Coalition-Proof Mechanisms Under Correlated Information

Huiyi Guo

Department of Economics, Texas A&M University

huiyiguo@tamu.edu

Preliminary Draft

Abstract

The paper considers two types of mechanisms that are immune from coalitional manipulations: standard mechanisms and ambiguous mechanisms. In finite-dimensional type spaces, we characterize the set of all information structures under which every efficient allocation rule is implementable via an interim coalitional incentive compatible, interim individually rational and ex-post budget-balanced standard mechanism. The requirement of coalition-proofness reduces the scope of implementability under a non-negligible set of information structures. However, when ambiguous mechanisms are allowed and agents are maxmin expected utility maximizers, coalition-proof implementation can be obtained under all most all information structures. Thus, the paper sheds light on how coalition-proofness can be achieved by engineering ambiguity in mechanism rules.

Keywords: Coalition-proofness; Ambiguous mechanism; Ambiguity aversion; Bayesian (partial) implementation; Correlated information; Budget balance; Individual rationality

1 Introduction

In mechanism design theory, most canonical models assume that agents behave non-cooperatively. However, there are many real-life environments, ranging from auctions, matching, to voting, where coalitional manipulation is a common practice. Hence, when agents can gain by jointly manipulate their behaviors, focusing on individual incentives alone is not enough to guarantee a mechanism designer's desired outcome.

This paper studies two budget-balanced mechanisms that may be used to implement efficient allocation rules in a way that is immune from coalitional manipulations: standard mechanisms and ambiguous mechanisms. Specifically, what are the information structures that guarantees implementability of every efficient allocation rules in every payoff environment via an interim coalitional incentive compatible, individually rational, and budget balanced mechanism? To answer this question, one must go beyond the independent belief case. This is because of the well-known conflict between efficiency, incentive compatibility, individual rationality, and budget balance under independent beliefs, even without the coalition proofness constraint. In the independent private value environment, the recent work of Safronov (2018) has obtained an optimistic result in implementing efficient allocation rule via a generalized expected externality mechanism, which is coalitional incentive compatible and budget-balanced. To go beyond the independent private value setting, the current work adopts a duality approach, which is not used in the generalized expected externality mechanism.

The interim coalitional incentive compatibility condition in this paper is based on three features of the model. Firstly, the mechanism designer is not required to know what coalitions can be formed. This is different from some papers in the literature assuming that a specific coalition, for instance, the grand coalition, is the only non-trivial coalition to be formed. Secondly, the mechanism designer may not know how members within a coalition transmit their private information. Some works assume that coalition members can only communicate with each other through their common knowledge. Alternatively, some assume that coalition members interact through a third-party incentive compatible mechanism. From a robustness point of view, Carroll (2019) points out that the designer may wish to consider the worst-case scenario where members in a coalition learn each other's' preference. Following this spirit, the current paper allows coalition members to pool their private information and to cooperate in the maximal possible manner. Thirdly, the paper allows members of a coalition to make budget balanced contingent transfers within group to facilitate cooperative deviations. Hence, our interim coalitional incentive compatibility condition requires that no coalition, after pooling their private information, can find a joint (mis)report and a within group transfer rule, so that each member is strictly better-off by following the group manipulation. Given a transferrable utility set up, the condition essentially imposes that no coalition can improve its members' joint payoffs by misreporting its type profile, which coincides with the one in Safronov (2018).

In the first main result, we characterize the set of information structures under which all efficient allocation rules are implementable via interim coalitional incentive compatible, interim individually rational, and ex-post budget balanced standard mechanisms. The information structure guaranteeing implementability in this sense satisfies a Strong Identifiability condition.

The Strong Identifiability condition is not a weak condition. When the coalition pattern is rich, i.e., when many coalitions could emerge, the condition has a non-negligible bite over the Convex Independence condition of Crémer and McLean (1988) and the Identifiability condition of Kosenok and Severinov (2008), which are necessary and sufficient for the same implementation question without coalition proofness concern. Hence, coalition proofness cannot be prevented easily under standard mechanisms.

However, we show that coalition proofness can be very promising if the mechanism designer can use ambiguous mechanisms and if agents are maxmin expected utility maximizers as in Gilboa and Schmeidler (1989). In an ambiguous mechanism, the designer commits to one standard mechanism, but she strategically tells agents that there are a few potential mechanisms rules. Facing ambiguity, agents are assumed to make decisions based on the worst-case expected payoff. Engineering ambiguity in mechanisms can help the designer in obtaining higher revenue or achieving efficient outcomes, as has been established by a few works, e.g., Bose and Renou (2014), Di Tillio et al. (2017), Guo (2018), Tang and Zhang (2018). As in Guo (2018), since it is publicly known that the designer wishes to implement the surplus-maximizing allocation rule, which is often unique, the paper only allows the designer to play with ambiguous transfer rules.

Specifically, in the second main result, we show that all efficient allocation rules are implementable via interim coalitional incentive compatible, interim individually rational, and ex-post budget balanced ambiguous mechanisms, if and only if the common prior satisfies a Strong Beliefs Determine Preferences (SBDP) Property.

The SBDP property requires that for any non-grand coalition, its different type profiles have distinct beliefs towards the distribution of the types of the complementary coalition, and it is a strengthening of the Beliefs Determines Preferences in the literature. The SBDP property is strictly weaker than the Strong Identifiability condition. Indeed, in any finite dimension type space, the SBDP property holds for almost all priors. Hence, introducing ambiguous mechanisms helps to extend the scope of implementability compared to unambiguous mechanisms.

Ambiguous mechanisms work in this paper as follows. When the SBDP property holds, we can design an ambiguous mechanism with two features. Firstly, when a coalition truthfully reveals members' private information, each mechanism rule in the ambiguous mechanism gives the coalition the same joint payoff. Namely, ambiguity does not affect the coalition's payoff on path. Secondly, when a coalition manipulates members' information, the coalition's joint payoff becomes uncertain, which prevents ambiguity-averse agents to deviate.

The fact that strategic engineering ambiguity helps the mechanism designer in obtaining coalition proofness may be surprising. When a coalition is formed and tries to deviate from truthful revelation, the paper allows for contingent transfers within the group. Hence, one may postulate that ambiguity can be hedged against by a proper design of within-group transfers. However, the second result shows that within-group transfers are not as powerful. They may eliminate ambiguity associated with misreporting in some information structures, but such structures belong to a set of measure zero. The systematic uncertainty that cannot be hedged against guarantees coalition proofness.

The rest of the paper proceeds as follows. Section 2 reviews related literature. Section 3 introduces the environment. Section 4 characterizes the information structure so that all efficient allocations are implementable via coalitional incentive compatible, individually rational and budget balanced mechanism. Section 6 overcomes the negative result in iSection 4 by allowing for ambiguous mechanisms. Sections 5 and 7 provide examples. Section 8 concludes the paper.

2 Literature Review

The current paper is related to three strands of the literature.

Firstly, the paper is related to the literature on first-best mechanism design under correlated information.

The papers by Crémer and McLean (1985, 1988) initiate this strand of research by characterizing the information structures that guarantee full surplus extraction. In particular, Crémer and McLean (1988) have shown that full surplus extraction can be guaranteed as a Bayesian Nash equilibrium if and only if the Convex Independence condition holds for all agents. A few later papers study how weak/strong the conditions for full surplus extraction are, among which, Neeman (2004) introduces an important necessary condition of Convex Independence, the Beliefs Determine Preferences property. The characterization result of the current paper is related to the Convex Independence and the Beliefs Determine Preferences property. However, the current paper studies how a benevolent social planner implements an efficient allocation rule via an individually rational and ex-post budget balanced mechanism rather than how an auctioneer can extract the full surplus from all agents.

The implementation question in the paper is related to the works of Matsushima (1991, 2007), Aoyagi (1998), Chung (1999), d'Aspremont et al. (2004), McLean and Postlewaite (2004, 2015) and Kosenok and Severinov (2008), among others. These papers obtain positive results on implementating efficient allocation rules via individually rational or/and budget balanced mechanisms beyond independent belief environments. In particular, Kosenok and Severinov (2008) characterize the information structure so that all efficient allocation rules are implementable via interim individually rational and ex-post budget balanced mechanisms. The current paper builds on the methodology of these works, especially the one of Kosenok and Severinov (2008), and studies the more demanding question that when there exists an coalitional incentive compatible, individually rational, and budget balanced mechanism to implement the efficient allocation rule.

Secondly, the paper adds to the literature on coalition proof mechanisms. Since Aumann (1959)'s strong Nash equilibrium, the concern of coalition manipulations in games has attracted economists' attention. With asymmetric information, various approaches has been adopted to study information transmission within a coalition.

One way is to assume that members in the coalition can pool their private information perfectly. The fine core of Wilson (1978) adopts this model. So are the coalitional incentive compatibility conditions of Chen and Micali (2012), Safronov (2018), among others. Requiring a mechanism to be coalition proof when its members pool private information imposes a relatively demanding requirement on the mechanism. However, the idea is consistent with the robustness spirit as in Carroll (2019), since the worst-case scenario to a mechanism designer is that agents can jointly manipulate information. The current paper adopts this assumption as well. In terms of the question being studied, the current paper is most related to Safronov (2018). He redesigns the expected externality mechanism in an independent private value environment by having each coalition endogenizing the externality imposed on the society. Thus, the new expected externality mechanism is not only interim incentive compatible, ex-post budget balanced, but also interim coalitional incentive compatible. The current paper goes beyond the independent private value environment, and explores which information structures can guarantee coalition-proof implementation of efficient allocation rules in all payoff environments, including interdependent value ones.

A few other papers provide interesting insights by modeling information transmission differently, but their coalition proofness imposes weaker restrictions on institutions. Some notions assume that coalition members can only use common knowledge, like the coarse core of Wilson (1978) and the Bayesian coalitional rationalization of Luo et al. (2017). In addition, some papers assume that coalition members can only communicate with each other through a third-party mechanism so that the collusive term itself has to be incentive compatible. For example, the works of Laffont and Martimort (1997, 1998, 2000), Forges et al. (2002), Che and Kim (2006, 2009), Liu (2018) explicitly modeled the third-party intermediary mechanisms. It is worth mentioning that agents can have correlated information in Laffont and Martimort

(2000) and Che and Kim (2006), who have obtained negative and positive results on coalition-proof implementation under their respetive setups. Similar to Che and Kim (2006), the current paper delivers promising news on coalition-proof implementation. However, the current paper does not assume that agents can only communicate via incentive compatible mechanisms, nor it assume that the grand coalition is the only coalition that can be formed.

Thirdly, the paper connects closely to the literature on mechanism design with ambiguity-averse agents.

Some papers in the literature explore how to engineer ambiguity in the mechanism rules so that the performance of an ambiguous mechanism is better than a standard unambiguous mechanism. Bose and Renou (2014) introduce an ambiguous communication device into the mechanism, which generates ambiguous beliefs and enlarges the set of implementable social choice functions. Di Tillio et al. (2017) design ambiguous allocation rules and transfer rules to improve an auctioneer's second-best revenue. Guo (2018) allows for ambiguity in transfer rules and show that all efficient allocations are implementable via individually rational and budget balanced mechanisms if and only if the Belief Determine Preferences property is satisfied. Tang and Zhang (2018) allow for ambiguity in mechanism rules to implement more social choice correspondences compared to a standard unambiguous mechanism. The current paper introduces ambiguous mechanisms for a similar reason to above papers, i.e., to improve the performance of mechanisms. However, since the current paper focuses on the issue of coalitional manipulations, it delivers the new message that ambiguous mechanisms helps to guarantee coalition proof implementation under more information structures.

In some other papers, the mechanism is unambiguous and agents are assumed to hold ambiguous beliefs about other agents' private information exogenously. This strand of the literature studies how ambiguous beliefs affects the design of an unambiguous mechanism and how ambiguity may help or hurt the designer in implementing efficient outcomes or maximize revenue. See, for example, Wolitzky (2016), Bose et al. (2006), Bose and Daripa (2009), Song (2016), De Castro and Yannelis (2018), Lopomo et al. (2019). However, in the current paper, agents do not hold ambiguous beliefs about other agents' information. Instead, ambiguity in this paper stems from the unknown mechanism rule, which differentiates the current paper from these papers.

3 Setup

We study an asymmetric information environment with $n \geqslant 2$ agents. The set of all agents is denoted by I and a generic element is i. A **coalition** S is a nonempty set of agents in I.

Each agent in i has a piece of payoff-relevant private information $\theta_i \in \Theta_i$. We call θ_i agent i's **type** and Θ_i agent i's **type set**. Let Θ denote $\prod_{i \in I} \Theta_i$, which is called the **type space**. For each coalition S, denote $\prod_{i \in S} \Theta_i$ by Θ_S for simplicity. Assume that the cardinality of Θ_i, denoted by $|\Theta_i|$, satisfies $2 \leqslant |\Theta_i| < \infty$. There is a **common prior** $p \in \Delta(\Theta)$ on the type space. We impose the **full support assumption**, i.e., $p(\theta) > 0$ for each $\theta \in \Theta$. For each coalition S, let $p(\theta_S)$ represent the marginal distribution of p on θ_S. When agent i has type θ_i, her **belief** is derived by updating p, and we denote the belief by $p(\cdot|\theta_i)$, which is a vector over Θ_{-i}. The pair (Θ, p) is called an **information structure**.

Each agent i has a quasi-linear utility function $u_i(a, \theta) + b$, where $a \in A$ is an element in the set of all feasible outcomes and $b \in \mathbb{R}$ is a monetary transfer. Notice that we allow the utility function to have interdependent values.

An **allocation rule** $q : \Theta \to A$ is a plan to assign an outcome contingent on agents' realized type profile. In this paper, a mechanism designer (MD) wishes to implement an ex-post **efficient** allocation rule q, i.e., one such that

$$\sum_{i \in I} u_i\big(q(\theta), \theta\big) \geqslant \sum_{i \in I} u_i\big(a, \theta\big) \forall a \in A, \theta \in \Theta.$$

A **coalition pattern** $\mathcal{S}$ is a collection of coalitions that can be formed. A **maximal coalition pattern** is the collection all non-empty subsets of I. Under this coalition pattern, all coalitions are permissible. Because of reasons such as geographic isolation or language barriers, it may be of interest to consider some other coalition patterns. For instance, the **minimal coalition pattern** is the collection all singletons of I, where no group of non-trivial (non-singleton) coalition can be formed. Consider another example where $I = \{1, 2, 3, 4\}$ and $\mathcal{S} = \{\{1\}, \{2\}, \{3\}, \{4\}, \{1, 2\}, \{3, 4\}\}$, agents 1 and 2 can talk to each other, agents 3 and 4 can also talk, but neither of the two pairs talk to the other pair. For simplicity of notation, for each $\mathcal{S}$, denote $\mathcal{S}^* = \{S \in \mathcal{S} | 1 \leqslant |S| < n-1\}$, the subset of $\mathcal{S}$ without the grand coalition, and $\mathring{\mathcal{S}} = \{S \in \mathcal{S} | 1 < |S| < n-1\}$, the subset of $\mathcal{S}$ without the grand coalition and singletons.

