Coalition-proof Ambiguous Mechanism

Huiyi Guo*

August 7, 2021

Abstract

The paper studies when efficient allocations are implementable via coalition-proof ambiguous mechanisms. We consider two coalition-proofness notions. When the mechanism designer is restricted to use simple mechanisms, under every prior, there exists an efficient allocation that is not coalition-proof implementable under either of our coalition-proofness notions. However, when ambiguous mechanisms are allowed and agents are maxmin expected utility maximizers, we prove that all efficient allocations are coalition-proof implementable if and only if the prior distribution of agents' types satisfies the Coalition Beliefs Determine Preferences (CBDP) property. This result holds under both of our coalition-proofness notions. We then extend our method to obtain a positive result on coalition-proof full surplus extraction and to incorporate a model with alternative ambiguity preferences. Thus, the paper sheds light on how coalition-proofness may be achieved by engineering ambiguity in mechanism rules.

Keywords: Coalition-proofness; Ambiguous mechanism; Implementation; Correlated information; Coalition incentive compatibility; Full surplus extraction

JEL: C71; C72; D81; D82

*Texas A&M University, Department of Economics, 4228 TAMU, College Station, TX 77843. Email address: huiyiguo@tamu.edu. The author is grateful for helpful discussions with Samson Alva, Subir Bose, Gabriel Carroll, Jimmy Chan, Hülya Eraslan, Wei He, Rich McLean, Xiaosheng Mu, Harry Pei, Mike Peters, Miklós Pintér, Mikhail Safronov, Vasiliki Skreta, Chris Shannon, Sergei Severinov, Balázs Szentes, as well as participants at the Workshop on Microeconomic Theory at the Chinese University of Hong Kong, 2019 Asian Meeting of the Econometric Society, 2019 China Meeting of the Econometric Society, the 19th SAET Conference on Current Trends in Economics, the 30th International Conference on Game Theory, 2019 Texas Economic Theory Camp, and 2020 North American Winter Meeting of the Econometric Society.

1 Introduction

In mechanism design theory, most works assume that agents behave non-cooperatively when revealing their private information to the mechanism designer. However, there are many real-life mechanisms, such as auctions, matching, and voting, where coalition manipulations arise as common practices.[1] When agents can gain by jointly manipulate their private information, imposing individual incentive constraints alone may not ensure the implementation of the mechanism designer' desired outcome. Thus, we would like to design mechanisms free from coalition manipulations.

The current paper aims to achieve coalition-proof implementation of efficient allocations with the assistance of a broad set of tools called ambiguous mechanisms. Ambiguous mechanisms have vaguely specified rules. Facing an ambiguous mechanism, agents do not know the true rule that will be enforced and have no information on the probability for each potential rule to be realized. In reality, one can interpret some mechanisms as ambiguous ones, for example, secret reserve price auctions and vague tax audit schemes. In our theoretical model, to implement an efficient allocation rule, the mechanism designer can engineer ambiguity by secretly committing to a simple mechanism (composed of the exogenous allocation rule and a transfer rule), but strategically announcing a set of potential simple mechanisms as in Bose and Renou (2014), Di Tillio et al. (2017), Guo (2019), Tang and Zhang (2021), etc. Starting with Ellsberg (1961), it has been argued that decision-makers often exhibit ambiguity aversion when lacking information to form a unique probability about the underlying uncertainty. We thus assume that agents are ambiguity-averse and make reporting decisions with the maxmin expected utility as in Gilboa and Schmeidler (1989). The paper will show that ambiguous mechanisms can facilitate coalition-proof efficient implementation by taking advantage of agents' ambiguity aversion.

We adopt two coalition-proofness notions in the paper. The stronger notion, the interim coalition incentive compatibility (CIC) condition, assumes that all coalitions may be formed and implicitly views each coalition as a pseudo agent. This pseudo agent's type is the profile of types of its members and the pseudo agent's utility is the aggregate utility of its members. According to this condition, no coalition can increase its utility by having its members coordinately misreport their private information. As every individual is a singleton coalition, the interim CIC condition is more demanding than imposing individual incentive constraints only. This coalition-proofness notion is inspired by the strong Nash equilibrium under complete information and is similar to the notions of Green and Laffont (1979), Chen and Micali

[1] See Che and Kim (2006) for a review of the literature that studies how agents collude against exogenously given institutions.

(2012), and Safronov (2018) among others under incomplete information, except that we consider maxmin expected utility of each coalition rather than ex-post utility or subjective expected utility. Our second coalition-proofness notion, the interim weak coalition incentive compatibility (WCIC) condition, allows all coalitions to be formed, but assumes that members of a coalition can only transmit information via an ambiguous collusive mechanism. This condition requires that no coalition can implement a profitable joint deviation via an incentive compatible and budget balanced ambiguous collusive mechanism. Since the information asymmetry within a coalition restricts the scope of coalition manipulations, there are weakly less coalition manipulations to worry about from the mechanism designer's perspective, which makes the interim WCIC condition weaker than the interim CIC condition. Our interim WCIC condition is related to coalition-proofness notions adopted by Laffont and Martimort (1997, 2000), Che and Kim (2006), etc.

To further explore the relationship between the two coalition-proofness notions, Proposition 1 shows that if an ambiguous mechanism violates the interim CIC condition but the prior satisfies the Belief Determine Preferences (BDP) property, then there exists a coalition and an ambiguous collusive mechanism within the coalition to implement a profitable collusive deviating strategy. The collusive mechanism is incentive compatible and budget balanced, and thus the main mechanism does not satisfy the interim WCIC condition. Under some additional condition, we can also make the collusive mechanism individually rational.

When only one potential mechanism rule is announced, an ambiguous mechanism degenerates into a simple mechanism. To motivate the use of ambiguous mechanisms, we explore what happens if the mechanism designer is restricted to use simple mechanisms only. Our Proposition 2 shows that regardless of the joint distribution of agents' types, there exists a payoff structure under which an efficient allocation rule is not implementable via a coalition-proof simple mechanism, under either of our coalition-proofness notions. This negative result contrasts with previous works by, for example, Crémer and McLean (1985, 1988) and Kosenok and Severinov (2008), which impose no coalition-proofness requirement but obtain positive results on full surplus extraction or efficient implementation under a broad set of priors with correlated beliefs. This negative result motivates the use of non-simple mechanisms to address the coalition-proof implementation problem.

In contrast with the negative result under simple mechanisms, coalition-proof efficient implementation is possible under ambiguous mechanisms. In Theorem 1, we characterize the set of all priors under which coalition-proof efficient implementation can be guaranteed. In particular, we show that all efficient allocation rules under all payoff structures are implementable via coalition-proof ambiguous mechanisms, if and only if the prior satisfies the Coalition Beliefs Determine Preferences (CBDP) property. This is true under both coalition-proofness notions adopted by the paper. In fact, when the CBDP property holds, one can design an ambiguous mechanism not only satisfying coalition-proofness, but also satisfying the interim coalition rationality condition (which ensures that all coalitions are willing to participate), the interim ambiguity insurance condition (which requires that on the equilibrium path, each coalition receives the same interim utility across different potential simple mechanisms), and the ex-post budget balance condition. The CBDP property strengthens the Beliefs Determine Preferences (BDP) property in the literature and requires that for any non-grand coalition, its different type profiles form distinct posterior beliefs towards the type profiles of agents out of the coalition. In any fixed finite type space, the CBDP property is a weak condition, although it rules out independent beliefs and entails correlation. Hence, we can claim that there is a broad set of priors where ambiguous mechanisms can help to fulfill coalition-proof efficient implementation that is impossible via simple mechanisms.

The reason why ambiguous mechanisms may outperform simple ones in coalition-proof efficient implementation is that different simple mechanisms in an ambiguous mechanism may weaken different coalitions' incentive to misreport. In the ambiguous mechanism we construct, a coalition earns the same on-path interim utility from all potential simple mechanisms. But when a coalition misreports, there is at least one unprofitable simple mechanism, which prevents the misreporting from being profitable. It is worth noting that our ambiguous mechanism is immune from deviation in mixed strategies, and thus the critique that mixed strategies can hedge against ambiguity does not apply to our mechanism.

We consider two extensions of our main results. First, we show that coalition-proof full surplus extraction can be guaranteed via ambiguous mechanisms if and only if the CBDP property holds. Then, we consider a less extreme ambiguity model, the smooth ambiguity model, and show that when all coalitions are ambiguity averse, the interim CIC condition can still be guaranteed under the CBDP property.

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

To begin with, the paper is related to the literature on mechanism design under correlated beliefs. The papers by Crémer and McLean (1985, 1988) and McAfee and Reny (1992) characterize the set of priors under which full surplus extraction can be guaranteed. A few later papers study how weak/strong the condition for full surplus extraction is, among which, Neeman (2004) introduces the BDP property as a necessary condition. The implementation question in our paper is related to d'Aspremont et al. (2004), McLean and Postlewaite (2004, 2015), Matsushima (2007), and Kosenok and Severinov (2008), among others. These papers obtain positive results on implementing efficient allocations via individually rational and/or budget balanced mechanisms beyond independent belief environments. In particular, Kosenok and Severinov (2008) characterize priors under which all efficient allocations are implementable via individually rational and budget balanced mechanisms. The methodology adopted in the current paper is related to the above-mentioned papers, especially the one of Crémer and McLean (1988) and Kosenok and Severinov (2008): we also rely on a theorem of the alternative to establish the existence of a transfer rule satisfying some incentive constraints. A key difference between the current paper and all above works is that we impose coalition-proofness requirements that are not considered by the papers above. Moreover, we show that without ambiguous mechanisms, coalition-proof implementation is hard to achieve even if agents' beliefs are correlated, which motivates the use of ambiguous mechanisms.

The paper also adds to the literature on coalition-proof mechanisms. One approach in this literature considers all possible coalitions, imposes coalition-proofness requirement on the mechanism axiomatically, and ignores strategic interactions due to information asymmetry within the coalition. Our interim CIC condition, and papers including Bennett and Conn (1977), Green and Laffont (1979), Krasa and Yannelis (1994), Chen and Micali (2012), Bierbrauer and Hellwig (2016), Safronov (2018), etc, follow this approach. Overall, the axiomatic approach leads to strong coalition-proofness notions, but it provides a benchmark to study coalition-proof implementation since the worst-case scenario from the mechanism designer's perspective is that agents can collude without facing within-coalition incentive constraints. However, there are environments where positive results can be obtained. For example, Bierbrauer and Hellwig (2016) show that the only belief-free coalition-proof mechanism for public good provision under private valuations is a voting mechanism. In a private value auction setup where agents partition themselves into arbitrarily many coalitions, Chen and Micali (2012) design an indirect mechanism where each agent truthfully reports the coalition he belongs to and the type profile of this coalition. Our interim CIC condition is directly related to the coalition-proofness notion of Safronov (2018), except that a coalition in the current paper faces multiple mechanism rules and uses maxmin expected utility. In a private value environment with independent beliefs, Safronov (2018) shows that a redesigned expected externality mechanism (which is budget balanced) can meet the additional coalition-proofness requirement. Proposition 2 of the current paper establishes a negative result on coalition-proof implementation via simple mechanisms beyond private valuations, regardless of whether the beliefs are independent or correlated. Then Theorem 1 provides positive results under ambiguous mechanisms when the prior satisfies the CBDP property. Our results thus complement the above-mentioned papers and show that coalition-proofness may not be hard to obtain in environments with correlated beliefs and any payoff structure, if the mechanism designer is allowed to engineer ambiguity in the mechanism rules.

Another approach to study coalition-proof mechanism design explicitly considers within-coalition information asymmetry and requires the collusive mechanism itself to be incentive compatible. Information friction within a coalition may undermine the coalition's ability to collude, which leads to weaker coalition-proofness notions compared to the ones studied under the axiomatic approach above. Laffont and Martimort (1997, 2000) and Che and Kim (2006), as well as our interim WCIC condition, follow this approach. Laffont and Martimort (1997, 2000) find that in a two-agent environment, coalition-proofness is not costly to obtain if agents have independent beliefs, but it is costly if agents have correlated beliefs. Che and Kim (2006) consider a multiple-agent environment where a certain coalition (but not all coalitions) can collude. They generalize the positive result on coalition-proof mechanism design under independent beliefs and identify a group of sufficient conditions for coalition-proof mechanism design under correlated beliefs. An observation that is new in this paper is that under the CBDP property, even if all coalitions can consider deviating via collusive mechanisms, efficient implementation and full surplus extraction are still possible via ambiguous mechanisms. Our result holds in all environments with at least two agents.

At last, the paper fits in the literature on mechanism design with ambiguity-averse agents. Some works in the literature explore if it is possible to strategically engineer ambiguity in the mechanism so that the performance of an ambiguous mechanism is better than that of a simple one. Bose and Renou (2014) embed an ambiguous communication device to 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. Bose and Daripa (2017a,b) study how the principal can adopt engineered ambiguous acts to elicit an agent's ambiguous belief in the α-MEU model and the smooth ambiguity model. Tang and Zhang (2021) show that ambiguous mechanisms are more potent in implementing social choice correspondences than simple mechanisms. In a game theory setup, Pintér (2021) introduces a mathematical construction with which players of a game can make ambiguous strategies. Within this branch of the literature, the current work is most related to Guo (2019), which shows that full surplus extraction and efficient implementation via ambiguous mechanisms be guaranteed if and only if the Beliefs Determine Preferences (BDP) property holds. Despite a close relationship between the BDP and the CBDP properties, results in the current paper are neither corollaries nor trivial extensions of the earlier results. The technical difficulty comes from the fact that all coalitions can be formed: overlapping coalitions make it challenging to incentivize truthful reporting of one coalition without affecting incentives of any other coalition. Also, the current paper considers the use of mixed strategies, which is a nontrivial question in presence of ambiguity but omitted by many of the above works, including Guo (2019). To the best of our knowledge, the current paper is the first one that studies how strategic ambiguity in mechanisms can be introduced to deter coalition manipulations.

