theta_S) p(\bar{\theta}_S, \theta_{-S}) = \tilde{a}_{\theta_j} p(\theta). \tag{25}$$

Recall that θ_i and θ_j are parts of θ, and $|S| \geqslant 2$. By varying i within S, (25) implies that $\tilde{a}_{\theta_i} = \tilde{a}_{\theta_{i'}}$ for all $i, i' \in S$. Then, by varying θ within Θ, we conclude that there exists $\kappa \in \mathbb{R}$ such that $\kappa = \tilde{a}_{\theta_i}$ for all $i \in S$ and $\theta_i \in \Theta_i$.

Fix any $i \in S$, $j \notin S$, $\theta_S \in \Theta_S$, and a pair of different type profiles $\theta_{-S} \neq \theta'_{-S} \in \Theta_{-S}$ for now. By expression (25), since $\tilde{a}_{\theta_i} \in \mathbb{R}_{++}$ and $\delta^S[\bar{\theta}_S](\theta_S) \in \mathbb{R}_+$ for all $\bar{\theta}_S \in \Theta_S$,

$$\begin{aligned}\frac{\tilde{b}_{(\theta_S, \theta_{-S})}}{\tilde{b}_{(\theta_S, \theta'_{-S})}} = \frac{\tilde{a}_{\theta_j} p(\theta_S, \theta_{-S})}{a_{\theta'_j} p(\theta_S, \theta'_{-S})} &= \frac{\tilde{a}_{\theta_i} p(\theta_S, \theta_{-S}) + \sum_{\bar{\theta}_S \in \Theta_S} \delta^S[\bar{\theta}_S](\theta_S) p(\bar{\theta}_S, \theta_{-S})}{\tilde{a}_{\theta_i} p(\theta_S, \theta'_{-S}) + \sum_{\bar{\theta}_S \in \Theta_S} \delta^S[\bar{\theta}_S](\theta_S) p(\bar{\theta}_S, \theta'_{-S})} \\ &\in \mathrm{con}\{\frac{p(\tilde{\theta}_S, \theta_{-S})}{p(\tilde{\theta}_S, \theta'_{-S})} | \tilde{\theta}_S \in \Theta_S\}.\end{aligned}$$

Then by varying $\theta_S \in \Theta_S$, we have

$$\mathrm{con}\{\frac{\tilde{a}_{\theta_j} p(\theta_S, \theta_{-S})}{\tilde{a}_{\theta'_j} p(\theta_S, \theta'_{-S})} | \theta_S \in \Theta_S\} \subseteq \mathrm{con}\{\frac{p(\tilde{\theta}_S, \theta_{-S})}{p(\tilde{\theta}_S, \theta'_{-S})} | \tilde{\theta}_S \in \Theta_S\}.$$

As a result, $\tilde{a}_{\theta_j} = \tilde{a}_{\theta'_j}$. Then by varying θ_{-S} and θ'_{-S}, we know there exists $\kappa' \in \mathbb{R}$ such that $\kappa' = \tilde{a}_{\theta_j}$ for all $\theta_j \in \Theta_j$.

From (25), we conclude that for all $\theta \in \Theta$, $(\kappa' - \kappa) p(\theta) = \sum_{\bar{\theta}_S \in \Theta_S} p(\bar{\theta}_S, \theta_{-S}) \delta^S[\bar{\theta}_S](\theta_S)$. The only way for this to hold is $\kappa' - \kappa = 1$, i.e., for all $\theta \in \Theta$, $p(\theta) = \sum_{\bar{\theta}_S \in \Theta_S} p(\bar{\theta}_S, \theta_{-S}) \delta^S[\bar{\theta}_S](\theta_S)$.

Step 2. Show that $\delta^S = \bar{\delta}^S$, which leads to a contradiction.

The conclusion from Step 1 implies that

$$p(\cdot|\theta_S) = \sum_{\bar{\theta}_S \in \Theta_S} \frac{p(\bar{\theta}_S)}{p(\theta_S)} \delta^S[\bar{\theta}_S](\theta_S) p(\cdot|\bar{\theta}_S), \forall \theta_S \in \Theta_S. \tag{26}$$