Let $K = |\mathcal{S}^*|$ and $K_i = |\{S \in \mathring{\mathcal{S}} | S \ni i\}|$ for all $i \in I$. When not specified, we assume the mechanism designer knows the coalition pattern $\mathcal{S}$ of the agents.

4 Simple Mechanisms

In this paper, we follow Safronov (2018) to focus on direct mechanisms. A (direct) **mechanism** is a pair (q, t), where q is an allocation rule and $t = (t_i : \Theta \to \mathbb{R})_{i \in I}$ is a **transfer rule**, which defines a monetary transfer to agent i at each type profile. A mechanism is interim **incentive compatible** (IC) if for all $i \in I$, $\bar{\theta}_i, \hat{\theta}_i \in \Theta_i$,

$$\sum_{\theta_{-i} \in \Theta_{-i}} [u_i\big(q(\bar{\theta}_i, \theta_{-i}), (\bar{\theta}_i, \theta_{-i})\big) + t_i(\bar{\theta}_i, \theta_{-i})] p(\theta_{-i}|\bar{\theta}_i)$$
$$\geqslant \sum_{\theta_{-i} \in \Theta_{-i}} [u_i\big(q(\hat{\theta}_i, \theta_{-i}), (\bar{\theta}_i, \theta_{-i})\big) + t_i(\hat{\theta}_i, \theta_{-i})] p(\theta_{-i}|\bar{\theta}_i). \quad (1)$$

We assume that when coalitions are formed, members within a coalition can perfect learn each other's private information, transfer money within the coalition, and jointly manipulate their reports to the designer. Hence, we need to consider a stronger notion of incentive compatibility. Following Safronov (2018), a mechanism is said to be interim **coalitional incentive compatible** (CIC) if

$$\sum_{i \in S} \sum_{\theta_{-S} \in \Theta_{-S}} [u_i(q(\bar{\theta}_S, \theta_{-S}), (\bar{\theta}_S, \theta_{-S})) + t_i(\bar{\theta}_S, \theta_{-S})] p(\theta_{-S}|\bar{\theta}_S)$$
$$\geqslant \sum_{i \in S} \sum_{\theta_{-S} \in \Theta_{-S}} [u_i(q(\hat{\theta}_S, \theta_{-S}), (\bar{\theta}_S, \theta_{-S})) + t_i(\hat{\theta}_S, \theta_{-S})] p(\theta_{-S}|\bar{\theta}_S) \quad (2)$$

for all $S \in \mathcal{S}$ and $\bar{\theta}_S \neq \hat{\theta}_S$.

In this notion, we require that no coalition can pool and jointly manipulate private information of agents within it to increase the aggregated utility derived from the posterior belief. Without loss of generality, we focus on misreporting in pure strategies.

One underlying assumption is that agents in a coalition pool their private information. This is a useful benchmark situation since pooling information helps the coalition to achieve the highest efficiency level. This is also reasonable when agents can disclose private information to each other. A few papers in the literature focus on the strategic interaction of agents within the coalition or assume that they interact with each other through a third-party incentive compatible mechanism, e.g. Che and Kim (2006). By doing so, these papers restrict the set of feasible deviations that a coalition can take. The current paper enlarges the set of possible deviations, and in this sense the coalition incentive compatibility notion in the current paper is demanding.

The other underlying assumption is that agents in a coalition can write contracts to redistribute wealth. In particular, when a coalition with type profile θ_S decides to report θ'_S (θ_S may or may not be equal to θ'_S), they can use a balanced within-coalition transfer rule, i.e., $\tau^{\theta_S;\theta'_S} \equiv (\tau_i^{\theta_S;\theta'_S}(\theta_{-S}))_{i\in S,\theta_{-S}\in\Theta_{-S}}$ satisfying $\sum_{i\in S}\tau_i^{\theta_S;\theta'_S}(\theta_{-S}) = 0$ for all $\theta_{-S} \in \Theta_{-S}$. For convention, when $S = I$, i.e., when $-S$ is an empty set, let $\tau_i^{\theta_S;\theta'_S}$ be an element in $\mathbb{R}$ instead of a mapping from Θ_{-S} to $\mathbb{R}$. It is easy to show that a mechanism satisfies the CIC condition, if and only if there does not exist a coalition $S \in \mathcal{S}$, two type profiles $\bar{\theta}_S$, $\hat{\theta}_S$, and two balanced within-coalition transfer rules $\tau^{\bar{\theta}_S;\hat{\theta}_S}$ and $\tau^{\bar{\theta}_S;\hat{\theta}_S}$, such that for all $i \in S$,

$$\sum_{\theta_{-S}\in\Theta_{-S}}[u_i(q(\hat{\theta}_S,\theta_{-S}),(\bar{\theta}_S,\theta_{-S})) + t_i(\hat{\theta}_S,\theta_{-S}) + \tau_i^{\bar{\theta}_S;\hat{\theta}_S}(\theta_{-S})]p(\theta_{-S}|\bar{\theta}_S)$$

$$> \sum_{\theta_{-S}\in\Theta_{-S}}[u_i(q(\bar{\theta}_S,\theta_{-S}),(\bar{\theta}_S,\theta_{-S})) + t_i(\bar{\theta}_S,\theta_{-S}) + \tau_i^{\bar{\theta}_S;\bar{\theta}_S}(\theta_{-S})]p(\theta_{-S}|\bar{\theta}_S).$$

A mechanism is ex-post **budget balanced** (BB) if for all $i \in I$ and $\theta \in \Theta$,

$$\sum_{i\in I} t_i(\theta) = 0.$$

When the mechanism designer wishes to guarantee implementability of all efficient allocation rules via coalitional incentive compatible and budget balanced simple mechanisms, the key condition on the information structure is the Coalitional Identifiability. To give an intuitive interpretation of the Coalitional Identifiability condition, we will introduce a few notations first. This interpretation assumes that at most one non-singleton non-grand coalition can emerge in the environment and manipulate members' information jointly.[1]

Formally, consider a distribution ξ over $\mathring{\mathcal{S}} \cup \{\varnothing\}$, which is called a **coalition emerging probability**. Recall that $S \in \mathring{\mathcal{S}}$ is a non-singleton, non-grand coalition given a coalition

[1]In this paper, the grand coalition can jointly manipulate its members' information. However, since the paper focuses on implementation of ex-post efficient allocation rule, the grand coalition does not have a profitable deviation. Hence, it is without loss to generality to focus on deviation of non-grand coalitions.

pattern $\mathcal{S}$. The notation $\xi(\varnothing)$ represents the probability that no non-trivial coalition is formed. For $S \in \mathring{\mathcal{S}}$, $\xi(S)$ is the probability that coalition S is formed.

For each $i \in I$, let $\sigma_i : \Theta_i \to \Delta(\Theta_i)$ be a mixed strategy of agent i in the direct mechanism. For $S \in \mathcal{S}$, define $\sigma_S = (\sigma_i)_{i \in S}$, the strategy profile of agents in S. When $\sigma_i(\theta_i)$ has weight one at θ_i for all $\theta_i \in \Theta_i$ and $i \in S$, σ_S is said to be a truthful strategy profile of coalition S. A truthful strategy profile is denoted by σ^*_S.

For any $S \in \mathcal{S}$, let $\delta_S : \Theta_S \to \Delta(\Theta_S)$ be a strategy of coalition S after members in it sharing private information. Intuitively, we can view δ_S as a strategy of a fictional agent with type set Θ_S. When $\delta_S(\theta_S)$ has weight one at θ_S for each $\theta_S \in \Theta_S$, $\delta_S(\theta_S)$ is a truthful strategy. Notice that for a non-singleton S and type profile θ_S, σ_S is an independent distribution while $\delta_S(\theta_S)$ can be correlated. Let $(\delta_S)_{S \in \mathcal{S}}$ be a profile of strategies of all coalitions in $\mathcal{S}$. When there exists $S \in \mathcal{S}$ such that δ_S is not truthful, the profile of strategies $(\delta_S)_{S \in \mathcal{S}}$ is non-truthful. For $S \in \mathcal{S}$, the distribution generated by δ_S and σ^*_{-S} is given by $\pi(\delta_S, \sigma^*_{-S})$, where

$$\pi(\delta_S, \sigma^*_{-S})(\theta) = \sum_{\bar{\theta}_S \in \Theta_S} p(\bar{\theta}_S, \theta_{-S}) \delta_S(\bar{\theta}_S)[\theta_S].$$

Consider a distribution generated by the following strategy profile. Agent i joins each non-singleton non-grand coalition $S \in \mathcal{S}$ containing i with probability $\xi(S)$ and follows strategy δ_S to report the type profile on behalf of the coalition S. In this case, agents out of S truthfully report. Agent i abstains from any non-trivial coalition with probability $1 - \sum_{S \in \mathring{\mathcal{S}}, S \ni i} \xi(S)$ and follows the strategy δ_i to report his own type. In this case, other agents truthfully report. Hence, a probability distribution over Θ,

$$(1 - \sum_{S \in \mathring{\mathcal{S}}, S \ni i} \xi(S)) \pi(\delta_i, \sigma^*_{-i}) + \sum_{S \in \mathring{\mathcal{S}}, S \ni i} \xi(S) \pi(\delta_S, \sigma^*_{-S}),$$

is generated.

These notations allow us to define the Coalitional Identifiability condition in an intuitive way. The condition requires that for any coalition emerging probability ξ, any distribution μ over Θ, and any profile of non-truthful strategies $(\delta_S)_{S \in \mathcal{S}}$, there exists an agent i who cannot have generated μ by following $(\delta_S)_{S \in \mathcal{S}, S \ni i}$.

Definition 1. *The **Coalitional Identifiability** (CI) condition holds if there does not exist a coalition emerging probability $\xi \in \Delta(\mathring{\mathcal{S}} \cup \{\varnothing\})$, a distribution function $\mu : \Theta \to \mathbb{R}$, and a profile of non-truthful strategies $(\delta_S)_{S \in \mathcal{S}}$, such that for all $i \in I$ and $\theta \in \Theta$,*

$$(1 - \sum_{S \in \mathring{\mathcal{S}}, S \ni i} \xi(S))\pi(\delta_i, \sigma^*_{-i})(\theta) + \sum_{S \in \mathring{\mathcal{S}}, S \ni i} \xi(S)\pi(\delta_S, \sigma^*_{-S})(\theta) = \mu(\theta). \tag{3}$$

As the first main result of the paper, we characterize the information structure under which coalitional incentive compatible and budget balanced implementation can be guaranteed.

Theorem 1. *Given any information structure (Θ, p), the following statements are equivalent:*

1. *The CI condition holds.*
2. *Any ex-post efficient allocation rule q under any profile of utility functions is implementable via an interim coalitional incentive compatible and ex-post budget-balanced mechanism.*

When the coalition pattern $\mathcal{S}$ is rich, the CI condition can be demanding. In fact, the following corollary shows that no information structure can guarantee implementability of all efficient allocation rules if the coalition pattern allows for two complementary coalitions.

Corollary 1. *When there exist two non-empty sets $S^1 \in \mathcal{S}$ and $S^2 \in \mathcal{S}$ such that $S^1 \cup S^1 = I$ and $S^1 \cap S^2 = \varnothing$, the CI condition fails under all information structures.*

According to Theorem 1 and the above result, as long as there exist two complementary coalitions, there always exists an efficient allocation rule that cannot be implementable via an interim CIC and ex-post BB simple mechanism. Hence, when the coalition pattern is rich enough, it is impossible to guarantee implementability of all efficient allocation rules via interim CIC and ex-post BB simple mechanisms. In particular, under the commonly adopted assumption that all coalitions could be formed, i.e., $\mathcal{S}$ is the maximal coalition pattern, the negative result holds, which calls for alternative approaches to solve the coalition-proof implementation problems.

Corollary 1 can be viewed as a corollary of a canonical impossibility result in the literature. The canonical impossibility result claims that when $n = 2$, under each information structure, there exists a profile of utility functions and an efficient allocation rule that is not implementable via an interim incentive compatible and ex-post BB simple mechanism. See d'Aspremont et al. (2004) for one reference. Notice that the canonical impossibility result in the literature relies on the assumption that $n = 2$ and does not require the interim CIC condition imposed in the current paper. To prove Corollary 1, we can treat the two complementary coalitions as two fictional agents. As in this two-agent problem, interim IC and ex-post BB cannot be guaranteed, then the more demanding problem of interim CIC and ex-post BB cannot be guaranteed in the multi-agent problem. For completeness, we provide the proof in the Appendix.

We remark that in Theorem 1, the mechanism does not need to satisfy any participation condition. One can consider imposing interim individual rationality (IR) condition, which requires that the left-hand side of expression (1) is non-negative, or interim coalitional rationality (CR) condition, which requires that the left-hand side of expression (2) is non-negative. It is possible to characterizes information structure under which interim CIC, ex-post BB, and interim IR/CR mechanism exists. However, imposing any of the two conditions requires a strengthening of the already demanding Coalitional Identifiability condition. Because of this, we do not present the two additional results in this paper.

5 An Example of Coalition-Proof Simple Mechanism

Let $I = \{1, 2, 3, 4\}$, where each agent i has a type set $\Theta_i = \{\theta_i^1, \theta_i^2\}$. The common prior p is defined in the following table:

$p(\theta)$	θ_3^1, θ_4^1	θ_3^1, θ_4^2	θ_3^2, θ_4^1	θ_3^2, θ_4^2
θ_1^1, θ_2^1	0.05	0.05	0.1	0.05
θ_1^1, θ_2^2	0.05	0.05	0.05	0.1
θ_1^2, θ_2^1	0.1	0.05	0.1	0.05
θ_1^2, θ_2^2	0.05	0.05	0.05	0.05

Let the feasible set of alternatives $A = \{x_0, x_1, x_2\}$, where x_0 gives all agents zero payoffs at all types. The payoffs of x_1 and x_2 are given in the table below.

	$u_1(x_1,\theta)$	$u_j(x_1,\theta)$ where $j \neq 1$	$u_1(x_2,\theta)$	$u_j(x_2,\theta)$ where $j \neq 1$
$\theta_1 = \theta_1^1$	1	1	6	-1
$\theta_1 \neq \theta_1^1$	0	1	1	1

Notice that $u_i(a,(\theta_1,\theta_{-1})) = u_i(a,(\theta_1,\theta'_{-1}))$ for all $i \in I$, $a \in A$, $\theta_1 \in \Theta_1$, and $\theta_{-1}, \theta'_{-1} \in \Theta_{-1}$. Namely, given the allocation, agent 1's type determines every agent's payoff.