In some other works, agents are assumed to hold ambiguous beliefs about other agents' private information exogenously, and the MD designs simple mechanisms for efficient implementation or revenue maximization. For example, Wolitzky (2016), De Castro et al. (2017), Song (2018, 2020), De Castro and Yannelis (2018), Kocherlakota and Song (2019), and Guo and Yannelis (2020, 2021) show that the conflict between efficiency and incentive compatibility may be softened when agents have ambiguous beliefs. Bose et al. (2006), Bose and Daripa (2009), Bodoh-Creed (2012), and Lopomo et al. (2020) show how ambiguous beliefs bring new insights to revenue maximization mechanisms or full surplus extraction. In the current paper, we do not assume that agents hold ambiguous beliefs about other agents' private information, which differentiates the current paper from the above-mentioned ones. Notice that our paper assumes that the prior is common knowledge between agents and the MD, which is different from the robust mechanism design works where the MD has no information (e.g., Bergemann and Morris (2005), Bierbrauer and Hellwig (2016)) or limited/imprecise information (e.g., Ollár and Penta (2017), Lopomo et al. (2021)) on agents' belief structure.

The rest of the paper proceeds as follows. Section 2 sets up the model and formalizes ambiguous mechanisms and ambiguous collusive mechanisms. Section 3 provides an example of ambiguous mechanism. The main results of the paper are presented in Section 4. In Section 5, we consider extensions on coalition-proof full surplus extraction and on alternative ambiguity aversion models. Section 6 concludes. All proofs are relegated to the Appendix.

2 Set-up

We study an environment with one mechanism designer (MD) and finitely many agents. For convenience, assume that the MD is female and each agent is male. The set of all agents is denoted by I, which has cardinality n with $2 \leqslant n < +\infty$, and an agent is indexed by i.

Each agent i privately observes his type $\theta_i \in \Theta_i$. Let Θ_i be agent i's type set and $\Theta \equiv \prod_{i \in I} \Theta_i$ be the type space. 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 $\theta \in \Theta$. An important aspect of the paper is that we consider profitable deviations of coalitions. A coalition, denoted by S, is a nonempty set of agents in I: for example, $S = \{i\}$ is a singleton coalition where agent i is the only member; $S = I$ is the grand coalition. Given that a non-grand coalition S has type profiles $\theta_S \equiv (\theta_i)_{i \in S} \in \Theta_S \equiv \prod_{i \in S} \Theta_i$, we let $p(\cdot|\theta_S) \equiv (p(\theta_{-S}|\theta_S))_{\theta_{-S} \in \Theta_{-S}}$ denote the posterior belief of types of agents out of the coalition, where $p(\theta_{-S}|\theta_S) \equiv p(\theta)/p(\theta_S)$, and $p(\theta_S) \equiv \sum_{\theta'_{-S} \in \Theta_{-S}} p(\theta_S, \theta'_{-S})$ represents the marginal distribution of prior p on θ_S. In the special case that $S = \{i\}$, $p(\cdot|\theta_i)$ is the belief of type-θ_i agent over other agents' types.

Each agent's quasi-linear utility function is of the form: $u_i(a, \theta) + b_i$, where $a \in A$ is an element in the set of feasible outcomes, θ is the profile of all agents' types, $u_i(a, \theta)$ is agent i's payoff from the outcome a, and $b_i \in \mathbb{R}$ is the monetary transfer received by i. Notice that the utility functions may have interdependent valuations.

An **allocation rule** $q : \Theta \to A$ is a plan to assign outcomes contingent on agents' type profiles. In this paper, the MD wishes to implement an exogenously given allocation rule q satisfying the ex-post **efficiency** condition, i.e., one such that $\sum_{i \in I} u_i\big(q(\theta), \theta\big) \geqslant \sum_{i \in I} u_i\big(a, \theta\big)$ for all $a \in A$ and $\theta \in \Theta$.[2] We assume that there is a feasible outcome giving each agent zero payoff. Hence, an ex-post efficient q has non-negative ex-post social surplus $\sum_{i \in I} u_i\big(q(\theta), \theta\big) \geqslant 0$ for all $\theta \in \Theta$.

2.1 Ambiguous Mechanism

The paper focuses on the use of direct mechanisms. We allow the MD to use a broad set of tools called ambiguous mechanisms. Formally, an **ambiguous mechanism** to implement q is a pair (q, T), where T is a compact set of potential transfer rules. An element of T is denoted by $t \equiv (t_i : \Theta \to \mathbb{R})_{i \in I}$. The MD secretly commits to an arbitrary transfer rule $t \in T$ and publicly announces the allocation rule q as well as the set of potential transfer rules T. The use of ambiguous mechanisms by the MD engineers ambiguity to agents. We assume that each agent has the maxmin expected utility (MEU) of Gilboa and Schmeidler (1989) and only cares about his worst-case interim utility across all potential transfer rules.[3] When the set of transfer rules T is a singleton, there is no non-trivial ambiguity. In this case, the ambiguous mechanism reduces to a **simple mechanism**, and an agent's interim utility is consistent with the subjective expected utility.

To prevent a coalition from collectively manipulating information submitted to the MD, we adopt the interim coalition incentive compatibility condition as one of our coalition-proofness notions. When a coalition S is formed, this coalition-proofness notion views the coalition as a pseudo agent. In particular, when each agent $i \in S$ has type θ_i, the pseudo agent's type is θ_S, his utility is the aggregate utilities of all members in S, and he forms the interim belief $p(\cdot|\theta_S)$ towards types of agents out of the coalition. When the pseudo agent follows the strategy $\delta_S : \Theta_S \to \Delta(\Theta_S)$ to misreport (the probability for type-θ_S coalition S to report θ'_S is denoted by $\delta_S[\theta_S](\theta'_S)$) and when the true transfer rule is $t \in T$, the coalition's

[2] In the main parts of the paper, we assume that the MD does not care about her revenue. In Section 5, we consider the revenue problem and study coalition-proof full surplus extraction via ambiguous mechanisms.

[3] Under this extreme version of maxmin expected utility model, agents care about the worst-case transfer rule only. Section 6 relaxes this assumption and explores alternative ambiguity aversion models.

expected utility is equal to

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

where $\theta \equiv (\theta_S, \theta_{-S})$. When $S = I$, the above expression should read as

$$V_I[\delta_I](t,\theta) \equiv \sum_{i\in I}\sum_{\theta'\in\Theta}[u_i\big(q(\theta'),\theta\big)+t_i(\theta')]\delta_I[\theta](\theta')$$

since there is no agent out of the coalition I. Notice that δ_S should be viewed as a strategy of the pseudo agent rather than a profile of strategies of agents in S because the reported types of agents in S could be correlated. In the special case that $\delta^*_S : \Theta_S \to \Delta(\Theta_S)$ is the distribution satisfying $\delta^*_S[\theta_S](\theta_S) = 1$ for all $\theta_S \in \Theta_S$, i.e., agents in S truthfully report, we denote $V_S[\delta^*_S](t,\theta_S)$ by $V_S(t,\theta_S)$ for simplicity. This coalition-proofness notion mandates that no coalition can jointly manipulate members' types and earn a higher MEU.

Formally, the ambiguous mechanism (q,T) is said to satisfy the interim **coalition incentive compatibility** (CIC) condition, if

$$\min_{t\in T} V_S(t,\theta_S) \geqslant \min_{t\in T} V_S[\delta_S](t,\theta_S)$$

for any $S \in 2^I\backslash\{\varnothing\}$, $\theta_S \in \Theta_S$, and $\delta_S : \Theta_S \to \Delta(\Theta_S)$.

We have a few remarks on the above condition.

This coalition-proofness notion views coalitions as pseudo agents, similar to those of Bennett and Conn (1977), Green and Laffont (1979), Krasa and Yannelis (1994), Chen and Micali (2012), and Safronov (2018), except that our pseudo agents are MEU maximizers. In particular, when (q,T) is a simple mechanism, the interim CIC condition coincides with the coalition-proofness notion of Safronov (2018). This notion implicitly imposes a few assumptions: first, members of a coalition can make side payments to each other to hedge against the individual-level ambiguity and thus perceive the same worst-case transfer rule; in addition, members of a coalition disclose their private information to each other without encountering incentive issues. As strategic interactions within a coalition may undermine the coalition's incentive to collude, our interim CIC condition imposes a strong stability requirement on the mechanism. In Section 2.2, we introduce a weaker coalition-proofness notion by allowing for strategic interaction between members of a coalition.

As all coalitions, including singletons, should truthfully report, our interim CIC condition is stronger than the interim **incentive compatibility** (IC) condition, which can be defined by requiring the inequality in the interim CIC condition to hold for all singleton coalitions.

Deviation in mixed strategies of a coalition S is explicitly taken into account in the interim CIC condition. In fact, when an ambiguous mechanism is adopted, focusing on pure strategy incentive compatibility is not without loss of generality, as agents may follow mixed strategies to hedge against ambiguity.

We also remark that agents cannot predict the transfer rule adopted by the MD. In our problem, it is common knowledge that the MD's goal is merely to implement the exogenous allocation rule q. As long as q can be implemented with the help of ambiguity on transfer rules, the MD has no incentive to choose one transfer rule over another.

Other than the interim CIC condition, the ambiguous mechanism we construct in our Theorem 1 also satisfies the three conditions defined below.

The ambiguous mechanism (q, T) is said to satisfy the interim **coalition rationality** (CR) condition if $\min_{t \in T} V_S(t, \theta_S) \geqslant 0$ for all coalition $S \in 2^I \backslash \{\varnothing\}$ and type profile $\theta_S \in \Theta_S$. Namely, under any potential transfer rule, the interim utility of each coalition is non-negative on the equilibrium path. As this inequality holds for all singleton coalitions, the familiar interim **individual rationality** (IR) condition is implied by the interim CR condition.

The ambiguous mechanism (q, T) satisfies the interim **ambiguity insurance** (AI) condition if for each $S \in 2^I \backslash \{\varnothing\}$, $V_S(t, \theta_S)$ is independent of $t \in T$. Under this condition, it is as if coalitions perceive no ambiguity on the equilibrium path. Notice that simple mechanisms satisfy the interim AI condition trivially.

The ambiguous mechanism (q, T) is said to satisfy the ex-post **budget balance** (BB) condition if $\sum_{i \in I} t_i(\theta) = 0$ for all $\theta \in \Theta$ and $t \in T$. This condition requires that each potential transfer rule in the ambiguous mechanism is budget balanced. We remark that when the allocation rule q is ex-post efficient and the ambiguous mechanism (q, T) satisfies the ex-post BB condition, the grand coalition I cannot benefit from misreporting.

2.2 Ambiguous Collusive Mechanism

In Section 2.1, we adopt the interim CIC condition as a strong axiomatic coalition-proofness notion. One may wonder if the interim CIC condition is too demanding when agents can only interact with each other via a third-party collusive mechanism. In this section, we formalize such an ambiguous collusive mechanism. If collusive deviations are only implementable via incentive compatible and budget balanced ambiguous collusive mechanisms, then some unimplementable collusive deviations, albeit profitable, should not be of a concern to the main mechanism (q, T). Thus, the MD can consider a weaker coalition-proofness notion than the interim CIC condition. In Section 4, we will further investigate the relationship between the two coalition-proofness notions and show that the weaker coalition-proofness notion implies the stronger notion under a mild restriction on the common prior.

Before types are reported in the main mechanism (q, T), a third-party mechanism designer (mediator) can approach a coalition and offer an ambiguous collusive mechanism. An **ambiguous collusive mechanism** of coalition S is a pair (δ_S, T^S), where $\delta_S : \Theta_S \to \Delta(\Theta_S)$ is the collusive deviating strategy of coalition S in (q, T), and the set T^S represents a set of potential $\tau \equiv (\tau_j : T \times \Theta \to \mathbb{R})_{j \in S}$ that further redistributes wealth within S and incentivizes truthful revelation within S. It is not surprising that each τ is contingent on θ_S, since the mediator tries to elicit the truthful θ_S from members of S. Allowing each $\tau \in T^S$ to be contingent on $\theta_{-S} \in \Theta_{-S}$ reported in the main mechanism can take better advantage of potential information correlation across all agents. Allowing τ to be contingent on $t \in T$ helps members in S to (at least partially) hedge against the individual level ambiguity induced by the main mechanism. After $\theta_S \in \Theta_S$ is elicited in the collusive mechanism and $t \in T$ and $\theta_{-S} \in \Theta_{-S}$ are realized from the main mechanism, the mediator reveals $\tau \in T^S$ and redistributes wealth within S accordingly. Given $t \in T$ and $\tau \in T^S$, when type-θ_i agent $i \in S$ misreports $\hat{\theta}_i$ in the collusive mechanism, not only the redistribution assigned by τ is affected, so are the allocation and transfer assigned in the main mechanism: the main mechanism assigns to θ_i the expected utility of

$$V_i[\delta_S](t, \theta_i, \hat{\theta}_i) \equiv \sum_{\theta'_S \in \Theta_S} \sum_{\theta_{-i} \in \Theta_{-i}} [u_i\big(q(\theta'_S, \theta_{-S}), \theta\big) + t_i(\theta'_S, \theta_{-S})] p(\theta_{-S}|\theta_S) \delta_S[\hat{\theta}_i, \theta_{S \setminus \{i\}}](\theta'_S),$$

and the collusive mechanism assigns to θ_i the additional expected transfer

$$\sum_{\theta_{-i} \in \Theta_{-i}} \tau_i(t, (\hat{\theta}_i, \theta_{-i})) p(\theta_{-i}|\theta_i).$$

Members in S should be provided the correct incentive to report in (δ_S, T^S) in order to implement the collusive deviating strategy δ_S. Given an ambiguous mechanism (q, T), we say the ambiguous collusive mechanism (δ_S, T^S) satisfies the interim ***S* incentive compatibility** (S-IC) condition if for all agent $i \in S$, true type $\theta_i \in \Theta_i$, and potentially mixed deviating strategy $\sigma_i : \Theta_i \to \Delta(\Theta_i)$,

$$\begin{aligned} \min_{\substack{t \in T, \\ \tau \in T^S}} \{V_i[\delta_S](t, \theta_i, \theta_i) + \sum_{\theta_{-i} \in \Theta_{-i}} \tau_i(t, (\theta_i, \theta_{-i})) p(\theta_{-i}|\theta_i)\} \\ \geqslant \min_{\substack{t \in T, \\ \tau \in T^S}} \{ \sum_{\hat{\theta}_i \in \Theta_i} [V_i[\delta_S](t, \theta_i, \hat{\theta}_i) + \sum_{\theta_{-i} \in \Theta_{-i}} \tau_i(t, (\hat{\theta}_i, \theta_{-i})) p(\theta_{-i}|\theta_i)] \sigma_i[\theta_i](\hat{\theta}_i)\}. \end{aligned}$$

It is reasonable to assume that the mediator offers no financial assistance to the coalition in the ambiguous collusive mechanism. The ambiguous collusive mechanism (δ_S, T^S) is said to satisfy the ex-post ***S* budget balance** (S-BB) condition if $\sum_{i \in S} \tau_i(t, \theta) = 0$ for all $t \in T$, $\theta \in \Theta$, and $\tau \in T^S$.