Since the S-BDP property holds, there is a one-to-one correspondence between elements of Θ_S and $\{p(\cdot|\theta_S) | \theta_S \in \Theta_S\}$. Let $\Theta_S^1 \subseteq \Theta_S$ be the set such that $\{p(\cdot|\theta_S) | \theta_S \in \Theta_S^1\}$ is the set of all extreme points of $\{p(\cdot|\theta_S) | \theta_S \in \Theta_S\}$. Recursively, for $k = 2, ..., K$, let $\Theta_S^k \subseteq \Theta_S \backslash (\Theta_S^1 \cup ... \cup \Theta_S^{k-1})$ be the set such that $\{p(\cdot|\theta_S) | \theta_S \in \Theta_S^k\}$ is the set of all extreme points of $\{p(\cdot|\theta_S)|\theta_S \in \Theta_S \backslash (\Theta_S^1 \cup ... \cup \Theta_S^{k-1})\}$. Since Θ_S is finite, it takes $K < \infty$ rounds, such that $\Theta_S^1 \cup ... \cup \Theta_S^K = \Theta_S$. Then $\{\Theta_S^1, \Theta_S^2, ..., \Theta_S^K\}$ is a finite partition of Θ_S.

Fix any $\theta_S \in \Theta_S^1$. Since $\{p(\cdot|\theta_S)|\theta_S \in \Theta_S^1\}$ is the set of all extreme points of $\{p(\cdot|\theta_S)|\theta_S \in \Theta_S\}$, by expression (26), $\frac{p(\bar{\theta}_S)}{p(\theta_S)}\delta^S[\bar{\theta}_S](\theta_S) = 0$ for any $\bar{\theta}_S \in \Theta_S \backslash \{\theta_S\}$, and $\frac{p(\theta_S)}{p(\theta_S)}\delta^S[\theta_S](\theta_S) = 1$. The latter also means that $\delta^S[\theta_S](\hat{\theta}_S) = 0$ for all $\hat{\theta}_S \in \Theta_S \backslash \{\theta_S\}$.

Fix any $\theta_S \in \Theta_S^2$. We have shown that $\delta^S[\bar{\theta}_S](\theta_S) = 0$ for all $\bar{\theta}_S \in \Theta_S^1$. Furthermore, since $\{p(\cdot|\theta_S)|\theta_S \in \Theta_S^2\}$ is the set of all extreme points of $\{p(\cdot|\theta_S)|\theta_S \in \Theta_S \backslash \Theta_S^1\}$, by expression (26), $\frac{p(\bar{\theta}_S)}{p(\theta_S)}\delta^S[\bar{\theta}_S](\theta_S) = 0$ for any $\bar{\theta}_S \in (\Theta_S \backslash \Theta_S^1) \backslash \{\theta_S\}$, and $\frac{p(\theta_S)}{p(\theta_S)}\delta^S[\theta_S](\theta_S) = 1$.

Recursively, we can show that $\delta^S[\theta_S](\theta_S) = 1$ for all $\theta_S \in \Theta_S$. As such, $\delta^S = \bar{\delta}^S$. $\square$

Lemma 4. *When the CBDP property fails, there exists a payoff structure, under which there does not exist any FSE ambiguous mechanism satisfying the CIC condition.*

Proof. Suppose there exists $S \in 2^I \backslash \{\emptyset, I\}$ and type profiles $\bar{\theta}_S \neq \hat{\theta}_S$ such that $p(\cdot|\bar{\theta}_S) = p(\cdot|\hat{\theta}_S)$. Fix any agent $i \in S$ for whom $\bar{\theta}_i \neq \hat{\theta}_i$. Consider the payoff structure in the proof of Proposition 1 with the modification that $\epsilon \in (0, \frac{1}{n-1})$ and the same efficient allocation rule q. Suppose a feasible ambiguous mechanism (q, T) satisfies the CIC condition. Then the following two CIC (or IC when $|S| = 1$) constraints must hold:

$$\begin{aligned}
\text{CIC}(\bar{\theta}_S; \hat{\theta}_S): \quad & \min_{t \in T}\{|S| + \sum_{j \in S} \sum_{\theta_{-S} \in \Theta_{-S}} t_j(\bar{\theta}_S, \theta_{-S}) p(\theta_{-S}|\bar{\theta}_S)\} \\
& \geqslant \min_{t \in T}\{2 - \epsilon + (|S| - 1) \cdot \frac{n-2}{n-1} + \sum_{j \in S} \sum_{\theta_{-S} \in \Theta_{-S}} t_j(\hat{\theta}_S, \theta_{-S}) p(\theta_{-S}|\bar{\theta}_S)\}, \\
\text{CIC}(\hat{\theta}_S; \bar{\theta}_S): \quad & \min_{t \in T}\{|S| + \sum_{j \in S} \sum_{\theta_{-S} \in \Theta_{-S}} t_j(\hat{\theta}_S, \theta_{-S}) p(\theta_{-S}|\hat{\theta}_S)\} \\
& \geqslant \min_{t \in T}\{2 - \epsilon + (|S| - 1) \cdot \frac{n-2}{n-1} + \sum_{j \in S} \sum_{\theta_{-S} \in \Theta_{-S}} t_j(\bar{\theta}_S, \theta_{-S}) p(\theta_{-S}|\hat{\theta}_S)\}.
\end{aligned}$$

In these inequalities, the terms independent of t can be moved out of the minimum operators.