Suppose the mechanism designer wishes to implement an efficient allocation rule defined by $q(\theta_1^1, \theta_{-1}) = x_1$ for all $\theta_{-1} \in \Theta_{-1}$, and $q(\theta_1, \theta_{-1}) = x_2$ for all $\theta_1 \neq \theta_1^1$ and all $\theta_{-1} \in \Theta_{-1}$. Notice that agent 1 is the dictator as his report determines the outcome.

We will consider two coalition patterns. Under the first coalition pattern, there exists an interim CIC and ex-post BB simple mechanism to implement q. Under the second one, there does not. This example also implies that not only the number of admissible coalitions matter for the implementation result, but also what these coalitions are.

5.1 $\mathcal{S} = \{\{1\},\{2\},\{3\},\{4\},\{1,2\},\{2,3\}\}$

In the Appendix, we verify that the CI condition holds under this coalition pattern. Thus, q is implementable via an interim CIC and ex-post BB simple mechanism. We provide a profile of utility functions, an efficient allocation rule, and a simple mechanism to implement the allocation rule.

We can consider the following transfer rule t:

t_1, t_2, t_3, t_4	θ_3^1, θ_4^1	θ_3^1, θ_4^2	θ_3^2, θ_4^1	θ_3^2, θ_4^2
θ_1^1, θ_2^1	-29, 107, -57, -21	0, -196, 0, 196	-15, 36, 216, -237	-69, 61, -257, 265
θ_1^1, θ_2^2	28, 153, 0, -181	-44, -251, -44, 339	0, -80, 231, -151	72, 20, -116, 24
θ_1^2, θ_2^1	378, -99, 20, -299	0, -296, 520, -224	-304, 225, -404, 483	-7, 0, 124, -117
θ_1^2, θ_2^2	0, 180, -459, 279	-420, 524, 0, -104	0, -180, 0, 180	279, -386, 510, -403

We claim that the simple mechanism (q,t) satisfies the conditions of ex-post BB and interim CIC.

It is easy to verify the ex-post BB condition. For instance, at state $\theta = (\theta_1^1, \theta_2^1, \theta_3^1, \theta_4^1)$,

$$BB(\theta_1^1, \theta_2^1, \theta_3^1, \theta_4^1): \qquad \sum_{i \in I} t_i(\theta) = -29 + 107 - 57 - 21 = 0.$$

One can check all the interim CIC constraints as well. Below we present the verification of two constraints as illustration. First, consider whether type-θ_1^1 agent 1 has the incentive to misreport type-θ_1^2. It is easy to show that

$$IC(\theta_1^1; \theta_1^2): \quad 1 + \sum_{\theta_{-1} \in \Theta_{-1}} t_1(\theta_1^1, \theta_{-1}) p(\theta_{-1}|\theta_1^1) = 1 > 6 + \sum_{\theta_{-1} \in \Theta_{-1}} t_1(\theta_1^2, \theta_{-1}) p(\theta_{-1}|\theta_1^1) \approx -1,$$

which establishes $IC(\theta_2^1; \theta_2^2)$. Then, consider whether type-θ_1^1 agent 1 and type-θ_2^1 agent 2 can form a coalition and have the incentive to misreport type profile (θ_1^1, θ_2^1).

$$IC(\theta_1^1, \theta_2^1; \theta_1^2, \theta_2^2): \quad 1 + 1 + \sum_{\theta_{-1-2} \in \Theta_{-1-2}} [t_1(\theta_1^1, \theta_2^1, \theta_{-1-2}) + t_2(\theta_1^1, \theta_2^1, \theta_{-1-2})] p(\theta_{-3-4}|\theta_1^1, \theta_2^1) \approx -15$$
$$> 6 - 1 + \sum_{\theta_{-1-2} \in \Theta_{-1-2}} [t_1(\theta_1^2, \theta_2^2, \theta_{-1-2}) + t_2(\theta_1^2, \theta_2^2, \theta_{-1-2})] p(\theta_{-3-4}|\theta_1^1, \theta_2^1) \approx -32.$$

5.2 $\mathcal{S} = \{\{1\}, \{2\}, \{3\}, \{4\}, \{1, 2\}, \{3, 4\}\}$

According to Theorem 2 and Corollary 1, the WCI condition fails under this coalition pattern, and thus there exists an efficient allocation rule that is not implementable via an interim CIC and ex-post BB simple mechanism. In fact, the efficient allocation rule q defined above is not implementable.

To see this, suppose by way of contradiction that there exists an interim CIC and ex-post BB transfer rule ϕ to implement q. Then for each type profile $\theta \in \Theta$, the ex-post BB condition requires that

$$BB(\theta) \qquad t_1(\theta) + t_2(\theta) + t_3(\theta) + t_4(\theta) = 0.$$

For each type θ_2, the interim CIC constraints of $S = \{1, 2\}$ require that

$$IC(\theta_1^1, \theta_2; \theta_1^2, \theta_2) \quad \sum_{\theta_3, \theta_4} [t_1(\theta_1^1, \theta_{-1}) + t_2(\theta_1^1, \theta_{-1}) - t_1(\theta_1^2, \theta_{-1}) - t_2(\theta_1^2, \theta_{-1})] p(\theta_3, \theta_4|\theta_1^1, \theta_2) \geqslant 5,$$
$$IC(\theta_1^2, \theta_2; \theta_1^1, \theta_2) \quad \sum_{\theta_3, \theta_4} [t_1(\theta_1^2, \theta_{-1}) + t_2(\theta_1^2, \theta_{-1}) - t_1(\theta_1^1, \theta_{-1}) - t_2(\theta_1^1, \theta_{-1})] p(\theta_3, \theta_4|\theta_1^1, \theta_2) \geqslant -1.$$

Any other interim CIC constraint with respect to $S = \{1,2\}$ or $S = \{3,4\}$ requires that

$$IC(\bar{\theta}_S;\hat{\theta}_S) \qquad \sum_{\theta_{-S}\in\Theta_{-S}} [t_1(\bar{\theta}_S,\theta_{-S}) + t_2(\bar{\theta}_S,\theta_{-S}) - t_1(\hat{\theta}_S,\theta_{-S}) - t_2(\hat{\theta}_S,\theta_{-S})]p(\theta_{-S}|\bar{\theta}_S) \geqslant 0.$$

For each type profile $\theta \in \Theta$, we multiply $BB(\theta)$ by $p(\theta_1,\theta_2)p(\theta_3,\theta_4) - p(\theta)$. For $S = \{1,2\}$ or $\{3,4\}$ and any pair of $\bar{\theta}_S \neq \hat{\theta}_S$, multiply $IC(\bar{\theta}_S;\hat{\theta}_S)$ by $p(\bar{\theta}_S)p(\hat{\theta}_S)$, which is positive by the full support assumption. Then add up the weighted expressions. We will have $0 \geqslant 4p(\theta_1^1)p(\theta_1^2) > 0$, a contradiction.

6 Ambiguous Mechanisms

In this section, we allow the MD to strategically enginering ambiguity in mechanism rules as in Di Tillio et al. (2017), Guo (2018), and Tang and Zhang (2018). This will bring a positive result to the question studied in previous sections.

In this paper, it is common knowledge that the MD wishes to implement the efficient allocation rule q. Hence, the paper allows the mechansim designer to introduce ambiguity in transfer rules only. Formally, an **ambiguous mechanism** is a pair (q, T), where $T = \{t = (t_i : \Theta \to \mathbb{R})\}$ is a set of potential transfer rules. The MD commits to the allocation rule q and an arbitrary transfer rule $t \in T$. She publicly announces q and the set of potentially transfer rules T to agents and conceals the transfer rule q adopted.

Agents face two sources uncertainty. An agent merely knows the distribution of other agents' types, which can be interpreted as risk. An agent does not even know the distribution under which the MD chooses the transfer rule, which we interpret as ambiguity. We assume that agents have maxmin expected utility and evaluate his interim payoffs in the following way:

$$\min_{t\in T} \sum_{\theta_{-i}\in\Theta_{-i}} [u_i\big(q(\theta),\theta\big) + t_i(\theta)]p(\theta_{-i}|\theta_i).$$

When a coalition is formed and an ambiguous mechanism is adopted, we allow coalition members to make contingent payments to each other to hedge against both sources of uncertainty. Accordingly, when agents within a coalition S share their private information θ_S and jointly adopt reporting strategy $\delta_S(\theta_S)$ in a direct mechanism ($\delta_S(\theta_S)$ may or may not be truthful reporting), a typical balanced within-coalition transfer rule can be denoted by

$\tau^{\theta_S;\delta_S(\theta_S)} = (\tau_i^{\theta_S;\delta_S(\theta_S)} : T \times \Theta_{-S} \to \mathbb{R})_{i\in S}$ satisfying $\sum_{i\in S} \tau_i^{\theta_S;\delta_S(\theta_S)}(t, \theta_{-S}) = 0$ for all $t \in T$ and $\theta_{-S} \in \Theta_{-S}$. In this case, agent $i \in S$ has MEU

$$\underline{U}_i(\theta_S, \tau) \equiv \min_{t\in T} \sum_{\theta_{-S}\in\Theta_{-S}} [u_i(q(\theta_S, \theta_{-S}), (\theta_S, \theta_{-S})) + t_i(\theta_S, \theta_{-S}) + \tau_i(t, \theta_{-S})]p(\theta_{-S}|\theta_S).$$

Remark 1. *A balanced within-coalition transfer rule τ is* ***efficient within the coalition****, i.e., there does not exist another balanced within-coalition transfer rule τ^* such that $\underline{U}_i(\theta_S, \tau^*) > \underline{U}_i(\theta_S, \tau)$ for all $i \in S$, if and only if the aggregated MEU is identical to the maxmin aggregated expected utility, i.e.,*

$$\sum_{i\in S} U_i(\theta_S, \tau) = \min_{t\in T} \sum_{i\in S} \sum_{\theta_{-S}\in\Theta_{-S}} [u_i(q(\theta_S, \theta_{-S}), (\theta_S, \theta_{-S})) + t_i(\theta_S, \theta_{-S})]p(\theta_{-S}|\theta_S).$$

This allows us to consider the maxmin aggregated utility of the coalition instead of treating each member's MEU separately.

The ambiguous mechanism is said to satisfy the interim **coalitional incentive compatibility** (CIC) condition, if there does not exist $S \in \mathcal{S}$, $\bar{\theta}_S \in \Theta_S$, $\tau = (\tau_i : T \times \Theta_{-S} \to \mathbb{R})_{i\in S}$, and a strategy of S denoted by $\delta_S : \Theta_S \to \Delta(\Theta_S)$ such that $\sum_{\theta_{-i}\in\Theta_{-i}} \tau_i(t, \theta_{-S}) = 0$ for all $t \in T$ and $\theta_{-S} \in \Theta_{-S}$, and for all $i \in S$,

$$\min_{t\in T} \sum_{\hat{\theta}_S\in\Theta_S} \sum_{\theta_{-S}\in\Theta_{-S}} [u_i\big(q(\hat{\theta}_S, \theta_{-S}), (\bar{\theta}_S, \theta_{-S})\big) + t_i(\hat{\theta}_S, \theta_{-S}) + \tau_i(t, \hat{\theta}_S, \theta_{-S})]p(\theta_{-S}|\bar{\theta}_S)\delta_S(\bar{\theta}_S)[\hat{\theta}_S]$$
$$> \min_{t\in T} \sum_{\hat{\theta}_S\in\Theta_S} \sum_{\theta_{-S}\in\Theta_{-S}} [u_i\big(q(\bar{\theta}_S, \theta_{-S}), (\bar{\theta}_S, \theta_{-S})\big) + t_i(\bar{\theta}_S, \theta_{-S})]p(\theta_{-S}|\bar{\theta}_S).$$

The ambiguous mechanism is said to satisfy the interim **coalitional incentive compatibility** (CIC) condition, if

$$\min_{t\in T} \sum_{i\in S} \sum_{\hat{\theta}_S\in\Theta_S} \sum_{\theta_{-S}\in\Theta_{-S}} [u_i\big(q(\bar{\theta}_S, \theta_{-S}), (\bar{\theta}_S, \theta_{-S})\big) + t_i(\bar{\theta}_S, \theta_{-S})]p(\theta_{-S}|\bar{\theta}_S)$$
$$\geqslant \min_{t\in T} \sum_{i\in S} \sum_{\hat{\theta}_S\in\Theta_S} \sum_{\theta_{-S}\in\Theta_{-S}} [u_i\big(q(\hat{\theta}_S, \theta_{-S}), (\bar{\theta}_S, \theta_{-S})\big) + t_i(\hat{\theta}_S, \theta_{-S})]p(\theta_{-S}|\bar{\theta}_S)\delta_S(\bar{\theta}_S)[\hat{\theta}_S] \quad (4)$$

for all $S \in \mathcal{S}$, $\bar{\theta}_S \in \Theta_S$, and strategy of S denoted by $\delta_S : \Theta_S \to \Delta(\Theta_S)$.

Notice that when ambiguous mechanism instead of simple mechanism is adopted, focusing on pure strategy deviations is no longer without loss of generality. Hence, we explicitly takes into account the possibility of deviation with mixed strategy in the above definition.

Ex-post **budget balance** (BB) of the ambiguous mechanism imposes that $\sum_{i \in I} t_i(\theta) = 0$ for all $\theta \in \Theta$ and $t \in T$. Namely, each potential transfer rule should be balanced.

Readers may wonder if agents are able to predict the transfer rule adopted by the MD. In this problem, since the MD wishes to implement an ex-post efficient allocation rule q via an ex-post BB ambiguous mechanism (q, T), each transfer rule leads to the same highest level of ex-post social surplus on path. Hence, agents cannot exclude any of the potential transfer rules in T.

In fact, under the necessary and sufficient condition to guarantee interim CIC and ex-post BB, one can further obtain a participation condition for free. We say the interim **coalitional rationality** (CR) condition holds for the ambiguous mechanism, if

$$\min_{t \in T} \sum_{i \in S} \sum_{\theta_{-S} \in \Theta_{-S}} [u_i\big(q(\theta), \theta\big) + t_i(\theta)] p(\theta_{-S}|\theta_S) \geqslant 0, \forall S \in \mathcal{S} \text{ and } \theta_S \in \Theta_S.$$

When the above inequality holds for all singleton coalitions, we say the interim **individual rationality** (IR) condition holds for the ambiguous mechanism.

Our necessary and sufficient condition, the Coalitional Beliefs Determine Preferences condition, is a strengthening of the Beliefs Determine Preferences (BDP) property introduced by Neeman (2004). It requires that for every coalition that can be formed, different type profiles should lead to different posterior beliefs of other agents' types. When the coalition pattern in the minimal one, i.e., when only singleton coalitions can be formed, the CBDP property reduces to the BDP property in the literature.