We say the ambiguous mechanism (q, T) satisfies the interim **weak coalition incentive compatibility** (WCIC) condition, if there does not exist any coalition $S \in 2^I \setminus \{\varnothing\}$, $\theta_S \in \Theta_S$, and ambiguous collusive mechanism (δ_S, T^S) satisfying the interim S-IC condition and the ex-post S-BB condition such that

$$\min_{t\in T} V_S[\delta_S](t,\theta_S) > \min_{t\in T} V_S(t,\theta_S).$$

From the definitions of the two coalition-proofness notions, it is clear that if (q,T) satisfies the interim CIC condition, then it also satisfies the interim WCIC condition.

One may also impose participation constraints on the collusive mechanism. Given an ambiguous mechanism (q,T), an ambiguous collusive mechanism (δ_S, T^S) of coalition S is said to satisfy the interim ***S* individual rationality** (S-IR) condition if for all $i \in S$ and $\theta_i \in \Theta_i$,

$$\min_{\substack{t\in T,\\ \tau\in T^S}} \{V_i[\delta_S](t,\theta_i,\theta_i) + \sum_{\theta_{-i}\in\Theta_{-i}} \tau_i(t,\theta)p(\theta_{-i}|\theta_i)\} \geqslant \min_{t\in T} V_i(t,\theta_i).$$

To directly connect the two coalition-proofness notions, we do not impose the S-IR condition on the collusive mechanism in the definition of the interim WCIC condition. However, Propositions 1 and 2 provide conditions under which the profitable ambiguous collusive mechanism also satisfies the S-IR condition.

3 An Example

In this section, we focus on the stronger coalition-proofness notion adopted by the paper and provide an example where the efficient allocation rule q is not implementable by any simple mechanism satisfying the interim CIC condition, but it is implementable by an ambiguous mechanism satisfying the interim CIC condition.

There are three agents in the environment. The MD can choose a feasible public good provision level $a \in A = \{0,1\}$. Assume the production of the public good is costless. Each agent $i \in I = \{1,2,3\}$ privately observes a type in the set $\Theta_i = \{\theta_i^H, \theta_i^L\}$, where $\theta_1^H = 5$, $\theta_1^L = 1$, $\theta_j^H = 2$, $\theta_j^L = 1$ for $j \in \{2,3\}$. The prior $p \in \Delta(\Theta)$ is given in the table below. Agents' payoff structure is given as follows: $u_1(a,\theta) = \theta_1 a$ and $u_j(a,\theta) = (\theta_j - \theta_1)a$ for all $a \in A$, $\theta \in \Theta$, and $j \in \{2,3\}$. Notice that agent 1 has private valuation; for agent $j \in \{2,3\}$, j's utility function depends on his own type, but also negatively depends on agent 1's type.

$p(\theta_1^H,\theta_2^H,\theta_3^H)$	$p(\theta_1^H,\theta_2^H,\theta_3^L)$	$p(\theta_1^H,\theta_2^L,\theta_3^H)$	$p(\theta_1^H,\theta_2^L,\theta_3^L)$
0.2	0.15	0.1	0.05
$p(\theta_1^L,\theta_2^H,\theta_3^H)$	$p(\theta_1^L,\theta_2^H,\theta_3^L)$	$p(\theta_1^L,\theta_2^L,\theta_3^H)$	$p(\theta_1^L,\theta_2^L,\theta_3^L)$
0.05	0.1	0.15	0.2

Table 1: common prior p

The ex-post social surplus is equal to $(\theta_1 + (\theta_2 - \theta_1) + (\theta_3 - \theta_1))a = (\theta_2 + \theta_3 - \theta_1)a$. Hence, the efficient allocation rule q is given by $q(\theta) = 1$ if and only if $\theta_1 = \theta_1^L$, and $q(\theta) = 0$ elsewhere. Notice that agent 1 has a conflict of interest with the social surplus, but his report determines the public good provision level completely. The payoff structure and the allocation rule q are deliberately constructed to be straightforward. In spite of the straightforward q, we prove the following claim in the Appendix.

Claim 1. *There does not exist any simple mechanism (q, t) satisfying the interim CIC condition.*

As such, one has to seek alternative solutions to achieve interim CIC. One solution is to adopt ambiguous mechanisms. In this example, one can come up with multiple ways to construct an ambiguous mechanism (q, T) that satisfies not only the interim CIC condition but also conditions of interim CR, interim AI, and ex-post BB. Below we present one way with four potential transfer rules: $T = \{t^1, t^2, t^3, t^4\}$.

In the ambiguous mechanism, $t^1 \equiv (t_1^1, t_2^1, t_3^1)$ and $t^3 \equiv (t_1^3, t_2^3, t_3^3) \in T$ are given in the tables below. The other two transfer rules are defined by $t^2 = -t^1$ and $t^4 = -t^3$.

$t^1(\theta_1^H, \theta_2^H, \theta_3^H)$	$t^1(\theta_1^H, \theta_2^H, \theta_3^L)$	$t^1(\theta_1^H, \theta_2^L, \theta_3^H)$	$t^1(\theta_1^H, \theta_2^L, \theta_3^L)$
$(18.75, -11.25, -7.5)$	$(-5, -5, 10)$	$(7.5, 22.5, -30)$	$(-75, 15, 60)$
$t^1(\theta_1^L, \theta_2^H, \theta_3^H)$	$t^1(\theta_1^L, \theta_2^H, \theta_3^L)$	$t^1(\theta_1^L, \theta_2^L, \theta_3^H)$	$t^1(\theta_1^L, \theta_2^L, \theta_3^L)$
$(-75, 15, 60)$	$(7.5, 22.5, -30)$	$(-5, -5, 10)$	$(18.75, -11.25, -7.5)$

Table 2: transfer rule t^1

$t^3(\theta_1^H, \theta_2^H, \theta_3^H)$	$t^3(\theta_1^H, \theta_2^H, \theta_3^L)$	$t^3(\theta_1^H, \theta_2^L, \theta_3^H)$	$t^3(\theta_1^H, \theta_2^L, \theta_3^L)$
$(12, 0, -12)$	$(16, -32, 16)$	$(-24, 0, 24)$	$(-48, 96, -48)$
$t^3(\theta_1^L, \theta_2^H, \theta_3^H)$	$t^3(\theta_1^L, \theta_2^H, \theta_3^L)$	$t^3(\theta_1^L, \theta_2^L, \theta_3^H)$	$t^3(\theta_1^L, \theta_2^L, \theta_3^L)$
$(-48, 48, 0)$	$(-24, 24, 0)$	$(16, -16, 0)$	$(12, -12, 0)$

Table 3: transfer rule t^3

Claim 2. *The ambiguous mechanism (q, T) satisfies the conditions of interim CIC, interim CR, interim AI, and ex-post BB.*

To demonstrate the mechanics of (q, T), we verify the interim CIC constraints for type-θ_1^H agent 1 and type-(θ_1^H, θ_2^L) coalition $\{1, 2\}$. The complete proof is relegated to the Appendix.

For type-θ_1^H agent 1, his on-path interim utility under t^1 is equal to

$$0.4 \cdot (0 + 18.75) + 0.3 \cdot (0 - 5) + 0.2 \cdot (0 + 7.5) + 0.1 \cdot (0 - 75) = 0.$$

Similarly, his interim utility levels under t^2, t^3, and t^4, are all equal to 0. Thus, his on-path MEU is equal to 0. The MEU of misreporting θ_1^L is $\min\{-21.875, 31.875, -17, 27\} = -21.875 < 0$. It is easy to see that deviating in mixed strategies is not profitable either. As a result, the interim CIC constraints of θ_1^H are satisfied.

Type-(θ_1^H, θ_2^L) coalition $\{1, 2\}$ has an on-path interim utility of

$$\frac{2}{3}\big((0 + 7.5) + (0 + 22.5)\big) + \frac{1}{3}\big((0 - 75) + (0 + 15)\big) = 0$$

under transfer rule t^1. Similarly, we can verify that the coalition's interim utility under each $t \in T$ is equal to 0 and thus its MEU is equal to 0. Then notice that its MEU levels by misreporting (θ_1^H, θ_2^H), (θ_1^L, θ_2^H), and (θ_1^L, θ_2^L) are:

$$\begin{aligned}
&\min\{1.667, -1.667, 2.667, -2.667\} = -2.667, \\
&\min\{-29, 31, 1, 1\} = -29, \text{and} \\
&\min\{-3.167, 5.167, 1, 1\} = -3.167
\end{aligned}$$

respectively. One can also check that there is no profitable misreporting in mixed strategies here, since no mixed strategy can make the coalition's interim utility higher than 0 for each $t \in T$. Hence, the interim CIC constraints of this coalition are also satisfied.

We remark that when different coalitions or different types of the same coalition misreport in different ways, the worst-case transfer rules may not be the same. One can hence claim that different interim CIC constraints are satisfied by different transfer rules. For the existence of an interim coalition incentive compatible simple mechanism, there has to be one transfer rule that simultaneously satisfying all CIC constraints, which could be too demanding. Since ambiguous mechanisms can relax this restriction, ambiguous mechanisms may outperform simple mechanisms in guaranteeing the interim CIC condition.

4 Results

In this section, we present three results. Proposition 1 further explores the relationship between the two coalition-proofness notions adopted by the paper. Proposition 2 shows an impossibility result on coalition-proof implementation via simple mechanisms. Theorem 1 shows a possibility result on coalition-proof implementation via ambiguous mechanisms.

The results of the paper rely on two important properties of the common prior p. The Beliefs Determine Preferences (BDP) property is introduced by Neeman (2004), which requires that for every agent, different types have different beliefs over other agents' types. We introduce the Coalition Beliefs Determine Preferences (CBDP) property, which is a strengthening of the BDP property and requires that for every non-grand coalition, different type profiles should form distinct posterior beliefs over types of agents out of the coalition.

Definition 1. *For a common prior p, we say*

1. *the **Beliefs Determine Preferences** (BDP) property holds if $p(\cdot|\theta_i) \neq p(\cdot|\theta'_i)$ for each $i \in I$ and each pair of types θ_i, $\theta'_i \in \Theta_i$ with $\theta_i \neq \theta'_i$.*

2. *the **Coalition Beliefs Determine Preferences** (CBDP) property holds if $p(\cdot|\theta_S) \neq p(\cdot|\theta'_S)$ for all non-grand coalition S and type profiles $\theta_S, \theta'_S \in \Theta_S$ with $\theta_S \neq \theta'_S$.*

When agents have independent beliefs, the prior p neither satisfies the BDP property nor the CBDP property. That said, the BDP property and the CBDP property impose weak restrictions on the prior over the fixed finite type space Θ: among all distributions over Θ, the ones for which the BDP or CBDP property fails constitute a null set.[4]

The prior in the motivating example satisfies the CBDP property. For example, when coalition $\{1, 2\}$ has type profile (θ_1^H, θ_2^L) verses (θ_1^H, θ_2^H), the coalition believes that it is more likely that agent 3 has type θ_3^H.

We first examine the relationship between our two coalition-proofness notions. Recall that by definition, the interim CIC condition implies the interim WCIC condition. Under the BDP property, we further show that if the interim CIC condition is not satisfied by (q, T), then the interim WCIC condition also fails, and thus interim CIC and interim WCIC are equivalent. This is because a coalition S can use an incentive compatible and budget balanced ambiguous collusive mechansim to implement a profitable collusive deviation. Moreover, if (q, T) also satisfies the interim AI condition, then the collusive mechanism can be designed to satisfy the interim S-IR condition.

Proposition 1. *Suppose the BDP property holds for the prior p.*

1. *An ambiguous mechanism (q, T) satisfies the interim CIC condition, if and only if it satisfies the interim WCIC condtion.*

[4] When one relaxes the assumption of a fixed finite type space and looks at the universal type space instead, the literature has documented various results on if the BDP property is generic. For instance, Heifetz and Neeman (2006) have shown that the BDP property is non-generic geometrically and measure-theoretically, and Chen and Xiong (2011) show that the BDP property is generic topologically.

2. *If an ambiguous mechanism (q, T) satisfies the interim AI condition but not the interim CIC condition, then there exists a coalition S, a type profile θ_S, and an ambiguous collusive mechanism (δ_S, T^S) satisfying the conditions of interim S-IC, ex-post S-BB, and interim S-IR such that* $\min_{t \in T} V_S[\delta_S](t, \theta_S) > \min_{t \in T} V_S(t, \theta_S)$.

The proof is relegated to the Appendix.

We remark that when the BDP property does not hold, there are payoff environments under which efficient allocations are not implementable via incentive compatible ambiguous mechanisms even if no coalition-proofness requirement is imposed. This can be seen from the proof of the current Lemma 3 (Case 2).

We then introduce a negative result on coalition-proof implementation via simple mechanisms. Given any prior, with independent beliefs or not, there always exists a payoff structure and an efficient allocation rule, such that the allocation rule is not implementable via coalition-proof simple mechanisms. This is different from the positive results on full surplus extraction or efficient implementation under correlated beliefs (see, for example, Crémer and McLean (1985, 1988) and Kosenok and Severinov (2008)).