The range of ϵ and the fact that $1 \leqslant |S| \leqslant n-1$ imply $2 - \epsilon + (|S| - 1) \cdot \frac{n-2}{n-1} > |S|$. Recall $p(\cdot|\bar{\theta}_S) = p(\cdot|\hat{\theta}_S)$. We can sum up the two CIC constraints, cancel all terms containing t, and reach a contradiction. $\square$

Lemma 5. *Suppose the BDP property holds for the prior p. Let (q, T) be an ambiguous mechanism that extracts the full surplus. Then if there exists $S \in 2^I \backslash \{\emptyset\}$ with $|S| \geqslant 2$ and δ^S such that $\sum_{\theta_S \in \Theta_S} V_S[q, t](\theta_S, \delta^S) p(\theta_S) > 0$ for all $t \in T$, then (q, T) is not RCP*.*

Proof. **Step 1.** For each $t \in T$, show that there exists $\zeta^t \equiv (\zeta_i^t : \Theta_S \to \mathbb{R})_{i\in S}$ such that

(a) $\sum_{i\in S} \zeta_i^t(\theta_S) = 0$ for all $\theta_S \in \Theta_S$;

(b) $\sum_{\theta_{S\setminus\{i\}}\in\Theta_{S\setminus\{i\}}} \zeta_i^t(\theta_i, \theta_{S\setminus\{i\}})p(\theta_{S\setminus\{i\}}|\theta_i) = w_i(t, \theta_i)$ for all $i \in S$ and $\theta_i \in \Theta_i$, where

$$w_i(t, \theta_i) \equiv - V_i[q^{\delta^S}, t^{\delta^S}](\theta_i, \theta_i) + \frac{1}{|S|} \sum_{\tilde{\theta}_S\in\Theta_S} V_S[q, t](\tilde{\theta}_S, \delta^S)p(\tilde{\theta}_S). \tag{27}$$

As $\sum_{i\in S} \sum_{\theta_i\in\Theta_i} w_i(t, \theta_i)p(\theta_i) = 0$, we apply Lemma A3 of Kosenok and Severinov (2008) in a sub-environment with agents in S and type space Θ_S to establish the existence of ζ^t.

Step 2. Construct an ambiguous S-side contract (δ^S, Ψ^S).

Let S^1,..., S^K in Lemma 3 be $|S|$ singletons such that $S^1 \cup ... \cup S^K = S$. Since the BDP property holds, for each $j \in S$ and $\bar{\theta}_j \in \Theta_j$, there exists a transfer rule $\xi^{\bar{\theta}_j} \equiv (\xi_i^{\bar{\theta}_j} : \Theta \to \mathbb{R})_{i\in S}$, satisfying the conditions in Lemma 3. Fix any $\lambda \in \mathbb{R}_+$ that is weakly larger than

$$\max_{\substack{t\in T, j\in S, \bar{\theta}_j\in\Theta_j,\\ \hat{\theta}_j\in\Theta_j\setminus\{\bar{\theta}_j\}}} \frac{V_j[q^{\delta^S}, t^{\delta^S}](\bar{\theta}_j, \bar{\sigma}_j) - V_j[q^{\delta^S}, t^{\delta^S}](\bar{\theta}_j, \hat{\theta}_j) + w_j(t, \bar{\theta}_j) - \sum_{\theta_{S\setminus\{j\}}\in\Theta_{S\setminus\{j\}}} \zeta_j^t(\hat{\theta}_j, \theta_{S\setminus\{j\}})p(\theta_{S\setminus\{j\}}|\bar{\theta}_j)}{\sum_{\theta_{-j}\in\Theta_{-j}} \xi_j^{\bar{\theta}_j}(\hat{\theta}_j, \theta_{-j})p(\theta_{-j}|\bar{\theta}_j)}.$$

For each $j \in S$ and $\bar{\theta}_j \in \Theta_i$, define $\psi^{\bar{\theta}_j} \equiv (\psi_i^{\bar{\theta}_j} : T \times \Theta \to \mathbb{R})_{i\in S}$ by $\psi_i^{\bar{\theta}_j}(t, \theta) \equiv \zeta_i^t(\theta_S) + \lambda\xi_i^{\bar{\theta}_j}(\theta)$ for all $t \in T$, $\theta \in \Theta$, and $i \in S$. Define $\Psi^S \equiv \{\psi^{\bar{\theta}_j} | j \in S, \bar{\theta}_j \in \Theta_j\}$, where each transfer rule is ex-post budget balanced within S due to (a) from Step 1 and (i) in Lemma 3.