Definition 2. *The **Coalitional Beliefs Determine Preferences** (CBDP) property holds if for all $S \in \mathcal{S}^*$, there does not exist $\theta_S, \theta'_S \in \Theta_S$ with $\theta_S \neq \theta'_S$ such that*

$$p(\theta_{-S}|\theta_S) = p(\theta_{-S}|\theta'_S), \forall \theta_{-S} \in \Theta_{-S}. \tag{5}$$

The CBDP condition is a weak condition. First, it is weaker than the CWI condition. To see this, suppose the CBDP condition fails because there exists a coalition $S \in \mathcal{S}^*$ and a pair of type profiles $\bar{\theta}_S \neq \hat{\theta}_S$ such that expression (5) holds. Let a profile of strategies δ be identical with the truthful one except in $\delta_S(\theta_S)$. We require that $\delta_S(\theta_S)[\theta'_S] = 1$. Furthermore, let $\mu = p$. Then it is easy to see that expression (3) holds and thus the CWI condition fails. Second, the CBDP condition holds generically given any finite dimension of the type space Θ.[2] This implies that among the set of all distributions over Θ, the one such that the CBDP condition fails is a null set.

Now we are ready to present the second main result of the paper. The first two parts of the theorem says that the CBDP condition is necessary and sufficient to guarantee the implementability of all efficient allocation rules via interim CIC and ex-post BB ambiguous mechanisms. The third part of the theorem says that under the CBDP condition, the ambiguous mechanism further satisfies the interim CR condition.

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

1. *The CBDP condition holds.*
2. *Any ex-post efficient allocation rule q is implementable via an interim CIC and ex-post BB ambiguous mechanism.*
3. *Any ex-post efficient allocation rule q is implementable via an interim CIC, interim CIR, and ex-post BB ambiguous mechanism.*

When the mechanism designer does not know what the coalition pattern is, she wishes to design an ambiguous mechanism that is coalition-proof under all coalition patterns. This is equivalent to requiring implementability under the maximal coalition pattern. Under this pattern, the CBDP condition also holds under almost all priors. Hence, in contrast to the negative result in Corollary 1, ambiguous mechanisms bring a very promising result on coalition-proof implementation, even though the coalition pattern is the maximal one or is not known by the MD.

7 An Example of Coalition-Proof Ambiguous Mechanism

Consider the negative example in Section 5.2. There does not exist an interim CIC and ex-post BB simple mechanism to implement the efficient allocation rule q. However, we

[2] This result follows from Fact 1 in the Appendix of Kosenok and Severinov (2008).

demonstrate that q is implementable via an interim CIC and ex-post BB ambiguous mechanism. Furthermore, the ambiguous mechanism can also satisfy the interim CR condition.

Consider the following transfer rule ϕ:

$\phi_1, \phi_2, \phi_3, \phi_4$	θ_3^1, θ_4^1	θ_3^1, θ_4^2	θ_3^2, θ_4^1	θ_3^2, θ_4^2
θ_1^1, θ_2^1	-120, -120, 120, 120	-60, 0, 180, -120	60, 60, -60, -60	180, -120, -180, 120
θ_1^1, θ_2^2	-180, 120, -60, 120	-60, 0, 180, -120	120, -120, 60, -60	0, 60, -90, 30
θ_1^2, θ_2^1	60, 0, -60, 0	60, -120, 60, 0	-60, 60, 60, -60,	-180, 120, -60, 120
θ_1^2, θ_2^2	180, 0, -240, 60	60, 120, -120, -60	-120, -120, 240, 0	0, -120, 120, 0

Let the set of potential transfers be $T = \{t^1 \equiv \phi, t^2 \equiv -\phi\}$.

Since each of the transfer rule satisfies the ex-post BB condition, the ambiguous mechanism is also BB.

The interim CR condition holds in this problem. To see this, consider any coalition $S \in \mathcal{S} = \{\{1\}, \{2\}, \{3\}, \{4\}, \{1, 2\}, \{3, 4\}\}$ with cardinality k. Notice that on-path, (i) the allocation rule q gives S an total ex-post utility of ka and (ii) the total transfer received by S has a zero expected value under each potential transfer rule. Thus, the CR condition holds.

As an illustration, we compute the MEU of agent 1 with type θ_1^1 and that of coalition $\{1, 2\}$ with type profile (θ_1^1, θ_2^1) below.

Under transfer rule ϕ, the expected transfer received by type-θ_1^1 agent 1 is $0.1(-120) + 0.1(-60) + 0.2(60) + 0.1(80) + 0.1(-180) + 0.1(-60) + 0.1(120) + 0.2(0) = 0$. By the definition of T, we know that

$$CR(\theta_1^1) \qquad \min_{t \in T}\{1 + \sum_{\theta_{-1}} t_1(\theta_1^1, \theta_{-1})p(\theta_{-1}|\theta_1^1)\} = \{1 + 0, 1 - 0\} = 1.$$

Under transfer rule ϕ, the expected transfer received by type-(θ_1^1, θ_2^1) coalition $\{1, 2\}$ is $0.2(-120 - 120) + 0.2(-60 + 0) + 0.4(60 + 60) + 0.2(180 - 120) = 0$. By the definition of T, we know that

$$CR(\theta_1^1, \theta_2^1) \qquad \min_{t \in T}\{2 + \sum_{\theta_{-1}} (t_1 + t_2)(\theta_1^1, \theta_2^1, \theta_{-1-2})p(\theta_{-1-2}|\theta_1^1, \theta_2^1)\} = \{2 + 0, 2 - 0\} = 2.$$

The interim CIC condition holds, because when any type-θ_S coalition $S \in \mathcal{S}$ unilaterally deviating from truthfully revelation, one of the potential transfer rule in the ambiguous mechanism gives the deviator a sufficiently large strictly positive interim value, and the other transfer rule gives a sufficiently negative interim value. The transfer rule with negative interim value determines the MEU of misreport, which disincentize the agent from misreporting.

As an illustration, we compute the MEU for type-θ_1^1 agent 1 to report with strategy δ_1 that that for type-(θ_1^1, θ_2^1) coalition S to report with strategy δ_S.

The expected transfer received by type-θ_1^1 agent 1 under transfer rule ϕ is

$$\begin{aligned}&\delta_1(\theta_1^1)[\theta_1^1][0.1(-120) + 0.1(-60) + 0.2(60) + 0.1(180) + 0.1(-180) + 0.1(-60) + 0.1(120) + 0.2(0)]\\&+\delta_1(\theta_1^1)[\theta_1^2][0.1(60) + 0.1(60) + 0.2(-60) + 0.1(-180) + 0.1(180) + 0.1(60) + 0.1(-120) + 0.2(0)]\\&= -6\delta_1(\theta_1^1)[\theta_1^2].\end{aligned}$$

Hence, the MEU of following δ_1 is

$$\begin{aligned}&\min_{t\in T}\{\delta_1(\theta_1^1)[\theta_1^1][1 + \sum_{\theta_{-1}} t_1(\theta_1^1, \theta_{-1})p(\theta_{-1}|\theta_1^1)] + \delta_1(\theta_1^1)[\theta_1^2][7 + \sum_{\theta_{-1}} t_1(\theta_1^2, \theta_{-1})p(\theta_{-1}|\theta_1^1)]\}\\&=1 \leqslant 1.\end{aligned}$$

The expected transfer received by type-(θ_1^1, θ_2^1) coalition $\{1, 2\}$ under transfer rule ϕ is

$$\begin{aligned}&\delta_{\{1,2\}}(\theta_1^1, \theta_2^1)[\theta_1^1, \theta_2^1][0.2(-120 - 120) + 0.2(-60 + 0) + 0.2(60 + 60) + 0.6(180 - 120)]\\&+\delta_{\{1,2\}}(\theta_1^1, \theta_2^1)[\theta_1^1, \theta_2^2][0.2(-180 + 120) + 0.2(-60 + 0) + 0.2(120 - 120) + 0.6(0 + 60)]\\&+\delta_{\{1,2\}}(\theta_1^1, \theta_2^1)[\theta_1^2, \theta_2^1][0.2(60 + 0) + 0.2(60 - 120) + 0.2(-60 + 60) + 0.6(-180 + 120)]\\&+\delta_{\{1,2\}}(\theta_1^1, \theta_2^1)[\theta_1^2, \theta_2^2][0.2(120 + 0) + 0.2(60 + 120) + 0.2(-120 - 120) + 0.6(0 - 120)]\\&=\delta_{\{1,2\}}(\theta_1^1, \theta_2^1)[\theta_1^1, \theta_2^2](-12) + \delta_{\{1,2\}}(\theta_1^1, \theta_2^1)[\theta_1^2, \theta_2^1](-12) + \delta_{\{1,2\}}(\theta_1^1, \theta_2^2)[\theta_1^1, \theta_2^2](-48)\\&\leqslant(1 - \delta_{\{1,2\}}(\theta_1^1, \theta_2^1)[\theta_1^1, \theta_2^2])(-12).\end{aligned}$$

Hence, the MEU of following $\delta_{\{1,2\}}$ is

$$\begin{aligned}&\min_{t\in T}\{\sum_{\theta_1,\theta_2}\{\delta_{\{1,2\}}(\theta_1^1,\theta_2^1)[\theta_1,\theta_2]\sum_{\theta_{-1-2}}[(u_1+u_2)(q(\theta),(\theta_1^1, \theta_2^1,\theta_{-1-2}))+(t_1+t_2)(\theta)]p(\theta_{-1-2}|\theta_1^1,\theta_2^1)\}\}\\&\leqslant 2 - 6(1 - \delta_{\{1,2\}}(\theta_1^1, \theta_2^1)[\theta_1^1, \theta_2^2]) \leqslant 2.\end{aligned}$$

In this example, if we enrich the coalition pattern to include all coalitions no more than carnality 2, we can still implement q by adding more transfers to the ambiguous mechanism. However, notice that the CBDP property fails to hold under the current information structure when $\mathcal{S}$ includes a coalition with cardinality 2. For instance, to prevent coalition $\{1, 2, 3\}$ with type profile $\{\theta_1^1, \theta_2^1, \theta_3^1\}$ from reporting $\{\theta_1^2, \theta_2^2, \theta_3^2\}$ and vice versa, we need the following two inequalities:

$$
\begin{aligned}
&\min_{t\in T}\{\sum_{\theta_4}[(u_1+u_2+u_3)(q(\theta_1^1,\theta_2^1,\theta_3^1,\theta_4),(\theta_1^1,\theta_2^1,\theta_3^1,\theta_4))+(t_1+t_2+t_3)(\theta_1^1,\theta_2^1,\theta_3^1,\theta_4)]p(\theta_4|\theta_1^1,\theta_2^1,\theta_3^1)\}\\
=&1+1+1+\min_{t\in T}\{\sum_{\theta_4}[(t_1+t_2+t_3)(\theta_1^1,\theta_2^1,\theta_3^1,\theta_4)]p(\theta_4|\theta_1^1,\theta_2^1,\theta_3^1)\}\\
\geqslant&\min_{t\in T}\{\sum_{\theta_4}[(u_1+u_2+u_3)(q(\theta_1^2,\theta_2^2,\theta_3^2,\theta_4),(\theta_1^1,\theta_2^1,\theta_3^1,\theta_4))+(t_1+t_2+t_3)(\theta_1^2,\theta_2^2,\theta_3^2,\theta_4)]p(\theta_4|\theta_1^1,\theta_2^1,\theta_3^1)\}\\
=&7-1-1+\min_{t\in T}\{\sum_{\theta_4}[(t_1+t_2+t_3)(\theta_1^2,\theta_2^2,\theta_3^2,\theta_4)]p(\theta_4|\theta_1^1,\theta_2^1,\theta_3^1)\}\\
&\min_{t\in T}\{\sum_{\theta_4}[(u_1+u_2+u_3)(q(\theta_1^2,\theta_2^2,\theta_3^2,\theta_4),(\theta_1^2,\theta_2^2,\theta_3^2,\theta_4))+(t_1+t_2+t_3)(\theta_1^2,\theta_2^2,\theta_3^2,\theta_4)]p(\theta_4|\theta_1^2,\theta_2^2,\theta_3^2)\}\\
=&1+1+1+\min_{t\in T}\{\sum_{\theta_4}[(t_1+t_2+t_3)(\theta_1^2,\theta_2^2,\theta_3^2,\theta_4)]p(\theta_4|\theta_1^2,\theta_2^2,\theta_3^2)\}\\
\geqslant&\min_{t\in T}\{\sum_{\theta_4}[(u_1+u_2+u_3)(q(\theta_1^1,\theta_2^1,\theta_3^1,\theta_4),(\theta_1^2,\theta_2^2,\theta_3^2,\theta_4))+(t_1+t_2+t_3)(\theta_1^1,\theta_2^1,\theta_3^1,\theta_4)]p(\theta_4|\theta_1^2,\theta_2^2,\theta_3^2)\}\\
=&0+1+1+\min_{t\in T}\{\sum_{\theta_4}[(t_1+t_2+t_3)(\theta_1^1,\theta_2^1,\theta_3^1,\theta_4)]p(\theta_4|\theta_1^2,\theta_2^2,\theta_3^2)\}
\end{aligned}
$$

Notice that the two distribution over Θ_4 $p(\cdot|\theta_1^1, \theta_2^1, \theta_2^1)$ and $p(\cdot|\theta_1^1, \theta_2^1, \theta_2^1)$ are identical. Hence, by adding the above two expressions, we have $3 + 3 \geqslant 5 + 2$, a contradiction.

8 Conclusion

The paper studies the information structure under which coalition proof, interim individually rational, and ex-post budget balanced mechanisms exist to implement ex-post efficient allocation rules. The Strong Identifiability condition is necessary and sufficient condition to guarantee the existence of such mechanisms. Compared to individually rational and budget balanced implementation under a non-cooperative framework, coalition proofness may be difficulty to guarantee for a non-negaligle set of information structures.

However, by adopting ambiguous mechanisms, we prove that coalition proofness, individually rational, and budget balance implementation can be guaranteed if and only if the Strong Beliefs Determine Preferences property is satisfied. Among the set of all possible priors, the priors such that the SBDP fails is a null set. Hence, coalition proofness can usually be guaranteed when ambiguous mechanisms can be used.

A Appendix

Proof of Corollary 1. Depending on the cardinality of S^1 and S^2, we discuss the following three cases.