Proposition 2. *Under any prior p, there exists a profile of quasi-linear utility functions and and an ex-post efficient allocation rule q such that for any simple mechanism (q, t), there exists a coalition S, a type profile θ_S, and an ambiguous collusive mechanism (δ_S, T^S) satisfying the conditions of interim S-IC, ex-post S-BB, and interim S-IR such that $V_S[\delta_S](t, \theta_S) > V_S(t, \theta_S)$.*

Since the collusive mechanism above satisfies the S-IC and S-BB conditions, the simple mechanism (q, t) does not satisfy the interim WCIC condition, and a fortiori, it does not satisfy the interim CIC condition. According to this proposition, under any p, there always exists at least one payoff structure under which at least one ex-post efficient allocation rule is not implementable via a coalition-proof simple mechanism. Notice that the proposition does not rule out the possibility that some efficient allocation rule under some payoff structure is implementable via coalition-proof simple mechanisms. For example, Safronov (2018) has demonstrated that under a prior with *independent beliefs*, every ex-post efficient allocation rule *under private valuation utility functions* is implementable via a redesigned expected externality mechanism, which is a simple mechanism satisfying the conditions of interim CIC and ex-post BB. Safronov (2018)'s mechanism does not apply to the situation with interdependent valuations or correlated beliefs though.

The formal proof of the proposition is relegated to the Appendix. We briefly explain why simple mechanisms may not guarantee the interim CIC condition here. The interim CIC condition requires the simple mechanism (q, t) to satisfy the CIC constraints of all coalitions, which overlap with each other. Thus, even if the MD can construct a transfer rule to incentivize the truthful reporting of one coalition by following proper scoring rule (see, for example, Börgers et al. (2015)) or the approach in Crémer and McLean (1985, 1988), it is impossible to do so without affecting the incentives of other coalitions that overlap with this coalition. In the proof, we fix any agent $i \in I$ and divide I into complementary coalitions: $\{i\}$ and $I\backslash\{i\}$. We then construct a payoff structure with interdependent valuations and an efficient allocation rule q, so that there is no simple mechanism under which coalitions $\{i\}$, $I\backslash\{i\}$, and I have the incentive to truthfully report. This implies that q is not implementable via a simple mechanism satisfying the interim CIC condition.

The negative result of Proposition 2 calls for alternative approaches to solve the coalition-proof implementation problem. We offer ambiguous mechanisms as a solution.

Theorem 1. *Given any prior p, the following statements are equivalent:*

1. *The CBDP property holds for p.*
2. *Any ex-post efficient allocation rule under any profile of quasi-linear utility functions is implementable via an ambiguous mechanism that satisfies the interim WCIC condition.*
3. *Any ex-post efficient allocation rule under any profile of quasi-linear utility functions is implementable via an ambiguous mechanism satisfying the interim CIC condition.*
4. *Any ex-post efficient allocation rule under any profile of quasi-linear utility functions is implementable via an ambiguous mechanism satisfying the conditions of interim CIC, interim CR, interim AI, and ex-post BB.*

According to the first three statements of the theorem, the CBDP property is the necessary and sufficient condition on the prior to ensure coalition-proof implementation of all efficient allocation rules via ambiguous mechanisms. Similar to Proposition 2, when the CBDP property does not hold, there exists some payoff structure under which some efficient allocation rule is not implementable via coalition-proof ambiguous mechanisms, although coalition-proof implementable ones may also exist. But when the CBDP property is satisfied, in contrast to the negative result in Proposition 2, all efficient allocation rules are coalition-proof implementable via ambiguous mechanisms. In this sense, ambiguous mechanisms outperform simple mechanisms.

The equivalence between the third and fourth statements of the theorem may give readers the impression that the interim CR, interim AI, and ex-post BB conditions are obtained for free under ambiguous mechanisms. We emphasize that this is correct only when the CBDP property holds. For instance, Safronov (2018) has shown that in an independent private value setting (where the CBDP property fails), every efficient allocation is implementable via an interim coalition incentive compatible and ex-post budget balanced simple mechanism (a degenerate ambiguous mechanism). In the same setting, it is easy to show that the additional interim IR condition, which is implied by the interim CR condition, cannot be guaranteed even if ambiguous mechanisms are used.

4.1 Sketch of the Proof of Theorem 1

We establish the result by proving that Statement 4 ⇒ Statement 3 ⇒ Statement 2 ⇒ Statement 1 ⇒ Statement 4. The proof that Statement 4 ⇒ Statement 3 ⇒ Statement 2 is trivial. To prove that Statement 2 ⇒ Statement 1, we assume that the CBDP property fails and then construct a payoff structure with interdependent valuations and an efficient allocation rule that is not implementable via any interim WCIC ambiguous mechanism. The main difficulty is in proving Statement 1 ⇒ Statement 4. We sketch a construction below.[5]

Lemma 4 in the Appendix shows that if the CBDP property holds, then for each non-grand coalition S with type profile $\bar{\theta}_S$, there exists a transfer rule $\phi^{\bar{\theta}_S} \equiv (\phi_i^{\bar{\theta}_S} : \Theta \to \mathbb{R})_{i \in I}$ satisfying the following three conditions. First, $\phi^{\bar{\theta}_S}$ satisfies the ex-post BB condition. Second, $\phi^{\bar{\theta}_S}$ gives every non-grand coalition (not only S or singleton coalitions) with every possible type profile zero interim transfer on the equilibrium path. Third, $\phi^{\bar{\theta}_S}$ gives type-$\bar{\theta}_S$ coalition S negative interim transfer whenever this coalition misreports in pure strategies. We emphasize that $\phi^{\bar{\theta}_S}$ itself is not an element of our set of potential transfer rules T, but will be used to construct an element of T later. According to the third condition, $\bar{\theta}_S$ also earns negative interim transfer under $\phi^{\bar{\theta}_S}$ when $\bar{\theta}_S$ deviates in mixed strategies, which is important when we demonstrate the interim CIC condition of the ambiguous mechanism. We remark that $\phi^{\bar{\theta}_S}$ does not impose restrictions on any other type/agent's off-path interim transfers. The first two conditions above will be used to establish the conditions of ex-post BB, interim CR, and interim AI of our ambiguous mechanism.

To complete the design of the ambiguous mechanism, consider the following budget balanced transfer rule η defined by $\eta_i(\theta) = \frac{1}{n} \sum_{j \in I} u_j(q(\theta), \theta) - u_i(q(\theta), \theta)$ for all $i \in I$ and $\theta \in \Theta$, which is used to equally distribute ex-post social surplus among all agents. Fix a sufficiently large constant M. Let $T \equiv \{\eta + M\phi^{\bar{\theta}_S} : S \in 2^I \backslash \{\varnothing, I\}, \bar{\theta}_S \in \Theta_S\}$. The mechanism (q, T) satisfies the ex-post BB condition as η and each $\phi^{\bar{\theta}_S}$ are budget balanced.

[5]There are multiple ways to construct an ambiguous mechanism satisfying the conditions in Statement 4. In fact, the ambiguous mechanism presented in Section 3 does not follow mechanically from the general approach we will introduce, but is tailored to have less transfer rules and a simpler exposition, given the specific payoff structure there.

To verify the interim CR and AI conditions, first notice that for each agent $i \in I$, the u_i part of the utility function and the η_i part in the transfer rule give i non-negative ex-post utility on path. Also, recall that for each non-grand coalition S and type profile $\bar{\theta}_S \in \Theta_S$, $\phi^{\bar{\theta}_S}$ gives every coalition zero expected transfer on path. Hence, every coalition's on-path interim utility is non-negative and independent of the transfer rule $t \in T$.

To verify the interim CIC condition, first notice that grand coalition's interim CIC constraints are satisfied trivially due to ex-post efficiency of q and ex-post budget balance of (q, T). Then consider any non-grand coalition S with type profile $\bar{\theta}_S$. When this coalition misreports in pure strategy or in mixed strategy, different transfers in T may give S distinct interim utility levels. However, when M is sufficiently large, the transfer rule $\eta + M\phi^{\bar{\theta}_S}$ gives the coalition a sufficiently low interim utility level due to the third condition of $\phi^{\bar{\theta}_S}$ established in Lemma 4. Since there exists a potential transfer rule under which misreporting is never profitable, the MEU-maximizing coalition has no incentive to misreport.

The key feature of this mechanism is that different ambiguity levels are perceived by coalitions on and off the equilibrium path. No ambiguity is perceived on path due to the interim AI condition. However, off path, the interim utility levels under different transfer rules can be unequal and at least one transfer rule leads to a low interim utility level. The concern that the worst-case transfer rule may be realized prevents coalitions from misreporting.

4.2 Comparison with Guo (2019)

In this section, we compare Proposition 1 and Theorem 1 of the current paper with Theorem 3.1 of Guo (2019). By Theorem 3.1(2), any ex-post efficient allocation rule under any profile of utility functions is implementable via an interim IR and ex-post BB ambiguous mechanism if and only if the BDP property holds for the prior p.

Proposition 1 here does not follow from Theorem 3.1(2). Hypothetically, one might treat a coalition S as a sub-environment, hoping to construct an ambiguous mechanism within S to implement profitable collusive deviations by following Theorem 3.1(2). However, the hypothetical approach does not lead to Proposition 1: if we follow it, then the corresponding BDP property in this sub-environment S requires that for each $i \in S$ and types $\theta_i \neq \theta_i'$, marginal beliefs should satisfy $(p(\theta_{S\setminus\{i\}}|\theta_i))_{\theta_{S\setminus\{i\}} \in \Theta_{S\setminus\{i\}}} \neq (p(\theta_{S\setminus\{i\}}|\theta_i'))_{\theta_{S\setminus\{i\}} \in \Theta_{S\setminus\{i\}}}$. This requirement is stronger than the BDP property we adopt in Proposition 1. In our ambiguous collusive mechanism (δ_S, T^S), each $\tau \in T^S$ transfers money within S, but is allowed to be contingent on information out of S: $\theta_{-S} \in \Theta_{-S}$ and $t \in T$. This additional degree of freedom helps to ensure that the BDP property is sufficient for Proposition 1.

The connection between the CBDP property and the BDP property might misleadingly lead to the impression that Theorem 1 is a trivial extension of Theorem 3.1(2). However, this is not true since coalitions overlap with each other and have interwined incentive constraints. For example, if one views coalitions $\{1, 2\}$ and $\{3\}$ as the only (pseudo) agents in the motivating example, one can indeed follow the earlier paper to construct an incentive compatible, individually rational, and budget balanced ambiguous mechanism between the two (pseudo) agents. However, this mechanism merely specifies aggregate transfers of agents 1 and 2, rather than their respective transfers. Splitting the aggregate transfers in a way satisfying all CIC constraints is not easy, as the MEU of all coalitions that overlap with coalitions $\{1, 2\}$ may be affected. From a technical point of view, Lemma 4 (crucial for Theorem 1) is more involved than Lemma A.4 of Guo (2019) (crucial for Theorem 3.1(2) there). According to the second condition in Lemma 4, the on-path interim transfer under $\phi^{\bar{\theta}_S}$ should be zero for all non-grand coalitions. This is more demanding than imposing the same condition for all individuals or for coalition S only. In fact, when $n \geqslant 3$, the number of non-grand coalitions, $2^n - 2$, can be much larger than the number of agents, n. The complexity due to overlapping coalitions necessitates a new argument that does not exist in the literature and poses technical challenges in establishing Lemma 4.

Besides, the current paper considers deviations in mixed strategies, which are omitted by a few earlier works on ambiguous mechanism design, including Guo (2019), although it is known that mixed strategies may be played by a decision maker to hedge against ambiguity. To discourage misreporting in mixed strategies, the ambiguous mechanism constructed in the current paper consists of more potential transfer rules than that in the previous paper.

5 Extensions

5.1 Full Surplus Extraction

In our efficient implementation problem studied in the main text, revenue is not a concern of the MD. In fact, the ambiguous mechanism is designed to be ex-post budget balanced and thus generates zero revenue to the MD in Theorem 1. However, by modifying the proof of Theorem 1, one can obtain a positive result on coalition-proof full surplus extraction.

Given $(u_i)_{i \in I}$, we say full surplus extraction can be achieved, if there exists an ambiguous mechanism (q, T) satisfying the interim IC and interim IR conditions such that for each $t \in T$,

$$-\sum_{\theta \in \Theta} \sum_{i \in I} t_i(\theta) p(\theta) = \sum_{\theta \in \Theta} \sum_{i \in I} u_i(q(\theta), \theta) p(\theta) = \max_{\tilde{q}: \Theta \to A} \sum_{\theta \in \Theta} \sum_{i \in I} u_i(\tilde{q}(\theta), \theta) p(\theta).$$

Namely, there should exist an ex-post efficient (or equivalently, ex-ante efficient) allocation rule q and a set of transfer rules T such that under the ambiguous mechanism (q, T), the ex-ante revenue raised from each transfer rule is equal to the maximal ex-ante social surplus.

We show that coalition-proof full surplus extraction can be achieved if and only if the CBDP property is satisfied.

Proposition 3. *Given any prior p, the following statements are equivalent:*

1. *The CBDP property holds for p.*
2. *Full surplus extraction under any profile of quasi-linear utility functions can be guaranteed via an ambiguous mechanism that satisfies the interim WCIC condition.*
3. *Full surplus extraction under any profile of quasi-linear utility functions can be guaranteed via an ambiguous mechanism satisfying the interim CIC condition.*

In the ambiguous mechanism (q, T) we construct in the proof, each $t \in T$ raises the same ex-post revenue across all $\theta \in \Theta$, which is equal to the ex-ante social surplus. Thus, (q, T) is not ex-post budget balanced. Under each $t \in T$, the grand coalition's ex-post utility may be positive or negative, although the ex-ante utility is zero. Hence, the interim CR condition is not satisfied by (q, T). However, our construction of (q, T) satisfies the interim AI condition.