Step 3. Verify the S-feasibility of S-reallocational manipulation $(q^{\delta^S}, T^{\delta^S} + \Psi^S)$.

Fix any $i \in S$ and $\theta_i \in \Theta_i$ throughout this step.

By (27) and the supposition of this lemma, for each $t \in T$,

$$V_i[q^{\delta^S}, t^{\delta^S}](\theta_i, \theta_i) + w_i(t, \theta_i) = \frac{1}{|S|} \sum_{\tilde{\theta}_S\in\Theta_S} V_S[q, t](\tilde{\theta}_S, \delta^S)p(\tilde{\theta}_S) > 0. \tag{28}$$

Notice that for each $\psi^S \in \Psi^S$, there exists $j \in S$ and $\bar{\theta}_j \in \Theta_j$ such that $\psi_i^S(t, (\theta_i, \theta_{-i})) = \zeta_i^t(\theta_i, \theta_{S\setminus\{i\}}) + \lambda\xi_i^{\bar{\theta}_j}(\theta_i, \theta_{-i})$ for all $\theta_{-i} \in \Theta_{-i}$, $i \in S$, and $t \in T$. Hence,

$$\begin{aligned}
&\min_{t\in T}\{V_i[q^{\delta^S}, t^{\delta^S} + \psi^S(t, \cdot)](\theta_i, \theta_i)\} \\
&= \min_{t\in T}\{V_i[q^{\delta^S}, t^{\delta^S}](\theta_i, \theta_i) + \sum_{\theta_{-i}\in\Theta_{-i}} [\zeta_i^t(\theta_i, \theta_{S\setminus\{i\}}) + \lambda\xi_i^{\bar{\theta}_j}(\theta_i, \theta_{-i})]p(\theta_{-i}|\theta_i)\} && (29) \\
&= \min_{t\in T}\{ \underbrace{V_i[q^{\delta^S}, t^{\delta^S}](\theta_i, \theta_i) + w_i(t, \theta_i)}_{\text{independent of } \psi^S \in \Psi^S \text{ and positive by (28)}} \} = V_i[q^{\delta^S}, T^{\delta^S} + \Psi^S](\theta_i, \bar{\sigma}_i) > 0. && (30)
\end{aligned}$$

Notice that the second equality follows from Condition (b) of ζ^t established from Step 1 and Condition (ii) of $\xi^{\bar{\theta}_j}$ stated in Lemma 3. Thus, we have established the S-IR condition.

On the other hand, for each $\sigma_i : \Theta_i \to \Delta(\Theta_i)$ such that $\sigma_i \neq \bar{\sigma}_i$, since $\psi^{\theta_i} \in \Psi^S$,

$$\begin{aligned} & V_i[q^{\delta^S}, T^{\delta^S} + \Psi^S](\theta_i, \sigma_i) \\ \leqslant & \min_{t \in T}\{ \sum_{\hat{\theta}_i \in \Theta_i} [V_i[q^{\delta^S}, t^{\delta^S}](\theta_i, \hat{\theta}_i) + \sum_{\theta_{-i} \in \Theta_{-i}} [\zeta_i^t(\hat{\theta}_i, \theta_{S\backslash\{i\}}) + \lambda \xi_i^{\theta_i}(\hat{\theta}_i, \theta_{-i})] p(\theta_{-i}|\theta_i)] \sigma_i[\theta_i](\hat{\theta}_i)\} \\ \leqslant & \min_{t \in T}\{V_i[q^{\delta^S}, t^{\delta^S}](\theta_i, \theta_i) + w_i(t, \theta_i)\} \overset{(30)}{=} V_i[q^{\delta^S}, T^{\delta^S} + \Psi^S](\theta_i, \bar{\sigma}_i), \end{aligned}$$

where the second inequality follows from the choice of λ. To this end, we have established the S-IC condition.

Thus, $(q^{\delta^S}, T^{\delta^S} + \Psi^S)$ is an S-feasible ambiguous S-communicative manipulation.

Step 4. Show that (q, T) is not RCP*.