First, we consider the case that both S^1 and S^2 are non-singleton. Let the coalition emerging probability satisfy $\xi(S^1) = \xi(S^2) = 0.5$. Define a joint distribution $\mu \in \Delta(\Theta)$ by $\mu(\theta) = 0.5p(\theta_{S^1})p(\theta_{S^2}) + 0.5p(\theta)$ for all $\theta \in \Theta$. For $S = S^1$ or S^2, let the strategy of the coalition S be $\delta_S(\theta_S)[\theta'_S] = p(\theta'_S)$ for all $\theta_S, \theta'_S \in \Theta_S$. For $S \in \mathcal{S}$ and $S \neq S^1, S^2$, let $\delta_S(\theta_S)[\theta_S] = 1$ for all $\theta_S \in \Theta_S$. Then, for $i \in S^1$,

$$0.5\pi(\delta_i, \sigma^*_{-i})(\theta) + 0.5\pi(\delta_{S^1}, \sigma^*_{-S^1})(\theta) = \mu(\theta).$$

A similar analysis applies to any agent $i \in S^2$.

Second, we consider the case that both S^1 and S^2 are singletons. This is the two-agent framework. Let the coalition emerging probability satisfy $\xi(\varnothing) = 1$. Define a joint distribution $\mu \in \Delta(\Theta)$ by $\mu(\theta) = p(\theta_1)p(\theta_2)$ for all $\theta \in \Theta$. For $i = 1, 2$, let the strategy of agent i be $\delta_i(\theta_i)[\theta'_i] = p(\theta'_i)$ for all $\theta_i, \theta'_i \in \Theta_i$. Then, for $i = 1, 2$,

$$\pi(\delta_i, \sigma^*_{-i})(\theta) = \mu(\theta).$$

Third, we consider the case when $S^1 = \{j\}$ is a singleton and $S^2 = S\backslash\{j\}$ is not (or vice versa). Let the coalition emerging probability satisfy $\xi(S^2) = 1$. Define a joint distribution $\mu \in \Delta(\Theta)$ by $\mu(\theta) = p(\theta_j)p(\theta_{-j})$ for all $\theta \in \Theta$. Let the strategy of agent j be $\delta_j(\theta_j)[\theta'_j] = p(\theta'_j)$ for all $\theta_j, \theta'_j \in \Theta_j$ and that for coalition S^2 be $\delta_{S^2}(\theta_{S^2})[\theta'_{S^2}] = p(\theta'_{S^2})$ for all $\theta_{S^2}, \theta'_{S^2} \in \Theta_{S^2}$ Then, for $i \in S^1$, i.e., $i = j$,

$$\pi(\delta_i, \sigma^*_{-i})(\theta) = \mu(\theta).$$

For $i \in S^2$,

$$\pi(\delta_{S^2}, \sigma^*_{-S^2})(\theta) = \mu(\theta).$$

In all three cases, the CI condition fails under all information structures. $\square$

Lemma 1. *The following two statements are equivalent.*

1. *The CWI condition holds.*

2. *There exists $\psi : \Theta \to \mathbb{R}^n$ such that,*

   (a) *for all $\theta \in \Theta$, $\sum_{i \in I} \psi_i(\theta) = 0$;*

   (b) *for all $S \in \mathcal{S}^*$ and $\bar{\theta}_S, \hat{\theta}_S \in \Theta_S$ with $\bar{\theta}_S \neq \hat{\theta}_S$,*

$$\sum_{i \in S} \sum_{\theta_{-S} \in \Theta_{-S}} [\psi_i(\hat{\theta}_S, \theta_{-S}) - \psi_i(\bar{\theta}_S, \theta_{-S})] p(\theta_{-S}|\bar{\theta}_S) < 0.$$

Proof. Before proving this lemma, we define two notations first.

For any $S \in \mathcal{S}^*$ and any pair of $\bar{\theta}_S, \hat{\theta}_S \in \Theta_S$ ($\bar{\theta}_S$ and $\hat{\theta}_S$ may be the same), we define a vector $p_{\bar{\theta}_S;\hat{\theta}_S}$ below. This vector has $n \times |\Theta|$ dimensions and every dimension corresponds to an agent in I and a type profile in Θ. For each $i \in S$ and $\theta_{-S} \in \Theta_{-S}$, there is a unique dimension of $p_{\bar{\theta}_S;\hat{\theta}_S}$ corresponding to agent i and type profile $(\hat{\theta}_S, \theta_{-S})$. Let this dimension of $p_{\bar{\theta}_S;\hat{\theta}_S}$ be $p(\theta_{-S}|\bar{\theta}_S)$. Thus, $|S| \times |\Theta_{-S}|$ dimensions of $p_{\bar{\theta}_S;\hat{\theta}_S}$ are defined. Let all other dimensions of $p_{\bar{\theta}_S;\hat{\theta}_S}$ be 0.

For each $\theta \in \Theta$, define an $n \times |\Theta|$-dimensional vector e_θ below. Every dimension of e_θ corresponds to an agent in I and a type profile in Θ. Let the dimension of e_θ that corresponds to θ and any agent $i \in I$ be 1. In this way, n dimensions of e_θ have been defined. Let other dimensions of e_θ be 0.[3]

[3] As an illustration, consider $I = \{1, 2, 3\}$ and $\Theta_i = \{\theta_i^1, \theta_i^2\}$. List the elements of Θ in the following order: $\big((\theta_1^1, \theta_2^1, \theta_3^1), (\theta_1^1, \theta_2^1, \theta_3^2), (\theta_1^1, \theta_2^2, \theta_3^1), (\theta_1^1, \theta_2^2, \theta_3^2), (\theta_1^2, \theta_2^1, \theta_3^1), (\theta_1^2, \theta_2^1, \theta_3^2), (\theta_1^2, \theta_2^2, \theta_3^1), (\theta_1^2, \theta_2^2, \theta_3^2)\big)$. For each vector $p_{\bar{\theta}_S;\hat{\theta}_S}$ or e_θ, its first, second, and third block of eight dimensions corresponds to agent 1, 2, and 3 respectively. According to the definitions, the vector $p_{(\theta_1^1, \theta_2^1);(\theta_1^1, \theta_2^2)}$ is $(0, 0, p(\theta_3^1|\theta_1^1, \theta_2^1), p(\theta_3^2|\theta_1^1, \theta_2^1), 0, 0, 0, 0, 0, 0, p(\theta_3^1|\theta_1^1, \theta_2^1), p(\theta_3^2|\theta_1^1, \theta_2^1), 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)$. The vector $e_{(\theta_1^1, \theta_2^1, \theta_3^2)}$ is $(0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0)$.

Statement 1 $\Rightarrow$ Statement 2. We prove by contraposition again. By Motzkin's Transposition Theorem, there does not exist a transfer rule ψ satisfying the two conditions in Statement 2 if and only if there exists a vector $(b_\theta)_{\theta\in\Theta}$ and a non-zero vector $(c_{\bar{\theta}_S;\hat{\theta}_S} \geqslant 0)_{S\in\mathcal{S}^*,\bar{\theta}_S\neq\hat{\theta}_S}$ such that

$$\sum_{\theta\in\Theta} b_\theta e_\theta = \sum_{S\in\mathcal{S}^*} \sum_{\substack{\bar{\theta}_S,\hat{\theta}_S\in\Theta_S \\ \text{with } \bar{\theta}_S\neq\hat{\theta}_S}} c_{\bar{\theta}_S;\hat{\theta}_S}[p_{\bar{\theta}_S;\hat{\theta}_S} - p_{\bar{\theta}_S;\bar{\theta}_S}]. \tag{6}$$

Notice that both sides are vectors. By focusing on any agent i's blocks in each of the two vectors and adding up all elements in that vector, we know that $\sum\limits_{\theta\in\Theta} b_\theta = 0$.

Define $K_i = |\{S \in \mathring{\mathcal{S}} | S \ni i\}|$ for all $i \in I$. Notice that $\sum\limits_{\theta\in\Theta} p(\theta)e_\theta = \sum\limits_{\theta_S\in\Theta_S} p(\theta_S)p_{\theta_S;\theta_S}$ for all $S \in \mathcal{S}^*$, we have

$$\lambda K \sum_{\theta\in\Theta} p(\theta)e_\theta = \lambda(K - K_i)\sum_{i\in I}\sum_{\theta_i\in\Theta_i} p(\theta_i)p_{\theta_i;\theta_i} + \lambda \sum_{S\in\mathring{\mathcal{S}}}\sum_{\theta_S\in\Theta_S} p(\theta_S)p_{\theta_S;\theta_S} \tag{7}$$

for all $\lambda > 0$ and $K > |\mathring{\mathcal{S}}|$.

Add expressions (6) and (7), where $\lambda > 0$ is sufficiently large so that $b_\theta + \lambda K p(\theta) > 0$ for all $\theta \in \Theta$, and $\lambda p(\bar{\theta}_S) - \sum\limits_{\hat{\theta}_S\neq\bar{\theta}_S} c_{\bar{\theta}_S;\hat{\theta}_S} \geqslant 0$ for each $S \in \mathcal{S}^*$ and $\bar{\theta}_S \in \Theta_S$. Hence, we have

$$\begin{aligned}
&\sum_{\theta\in\Theta}[b_\theta + \lambda K p(\theta)]e_\theta \\
=&\sum_{i\in I} \sum_{\substack{\bar{\theta}_i,\hat{\theta}_i\in\Theta_i \\ \text{with } \bar{\theta}_i\neq\hat{\theta}_i}} c_{\bar{\theta}_i;\hat{\theta}_i}p_{\bar{\theta}_i;\hat{\theta}_i} + \sum_{i\in I}\sum_{\bar{\theta}_i\in\Theta_i}[\lambda(K-K_i)p(\bar{\theta}_i) - \sum_{\hat{\theta}_i\neq\bar{\theta}_i} c_{\bar{\theta}_i;\hat{\theta}_i}]p_{\bar{\theta}_i;\bar{\theta}_i} \\
+&\sum_{S\in\mathring{\mathcal{S}}} \sum_{\substack{\bar{\theta}_S,\hat{\theta}_S\in\Theta_S \\ \text{with } \bar{\theta}_S\neq\hat{\theta}_S}} c_{\bar{\theta}_S;\hat{\theta}_S}p_{\bar{\theta}_S;\hat{\theta}_S} + \sum_{S\in\mathring{\mathcal{S}}}\sum_{\bar{\theta}_S\in\Theta_S}[\lambda p(\bar{\theta}_S) - \sum_{\hat{\theta}_S\neq\bar{\theta}_S} c_{\bar{\theta}_S;\hat{\theta}_S}]p_{\bar{\theta}_S;\bar{\theta}_S}.
\end{aligned} \tag{8}$$

Then define $\mu(\bar{\theta}) \equiv \dfrac{b_{\bar{\theta}} + \lambda K p(\bar{\theta})}{\sum_{\theta\in\Theta}[b_\theta + \lambda K p(\theta)]} = \dfrac{b_{\bar{\theta}} + \lambda K p(\bar{\theta})}{\lambda K}$ for all $\bar{\theta} \in \Theta$, where the second equality comes from the fact that $\sum\limits_{\theta\in\Theta} b_\theta = 0$ and $\sum\limits_{\theta\in\Theta} p(\theta) = 1$. The function $\mu : \Theta \to \mathbb{R}$ is a distribution now.

Let the coalition emerging probability be $\xi(S) = \dfrac{1}{K}$ for all $S \in \mathring{\mathcal{S}}$ and $\xi(\varnothing) = 1 - \dfrac{|\mathring{\mathcal{S}}|}{K}$.

Define the strategy for agent $i \in I$ by $\delta_i(\bar{\theta}_i)[\hat{\theta}_i] = \dfrac{c_{\bar{\theta}_i;\hat{\theta}_i}}{\lambda p(\bar{\theta}_i)(K - K_i)}$ for any pairs of $\bar{\theta}_i \neq \hat{\theta}_i$ and $\delta_i(\bar{\theta}_i)[\bar{\theta}_i] = 1 - \dfrac{\sum_{\bar{\theta}_i\neq\hat{\theta}} c_{\bar{\theta}_i;\hat{\theta}_i}}{\lambda p(\bar{\theta}_i)(K - K_i)}$ for all $\bar{\theta}_i \in \Theta_i$.

Define the strategy for any coalition $S \in \mathring{\mathcal{S}}$ by $\delta_S(\bar{\theta}_S)[\hat{\theta}_S] = \dfrac{c_{\bar{\theta}_S;\hat{\theta}_S}}{\lambda p(\bar{\theta}_S)}$ for any pairs of $\bar{\theta}_S \neq \hat{\theta}_S$ and $\delta_S(\bar{\theta}_S)[\bar{\theta}_S] = 1 - \dfrac{\sum_{\hat{\theta}_S \neq \bar{\theta}_S} c_{\bar{\theta}_S;\hat{\theta}_S}}{\lambda p(\bar{\theta}_S)}$ for all $\bar{\theta}_S \in \Theta_S$.

The profile of strategies $(\delta_S)_{S \in \mathcal{S}^*}$ is non-truthful because the vector $(c_{\bar{\theta}_S;\hat{\theta}_S} \geqslant 0)_{S \in \mathcal{S}^*, \bar{\theta}_S \neq \hat{\theta}_S}$ is non-zero.

Hence, for each $i \in I$, by focusing on agent i's component of the vectors on both sides, expression (8) implies that

$$(1 - \sum_{S \in \mathring{\mathcal{S}}, S \ni i} \xi(S))\pi(\delta_i, \sigma^*_{-i})(\theta) + \sum_{S \in \mathring{\mathcal{S}}, S \ni i} \xi(S)\pi(\delta_S, \sigma^*_{-S})(\theta) = \mu(\theta), \forall \theta \in \Theta,$$

which means that the CWI condition fails to hold.

Statement 2 $\Rightarrow$ Statement 1.

We prove by contraposition. Suppose the CWI condition fails.