The result is related to Theorem 2 and Corollary 2 of Che and Kim (2006) on coalition-proof full surplus extraction, where simple mechanisms are adopted and within-coalition strategic interactions are considered. A main difference between their coalition-proofness notion and our interim WCIC condition is that they only consider collusion of the grand coalition. Under their Condition PI' and the Convex Independence condition of Crémer and McLean (1988), they show that full surplus extraction can be achievable in a way that is immune from manipulations of the grand coalition and individuals. Our CBDP property and their sufficient condition for coalition-proof full surplus extraction do not imply each other, but recall that our mechanism is immune from manipulations of all coalitions. If we assume instead that the grand coalition is the only non-trivial coalition that can be formed, we can modify the proofs of the current paper and show that the BDP property is necessary and sufficient for coalition-proof full surplus extraction/efficient implementation via ambiguous mechanisms. The BDP property is weaker than the Convex Independence condition and relaxes Condition PI', and thus coalition-proof full surplus extraction via ambiguous mechanisms can be obtained under a weaker constraint over p.

5.2 Alternative Preferences

The main parts of the paper adopt a particularly simple form of the MEU, where coalitions only care about the worst-case transfer rule. In this section, we look at a less extreme ambiguity aversion model: the smooth ambiguity model. The key insight that ambiguous mechanisms help to achieve the interim CIC condition under the CBDP property still holds.

In the smooth ambiguity aversion model of Klibanoff et al. (2005), a type-θ_S coalition S facing an ambiguous mechanism (q, T) has the following on-path interim utility:

$$\int_{\pi\in\Pi}\psi\Big(\int_{t\in T}V_S(t,\theta_S)d\pi(t)\Big)d\mu(\pi),$$

where π is a distribution over T representing a subjective belief on the true transfer rule, $\Pi \subseteq \Delta(T)$ is a set of such distributions that are relevant, $\psi : \mathbb{R} \to \mathbb{R}$ is a strictly increasing function characterizing ambiguity attitude, and μ is a subjective probability over Π. The coalition S, viewed as a pseudo agent, is ambiguity-averse (resp. ambiguity-neutral or ambiguity-seeking) if ψ is strictly concave (resp. affine or strictly convex). We assume that the MD does not know the $\theta \in \Theta$ that is realized, but has knowledge about other components of each agent's interim and ex-post utility functions.

For tractability, we focus on a class of smooth ambiguity aversion model. Let $\bar{\pi}$ denote the uniform distribution over T, $\epsilon \in (0, 1]$ be a constant, Π^ϵ be the following set of ϵ-contaminated distributions, $\Pi^\epsilon = \{(1-\epsilon)\bar{\pi} + \epsilon\pi : \pi \in \Delta(T)\}$, $\bar{\mu}$ be the uniform distribution over Π^ϵ, and $\psi : \mathbb{R} \to \mathbb{R}$ be a strictly increasing, strictly concave, and second order differentiable function. Then we have the following corollary.

Corollary 1. *Suppose for any ambiguous mechanism (q, T), coalition S, and type profile θ_S, type-θ_S coalition S has on-path interim utility $\int_{\pi\in\Pi^\epsilon}\psi\big(\int_{t\in T}V_S(t,\theta_S)d\pi(t)\big)d\bar{\mu}(\pi)$. When the CBDP property holds for p, then any ex-post efficient allocation rule under any profile of quasi-linear utility functions is implementable via an ambiguous mechanism satisfying the conditions of interim CIC, interim CR, interim AI, and ex-post BB.*

In the proof, we modify the ambiguous mechanism designed for Theorem 1 so that $T \equiv \{\eta + M\phi^{\theta_C} : C \in 2^I\backslash\{\varnothing, I\}, \theta_C \in \Theta_C\} \cup \{\eta - M\phi^{\theta_C} : C \in 2^I\backslash\{\varnothing, I\}, \theta_C \in \Theta_C\}$, where the scale factor M is a sufficiently large number. The $+M\phi^{\theta_C}$ and $-M\phi^{\theta_C}$ terms give the ambiguous mechanism a symmetric design. The smooth ambiguity aversion preference we adopt gives the worse ones heavier weights and makes misreporting unprofitable.

By Corollary 1, as long as coalitions are ambiguity averse, under the CBDP property, there always exists an ambiguous mechanism to achieve the CIC condition, and there is no restriction on how ambiguity averse coalitions should be. Notice that it is not the case that the same ambiguous mechanism can guarantee the interim CIC condition for all strictly concave ψ. In fact, for less concave ψ, we may need to enlarge the scale factor M to ensure the interim CIC condition. However, in the limiting case that ψ is affine, since coalitions are ambiguity neutral and reduce ambiguous mechanisms to simple mechanisms, ambiguous mechanisms do not have any advantage over simple mechanisms any more.

The corollary shows that the usefulness of ambiguous mechanisms goes beyond the extreme MEU model. There are other ambiguity aversion models under which the corollary holds. For instance, when type-θ_S coalition S has the following (more general) MEU of Gilboa and Schmeidler (1989): $\min_{\pi\in\Pi^\epsilon}\int_{t\in T} V_S(t,\theta_S)d\pi(t)$, or the following α-MEU of Ghirardato and Marinacci (2002): $\alpha\min_{\pi\in\Pi^\epsilon}\int_{t\in T} V_S(t,\theta_S)d\pi(t)+(1-\alpha)\max_{\pi\in\Pi^\epsilon}\int_{t\in T} V_S(t,\theta_S)d\pi(t)$ where $\alpha\in(0.5,1]$, we can adjust the scale factor M in the ambiguous mechanism constructed above to make the corollary go through. In fact, even if different coalitions have different degrees of (strict) ambiguity aversion, or follow different ambiguity aversion models we have discussed above, the corollary still holds.

6 Concluding Remarks

The paper considers two coalition-proofness notions, the interim CIC condition and the interim WCIC condition, and aims to implement efficient allocations via coalition-proof ambiguous mechanisms. The interim CIC condition is stronger in general, but is equivalent to the interim WCIC condition when the BDP property holds for the prior. When only simple mechanisms can be used by the MD, under any prior, there exists an efficient allocation that is not coalition-proof implementable, regardless of the coalition-proofness notion adopted. However, when ambiguous mechanisms are allowed, the Coalition Beliefs Determine Preferences property on the prior is necessary and sufficient to guarantee coalition-proof implementation of efficient allocations. Hence, the use of ambiguous mechanisms can help to fulfill coalition-proof efficient implementation that cannot be achieved under simple mechanisms.

A Appendix

A.1 Details of the Example

Proof of Claim 1. Suppose by way of contradiction that there is a simple mechanism (q,t) satisfying the interim CIC condition. For a coalition S and a pair of type profiles $\theta_S\neq\theta'_S$, we let $CIC(\theta_S;\theta'_S)$ denote the constraint that type-θ_S coalition S does not benefit from misreporting θ'_S. Then (q,t) should at least satisfy the following three groups of CIC constraints.

Coalition $S=\{1\}$ has two pure-strategy CIC constraints. For example, $CIC(\theta_1^H;\theta_1^L)$ requires

$$0.4t_1(\theta_1^H,\theta_2^H,\theta_3^H)+0.3t_1(\theta_1^H,\theta_2^H,\theta_3^L)+0.2t_1(\theta_1^H,\theta_2^L,\theta_3^H)+0.1t_1(\theta_1^H,\theta_2^L,\theta_3^L)$$

$$\geqslant 5 + 0.4t_1(\theta_1^L, \theta_2^H, \theta_3^H) + 0.3t_1(\theta_1^L, \theta_2^H, \theta_3^L) + 0.2t_1(\theta_1^L, \theta_2^L, \theta_3^H) + 0.1t_1(\theta_1^L, \theta_2^L, \theta_3^L).$$

Also, coalition $S = \{2, 3\}$ has twelve pure-strategy CIC constraints. For instance, type-(θ_2^H, θ_3^H) coalition $\{2, 3\}$'s posterior belief is that agent 1 has type θ_1^H with probability 0.8 and type θ_1^L with probability 0.2. It follows from $CIC(\theta_2^H, \theta_3^H; \theta_2^H, \theta_3^L)$ that

$$\begin{aligned} &0.8[t_2(\theta_1^H, \theta_2^H, \theta_3^H) + t_3(\theta_1^H, \theta_2^H, \theta_3^H)] + 0.2[3 + t_2(\theta_1^L, \theta_2^H, \theta_3^H) + t_3(\theta_1^L, \theta_2^H, \theta_3^H)] \\ &\geqslant 0.8[t_2(\theta_1^H, \theta_2^H, \theta_3^L) + t_3(\theta_1^H, \theta_2^H, \theta_3^L)] + 0.2[3 + t_2(\theta_1^L, \theta_2^H, \theta_3^L) + t_3(\theta_1^L, \theta_2^H, \theta_3^L)]. \end{aligned}$$

Furthermore, the grand coalition $S = I$ has fifty-six pure-strategy CIC constraints. For example, constraint $CIC(\theta_1^H, \theta_2^H, \theta_2^H; \theta_1^L, \theta_2^H, \theta_2^H)$ entails that

$$\begin{aligned} &t_1(\theta_1^H, \theta_2^H, \theta_3^H) + t_2(\theta_1^H, \theta_2^H, \theta_3^H) + t_3(\theta_1^H, \theta_2^H, \theta_3^H) \\ &\geqslant -1 + t_1(\theta_1^L, \theta_2^H, \theta_3^H) + t_2(\theta_1^L, \theta_2^H, \theta_3^H) + t_3(\theta_1^L, \theta_2^H, \theta_3^H). \end{aligned}$$

One can make a weighted sum of these inequalities to get a contradiction. In particular, for $S = \{1\}$ or $\{2, 3\}$ and each pair of type profiles $\bar{\theta}_S \neq \hat{\theta}_S$, multiply $CIC(\bar{\theta}_S; \hat{\theta}_S)$ by $p(\bar{\theta}_S)p(\hat{\theta}_S)$. Also, multiply the following inequalities by 0.075: $CIC(\theta_1^L, \theta_2^L, \theta_2^L; \theta_1^H, \theta_2^H, \theta_3^H)$, $CIC(\theta_1^H, \theta_2^L, \theta_3^L; \theta_1^L, \theta_2^L, \theta_2^L)$, and $CIC(\theta_1^L, \theta_2^H, \theta_3^H; \theta_1^L, \theta_2^L, \theta_2^L)$. Moreover, multiply these four by 0.025: $CIC(\theta_1^L, \theta_2^L, \theta_2^L; \theta_1^H, \theta_2^H, \theta_3^L)$, $CIC(\theta_1^H, \theta_2^L, \theta_3^H; \theta_1^L, \theta_2^L, \theta_2^L)$, $CIC(\theta_1^L, \theta_2^H, \theta_3^L; \theta_1^L, \theta_2^L, \theta_2^L)$, and $CIC(\theta_1^L, \theta_2^L, \theta_2^L; \theta_1^L, \theta_2^L, \theta_3^H)$. By summing up the scaled inequalities and rearranging terms, all terms containing $t_i(\theta)$ are canceled and we obtain that $0 \geqslant 1.625$, a contradiction. Thus, there is no simple mechanism satisfying the interim CIC condition. □

Proof of Claim 2. The ex-post BB condition is straightforward to verify.

We now demonstrate the interim CR condition and the interim AI condition. First, notice that $u_i(q(\theta), \theta) \geqslant 0$ for all $i \in I$ and $\theta \in \Theta$. Second, it is easy to verify that $\sum_{\theta_{-S} \in \Theta_{-S}} \sum_{i \in S} t_i(\theta) p(\theta_{-S} | \theta_S) = 0$ for all coalition $S \subseteq I$, $\theta_S \in \Theta_S$, and $t \in T$. Thus,

$$V_S(t, \theta_S) = \sum_{\theta_{-S} \in \Theta_{-S}} \sum_{i \in S} (u_i(q(\theta), \theta) + t_i(\theta)) p(\theta_{-S} | \theta_S) = \sum_{\theta_{-S} \in \Theta_{-S}} \sum_{i \in S} u_i(q(\theta), \theta) p(\theta_{-S} | \theta_S) \geqslant 0,$$

for all coalition S, $\theta_S \in \Theta_S$, and $t \in T$. Hence, the interim CR and AI conditions are satisfied.

At last, we verify the interim CIC condition. In this example, it suffices to focus on the interim CIC constraints of the following five coalitions: type-θ_1^H agent 1 (verified in the text), type-(θ_1^H, θ_2^H) coalition $(1, 2)$, type-(θ_1^H, θ_2^L) coalition-$(1, 2)$ (verified in the text), type-(θ_1^H, θ_3^H) coalition $(1, 3)$, and type-(θ_1^H, θ_3^L) coalition-$(1, 3)$. This is because for any other type-θ_S coalition S and deviating strategy $\delta_S : \Theta_S \to \Delta(\Theta_S)$,

$$\min_{t \in T} V_S(t, \theta_S) = \sum_{\theta_{-S} \in \Theta_{-S}} \sum_{i \in S} u_i(q(\theta), \theta) p(\theta_{-S} | \theta_S)$$

$$\geqslant \sum_{\theta'_S \in \Theta_S} \sum_{\theta_{-S} \in \Theta_{-S}} \sum_{i \in S} u_i(q(\theta'_S, \theta_{-S}), \theta) p(\theta_{-S}|\theta_S) \delta_S[\theta_S](\theta'_S) \geqslant \min_{t \in T} V_S[\delta_S](t, \theta_S),$$

where the first equality follows from the observation from the previous paragraph, the first inequality follows from our construction of q, and the last inequality comes from the symmetric design of the transfer rules, i.e., $t^3 = -t^1$ and $t^4 = -t^2$.

For type-(θ_1^H, θ_2^H) coalition $(1, 2)$, the on-path MEU is equal to 0. By misreporting (θ_1^H, θ_2^L), (θ_1^L, θ_2^H), and (θ_1^L, θ_2^L), the coalition's MEU levels are equal to

$$\begin{aligned}
&\min\{-8.571, 8.571, 6.857, -6.857\} = -8.571, \\
&\min\{-19.429, 23.429, 2, 2\} = -19.429, \text{and} \\
&\min\{-0.5, 4.5, 2, 2\} = -0.5
\end{aligned}$$

respectively. Transfer rule $t^1 \in T$ can disincentivize deviation in pure and mixed strategies.