Fix any $t \in T$, since $\sum_{\theta_S \in \Theta_S} V_S[q, t](\theta_S, \delta^S) p(\theta_S) > 0$, the efficiency of q implies that in this ambiguous S-communicative manipulation, either the MD does not extract the full surplus or the IR constraint of an agent out of S is violated. Hence, (q, T) is not RCP*. □

Lemma 6. *If the CBDP property fails, then there exists a payoff structure under which there does not exist an FSE ambiguous mechanism that is RCP*.*

Proof. It is without loss to assume that the BDP property holds, as otherwise, Lemma 4 has shown that FSE cannot be guaranteed. Suppose for some non-singleton $S \in 2^I \backslash \{\varnothing, I\}$, there are two type profiles $\theta_S^1 \neq \theta_S^2$ with $p(\theta_S^1) \leqslant p(\theta_S^2)$ such that $p(\cdot|\theta_S^1) = p(\cdot|\theta_S^2)$. Denote an agent in S whose types are different under θ_S^1 and θ_S^2 by i, and label his component in θ_S^1 by $\bar{\theta}_i$. Consider the payoff structure in Lemma 4.

Let δ^S be the truthful joint reporting strategy except that (i) $\delta^S[\theta_S^1](\theta_S^2) = 1$, (ii) $\delta^S[\theta_S^2](\theta_S^1) = p(\theta_S^1)/p(\theta_S^2)$ and $\delta^S[\theta_S^2](\theta_S^2) = 1 - p(\theta_S^1)/p(\theta_S^2)$. It is easy to see that $p(\theta) = \sum_{\bar{\theta}_S \in \Theta_S} p(\bar{\theta}_S, \theta_{-S}) \delta^S[\bar{\theta}_S](\theta_S)$ for all $\theta \in \Theta$.

Since (q, T) extracts the full surplus, $\sum_{\theta'_S \in \Theta_S} \sum_{\theta_{-S} \in \Theta_{-S}} \sum_{i \in S} t_i(\theta'_S, \theta_{-S}) p(\theta'_S, \theta_{-S}) = -|S|$ for each $t \in T$. Hence, by the observation established from the previous paragraph,

$$\sum_{\theta'_S \in \Theta_S} \sum_{\theta_{-S} \in \Theta_{-S}} \sum_{i \in S} t_i(\theta'_S, \theta_{-S}) \sum_{\theta_S \in \Theta_S} p(\theta_S, \theta_{-S}) \delta^S[\theta_S](\theta'_S) = -|S|, \forall t \in T. \tag{31}$$

Notice that for each $t \in T$, $\sum_{\theta_S \in \Theta_S} V_S[q,t](\theta_S, \delta^S)p(\theta_S)$ can be equivalently written as

$$\sum_{\theta_S \in \Theta_S} \sum_{i \in S} \sum_{\theta'_S \in \Theta_S} \sum_{\theta_{-S} \in \Theta_{-S}} [u_i\big(q(\theta'_S, \theta_{-S}), (\theta_S, \theta_{-S})\big) + t_i(\theta'_S, \theta_{-S})] p(\theta_S, \theta_{-S}) \delta^S[\theta_S](\theta'_S)$$

$$\overset{(31)}{=} \sum_{\theta_S \in \Theta_S} \sum_{i \in S} \sum_{\theta'_S \in \Theta_S} \sum_{\theta_{-S} \in \Theta_{-S}} u_i\big(q(\theta'_S, \theta_{-S}), (\theta_S, \theta_{-S})\big) p(\theta_S, \theta_{-S}) \delta^S[\theta_S](\theta'_S) - |S| > 0.$$

The inequality holds, because the payoff structure is the same as that in Lemma 4, where it is shown that $\sum_{i \in S} u_i\big(q(\theta'_S, \theta_{-S}), (\theta_S, \theta_{-S})\big) \geqslant |S|$ and the strict inequality occurs with positive probability under δ^S (since there is a misreport between type $\bar{\theta}_i$ and another type in $\Theta_i \backslash \{\bar{\theta}_i\}$). Hence, $\sum_{\theta_S \in \Theta_S} V_S[q,t](\theta_S, \delta^S)p(\theta_S) > 0$ for all $t \in T$. By Lemma 5, there does not exist any FSE ambiguous mechanism that is RCP*. □

Proof of Proposition 2. As we remark in the text, Statement 2 ⇒ Statement 1 and Statement 3 ⇒ Statement 1 follow from Lemma 4 and Lemma 6, respectively. It remains to establish Statement 1 ⇒ Statement 2 and Statement 1 ⇒ Statement 3.

Statement 1 ⇒ Statement 2.