Then there exists a coalition emerging probability ξ, a distribution function μ, and a profile of non-truthful strategies $(\delta_S)_{S \in \mathcal{S}}$, such that for all $i \in I$ and $\theta \in \Theta$,

$$(1 - \sum_{S \in \mathring{\mathcal{S}}, S \ni i} \xi(S))\pi(\delta_i, \sigma^*_{-i})(\theta) + \sum_{S \in \mathring{\mathcal{S}}, S \ni i} \xi(S)\pi(\delta_S, \sigma^*_{-S})(\theta) = \mu(\theta). \tag{9}$$

This means that for all $i \in I$ and $\theta \in \Theta$,

$$(1 - \sum_{S \in \mathring{\mathcal{S}}, S \ni i} \xi(S)) \sum_{\bar{\theta}_i \in \Theta_i} p(\bar{\theta}_i)\delta_i(\bar{\theta}_i)[\theta_i]p(\theta_{-i}|\bar{\theta}_i) + \sum_{S \in \mathring{\mathcal{S}}, S \ni i} \sum_{\bar{\theta}_S \in \Theta_S} \xi(S)p(\bar{\theta}_S)\delta_S(\bar{\theta}_S)[\theta_S]p(\theta_{-S}|\bar{\theta}_S) = \mu(\theta). \tag{10}$$

Notice that for all $i \in I$ and $\theta \in \Theta$,

$$(1 - \sum_{S \in \mathring{\mathcal{S}}, S \ni i} \xi(S))p(\theta_i)p(\theta_{-i}|\theta_i) \sum_{\theta'_i \in \Theta_i} \delta_i(\theta_i)[\theta'_i] + \sum_{S \in \mathring{\mathcal{S}}, S \ni i} \xi(S)p(\theta_S)p(\theta_{-S}|\theta_S) \sum_{\theta'_S \in \Theta_S} \delta_S(\theta_S)[\theta'_S] = p(\theta). \tag{11}$$

By subtracting expression (11) from (10), we know for all $i \in I$ and $\theta \in \Theta$,

$$\begin{aligned}
&(1 - \sum_{S \in \mathring{\mathcal{S}}, S \ni i} \xi(S)) \sum_{\bar{\theta}_i \neq \theta_i} [p(\bar{\theta}_i)\delta_i(\bar{\theta}_i)[\theta_i]p(\theta_{-i}|\bar{\theta}_i) - p(\theta_i)\delta_i(\theta_i)[\bar{\theta}_i]p(\theta_{-i}|\theta_i)] \\
&+ \sum_{S \in \mathring{\mathcal{S}}, S \ni i} \sum_{\bar{\theta}_S \neq \theta_S} \xi(S)[p(\bar{\theta}_S)\delta_S(\bar{\theta}_S)[\theta_S]p(\theta_{-S}|\bar{\theta}_S) - p(\theta_S)\delta_S(\theta_S)[\bar{\theta}_S]p(\theta_{-S}|\theta_S)] \\
&= \mu(\theta) - p(\theta).
\end{aligned} \tag{12}$$

Define $b_\theta = \mu(\theta) - p(\theta)$ for all $\theta \in \Theta$, $c_{\bar{\theta}_S;\hat{\theta}_S} = \xi(S)p(\bar{\theta}_S)\delta_S(\bar{\theta}_S)[\hat{\theta}_S]$ for $S \in \mathring{\mathcal{S}}$ and all pairs of $\bar{\theta}_S \neq \hat{\theta}_S$; $c_{\bar{\theta}_i;\hat{\theta}_i} = (1 - \sum_{S\in\mathring{\mathcal{S}}, S\ni i} \xi(S))p(\bar{\theta}_i)\delta_i(\bar{\theta}_i)[\hat{\theta}_i]$ for $i \in I$ and all pairs of $\bar{\theta}_i \neq \hat{\theta}_i$. From expression (12), we know that expression (6) holds. The fact that $(\delta_S)_{S\in\mathcal{S}^*}$ is a profile of non-truthful strategies implies that the vector $(c_{\bar{\theta}_S\hat{\theta}_S})_{S\in\mathcal{S}^*,\bar{\theta}_S,\hat{\theta}_S\in\Theta_S \text{ with } \bar{\theta}_S\neq\hat{\theta}_S}$ is non-zero and non-negative. By Motzkin's Transposition Theorem, there does not exist a solution ψ satisfying the two conditions in Statement 2. □

Proof of Theorem 1. **Statement 1 ⇒ Statement 2**.

Since the CWI condition is satisfied, there exists an ex-post BB transfer rule ψ satisfying the conditions in Statement 2 of Lemma 1. We define $t_i(\theta) = M\psi_i(\theta)$ for all $i \in I$ and $\theta \in \Theta$, where $M > 0$ is sufficiently large such that

$$M[\sum_{i\in S}\psi_i(\bar{\theta}_S, \theta_{-S})p(\theta_{-S}|\bar{\theta}_S) - \sum_{i\in S}\psi_i(\hat{\theta}_S, \theta_{-S})p(\theta_{-S}|\bar{\theta}_S)] \geqslant \sum_{i\in S}[u_i(q(\hat{\theta}_S, \theta_{-S}), (\bar{\theta}_S, \theta_{-S})) - u_i(q(\hat{\theta}_S, \theta_{-S}), (\hat{\theta}_S, \theta_{-S}))]p(\theta_{-S}|\bar{\theta}_S) \quad (13)$$

for all $S \in \mathcal{S}^*$, $\bar{\theta}_S, \hat{\theta}_S \in \Theta_S$ with $\bar{\theta}_S \neq \hat{\theta}_S$. Notice that such an M exists because

$$\sum_{i\in S}\psi_i(\bar{\theta}_S, \theta_{-S})p(\theta_{-S}|\bar{\theta}_S) - \sum_{i\in S}\psi_i(\hat{\theta}_S, \theta_{-S})p(\theta_{-S}|\bar{\theta}_S) > 0.$$

The mechanism is ex-post BB, since ψ satisfies the ex-post BB condition. To verify interim CIC, suppose a coalition S with type profile $\bar{\theta}_S$ wishes to misreport $\hat{\theta}_S$ jointly. By pooling information with all members in the coalition and truthfully revealing the type profile to the mechanism designer, the expected utility of the coalition's aggregated payoff is

$$\sum_{i\in S}u_i(q(\bar{\theta}_S, \theta_{-S}), (\bar{\theta}_S, \theta_{-S}))p(\theta_{-S}|\bar{\theta}_S) + M\sum_{i\in S}\psi_i(\bar{\theta}_S, \theta_{-S})p(\theta_{-S}|\bar{\theta}_S). \quad (14)$$

By pooling information with all members in the coalition and misreporting type profile $\hat{\theta}_S \neq \bar{\theta}_S$, the expected utility of the coalition's aggregated payoff is

$$\sum_{i\in S}u_i(q(\hat{\theta}_S, \theta_{-S}), (\bar{\theta}_S, \theta_{-S}))p(\theta_{-S}|\bar{\theta}_S) + M\sum_{i\in S}\psi_i(\hat{\theta}_S, \theta_{-S})p(\theta_{-S}|\bar{\theta}_S), \quad (15)$$

which is weakly lower the value of expression (14) as the choice of M satisfies expression (13). Hence, the interim CIC condition holds, too.

Statement 2 ⇒ Statement 1. We prove by contraposition. Suppose the WCI condition fails. Then by Lemma 1, there exists a vector $(b_\theta)_{\theta\in\Theta}$ and a non-zero vector $(c_{\bar{\theta}_S;\hat{\theta}_S} \geqslant 0)_{S\in\mathcal{S}^*,\bar{\theta}_S,\hat{\theta}_S\in\Theta_S \text{ with } \bar{\theta}_S\neq\hat{\theta}_S}$ such that

$$\sum_{\theta\in\Theta} b_\theta e_\theta = \sum_{S\in\mathcal{S}^*} \sum_{\substack{\bar{\theta}_S,\hat{\theta}_S\in\Theta_S \\ \text{with } \bar{\theta}_S\neq\hat{\theta}_S}} c_{\bar{\theta}_S;\hat{\theta}_S}[p_{\bar{\theta}_S;\hat{\theta}_S} - p_{\bar{\theta}_S;\bar{\theta}_S}]. \tag{16}$$

Fix one coalition $S \in \mathcal{S}^*$ and a pair $\bar{\theta}_S \neq \hat{\theta}_S$ such that $c_{\bar{\theta}_S;\hat{\theta}_S} > 0$. Then fix any agent $i \in S$ for whom $\bar{\theta}_i \neq \hat{\theta}_i$ for the rest of the argument. Now we construct a profile of utility functions and an ex-post efficient allocation rule q such that q is not implementable via an interim CIC and ex-post BB simple mechanism.

Let $A = \{x_0, x_1, x_2\}$ be the set of feasible outcomes. Outcome x_0 gives all agents zero payoffs under any type profile. The payoffs of outcome x_1 and x_2 are as follows, where $a, B > 0$ and B is sufficiently large such that

$$B(1-\tfrac{|S|-1}{n-1})\sum_{C\ni i}\sum_{\theta_{C\setminus\{i\}},\theta_i=\bar{\theta}_i,\theta'_i\neq\bar{\theta}_i} c_{\theta_C;\theta'_C} \geqslant a\sum_{C\ni i}\sum_{\theta_{C\setminus\{i\}},\theta_i\neq\bar{\theta}_i,\theta'_i=\bar{\theta}_i} c_{\theta_C;\theta'_C}.$$

	$u_i(x_1,\theta)$	$u_j(x_1,\theta)$ where $j\neq i$	$u_i(x_2,\theta)$	$u_j(x_2,\theta)$ where $j\neq i$
$\theta_i = \bar{\theta}_i$	a	a	a+B	a-$\frac{1}{n-1}B$
$\theta_i \neq \bar{\theta}_i$	0	a	a	a

Consider the ex-post efficient allocation rule q defined by $q(\bar{\theta}_i, \theta_{-i}) = x_1$ for all $\theta_{-i} \in \Theta_{-i}$, and $q(\theta) = x_2$ for all $\theta_i \neq \bar{\theta}_i$ and all $\theta_{-i} \in \Theta_{-i}$.

For each $\theta \in \Theta$, by $BB(\theta)$, the following equation should be satisfied:

$$\sum_{j\in I} t_j(\theta) = 0.$$

For each coalition $C \in \mathcal{S}$ and a pair of $\theta_C \neq \theta'_C$, we consider the three possible cases. Case 1, when $i \in C$ and $\theta_i = \bar{\theta}_i$. By $IC(\theta_C; \theta'_C)$,

$$\sum_{j\in C}\sum_{\theta_{-C}\in\Theta_{-C}}[t_j(\theta'_C,\theta_{-C})-t_j(\theta_C,\theta_{-C})]p(\theta_{-C}|\theta_C) \leqslant -a|C|-B(1-\tfrac{|S|-1}{n-1})+a|C| = -B(1-\tfrac{|C|-1}{n-1}).$$

Case 2, when $i \in C$, $\theta_i \neq \bar{\theta}_i$, and $\theta'_i = \bar{\theta}_i$. By $IC(\theta_C; \theta'_C)$,

$$\sum_{j \in C} \sum_{\theta_{-C} \in \Theta_{-C}} [t_j(\theta'_C, \theta_{-C}) - t_j(\theta_C, \theta_{-C})] p(\theta_{-C}|\theta_C) \leqslant -a(|C| - 1) + a|C| = a.$$

Case 3, any other situation. By $IC(\theta_C; \theta'_C)$,

$$\sum_{j \in C} \sum_{\theta_{-C} \in \Theta_{-C}} [t_j(\theta'_C, \theta_{-C}) - t_j(\theta_C, \theta_{-C})] p(\theta_{-C}|\theta_C) \leqslant -a|C| + a|C| = 0.$$

Multiply $BB(\theta)$ by $-b_\theta$ for each $\theta \in \Theta$ and $IC(\theta_C; \theta'_C)$ by $c_{\theta_C;\theta'_C}$ for all coalition $C \in \mathcal{S}^*$ and each pair of $\theta_C \neq \theta'_C$. Then add up the expressions. This gives us

$$0 \leqslant -B(1 - \tfrac{|S|-1}{n-1}) \sum_{C \ni i} \sum_{\theta_{C \setminus \{i\}}, \theta_i = \bar{\theta}_i, \theta'_i \neq \bar{\theta}_i} c_{\theta_C;\theta'_C} + a \sum_{C \ni i} \sum_{\theta_{C \setminus \{i\}}, \theta_i \neq \bar{\theta}_i, \theta'_i = \bar{\theta}_i} c_{\theta_C;\theta'_C} < 0.$$

This contradiction implies that q is not implementable via an interim CIC and ex-post BB simple mechanism. □

Lemma 2. *When $n = 2$ and $\mathcal{S} = \{1, 2\}$, the WCI condition fails under all information structures.*

Proof. Let (Θ, p) be an information structure. For each coalition $i \in I$ and types $\theta_i, \theta'_i \in \Theta_i$, let $\delta_i(\theta_i)[\theta'_i] = p(\theta'_i)$. As $\mathring{\mathcal{S}}$ is an empty set, let the coalition emerging probability be $\xi(\varnothing) = 1$. Define a distribution μ such that $\mu(\theta) = p(\theta_1)p(\theta_2)$ for all $\theta \in \Theta$. Then, it is easy to see that for both $i = 1, 2$ and any $\theta \in \Theta$,

$$(1 - \sum_{S \in \mathring{\mathcal{S}}, S \ni i} \xi(S)) \pi(\delta_i, \sigma^*_{-i})(\theta) + \sum_{S \in \mathring{\mathcal{S}}, S \ni i} \xi(S) \pi(\delta_S, \sigma^*_{-S})(\theta) = \pi(\delta_i, \sigma^*_{-i})(\theta) = p(\theta_1)p(\theta_2) = \mu(\theta).$$

Hence, the CWI condition fails for (Θ, p). □

Proof. Let (Θ, p) be an information structure with n agents. Suppose $S^1, S^2 \in \mathcal{S}$ satisfies that $S^1 \cup S^2 = I$ and $S^1 \cap S^2 = \varnothing$. The $n = 2$ case follows trivially from Lemma 2. Thus, we assume that $n \geqslant 3$. We discuss the following three cases.

Case 1: $|S^1| > 1$ and $|S^2| > 1$. We define the coalition emerging probability $\xi(S^1) = \xi(S^2) = \frac{1}{2}$, and $\xi(S) = 0$ for any other $S \in \mathring{\mathcal{S}}$ and $S = \varnothing$. For $S = S^1, S^2$, we define $\delta_S(\theta_S)[\theta'_S] = p(\theta'_S)$ for all θ_S, $\theta'_S \in \Theta_S$. Define δ_i as the truthful strategy for all $i \in I$. Let $\mu(\theta) = 0.5p(\theta_{S^1})p(\theta_{S^2}) + 0.5p(\theta)$. Thus, for each $i \in S^1$ and $\theta \in \Theta$, we have

$$\frac{1}{2}\pi(\delta_i, \sigma^*_{-i})(\theta) + \frac{1}{2}\pi(\delta_{S^1}, \sigma^*_{-S^1})(\theta) = \mu(\theta).$$

Similar analysis applies if $i \in S^2$. Hence, the CWI condition fails.

Case 2: $|S^1| = 1$ and $|S^2| > 1$. We define the coalition emerging probability $\xi(S^2) = 1$, and $\xi(S) = 0$ for any other $S \in \mathring{\mathcal{S}}$ and $S = \emptyset$. We define $\delta_S(\theta_S)[\theta'_S] = p(\theta'_S)$ for $S = S^1, S^2$ and all θ_S, $\theta'_S \in \Theta_S$. Let $\mu(\theta) = p(\theta_{S^1})p(\theta_{S^2})$. Thus, for i such that $S^1 = \{i\}$ and $\theta \in \Theta$, we have

$$\pi(\delta_i, \sigma^*_{-i})(\theta) = \mu(\theta).$$

For $i \in I$ such that $i \in S^2$ and $\theta \in \Theta$, we have

$$\pi(\delta_{S^2}, \sigma^*_{-S^2})(\theta) = \mu(\theta).$$

Hence, the CWI condition fails.