Type-(θ_1^H, θ_3^H) coalition $(1, 3)$'s on-path MEU is equal to 0. By misreporting (θ_1^H, θ_3^L), (θ_1^L, θ_3^H), and (θ_1^L, θ_3^L), the coalition's MEU levels are equal to

$$\begin{aligned}
&\min\{-1.667, 1.667, -10.667, 10.667\} = -10.667, \\
&\min\{-6.333, 10.333, -24.667, 28.667\} = -24.667, \text{and} \\
&\min\{-9.25, 13.25, -10, 14\} = -10
\end{aligned}$$

respectively. Transfer rule t^3 disincentivizes deviation in pure and mixed strategies.

The MEU of type-(θ_1^H, θ_3^L) coalition $\{1, 3\}$ is equal to 0. By misreporting (θ_1^H, θ_3^H), (θ_1^L, θ_3^H), and (θ_1^L, θ_3^L), the coalition's MEU levels are equal to

$$\begin{aligned}
&\min\{2.813, -2.813, 0, 0\} = -2.813, \\
&\min\{-9, 11, -31, 33\} = -31, \text{and} \\
&\min\{-13.063, 15.063, -14, 16\} = -14
\end{aligned}$$

respectively. Due to t^3, it is easy to verify that no deviation is profitable. □

A.2 Motzkin's Transposition Theorem and Notations

To unify the notation, for an m-dimensional vector x, if all of its dimensions are positive or non-negative, then we say $x \in \mathbb{R}^m_{++}$ or $x \in \mathbb{R}^m_{+}$ respectively. For $x, x' \in \mathbb{R}^m$, if $x - x' \in \mathbb{R}^m_{++}$ or $x - x' \in \mathbb{R}^m_{+}$, then we say $x > x'$ or $x \geqslant x'$ respectively. Let $\mathbf{0}$ denote a zero vector.

We demonstrate the existence/non-existence of a transfer rule satisfying certain linear constraints by showing that a system of linear equations and inequalities has a solution/no solution. We show the latter by applying the transposition theorem by Motzkin (1951).

Theorem 2. *Let $E \in \mathbb{R}^{m\times l}$, $B \in \mathbb{R}^{o\times l}$, and $D \in \mathbb{R}^{k\times l}$ be matrices. If the system $Ex < \boldsymbol{0}$, $Bx \geqslant \boldsymbol{0}$, $Dx = \boldsymbol{0}$ has a column vector solution $x \in \mathbb{R}^l$, then $E'y_1 + B'y_2 + D'y_3 = \boldsymbol{0}$, where column vectors $y_1 \in \mathbb{R}^m_+\backslash\{\boldsymbol{0}\}$, $y_2 \in \mathbb{R}^o_+$, and $y_3 \in \mathbb{R}^k$, has no solution, and vice versa.*

Motzkin (1951) has pointed out that B (or D) can be missing in this result. In the current paper, it suffices to focus on the case that B is missing. To apply the theorem, it is hence important to construct matrices E and D. As a preparation, we define a row vector $p_{\theta_C\theta'_C} \in \mathbb{R}^{n|\Theta|}_+\backslash\{\mathbf{0}\}$ for each non-grand coalition C and type profiles $\theta_C, \theta'_C \in \Theta_C$, as well as a row vector $e^S_\theta \in \mathbb{R}^{n|\Theta|}_+\backslash\{\mathbf{0}\}$ for each non-singleton coalition S and type profile $\theta_S \in \Theta_S$.

For convenience, we order and index the elements of Θ by θ^1, θ^2, ..., $\theta^{|\Theta|}$. For any row or column vector $x \in \mathbb{R}^{n|\Theta|}$, we divide its elements into n blocks of $|\Theta|$ elements. We say the first block of $|\Theta|$ elements correspond to agent 1, ..., and the last block of $|\Theta|$ elements correspond to agent n. Within each block, we say the $|\Theta|$ elements correspond to θ^1,..., $\theta^{|\Theta|}$ respectively. Hence, each element of $x \in \mathbb{R}^{n|\Theta|}$ corresponds to one agent and one type profile.

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

For each $\theta \in \Theta$ and coalition S, we also define a row vector $e^S_\theta \in \mathbb{R}^{n|\Theta|}_+\backslash\{\mathbf{0}\}$ as follows. For each agent $i \in S$, let the dimension of e^S_θ that corresponds to agent i and type profile θ be equal to 1. Thus, we have defined $|S|$ dimensions of e^S_θ. Let all other dimensions of e^S_θ be 0. When $S = I$, denote e^I_θ by e_θ for simplicity.

As an illustration, we look at a concrete example. Let $I = \{1, 2, 3\}$ and $\Theta_i = \{\theta^1_i, \theta^2_i\}$ for each $i \in I$. Order the eight elements of Θ by:

$$(\theta^1_1, \theta^1_2, \theta^1_3), (\theta^1_1, \theta^1_2, \theta^2_3), (\theta^1_1, \theta^2_2, \theta^1_3), (\theta^1_1, \theta^2_2, \theta^2_3), (\theta^2_1, \theta^1_2, \theta^1_3), (\theta^2_1, \theta^1_2, \theta^2_3), (\theta^2_1, \theta^2_2, \theta^1_3), (\theta^2_1, \theta^2_2, \theta^2_3).$$

For each vector in $\mathbb{R}^{24}$, its first, second, and third blocks of eight dimensions correspond to agents 1, 2, and 3 respectively. Let $\mathbf{0}_8$ denote a zero row vector in $\mathbb{R}^8$. We present the explicit forms of the following three vectors below as an illustration:

$$p_{(\theta^1_1,\theta^1_2)(\theta^1_1,\theta^2_2)} = (\underbrace{0, 0, p(\theta^1_1, \theta^1_2, \theta^1_3), p(\theta^1_1, \theta^1_2, \theta^2_3), 0, ..., 0}_{\text{eight dimensions}}, \underbrace{0, 0, p(\theta^1_1, \theta^1_2, \theta^1_3), p(\theta^1_1, \theta^1_2, \theta^2_3), 0, ..., 0}_{\text{eight dimensions}}, \mathbf{0}_8);$$

$$e_{(\theta^1_1,\theta^1_2,\theta^2_3)} = (\underbrace{0, 1, 0, ..., 0}_{\text{eight dimensions}}, \underbrace{0, 1, 0, ..., 0}_{\text{eight dimensions}}, \underbrace{0, 1, 0, ..., 0}_{\text{eight dimensions}});$$

$$e^{\{1,2\}}_{(\theta^1_1,\theta^1_2,\theta^2_3)} = (\underbrace{0, 1, 0, ..., 0}_{\text{eight dimensions}}, \underbrace{0, 1, 0, ..., 0}_{\text{eight dimensions}}, \mathbf{0}_8).$$

A.3 Proof of Proposition 1

To establish Proposition 1, we first establish Lemma 1, which states that for each non-singleton coalition S, agent $i \in S$, and type $\bar{\theta}_i \in \Theta_i$, if $p(\cdot|\hat{\theta}_i) \neq p(\cdot|\bar{\theta}_i)$ for any $\hat{\theta}_i \neq \bar{\theta}_i$, then there exists a transfer rule within coalition S, $\xi^{\bar{\theta}_i}$, satisfying the stated conditions.[6] When (q, T) violates the interim CIC condition, but the BDP property holds, Proposition 1 constructs an ambiguous collusive mechanism within some coalition to implement a profitable collusive deviating strategy by utilizing transfer rules $(\xi^{\bar{\theta}_i})_{i \in I, \bar{\theta}_i \in \Theta_i}$.

Lemma 1. *For each coalition S with $|S| \geqslant 2$, agent $i \in S$, and type $\bar{\theta}_i \in \Theta_i$, if there does not exist $\hat{\theta}_i \in \Theta_i \backslash \{\bar{\theta}_i\}$ such that $p(\cdot|\hat{\theta}_i) = p(\cdot|\bar{\theta}_i)$, then there exists a transfer rule $\xi^{\bar{\theta}_i} \equiv (\xi_j^{\bar{\theta}_i} : \Theta \to \mathbb{R})_{j \in S}$, such that*

1. $\sum_{j \in S} \xi_j^{\bar{\theta}_i}(\theta) = 0$ *for all* $\theta \in \Theta$,
2. $\sum_{\theta_{-j} \in \Theta_{-j}} \xi_j^{\bar{\theta}_i}(\theta) p(\theta_{-j}|\theta_j) = 0$ *for all* $j \in S$ *and* $\theta_j \in \Theta_j$,
3. $\sum_{\theta_{-i} \in \Theta_{-i}} \xi_i^{\bar{\theta}_i}(\hat{\theta}_i, \theta_{-i}) p(\theta_{-i}|\bar{\theta}_i) < 0$ *for all* $\hat{\theta}_i \in \Theta_i \backslash \{\bar{\theta}_i\}$.

Proof. Fix any coalition S with $|S| \geqslant 2$, agent $i \in S$, and type $\bar{\theta}_i \in \Theta_i$. With the vectors defined in Section A.2, we construct matrices E, D^1, D^2, and D of dimensions $m \times l$, $k^1 \times l$, $k^2 \times l$, and $k \times l$ respectively, where $m = |\Theta_i| - 1$, $k^1 = \sum_{j \in S} |\Theta_j|$, $k^2 = |\Theta|$, $k = k^1 + k^2$, and $l = n|\Theta|$. Matrix E is obtained by stacking up m row vectors $p_{\bar{\theta}_i \hat{\theta}_i} \in \mathbb{R}_+^l$ for all $\hat{\theta}_i \in \Theta_i \backslash \{\bar{\theta}_i\}$ (the vertical order of these vectors does not matter). Construct matrix D^1 by stacking up k^1 row vectors $p_{\theta_j \theta_j} \in \mathbb{R}_+^l$ for all $j \in S$ and $\theta_j \in \Theta_j$. Matrix D^2 is obtained by stacking up k^2 row vectors $e_\theta^S \in \mathbb{R}_+^l$ for all $\theta \in \Theta$. Stack up D^1 and D^2 to get matrix D.

Suppose by way of contradiction that there is no transfer rule $\xi^{\bar{\theta}_i}$ satisfying the three conditions stated in the lemma. Notice that within each row of D and E, the dimensions that correspond to any agent $j \notin S$ are equal to zero. Then we can claim that $Ex < \mathbf{0}$, $Dx = \mathbf{0}$ has no column vector solution $x \in \mathbb{R}^l$. By Motzkin's transposition theorem, there are column vectors $y_1 \in \mathbb{R}_+^m \backslash \{\mathbf{0}\}$ and $y_3 \in \mathbb{R}^k$, such that $E'y_1 + D'y_3 = \mathbf{0}$, where both sides are column vectors in $\mathbb{R}^l$, or equivalently there are column vectors $y_1 \in \mathbb{R}_+^m \backslash \{\mathbf{0}\}$, $y_3^1 \in \mathbb{R}^{k^1}$, and $y_3^2 \in \mathbb{R}^{k^2}$, such that $-(y_3^1)'D^1 - (y_3^2)'D^2 = y_1'E$, where both sides are row vectors in $\mathbb{R}^l$.

By the construction of E, D^1, and D^2, the above observation implies that there exists a profile of numbers $(a_{\theta_j} \in \mathbb{R})_{j \in S, \theta_j \in \Theta_j}$, a profile of numbers $(b_\theta \in \mathbb{R})_{\theta \in \Theta}$, and a profile of non-negative numbers $(c_{\hat{\theta}_i} \in \mathbb{R}_+)_{\hat{\theta}_i \in \Theta_i \backslash \{\bar{\theta}_i\}}$ with $c_{\hat{\theta}_i} > 0$ for some $\hat{\theta}_i \in \Theta_i \backslash \{\bar{\theta}_i\}$ such that

$$\sum_{j \in S} \sum_{\theta_j \in \Theta_j} a_{\theta_j} p_{\theta_j \theta_j} + \sum_{\theta \in \Theta} b_\theta e_\theta^S = \sum_{\hat{\theta}_i \in \Theta_i \backslash \{\bar{\theta}_i\}} c_{\hat{\theta}_i} p_{\bar{\theta}_i \hat{\theta}_i}. \tag{1}$$

[6] Lemma 1 can be stated as an if and only if result. We adopt the current statement because only the if direction is used in Proposition 1.

Fix any $\hat{\theta}_i$ with $c_{\hat{\theta}_i} \neq 0$ and agent $j \in S\backslash\{i\}$. For each $\theta_{-i} \in \Theta_{-i}$, from the dimension corresponding to agent i and type profile $(\bar{\theta}_i, \theta_{-i})$ on both sides of expression (1), we have $a_{\bar{\theta}_i} p(\bar{\theta}_i, \theta_{-i}) + b_{(\bar{\theta}_i,\theta_{-i})} = 0$. From the dimension corresponding to agent j and type profile $(\bar{\theta}_i, \theta_{-i})$, we have $a_{\theta_j} p(\bar{\theta}_i, \theta_{-i}) + b_{(\bar{\theta}_i,\theta_{-i})} = 0$. Since $p(\bar{\theta}_i, \theta_{-i}) > 0$, $a_{\bar{\theta}_i} = a_{\theta_j}$ for all $\theta_j \in \Theta_j$.

For each $\theta_{-i} \in \Theta_{-i}$, by focusing on the dimensions corresponding to agent i and type profile $(\hat{\theta}_i, \theta_{-i})$ and corresponding to agent j and type profile $(\hat{\theta}_i, \theta_{-i})$ respectively, we know that $a_{\hat{\theta}_i} p(\hat{\theta}_i, \theta_{-i}) + b_{(\hat{\theta}_i,\theta_{-i})} = c_{\hat{\theta}_i} p(\bar{\theta}_i, \theta_{-i})$ and $a_{\theta_j} p(\hat{\theta}_i, \theta_{-i}) + b_{(\hat{\theta}_i,\theta_{-i})} = 0$. Since $a_{\bar{\theta}_i} = a_{\theta_j}$ for all $\theta_j \in \Theta_j$, we have $(a_{\hat{\theta}_i} - a_{\bar{\theta}_i}) p(\hat{\theta}_i, \theta_{-i}) = c_{\hat{\theta}_i} p(\bar{\theta}_i, \theta_{-i})$ for all $\theta_{-i} \in \Theta_{-i}$. Adding up this expression across all $\theta_{-i} \in \Theta_{-i}$ yields that $(a_{\hat{\theta}_i} - a_{\bar{\theta}_i}) p(\hat{\theta}_i) = c_{\hat{\theta}_i} p(\bar{\theta}_i) > 0$. The last two equations imply $p(\cdot|\hat{\theta}_i) = p(\cdot|\bar{\theta}_i)$, a contradiction with the BDP property. □

Proof of Proposition 1. **Part 1**.If (q, T) satisfies the interim CIC condition, it satisfies the interim WCIC condition by definition. We prove the other direction five steps.