Step 1. Fix any efficient allocation rule $q : \Theta \to A$, and construct (q, T).

For each $i \in I$ and $\theta_i \in \Theta_i$, define

$$w_i(\theta_i) \equiv \sum_{\theta_{-i} \in \Theta_{-i}} [u_i\big(q(\theta_i, \theta_{-i}), (\theta_i, \theta_{-i})\big) + \frac{1}{n} u_0\big(q(\theta_i, \theta_{-i})\big)] p(\theta_{-i}|\theta_i) - \frac{1}{n} FS.$$

It is clear that $\sum_{i \in I} \sum_{\theta_i \in \Theta_i} w_i(\theta_i)p(\theta_i) = 0$. Hence, by Lemma A.3 of Kosenok and Severinov (2008), there exists an ex-post budget balanced transfer rule $\tau : \Theta \to \mathbb{R}^n$ such that $\sum_{\theta_{-i} \in \Theta_{-i}} \tau_i(\theta_i, \theta_{-i}) p(\theta_{-i}|\theta_i) = w_i(\theta_i)$ for all $i \in I$ and $\theta_i \in \Theta_i$.

For each $i \in I$ and $\theta \in \Theta$, define $\eta_i(\theta) \equiv \frac{1}{n} u_0\big(q(\theta)\big) - \tau_i(\theta) - \frac{1}{n} FS$. Apparently, $\sum_{i \in I} \eta_i(\theta) = u_0\big(q(\theta)\big) - FS$ for all $\theta \in \Theta$. Also, for all $i \in I$ and $\theta_i \in \Theta_i$,

$$\begin{aligned}\sum_{\theta_{-i} \in \Theta_{-i}} \eta_i(\theta_i, \theta_{-i}) p(\theta_{-i}|\theta_i) =& \frac{1}{n} \sum_{\theta_{-i} \in \Theta_{-i}} u_0\big(q(\theta_i, \theta_{-i})\big) p(\theta_{-i}|\theta_i) - w_i(\theta_i) - \frac{1}{n} FS \\ =& - \sum_{\theta_{-i} \in \Theta_{-i}} u_i\big(q(\theta_i, \theta_{-i}), (\theta_i, \theta_{-i})\big) p(\theta_{-i}|\theta_i). \qquad (32)\end{aligned}$$

Fix any constant $\lambda \in \mathbb{R}_+$ that is weakly larger than

$$\max_{\substack{S \in 2^I \backslash \{\varnothing, I\}, \\ \bar{\theta}_S, \hat{\theta}_S \in \Theta_S \text{ with } \bar{\theta}_S \neq \hat{\theta}_S}} \frac{V_S[q,\eta](\bar{\theta}_S, \bar{\theta}_S) - V_S[q,\eta](\bar{\theta}_S, \hat{\theta}_S)}{\sum_{i \in S} \sum_{\theta_{-S} \in \Theta_{-S}} \phi_i^{\bar{\theta}_S}(\hat{\theta}_S, \theta_{-S}) p(\theta_{-S}|\bar{\theta}_S)},$$

where each $\phi^{\bar{\theta}_S}$ satisfies the conditions in Lemma 1.

Define $T = \{\eta + \lambda\phi^{\theta_S} | S \in 2^I\backslash\{\varnothing, I\}, \theta_S \in \Theta_S\}$.

Step 2. Verify that (q, T) is feasible, extracts the full surplus, and satisfies CIC.

Fix any $t \in T$. There is $C \in 2^I\backslash\{\varnothing, I\}$ and $\tilde{\theta}_C \in \Theta_C$ such that $t = \eta + \lambda\phi^{\tilde{\theta}_C}$. Notice that

$$u_0\big(q(\theta)\big) - \sum_{i\in I} t_i(\theta) = u_0\big(q(\theta)\big) - \sum_{i\in I} \eta_i(\theta) = u_0\big(q(\theta)\big) - [u_0\big(q(\theta)\big) - FS] = FS$$

for all $\theta \in \Theta$, i.e., the ex-post payoff of the MD is constant and equal to FS, where we use the fact that $\phi^{\tilde{\theta}_C}$ is ex-post budget balanced. Also, for all $i \in I$ and $\theta_i \in \Theta_i$,

$$V_i[q,t](\theta_i,\theta_i) = \sum_{\theta_{-i}\in\Theta_{-i}} [u_i\big(q(\theta_i,\theta_{-i}),(\theta_i,\theta_{-i})\big) + \eta_i(\theta_i,\theta_{-i}) + \lambda\phi_i^{\tilde{\theta}_C}(\theta_i,\theta_{-i})]p(\theta_{-i}|\theta_i) = 0,$$

where we use expression (32) and Condition (i) in Lemma 1. The above two equations imply that (7) is satisfied, and that IR holds, as $V_i[q,T](\theta_i,\theta_i) = 0$ for all $i \in I$ and $\theta_i \in \Theta_i$.