Case 3: $|S^1| > 1$ and $|S^2| = 1$. The analysis is symmetric to Case 2 and thus omitted.

□

To verify that under coalition pattern $\mathcal{S} = \{\{1\}, \{2\}, \{3\}, \{4\}, \{1,2\}, \{2,3\}\}$, the information structure defined in the example satisfies the CWI condition, we first define a few matrices.

For each $S \in \mathcal{S}^*$, we stack the row vectors $p_{\bar{\theta}_S;\hat{\theta}_S} - p_{\bar{\theta}_S;\bar{\theta}_S}$ for all $\bar{\theta}_S, \hat{\theta}_S \in \Theta_S$ with $\bar{\theta}_S \neq \hat{\theta}_S$ to form a matrix M_S. Hence, for $i = 1, 2, 3$, M_i is a 2×64 matrix, and for $S = \{1, 2\}$ or $\{2, 3\}$, M_S is a 12×64 matrix.

Let matrix A correspond to the 2×16 submatrix formed by the first 16 columns of M_1. Similarly, let matrices B, C, and D correspond to the second 16 columns of M_2, the third 16 columns of M_3, and the last 16 columns of M_4 respectively.

Let matrix E correspond to the 12×16 submatrix formed by the first (or equivalently, the second) 16 columns of $M_{\{1,2\}}$. Let F correspond to the second (or equivalently, the third) 16 columns of $M_{\{2,3\}}$.

Matrix $O_{m\times k}$ is a zero matrix of $m \times k$ dimensions.

With the above notations, we can define the following two matices:

$$M = \begin{pmatrix} -D & -D & -D \\ A & O_{2\times 16} & O_{2\times 16} \\ O_{2\times 16} & B & O_{2\times 16} \\ O_{2\times 16} & O_{2\times 16} & C \\ E & E & O_{12\times 16} \\ O_{12\times 16} & F & F \end{pmatrix}, \quad \tilde{M} = \begin{pmatrix} A & O_{2\times 16} & O_{2\times 16} & O_{2\times 16} \\ O_{2\times 16} & B & O_{2\times 16} & O_{2\times 16} \\ O_{2\times 16} & O_{2\times 16} & C & O_{2\times 16} \\ O_{2\times 16} & O_{2\times 16} & O_{2\times 16} & D \\ E & E & O_{12\times 16} & O_{12\times 16} \\ O_{12\times 16} & F & F & O_{12\times 16} \end{pmatrix}.$$

To verify the CWI condition, we suppose by way of contradiction that the CWI condition fails. According to the intermediate step in Lemma 1, this condition fails if and only if there exists $(b_\theta)_{\theta\in\Theta}$ and a non-zero vector $(c_{\bar{\theta}_S;\hat{\theta}_S} \geqslant 0)_{S\in\mathcal{S}*,\bar{\theta}_S,\hat{\theta}_S \text{ with } \bar{\theta}_S\neq\hat{\theta}_S}$ such that

$$\sum_{\theta\in\Theta} b_\theta e_\theta = \sum_{S\in\mathcal{S}*} \sum_{\substack{\bar{\theta}_S,\hat{\theta}_S\in\Theta_S \\ \text{with } \bar{\theta}_S\neq\hat{\theta}_S}} c_{\bar{\theta}_S\hat{\theta}_S}[p_{\bar{\theta}_S;\hat{\theta}_S} - p_{\bar{\theta}_S;\bar{\theta}_S}]. \tag{17}$$

Notice that the first, second, third, and fourth 16 elements of $\sum_{\theta\in\Theta} b_\theta e_\theta$ are the same. Let β be the row vector corresponding to its first 16 elements. we know from expression (17) that $\tilde{M}^T y^T = (\beta,\beta,\beta,\beta)^T$ has a non-zero solution y, which is a vector of 32 dimensions. This further implies that $M^T x^T = O_{48\times 1}$ has a non-zero solution x, which contradicts with the fact that matrix M has rank 32.

Proof of Remark 1. The **if** direction is trivial. We establish the **only if** direction. Let τ be efficient within the coalition. First, it is easy to see that

$$\sum_{i\in S} \underline{U}_i(\tau) \leqslant \min_{t\in T} \sum_{i\in S} \sum_{\theta_{-S}\in\Theta_{-S}} [u_i(q(\theta_S,\theta_{-S}),(\theta_S,\theta_{-S})) + t_i(\theta_S,\theta_{-S})]p(\theta_{-S}|\theta_S).$$

To prove that the equality holds, suppose by way of contradiction that $<$ holds instead.

Define a new within coalition transfer $\hat{\tau}$ such that

$$\hat{\tau}_i(t,\theta_{-S}) = \frac{1}{|S|}\sum_{i\in S}[u_i(q(\theta_S,\theta_{-S}),(\theta_S,\theta_{-S})) + t_i(\theta_S,\theta_{-S})] - u_i(q(\theta_S,\theta_{-S}),(\theta_S,\theta_{-S})) - t_i(\theta_S,\theta_{-S})$$

for each $\theta_{-S}\in\Theta_{-S}$, $i\in S$, and $t\in T$. We thus have

$$\begin{aligned} &\sum_{\theta_{-S}\in\Theta_{-S}} [u_i(q(\theta_S,\theta_{-S}),(\theta_S,\theta_{-S})) + t_i(\theta_S,\theta_{-S}) + \hat{\tau}_i(t,\theta_{-S})]p(\theta_{-S}|\theta_S) \\ =&\frac{1}{|S|}\sum_{\theta_{-S}\in\Theta_{-S}} [u_i(q(\theta_S,\theta_{-S}),(\theta_S,\theta_{-S})) + t_i(\theta_S,\theta_{-S})]p(\theta_{-S}|\theta_S)\ \forall t\in T. \end{aligned}$$

It follows immediately that

$$\sum_{i\in S}\underline{U}_i(\hat{\tau}) = \min_{t\in T}\sum_{i\in S}\sum_{\theta_{-S}\in\Theta_{-S}}[u_i(q(\theta_S,\theta_{-S}),(\theta_S,\theta_{-S}))+t_i(\theta_S,\theta_{-S})]p(\theta_{-S}|\theta_S).$$

Hence, there exists a vector $(c_i)_{i\in S}\in\mathbb{R}^{|S|}$ such that $\sum_{i\in S} c_i = 0$ and $c_i + \underline{U}_i(\hat{\tau}) > \underline{U}_i(\tau)$ for all $i\in S$. Then define $\tau^*(t,\theta_{-S}) = \hat{\tau}(t,\theta_{-S}) + c_i$ for all $t\in T$ and $\theta_{-S}\in\Theta_{-S}$, we know that $\underline{U}_i(\tau^*) > \underline{U}_i(\tau)$ for all $i\in S$. This contradicts with the supposition that τ is efficient within the coalition. □

Lemma 3. *When the CBDP property fails, there exists a profile of utility functions and an ex-post efficient allocation rule q, such that q is not implementable via an interim CIC and ex-post BB ambiguous mechanisim.*

Proof. When the CBDP property fails, there exists a coalition $S\in\mathcal{S}^*$ and a pair type profiles $\bar{\theta}_S\neq\hat{\theta}_S$ such that $p(\cdot|\bar{\theta}_S) = p(\cdot|\hat{\theta}_S)$. Assume without loss of generality that an agent $i\in S$ has $\bar{\theta}_i\neq\hat{\theta}_i$. Consider the feasible set of outcomes A, agents' utility functions $(u_i)_{i\in I}$, and the efficient allocation rule q in the proof of Theorem 1 except that $0 < a < \frac{n-|S|}{n-1}B$.

From $IC(\bar{\theta}_S;\hat{\theta}_S)$, $IC(\hat{\theta}_S;\bar{\theta}_S)$, we know the following two inequalities should hold:

$$\begin{aligned}
&a|S| + \min_{t\in T}[\sum_{j\in S}\sum_{\theta_{-S}\in\Theta_{-S}} t_j(\bar{\theta}_S,\theta_{-S})p(\theta_{-S}|\bar{\theta}_S)]\\
\geqslant& a|S| + \tfrac{n-|S|}{n-1}B + \min_{t\in T}[\sum_{j\in S}\sum_{\theta_{-S}\in\Theta_{-S}} t_j(\hat{\theta}_S,\theta_{-S})p(\theta_{-S}|\bar{\theta}_S)],\\
&a|S| + \min_{t\in T}[\sum_{j\in S}\sum_{\theta_{-S}\in\Theta_{-S}} t_j(\hat{\theta}_S,\theta_{-S})p(\theta_{-S}|\hat{\theta}_S)]\\
\geqslant& a(|S|-1) + \min_{t\in T}[\sum_{j\in S}\sum_{\theta_{-S}\in\Theta_{-S}} t_j(\bar{\theta}_S,\theta_{-S})p(\theta_{-S}|\hat{\theta}_S)].
\end{aligned}$$

Adding the two expressions up and taking into account that $p(\cdot|\bar{\theta}_S) = p(\cdot|\hat{\theta}_S)$, we have $a\geqslant\frac{n-|S|}{n-1}B$, a contradiction. □

Lemma 4. *If the CBDP condition holds, then for any $S\in\mathcal{S}^*$ and $\bar{\theta}_S$, there exists $\phi^{\bar{\theta}_S}:\Theta\to\mathbb{R}^n$ such that,*

1. $\sum_{i\in I}\phi_i^{\bar{\theta}_S}(\theta) = 0$ *for all* $\theta\in\Theta$,

2. $\sum_{i \in C} \sum_{\theta_{-C} \in \Theta_{-C}} \phi_i^{\bar{\theta}_S}(\theta_C, \theta_{-C}) p(\theta_{-C}|\theta_C) = 0$ *for all non-grand coalition* $C \subseteq I$ *and type profile* $\theta_C \in \Theta_C$*;*

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

Proof. Let the SBDP condition hold under the information structure. Suppose by way of contradiction that there exists $S \in \mathcal{S}^*$ and $\bar{\theta}_S$ such that there does not exist a transfer rule $\phi^{\bar{\theta}_S} : \Theta \to \mathbb{R}^n$ satisfying the three conditions. By Motzkin's transposition theorem, there exists a vector $(a_{\bar{\theta}_C})_{\text{non-grand coalition } C, \bar{\theta}_C \in \Theta_C}$, vector $(b_\theta)_{\theta \in \Theta}$, and a non-zero non-negative vector $(c_{\hat{\theta}_S})_{\hat{\theta}_S \neq \bar{\theta}_S}$, such that

$$\sum_{\text{non-grand coalition } C} \sum_{\theta_C \in \Theta_C} a_{\theta_C} p_{\theta_C;\theta_C} + \sum_{\theta \in \Theta} b_\theta e_\theta = \sum_{\hat{\theta}_S \neq \bar{\theta}_S} c_{\hat{\theta}_S} p_{\bar{\theta}_S;\hat{\theta}_S}, \tag{18}$$

where the notations are defined in Lemma 1.

Step 1. Define $c_{\bar{\theta}_S} = 0$. We want to prove that for each $\theta \in \Theta$ with c_{θ_S}=0, $i \in S$, $j \notin S$, $\theta'_i \neq \theta_i$, and $\theta'_j \neq \theta_j$,

$$c_{\theta'_i, \theta_{S \setminus \{i\}}} [\frac{p(\theta_{-S}|\theta_S)}{p(\theta_{-S}|\theta'_i, \theta_{S \setminus \{i\}})} - \frac{p(\theta'_j, \theta_{-S \cup \{j\}}|\theta_S)}{p(\theta'_j, \theta_{-S \cup \{j\}}|\theta'_i, \theta_{S \setminus \{i\}})}] = 0.$$

Consider the following four type profiles $(\theta_i, \theta_j, \theta_{-i-j})$, $(\theta_i, \theta'_j, \theta_{-i-j})$, $(\theta'_i, \theta_j, \theta_{-i-j})$, and $(\theta'_i, \theta'_j, \theta_{-i-j})$, and two agents i and j.

Since $c_{\theta_S} = 0$, the vector $\sum_{\hat{\theta}_S \neq \bar{\theta}_S} c_{\hat{\theta}_S} p_{\bar{\theta}_S;\hat{\theta}_S}$'s n dimensions corresponding to type profiles $(\theta_i, \theta_j, \theta_{-i-j})$ or $(\theta_i, \theta'_j, \theta_{-i-j})$ and agent i are equal to zero. Since $j \notin S$, this vector's n dimensions corresponding to any type profile and agent j are equal to zero.

For each $\tilde{\theta} \in \Theta$ and coalition $K \subseteq I$, define

$$A_K(\tilde{\theta}) \equiv \sum_{\text{non-grand coalition } C \supseteq K} a_{\tilde{\theta}_C} \frac{p(\tilde{\theta})}{p(\tilde{\theta}_C)}.$$

For an agent $k \in I$, denote $A_{\{k\}}(\tilde{\theta}) = A_k(\tilde{\theta})$ for simplicity. In addition, define

$$\delta_{ij}(\tilde{\theta}) \equiv A_i(\tilde{\theta}) - A_j(\tilde{\theta}).$$

Hence, when $\tilde{\theta} = (\theta_i, \theta_j, \theta_{-i-j})$, or $(\theta_i, \theta'_j, \theta_{-i-j})$, we know that $A_i(\tilde{\theta}) = -b_{\tilde{\theta}} = A_j(\tilde{\theta})$, which implies that

$$\delta_{ij}(\tilde{\theta}) \equiv A_i(\tilde{\theta}) - A_j(\tilde{\theta}) = 0. \tag{19}$$

When $\tilde{\theta} = (\theta'_i, \theta_j, \theta_{-i-j})$, or $(\theta'_i, \theta'_j, \theta_{-i-j})$, we know that $A_i(\tilde{\theta}) - c_{(\theta'_i, \theta_{S\setminus\{i\}})} \dfrac{p(\theta_S, \tilde{\theta}_{-S})}{p(\theta_S)} = -b_{\tilde{\theta}} = A_j(\tilde{\theta})$, which implies that

$$\delta_{ij}(\tilde{\theta}) \equiv A_i(\tilde{\theta}) - A_j(\tilde{\theta}) = c_{(\theta'_i, \theta_{S\setminus\{i\}})} \frac{p(\tilde{\theta}_{-S}|\theta_S)}{p(\tilde{\theta})}. \tag{20}$$