Step 1. Identify a coalition S and its most profitable collusive deviating strategy δ_S.

Suppose (q, T) does not satisfy the interim CIC condition. Then there exists a coalition S, $\theta_S \in \Theta_S$, and $\tilde{\delta}_S : \Theta_S \to \Delta(\Theta_S)$ such that $\min_{t \in T} V_S[\tilde{\delta}_S](t, \theta_S) > \min_{t \in T} V_S(t, \theta_S)$. Fix such a coalition S and let $\delta_S : \Theta_S \to \Delta(\Theta_S)$ be a most profitable deviating strategy, i.e.,

$$\delta_S[\theta_S] \in \arg \max_{\tilde{\delta}_S[\theta_S] \in \Delta(\Theta_S)} \min_{t \in T} V_S[\tilde{\delta}_S](t, \theta_S), \forall \theta_S \in \Theta_S.$$

In the special case that $S = \{i\}$ is a singleton, let the ambiguous collusive mechanism be $(\delta_i, T^i = \{\tau_i\})$, where $\tau_i : T \times \Theta \to \mathbb{R}$ is defined by $\tau_i(t, \theta) = 0$ for all $t \in T$ and $\theta \in \Theta$. As δ_i is a most profitable deviating strategy and τ_i offers zero additional transfer, the interim S-IC condition and the interim BB conditions are satisfied. In Step 2 through Step 4, we thus focus on the case where $|S| \geqslant 2$.

Step 2. Construct $\zeta^t \equiv (\zeta_i^t : \Theta_S \to \mathbb{R})_{i \in S}$ for each $t \in T$.

For each $i \in S$, $t \in T$, and $\theta_i \in \Theta_i$, define

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

Since $\sum_{i \in S} \sum_{\theta_i \in \Theta_i} w_i[\delta_S](t, \theta_i) p(\theta_i) = 0$ for each $t \in T$, if we replace I and Θ in Lemma A3 of Kosenok and Severinov (2008) by S and Θ_S, there exists $\zeta^t \equiv (\zeta_i^t : \Theta_S \to \mathbb{R})_{i \in S}$ such that

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

2. $\sum_{\theta_{S\backslash\{i\}} \in \Theta_{S\backslash\{i\}}} \zeta_i^t(\theta_S) p(\theta_{S\backslash\{i\}}|\theta_i) = w_i[\delta_S](t, \theta_i)$ for all $i \in S$ and $\theta_i \in \Theta_i$.

Step 3. Find a constant number M.

Since the BDP property holds, for each $i \in S$ and $\bar{\theta}_i \in \Theta_i$, there exists a transfer rule within S, $\xi^{\bar{\theta}_i} \equiv (\xi_j^{\bar{\theta}_i} : \Theta \to \mathbb{R})_{j \in S}$, satisfying the three conditions stated in Lemma 1. Fix any constant number $M > 0$ that is weakly larger than

$$\frac{V_i[\delta_S](t, \bar{\theta}_i, \bar{\theta}_i) - V_i[\delta_S](t, \bar{\theta}_i, \hat{\theta}_i) + w_i[\delta_S](t, \bar{\theta}_i) - \sum_{\theta_{-i} \in \Theta_{-i}} \zeta_i^t(\hat{\theta}_i, \theta_{S \setminus \{i\}}) p(\theta_{-i}|\bar{\theta}_i)}{\sum_{\theta_{-i} \in \Theta_{-i}} \xi_i^{\bar{\theta}_i}(\hat{\theta}_i, \theta_{-i}) p(\theta_{-i}|\bar{\theta}_i)}$$

for any $t \in T$, $i \in S$, $\bar{\theta}_i \in \Theta_i$, and $\hat{\theta}_i \in \Theta_i \setminus \{\bar{\theta}_i\}$.

Step 4. Define the set of potential collusive transfers T^S.

For each $i \in S$ and $\bar{\theta}_i \in \Theta_i$, define $\tau^{\bar{\theta}_i} \equiv (\tau_j^{\bar{\theta}_i} : T \times \Theta \to \mathbb{R})_{j \in S}$ by $\tau_j^{\bar{\theta}_i}(t, \theta) \equiv \zeta_j^t(\theta_S) + M \xi_j^{\bar{\theta}_i}(\theta)$ for all $t \in T$, $\theta \in \Theta$, and $j \in S$. By Condition 1 of ζ^t established in Step 2 and Condition 1 of $\xi^{\bar{\theta}_i}$ in Lemma 1, $\tau^{\bar{\theta}_i}$ is budget balanced within S. Let $T^S \equiv \{\tau^{\bar{\theta}_i} : i \in S \text{ and } \bar{\theta}_i \in \Theta_i\}$. The collusive mechanism (δ_S, T^S) satisfies the S-BB condition.

Step 5. Prove the interim S-IC condition for (δ_S, T^S).

For each $\tau \in T^S$, there exists $j \in S$ and $\tilde{\theta}_j \in \Theta_j$ such that $\tau_i(t, \theta) = \zeta_i^t(\theta_S) + M \xi_i^{\tilde{\theta}_j}(\theta)$ for all $t \in T$, $\theta \in \Theta$, and $i \in S$. Given this τ, for each $i \in S$ and $\bar{\theta}_i \in \Theta_i$,

$$\begin{aligned}
&\min_{t \in T}\{V_i[\delta_S](t, \bar{\theta}_i, \bar{\theta}_i) + \sum_{\theta_{-i} \in \Theta_{-i}} \tau_i(t, (\bar{\theta}_i, \theta_{-i})) p(\theta_{-i}|\bar{\theta}_i)\} \\
&= \min_{t \in T}\{V_i[\delta_S](t, \bar{\theta}_i, \bar{\theta}_i) + \sum_{\theta_{-i} \in \Theta_{-i}} [\zeta_i^t(\bar{\theta}_i, \theta_{S \setminus \{i\}}) + M \xi_i^{\tilde{\theta}_j}(\bar{\theta}_i, \theta_{-i})] p(\theta_{-i}|\bar{\theta}_i)\} \\
&= \min_{t \in T}\{V_i[\delta_S](t, \bar{\theta}_i, \bar{\theta}_i) + w_i[\delta_S](t, \bar{\theta}_i)\} \qquad (2) \\
&= \min_{t \in T}\{V_i(t, \bar{\theta}_i) + \frac{1}{|S|} \sum_{\tilde{\theta}_S \in \Theta_S} [V_S[\delta_S](t, \tilde{\theta}_S) - V_S(t, \tilde{\theta}_S)] p(\tilde{\theta}_S)\}. \qquad (3)
\end{aligned}$$

In the above expression, the first equality follows from the definition of τ. The second equality follows from Condition 2 of ζ^t established in Step 2 and Condition 2 of $\xi^{\tilde{\theta}_j}$ stated in Lemma 1. The third equality uses the definition of $w_i[\delta_S](t, \bar{\theta}_i)$. Notice that the value of expression (3), type-$\bar{\theta}_i$ agent i's MEU of participating and truthfully reporting in the ambiguous collusive mechanism, is independent of $\tau \in T^S$. By the choice of M, the value of expression (2) is weakly higher than

$$\begin{aligned}
&\min_{t \in T} \sum_{\hat{\theta}_i \in \Theta_i} [V_i[\delta_S](t, \bar{\theta}_i, \hat{\theta}_i) + \sum_{\theta_{-i} \in \Theta_{-i}} [\zeta_i^t(\hat{\theta}_i, \theta_{S \setminus \{i\}}) + M \xi_i^{\bar{\theta}_i}(\hat{\theta}_i, \theta_{-i})] p(\theta_{-i}|\bar{\theta}_i)] \sigma_i[\bar{\theta}_i](\hat{\theta}_i) \\
&\geqslant \min_{\substack{t \in T, \\ \tilde{\tau} \in T^S}} \sum_{\hat{\theta}_i \in \Theta_i} [V_i[\delta_S](t, \bar{\theta}_i, \hat{\theta}_i) + \sum_{\theta_{-i} \in \Theta_{-i}} \tilde{\tau}_i(t, (\hat{\theta}_i, \theta_{-i})) p(\theta_{-i}|\bar{\theta}_i)] \sigma_i[\bar{\theta}_i](\hat{\theta}_i),
\end{aligned}$$

where the inequality follows from $\tau \in T^S$. Hence, (δ_S, T^S) satisfies the interim S-IC condition.

Part 2. It suffices to show that (δ_S, T^S) constructed above also satisfies the interim S-IR condition when the (q, T) above satisfies the interim AI condition. We focus on $|S| \geqslant 2$, since the proof is trivial when $|S| = 1$. By the interim AI condition, for each $\theta_S \in \Theta_S$, $V_S(t, \theta_S)$ is independent of $t \in T$. As $\min_{t \in T} V_S[\delta_S](t, \theta_S) \geqslant \min_{t \in T} V_S(t, \theta_S)$ for all $\theta_S \in \Theta_S$, we must have $V_S[\delta_S](t, \theta_S) \geqslant V_S(t, \theta_S)$ for all $\theta_S \in \Theta_S$ and $t \in T$. Hence, $\sum_{\tilde{\theta}_S \in \Theta_S}[V_S[\delta_S](t, \tilde{\theta}_S) - V_S(t, \tilde{\theta}_S)]p(\tilde{\theta}_S) \geqslant 0$ for all $t \in T$, and expression (3) has a higher value than $\min_{t \in T} V_i(t, \bar{\theta}_i)$ for all $i \in S$ and $\bar{\theta}_i \in \Theta_i$. Hence, (δ_S, T^S) satisfies the interim S-IR condition. □

A.4 Proof of Proposition 2

To prove Proposition 2, we first establish Lemma 2, which shows that there exists an efficient allocation rule for which no simple mechanism satisfies the interim CIC condition. Then we take advantage of Proposition 1 and show that there is an ambiguous collusive mechanism to implement a profitable collusive deviation.

Lemma 2. *Under all common prior p, there exists a profile of quasi-linear utility functions and an ex-post efficient allocation rule q, such that q is not implementable via a simple mechanism satisfying the interim CIC condition.*

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

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

The set of feasible outcomes is $A = \{x_0, x_1, x_2\}$. Their payoffs are given in the following table. In the two-dimensional vectors, the first one is the payoff of agent i and the second one is that of each agent $j \in I \backslash \{i\}$. All agents' payoffs only depend on the type of agent i.

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

Table 4: payoffs of agent i and each agent $j \in I \backslash \{i\}$ from outcomes x_0, x_1, and x_2

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

Step 2. We show that q is not implementable via a simple mechanism satisfying the interim CIC condition.

Suppose by way of contradiction that there exists (q, t) satisfying the interim CIC condition. Among others, the following groups of CIC constraints have to be satisfied.

The first group of CIC constraints are imposed on the singleton coalition $\{i\}$. When either θ_i or θ_i' is equal to $\bar{\theta}_i$ and $\theta_i \neq \theta_i'$, we require

$$CIC(\theta_i;\theta_i') \qquad \sum_{\theta_{-i}\in\Theta_{-i}} t_i(\theta_i,\theta_{-i})p(\theta_{-i}|\theta_i) - \sum_{\theta_{-i}\in\Theta_{-i}} t_i(\theta_i',\theta_{-i})p(\theta_{-i}|\theta_i) \geqslant 2-\epsilon-1 = 1-\epsilon, \quad (4)$$

where the right-hand side of the inequality is equal to the change in agent i's payoff from the changed allocation. When neither θ_i nor θ_i' is equal to $\bar{\theta}_i$ and $\theta_i \neq \theta_i'$, it must hold that

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

We multiply each of the constraint $CIC(\theta_i;\theta_i')$ described above by $p(\theta_i)p(\theta_i')$.

The second group of CIC constraints deals with the incentives of the coalition $I\backslash\{i\}$. For each pair of type profiles $\theta_{-i},\theta_{-i}' \in \Theta_{-i}$ with $\theta_{-i} \neq \theta_{-i}'$, we require

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

Multiply each of the constraint $CIC(\theta_{-i};\theta_{-i}')$ by $p(\theta_{-i})p(\theta_{-i}')$.

The third group of CIC constraints addresses the grand coalition's incentives. For each pair of type profiles $\theta \neq \theta'$, when $\theta = \bar{\theta}$ and $\theta_i' \neq \bar{\theta}_i$, or when $\theta' = \bar{\theta}$ and $\theta_i \neq \bar{\theta}_i$, we require

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

For each $\theta_i' \neq \bar{\theta}_i$, we multiply $CIC(\bar{\theta};\theta')$ by $|p(\theta') - p(\theta_i')p(\theta_{-i}')| + p(\theta') - p(\theta_i')p(\theta_{-i}')$. Also, for each $\theta_i \neq \bar{\theta}_i$, multiply $CIC(\theta;\bar{\theta})$ by $|p(\theta) - p(\theta_i)p(\theta_{-i})|$.