To demonstrate CIC and IC, we discuss two cases.

Case 1, $S = I$. Since each $\phi^{\tilde{\theta}_C}$ is budget balanced, for each $t \in T$, $\theta \in \Theta$, and δ^I,

$$\begin{aligned} V_I[q,t](\theta,\delta^I) &= \sum_{\theta'\in\Theta}\sum_{i\in I}[u_i\big(q(\theta'),\theta\big) + \eta_i(\theta')]\delta^I[\theta](\theta') \\ &= \sum_{\theta'\in\Theta}[\sum_{i\in I} u_i\big(q(\theta'),\theta\big) + u_0\big(q(\theta')\big)]\delta^I[\theta](\theta') - FS \\ &\leqslant \sum_{i\in I} u_i\big(q(\theta),\theta\big) + u_0\big(q(\theta)\big) - FS = V_I[q,t](\theta,\bar{\delta}^I), \end{aligned} \tag{33}$$

where the inequality follows from the efficiency of q. As a result, $V_I[q,T](\theta,\bar{\delta}^I) \geqslant V_I[q,T](\theta,\delta^I)$.

Case 2, $S \in 2^I\backslash\{\varnothing, I\}$. For each $\theta_S \in \Theta_S$ and $t \in T$, since $t = \eta + \lambda\phi^{\tilde{\theta}_C}$ for some $C \in 2^I\backslash\{\varnothing, I\}$ and $\tilde{\theta}_C \in \Theta_C$, Condition (i) of $\phi^{\tilde{\theta}_C}$ in Lemma 1 implies that

$$V_S[q,t](\theta_S,\bar{\delta}^S) = V_S[q,\eta](\theta_S,\bar{\delta}^S) = V_S[q,T](\theta_S,\bar{\delta}^S).$$

For each $\theta_S \in \Theta_S$, as $\eta + \lambda\phi^{\theta_S} \in T$, $V_S[q,T](\theta_S,\hat{\theta}_S)$ is no higher than

$$V_S[q,\eta](\theta_S,\hat{\theta}_S) + \lambda\sum_{i\in S}\sum_{\theta_{-S}\in\Theta_{-S}} \phi_i^{\theta_S}(\hat{\theta}_S,\theta_{-S})p(\theta_{-S}|\theta_S) \leqslant V_S[q,\eta](\theta_S,\bar{\delta}^S), \forall\hat{\theta}_S \in \Theta_S,$$

where the inequality follows from the choice of λ and Conditions (i) and (iii) of ϕ^{θ_S} stated in Lemma 1. Hence, $V_S[q,T](\theta_S,\delta^S) \leqslant V_S[q,\eta + \lambda\phi^{\theta_S}](\theta_S,\delta^S) \leqslant V_S[q,T](\theta_S,\bar{\delta}^S)$ for any δ^S.

To this end, we have completed Step 2.

Statement 1 ⇒ Statement 3.

Step 1. Fix any efficient allocation rule $q : \Theta \to A$, and construct another (q, T).

For each $i \in I$ and $\bar{\theta}_i$, by Lemma 1 with $S = \{i\}$, there exists $\phi^{\bar{\theta}_i}$ satisfying the conditions stated in Lemma 1. Now, fix any $\lambda_1 \in \mathbb{R}_+$ that is weaky larger than

$$\max_{\substack{i\in I,\\ \bar{\theta}_i,\hat{\theta}_i\in\Theta_i \text{ with } \bar{\theta}_i\neq\hat{\theta}_i}} \frac{V_i[q,\eta](\bar{\theta}_i,\bar{\theta}_i) - V_i[q,\eta](\bar{\theta}_i,\hat{\theta}_i)}{\sum_{\theta_{-i}\in\Theta_{-i}} \phi_i^{\bar{\theta}_i}(\hat{\theta}_i,\theta_{-i})p(\theta_{-i}|\bar{\theta}_i)}.$$