By the definition of each $\delta_{ij}(\tilde{\theta})$, we know that

$$\begin{aligned}
&\delta_{ij}(\theta_i, \theta_j, \theta_{-i-j}) - \delta_{ij}(\theta_i, \theta'_j, \theta_{-i-j}) - \delta_{ij}(\theta'_i, \theta_j, \theta_{-i-j}) + \delta_{ij}(\theta'_i, \theta'_j, \theta_{-i-j}) \\
=& A_i(\theta_i, \theta_j, \theta_{-i-j}) - A_j(\theta_i, \theta_j, \theta_{-i-j}) - A_i(\theta_i, \theta'_j, \theta_{-i-j}) + A_j(\theta_i, \theta'_j, \theta_{-i-j}) \\
&- A_i(\theta'_i, \theta_j, \theta_{-i-j}) + A_j(\theta'_i, \theta_j, \theta_{-i-j}) + A_i(\theta'_i, \theta'_j, \theta_{-i-j}) - A_j(\theta'_i, \theta'_j, \theta_{-i-j}) \\
=& [A_i(\theta_i, \theta_j, \theta_{-i-j}) - A_i(\theta_i, \theta'_j, \theta_{-i-j})] + [A_i(\theta'_i, \theta'_j, \theta_{-i-j}) - A_i(\theta'_i, \theta_j, \theta_{-i-j})] \\
&+ [A_j(\theta'_i, \theta_j, \theta_{-i-j}) - A_j(\theta_i, \theta_j, \theta_{-i-j})] + [A_j(\theta_i, \theta'_j, \theta_{-i-j}) - A_j(\theta'_i, \theta'_j, \theta_{-i-j})] \\
=& \sum_{C \subseteq I\setminus\{i,j\}} a_{\theta_i, \theta_j, \theta_C} \frac{p(\theta_i, \theta_j, \theta_{-i-j})}{p(\theta_i, \theta_j, \theta_C)} - \sum_{C \subseteq I\setminus\{i,j\}} a_{\theta_i, \theta'_j, \theta_C} \frac{p(\theta_i, \theta'_j, \theta_{-i-j})}{p(\theta_i, \theta'_j, \theta_C)} \\
&+ \sum_{C \subseteq I\setminus\{i,j\}} a_{\theta'_i, \theta'_j, \theta_C} \frac{p(\theta'_i, \theta'_j, \theta_{-i-j})}{p(\theta'_i, \theta'_j, \theta_C)} - \sum_{C \subseteq I\setminus\{i,j\}} a_{\theta'_i, \theta_j, \theta_C} \frac{p(\theta'_i, \theta_j, \theta_{-i-j})}{p(\theta'_i, \theta_j, \theta_C)} \\
&+ \sum_{C \subseteq I\setminus\{i,j\}} a_{\theta'_i, \theta_j, \theta_C} \frac{p(\theta'_i, \theta_j, \theta_{-i-j})}{p(\theta'_i, \theta_j, \theta_C)} - \sum_{C \subseteq I\setminus\{i,j\}} a_{\theta_i, \theta_j, \theta_C} \frac{p(\theta_i, \theta_j, \theta_{-i-j})}{p(\theta_i, \theta_j, \theta_C)} \\
&+ \sum_{C \subseteq I\setminus\{i,j\}} a_{\theta_i, \theta'_j, \theta_C} \frac{p(\theta_i, \theta'_j, \theta_{-i-j})}{p(\theta_i, \theta'_j, \theta_C)} - \sum_{C \subseteq I\setminus\{i,j\}} a_{\theta'_i, \theta'_j, \theta_C} \frac{p(\theta'_i, \theta'_j, \theta_{-i-j})}{p(\theta'_i, \theta'_j, \theta_C)} = 0.
\end{aligned}$$

From expressions (19) and (20), we further know that

$$\begin{aligned}
&\delta_{ij}(\theta_i, \theta_j, \theta_{-i-j}) - \delta_{ij}(\theta_i, \theta'_j, \theta_{-i-j}) - \delta_{ij}(\theta'_i, \theta_j, \theta_{-i-j}) + \delta_{ij}(\theta'_i, \theta'_j, \theta_{-i-j}) \\
=& - c_{(\theta'_i, \theta_{S\setminus\{i\}})} \frac{p(\theta_{-S}|\theta_S)}{p(\theta'_i, \theta_{S\setminus\{i\}}, \theta_{-S})} + c_{(\theta'_i, \theta_{S\setminus\{i\}})} \frac{p(\theta'_j, \theta_{-S\cup\{j\}}|\theta_S)}{p(\theta'_i, \theta_{S\setminus\{i\}}, \theta'_j, \theta_{-S\cup\{j\}})}.
\end{aligned}$$

Hence, if we multiply both sides of the above expression by $p(\theta'_i, \theta_{S\setminus\{i\}})$, we have

$$c_{\theta'_i, \theta_{S\setminus\{i\}}} [\frac{p(\theta_{-S}|\theta_S)}{p(\theta_{-S}|\theta'_i, \theta_{S\setminus\{i\}})} - \frac{p(\theta'_j, \theta_{-S\cup\{j\}}|\theta_S)}{p(\theta'_j, \theta_{-S\cup\{j\}}|\theta'_i, \theta_{S\setminus\{i\}})}] = 0.$$

Step 2. Notice that θ_{-S} and $(\theta'_j, \theta_{-S\cup\{j\}})$ only differ in one agent. For each $\theta \in \Theta$ with $c_{\theta_S} = 0$, by repeating the argument in Step 1 recursively, we know that for each pair of $\theta_{-S} \neq \theta'_{-S}$,

$$c_{\theta'_i, \theta_{S\setminus\{i\}}} [\frac{p(\theta_{-S}|\theta_S)}{p(\theta_{-S}|\theta'_i, \theta_{S\setminus\{i\}})} - \frac{p(\theta'_{-S}|\theta_S)}{p(\theta'_{-S}|\theta'_i, \theta_{S\setminus\{i\}})}] = 0.$$

By the SBDP property, we know that $c_{\theta'_i, \theta_{S\setminus\{i\}}} = 0$.

Step 3. Notice that θ_S and $\theta_{\theta'_i, \theta_{S\setminus\{i\}}}$ only differ in one agent. Apply the argument of Step 2 recursively, we can prove that $c_{\hat{\theta}_S} = 0$ for all $\hat{\theta}_S \neq \bar{\theta}_S$, which contradict the fact that $(c_{\hat{\theta}_S})_{\hat{\theta}_S \neq \bar{\theta}_S}$ is non-zero. □

Proof of Theorem 2. **Statement 2** $\Rightarrow$ **Statement 3**. This direction is trivial.

Statement 3 $\Rightarrow$ **Statement 1**. This direction follows directly from Lemma 3.

Statement 1 $\Rightarrow$ **Statement 2**. Given any profile of utility functions and an ex-post efficient allocation rule q, we define

$$\eta_i(\theta) = \frac{1}{n} \sum_{j \in I} u_j(q(\theta), \theta) - u_i(q(\theta), \theta).$$

Pick a sufficiently large constant $M > 0$ such that

$$\begin{aligned} &\sum_{\theta_{-i} \in \Theta_{-i}} [u_i(q(\bar{\theta}_i, \theta_{-i}), (\bar{\theta}_i, \theta_{-i})) + \eta_i(\bar{\theta}_i, \theta_{-i})] p(\theta_{-i}|\bar{\theta}_i) \\ \geqslant &\sum_{\theta_{-i} \in \Theta_{-i}} [u_i(q(\hat{\theta}_i, \theta_{-i}), (\bar{\theta}_i, \theta_{-i})) + \eta_i(\hat{\theta}_i, \theta_{-i})] p(\theta_{-i}|\bar{\theta}_i) - M \sum_{\theta_{-i} \in \Theta_{-i}} \phi_i^{\theta_C}(\hat{\theta}_i, \theta_{-i}) p(\theta_{-i}|\bar{\theta}_i) \end{aligned}$$

for any $i \in I$, pair of $\bar{\theta}_i \neq \hat{\theta}_i$, non-grand coalition C, and $\theta_C \in \Theta_C$, where each ϕ^{θ_C} satisfies the conditions stated in Lemma 4.

Let $T = \{\eta + M\phi^{\theta_C} : C \in \mathcal{S}^*, \theta_C \in \Theta_C\}$. We verify that the ambiguous mechanism (q, T) satisfies the conditions of interim CIC, interim CR, and ex-post BB.

We first establish the ex-post BB condition. For each $t \in T$, t can be expressed as $\eta + M\phi^{\theta_C}$. Notice that both η and ϕ^{θ_C} satisfy the ex-post BB condition. Hence, t also satisfies the condition.

To verify the interim CIC condition for a coalition $S \in \mathcal{S}$, consider a coalition $S \in \mathcal{S}$ with type profile $\bar{\theta}_S$. After sharing their information within the coalition and truthful report their types to the MD, the maxmin aggregated expected utility of members in S is

$$\begin{aligned}
&\min_{t\in T}\sum_{i\in S}\sum_{\theta_{-S}\in\Theta_{-S}}[u_i(q(\bar{\theta}_S,\theta_{-S}),(\bar{\theta}_S,\theta_{-S}))+t_i(\bar{\theta}_S,\theta_{-S})]p(\theta_{-S}|\bar{\theta}_S)\\
=&\min_{C\in\mathcal{S}^*,\theta_S\in\Theta_S}\sum_{i\in S}\sum_{\theta_{-S}\in\Theta_{-S}}[\frac{1}{n}\sum_{j\in I}u_j(q(\bar{\theta}_S,\theta_{-S}),(\bar{\theta}_S,\theta_{-S}))+M\phi_i^{\theta_C}(\bar{\theta}_S,\theta_{-S})]p(\theta_{-S}|\bar{\theta}_S)\\
=&\min_{C\in\mathcal{S}^*,\theta_S\in\Theta_S}\sum_{\theta_{-S}\in\Theta_{-S}}[\frac{|S|}{n}\sum_{j\in I}u_j(q(\bar{\theta}_S,\theta_{-S}),(\bar{\theta}_S,\theta_{-S}))+\sum_{i\in S}M\phi_i^{\theta_C}(\bar{\theta}_S,\theta_{-S})]p(\theta_{-S}|\bar{\theta}_S)\\
=&\frac{|S|}{n}\sum_{j\in I}u_j(q(\bar{\theta}_S,\theta_{-S}),(\bar{\theta}_S,\theta_{-S}))p(\theta_{-S}|\bar{\theta}_S)\geqslant 0, \qquad (21)
\end{aligned}$$

where the first equality uses the definition of T, the second equality moves the summation across $i\in S$ into the square bracket, and the third equality applies the ex-post BB condition of each ϕ^{θ_C}. This establishes the interim CR condition.

To verify the interim CIC condition, we first notice that by ex-post efficiency of q, it is impossible that the grand coalition can profit from deviating. Hence, it suffices to focus on deviation of a coalition $S\in\mathcal{S}^*$. The maxmin aggregated expected utility for type-$\bar{\theta}_S$ coalition S to adopt strategy δ_S is

$$\begin{aligned}
&\min_{t\in T}\sum_{i\in S}\sum_{\hat{\theta}_S\in\Theta_S}\sum_{\theta_{-S}\in\Theta_{-S}}[u_i(q(\hat{\theta}_S,\theta_{-S}),(\bar{\theta}_S,\theta_{-S}))+t_i(\hat{\theta}_S,\theta_{-S})]p(\theta_{-S}|\bar{\theta}_S)\delta_S(\bar{\theta}_S)[\hat{\theta}_S]\\
\leqslant&\sum_{i\in S}\sum_{\hat{\theta}_S\in\Theta_S}\sum_{\theta_{-S}\in\Theta_{-S}}[u_i(q(\hat{\theta}_S,\theta_{-S}),(\bar{\theta}_S,\theta_{-S}))+\eta_i(\hat{\theta}_S,\theta_{-S})+M\phi_i^{\bar{\theta}_S}(\hat{\theta}_S,\theta_{-S})]p(\theta_{-S}|\bar{\theta}_S)\delta_S(\bar{\theta}_S)[\hat{\theta}_S]\\
\leqslant&\frac{|S|}{n}\sum_{j\in I}u_j(q(\bar{\theta}_S,\theta_{-S}),(\bar{\theta}_S,\theta_{-S}))p(\theta_{-S}|\bar{\theta}_S)\\
=&\min_{t\in T}\sum_{i\in S}\sum_{\theta_{-S}\in\Theta_{-S}}[u_i(q(\bar{\theta}_S,\theta_{-S}),(\bar{\theta}_S,\theta_{-S}))+t_i(\bar{\theta}_S,\theta_{-S})]p(\theta_{-S}|\bar{\theta}_S), \qquad (22)
\end{aligned}$$

where the first inequality comes from the fact that $\eta+M\phi^{\bar{\theta}_S}\in T$, the second inequality follows from the choice of M, and the equality follows from expression (21). Hence, the interim CIC condition holds. □

References

Aoyagi, M. (1998). Correlated types and Bayesian incentive compatible mechanisms with budget balance. *Journal of Economic Theory*, 79(1):142–151.

Aumann, R. (1959). Acceptable points in general cooperative n-person games. *Contributions to the Theory of Games (AM-4)*, 4.

Bose, S. and Daripa, A. (2009). A dynamic mechanism and surplus extraction under ambiguity. *Journal of Economic Theory*, 144(5):2084–2114.

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.

Carroll, G. (2019). Robustness in mechanism design and contracting. *Annual Review of Economics*. Forthcoming.

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

Che, Y.-K. and Kim, J. (2009). Optimal collusion-proof auctions. *Journal of Economic Theory*, 144(2):565–603.

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

Chung, K.-S. (1999). A note on Matsushima's regularity condition. *Journal of Economic Theory*, 87(2):429–433.

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

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*.

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

Forges, F., Mertens, J.-F., and Vohra, R. (2002). The ex ante incentive compatible core in the absence of wealth effects. *Econometrica*, 70(5):1865–1892.

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

Guo, H. (2018). Mechanism design with ambiguous transfers. Working paper.

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*, pages 875–911.

Laffont, J.-J. and Martimort, D. (1998). Collusion and delegation. *The Rand Journal of Economics*, pages 280–305.

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

Liu, H. (2018). Collusion-proof dynamic mechanisms. Working paper.

Lopomo, G., Rigotti, L., and Shannon, C. (2019). Uncertainty and robustness of surplus extraction. Working Paper.

Luo, X., Yongchuan, Q., and Yang, C.-C. (2017). Bayesian coalitional rationalizability and rationalizability. Working Paper.

Matsushima, H. (1991). Incentive compatible mechanisms with full transferability. *Journal of Economic Theory*, 54(1):198–203.

Matsushima, H. (2007). Mechanism design with side payments: Individual rationality and iterative dominance. *Journal of Economic Theory*, 133(1):1–30.

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.

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

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

Song, Y. (2016). Efficient implementation with interdependent valuations and maxmin agents. Working paper.

Tang, R. and Zhang, M. (2018). Maxmin implementation. Working paper.

Wilson, R. (1978). Information, efficiency, and the core of an economy. *Econometrica: Journal of the Econometric Society*, pages 807–816.

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