For each $S \in 2^I\backslash\{\varnothing, I\}$ with $2 \leqslant |S| \leqslant n-1$, let $\phi^S : \Theta \to \mathbb{R}^n$ be a transfer rule satisfying conditions stated in Lemma 2, and $\lambda_2 \in \mathbb{R}_+$ be strictly larger than

$$\max_{\substack{S\in 2^I\backslash\{\varnothing,I\} \text{ with } 2\leqslant|S|\leqslant n-1,\\ \text{deterministic } \delta^S\neq\bar{\delta}^S,\tilde{q}:\Theta\to A}} \frac{-\sum_{\theta_S\in\Theta_S} V_S[\tilde{q},\eta^{\delta^S}](\theta_S,\theta_S)p(\theta_S)}{\sum_{\theta_S\in\Theta_S}\sum_{\hat{\theta}_S\in\Theta_S}\sum_{i\in S}\sum_{\theta_{-S}\in\Theta_{-S}} \phi_i^S(\hat{\theta}_S,\theta_{-S})p(\theta_S,\theta_{-S})\delta^S[\theta_S](\hat{\theta}_S)}.$$

Thus, for all $S \in 2^I\backslash\{\varnothing, I\}$ with $2 \leqslant |S| \leqslant n-1$, deterministic $\delta^S \neq \bar{\delta}^S$, and $\tilde{q} : \Theta \to A$,

$$\sum_{\theta_S\in\Theta_S} V_S[\tilde{q},\eta^{\delta^S}](\theta_S,\theta_S)p(\theta_S) + \lambda_2 \sum_{\theta_S\in\Theta_S}\sum_{\hat{\theta}_S\in\Theta_S}\sum_{i\in S}\sum_{\theta_{-S}\in\Theta_{-S}} \phi_i^S(\hat{\theta}_S,\theta_{-S})p(\theta_S,\theta_{-S})\delta^S[\theta_S](\hat{\theta}_S) < 0. \quad (34)$$

As a result, the inequality also holds for stochastic $\delta^S \neq \bar{\delta}^S$ and $\tilde{q}$ satisfying (3) and (4).

Now define $T \equiv \{\eta + \lambda_1\phi^{\tilde{\theta}_i} | i \in I, \tilde{\theta}_i \in \Theta_i\} \cup \{\eta + \lambda_2\phi^S | S \in 2^I\backslash\{\varnothing, I\} \text{ with } 2 \leqslant |S| \leqslant n-1\}$.

Step 2. Following the same argument as in Statement 1 ⇒ Statement 2, it is easy to see that (q, T) is a feasible ambiguous mechanism such that the MD's ex-post revenue is constant and equal to FS. Also, every I-feasible ambiguous I-reallocational manipulation leads to a feasible ambiguous mechanism because there is no agent out of I. Hence, (q, T) is RCP* with respect to I.

To show that (q, T) is RCP* with respect to each non-grand coalition S, suppose that an ambiguous S-collusive mechanism (δ^S, Ψ^S) induces an S-feasible ambiguous S-reallocational manipulation $(\tilde{q}, T^{\delta^S} + \Psi^S)$. When $\delta^S = \bar{\delta}^S$, by (3) and (4) as well as the feasibility of (q, T), this manipulation leads to a feasible ambiguous mechanism. Hence, it is without loss to assume that $\delta^S \neq \bar{\delta}^S$.

Now let $t = \eta + \lambda_2\phi^S \in T$ and fix any $\psi^S \in \Psi^S$. By S-IR, for all $i \in S$, $\theta_i \in \Theta_i$, $V_i[\tilde{q}, t^{\delta^S} + \psi^S(t, \cdot)](\theta_i, \theta_i) \geqslant 0$, i.e.,

$$\sum_{\theta_{-i}\in\Theta_{-i}} \Big[u_i\big(\tilde{q}(\theta_i,\theta_{-i}),(\theta_i,\theta_{-i})\big) + \sum_{\theta'_S\in\Theta_S} t_i(\theta'_S,\theta_{-S})\delta^S[\theta_S](\theta'_S) + \psi_i^S\big(t,(\theta_i,\theta_{-i})\big)\Big]p(\theta_{-i}|\theta_i) \geqslant 0.$$

A weighted sum of the above inequalities and the budget balance of ψ^S within S imply that

$$\sum_{\theta_S \in \Theta_S} V_S[\tilde{q}, t^{\delta^S} + \psi^S(t, \cdot)](\theta_S, \theta_S)p(\theta_S) = \sum_{\theta_S \in \Theta_S} V_S[\tilde{q}, t^{\delta^S}](\theta_S, \theta_S)p(\theta_S) \geqslant 0,$$

which contradicts (34) given $t = \eta + \lambda_2 \phi^S$. Hence, there does not exist any S-feasible ambiguous S-reallocational manipulation with $\delta^S \neq \bar{\delta}^S$. We conclude that (q, T) is RCP*.

□