Aggregate all above-mentioned scaled CIC constraints. In the left-hand side of the weighted sum, for each $\theta \in \Theta\backslash\{\bar{\theta}\}$, the coefficient of $t_i(\theta)$ is equal to

$$\begin{aligned}&\sum_{\theta_i'\in\Theta_i\backslash\{\theta_i\}} p(\theta_i')p(\theta) - \sum_{\theta_i'\in\Theta_i\backslash\{\theta_i\}} p(\theta_i)p(\theta_i',\theta_{-i}) - [|p(\theta)-p(\theta_i)p(\theta_{-i})| + p(\theta) - p(\theta_i)p(\theta_{-i})]\\ &+ |p(\theta)-p(\theta_i)p(\theta_{-i})| = [1-p(\theta_i)]p(\theta) - p(\theta_i)[p(\theta_{-i}) - p(\theta)] - [p(\theta) - p(\theta_i)p(\theta_{-i})] = 0.\end{aligned}$$

Similarly, for each $j \in I\backslash\{i\}$ and $\theta \in \Theta\backslash\{\bar{\theta}\}$, the coefficient of $t_j(\theta)$ is equal to $[1-p(\theta_{-i})]p(\theta) - p(\theta_{-i})[p(\theta_i)-p(\theta)] - [p(\theta)-p(\theta_i)p(\theta_{-i})] = 0$. The coefficient of $t_i(\bar{\theta})$ is equal to $[1-p(\bar{\theta}_i)]p(\bar{\theta}) - p(\bar{\theta}_i)[p(\bar{\theta}_{-i})-p(\bar{\theta})] + \sum_{\theta'\in\Theta\backslash\{\bar{\theta}\}}[p(\theta')-p(\theta_i')p(\theta_{-i}')] = 0$. The coefficient of $t_j(\bar{\theta})$ where $j \in I\backslash\{i\}$ is equal to $[1-p(\bar{\theta}_{-i})]p(\bar{\theta}) - p(\bar{\theta}_{-i})[p(\bar{\theta}_i)-p(\bar{\theta})] + \sum_{\theta'\in\Theta\backslash\{\bar{\theta}\}}[p(\theta')-p(\theta_i')p(\theta_{-i}')] = 0$. Hence, the weighted sum leads to $0 \geqslant 2(1-\epsilon)p(\bar{\theta}_i)(1-p(\bar{\theta}_i)) - \epsilon\sum_{\theta_{-i}\in\Theta_{-i}}\sum_{\theta_i\in\Theta_i\backslash\{\bar{\theta}_i\}}[2|p(\theta)-p(\theta_i)p(\theta_{-i})| + p(\theta)-p(\theta_i)p(\theta_{-i})]$. The right-hand side value is higher than $2(1-\epsilon)p(\bar{\theta}_i)(1-p(\bar{\theta}_i)) - 3\epsilon|\Theta|$, which is positive since $\epsilon \in (0, \frac{2p(\bar{\theta}_i)(1-p(\bar{\theta}_i))}{2p(\bar{\theta}_i)(1-p(\bar{\theta}_i))+3|\Theta|})$. As a result, we get $0 > 0$, a contradiction. □

Proof of Proposition 2. Case 1: Suppose the BDP property holds for the prior p. By Lemma 2, there is an expost efficient allocation rule q for which no simple mechanism (q, t) satisfies the interim CIC condition. For any (q, t), as it satisfies the interim AI condition but not the interim CIC condition, we can follow Proposition 1 to complete the current proof.

Case 2: Suppose the BDP property fails for the prior p. Then there exists an agent $i \in I$ and types $\theta_i^1 \neq \theta_i^2$ such that $p(\cdot|\theta_i^1) = p(\cdot|\theta_i^2)$. We follow the construction of $(u_j)_{j \in I}$ and q as in Step 2 of the proof of Lemma 2, and let $\bar{\theta}_i = \theta_i^1$. Suppose there exists a simple mechanism (q, t) satisfying the (degenerate) interim CIC constraints $CIC(\theta_i^1; \theta_i^2)$ and $CIC(\theta_i^2; \theta_i^1)$. Then

$$CIC(\theta_i^1; \theta_i^2) \quad \sum_{\theta_{-i} \in \Theta_{-i}} t_i(\theta_i^1, \theta_{-i}) p(\theta_{-i}|\theta_i^1) - \sum_{\theta_{-i} \in \Theta_{-i}} t_i(\theta_i^2, \theta_{-i}) p(\theta_{-i}|\theta_i^1) \geqslant 2 - \epsilon - 1 = 1 - \epsilon,$$

$$CIC(\theta_i^2; \theta_i^1) \quad \sum_{\theta_{-i} \in \Theta_{-i}} t_i(\theta_i^2, \theta_{-i}) p(\theta_{-i}|\theta_i^2) - \sum_{\theta_{-i} \in \Theta_{-i}} t_i(\theta_i^1, \theta_{-i}) p(\theta_{-i}|\theta_i^2) \geqslant 2 - \epsilon - 1 = 1 - \epsilon.$$

Adding up the two inequalities leads to $0 \geqslant 2 - \epsilon > 0$, a contradiction. Define a trivial ambiguous collusive mechanism (δ_i, T^i) in a way similar to Proposition 1, where δ_i is the most profitable reporting strategy and T^i is a singleton that gives agent i zero transfer. It is easy to see that (δ_i, T^i) satisfies the S-IC, S-BB, and S-IR conditions and leads to $V_i[\delta_i](t, \theta_i^1) > V_i(t, \theta_i^1)$ or $V_i[\delta_i](t, \theta_i^2) > V_i(t, \theta_i^2)$.

Hence, Proposition 2 holds in both cases. □

A.5 Proof of Theorem 1

We follow the direction Statement 4 ⇒ Statement 3 ⇒ Statement 2 ⇒ Statement 1 ⇒ Statement 4 to establish Theorem 1. Statement 4 ⇒ Statement 3 ⇒ Statement 2 is trivial. Statement 2 ⇒ Statement 1 is proved by Lemma 3. To establish Statement 1 ⇒ Statement 4, the key step is to prove Lemma 4. In Lemma 4, we show that for any non-grand coalition S and type profile $\bar{\theta}_S \in \Theta_S$, if the posterior belief of $\bar{\theta}_S$ over Θ_{-S} is different from that of any $\hat{\theta}_S \in \Theta_S \backslash \{\bar{\theta}_S\}$, then there exists a transfer rule within I, $\phi^{\bar{\theta}_S}$, satisfying three conditions.[7] Then, we construct an ambiguous mechanism (q, T) with these transfers $(\phi^{\bar{\theta}_S})_{S \in 2^I \backslash \{\emptyset, I\}, \bar{\theta}_S \in \Theta_S}$.

Lemma 3. *When the CBDP property fails, there exists a profile of quasi-linear utility functions and an ex-post efficient allocation rule q, such that q is not implementable via an ambiguous mechanism satisfying the interim WCIC condition.*

Proof. To establish the lemma, we discuss two cases.

Case 1: Suppose the BDP property holds but the CBDP property fails. There exists a non-grand coalition S and type profiles $\bar{\theta}_S \neq \hat{\theta}_S$ such that $p(\cdot|\bar{\theta}_S) = p(\cdot|\hat{\theta}_S)$. Fix any agent

[7] Similar to Lemma 1, the only if direction of Lemma 4 is also true, but is not used in later proof.

$i \in S$ for whom $\bar{\theta}_i \neq \hat{\theta}_i$. Consider the same $(u_j)_{j\in I}$ and q constructed in the proof of Lemma 2 with the modification that $\epsilon \in (0, \frac{1}{n-1})$. Now suppose there is an ambiguous mechanism (q, T) satisfying the interim CIC condition. Constraints $CIC(\bar{\theta}_S; \hat{\theta}_S)$ and $CIC(\hat{\theta}_S; \bar{\theta}_S)$ require that:

$$\min_{t\in T}\{|S| + \sum_{j\in S}\sum_{\theta_{-S}\in\Theta_{-S}} t_j(\bar{\theta}_S, \theta_{-S})p(\theta_{-S}|\bar{\theta}_S)\}$$
$$\geqslant \min_{t\in T}\{2 - \epsilon + (|S|-1)\cdot\frac{n-2}{n-1} + \sum_{j\in S}\sum_{\theta_{-S}\in\Theta_{-S}} t_j(\hat{\theta}_S, \theta_{-S})p(\theta_{-S}|\bar{\theta}_S)\},$$
$$\min_{t\in T}\{|S| + \sum_{j\in S}\sum_{\theta_{-S}\in\Theta_{-S}} t_j(\hat{\theta}_S, \theta_{-S})p(\theta_{-S}|\hat{\theta}_S)\}$$
$$\geqslant \min_{t\in T}\{2 - \epsilon + (|S|-1)\cdot\frac{n-2}{n-1} + \sum_{j\in S}\sum_{\theta_{-S}\in\Theta_{-S}} t_j(\bar{\theta}_S, \theta_{-S})p(\theta_{-S}|\hat{\theta}_S)\}.$$

By moving all terms independent of $t \in T$ out of the minimization operators, summing up the two inequalities, and taking into account that $p(\cdot|\bar{\theta}_S) = p(\cdot|\hat{\theta}_S)$, we have $2|S| \geqslant 4 - 2\epsilon + 2(|S|-1)\cdot\frac{n-2}{n-1}$. This implies $0 \geqslant n - |S| - \epsilon(n-1) > 0$ (since S is a non-grand coalition), a contradiction. As (q, T) does not satisfy the interim CIC condition and the BDP property holds, by Proposition 1, (q, T) does not satisfy the interim WCIC condition.

Case 2: Suppose the BDP property fails. There exists an agent i and type profiles $\bar{\theta}_i \neq \hat{\theta}_i$ such that $p(\cdot|\bar{\theta}_i) = p(\cdot|\hat{\theta}_i)$. Consider the same $(u_j)_{j\in I}$, allocation rule q, and method as in Case 1. We can show that there does not exist an ambiguous mechanism satisfying the following (degenerate) interim CIC constraints: $CIC(\bar{\theta}_i; \hat{\theta}_i)$ and $CIC(\hat{\theta}_i; \bar{\theta}_i)$. Hence, for every (q, T), there exists a singleton coalition who can use a trivial ambiguous collusive mechanism (the once specified in the proof of Propositions 1 and 2) to profitably deviate. □

Lemma 4. *For any non-grand coalition $S \in 2^I\backslash\{\varnothing, I\}$ and type profile $\bar{\theta}_S \in \Theta_S$, if there does not exist $\hat{\theta}_S \in \Theta_S\backslash\{\bar{\theta}_S\}$ such that $p(\cdot|\hat{\theta}_S) = p(\cdot|\bar{\theta}_S)$, then there exists a transfer rule $\phi^{\bar{\theta}_S} \equiv (\phi_i^{\bar{\theta}_S} : \Theta \to \mathbb{R})_{i\in I}$ such that*

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

Proof. Fix any non-grand coalition S and type profile $\bar{\theta}_S$. Suppose by way of contradiction that there is no transfer rule $\phi^{\bar{\theta}_S}$ satisfying the three conditions above. Then one can apply Motzkin's transposition theorem in a way similar to Lemma 1 and claim that there exists a profile of numbers $(a_{\theta_C} \in \mathbb{R})_{C\in 2^I\setminus\{\varnothing,I\},\,\theta_C\in\Theta_C}$, a profile of numbers $(b_\theta \in \mathbb{R})_{\theta\in\Theta}$, and a profile of non-negative numbers $(c_{\theta_S} \in \mathbb{R}_+)_{\theta_S\in\Theta_S\setminus\{\bar\theta_S\}}$ with $c_{\theta_S} > 0$ for some $\theta_S \in \Theta_S\setminus\{\bar\theta_S\}$, such that

$$\sum_{C\in 2^I\setminus\{\varnothing,I\}}\sum_{\theta_C\in\Theta_C} a_{\theta_C} p_{\theta_C\theta_C} + \sum_{\theta\in\Theta} b_\theta e_\theta = \sum_{\theta_S\in\Theta_S\setminus\{\bar\theta_S\}} c_{\theta_S} p_{\bar\theta_S\theta_S} \equiv \sum_{\theta_S\in\Theta_S} c_{\theta_S} p_{\bar\theta_S\theta_S}, \tag{5}$$

where we define $c_{\bar\theta_S} \equiv 0$ for convenience of notation. Recall that row vectors $p_{\theta_C\theta_C}$, e_θ, and $p_{\bar\theta_S\theta_S} \in \mathbb{R}_+^{n|\Theta|}$ are defined in Section A.2.

We reach a contradiction by establishing four claims in the remainder of this proof.

Claim 3. *If a type profile $\theta_S \in \Theta_S$ is such that $c_{\theta_S} = 0$, then for any type profile $\theta'_S \in \Theta_S$ that is different from θ_S for only one agent, and any type profiles θ_{-S}, $\theta'_{-S} \in \Theta_{-S}$ that differ from each other for only one agent,*

$$c_{\theta'_S}\frac{p(\bar\theta_S)}{p(\theta'_S)}\Big[\frac{p(\theta_{-S}|\bar\theta_S)}{p(\theta_{-S}|\theta'_S)} - \frac{p(\theta'_{-S}|\bar\theta_S)}{p(\theta'_{-S}|\theta'_S)}\Big] = 0. \tag{6}$$

Proof of Claim 3. Fix any $\theta_S \in \Theta_S$ such that $c_{\theta_S} = 0$, $\theta'_S \in \Theta_S$ that is different from θ_S for only one agent (denoted by i), and type profiles θ_{-S}, $\theta'_{-S} \in \Theta_{-S}$ that differ from each other for only one agent (denoted by j). Thus, $\theta'_S = (\theta'_i, \theta_{S\setminus\{i\}})$ and $\theta'_{-S} = (\theta'_j, \theta_{I\setminus(S\cup\{j\})})$.

For each type profile $\tilde\theta \in \Theta$ and agent $k \in I$, define $A_k(\tilde\theta) \equiv \sum_{C\in 2^I\setminus\{\varnothing,I\},C\ni k} a_{\tilde\theta_C}$ and

$$\begin{aligned}\Delta \equiv [A_i(\theta) - A_i(\theta'_j,\theta_{-j})] + [A_i(\theta'_i,\theta'_j,\theta_{-i-j}) - A_i(\theta'_i,\theta_{-i})] \\ + [A_j(\theta'_i,\theta_{-i}) - A_j(\theta)] + [A_j(\theta'_j,\theta_{-j}) - A_j(\theta'_i,\theta'_j,\theta_{-i-j})].\end{aligned}$$

According to the definition of each $A_k(\tilde\theta)$, we can cancel the common terms in the first square bracket of the above expression. Thus,

$$\begin{aligned}A_i(\theta) - A_i(\theta'_j,\theta_{-j}) =& \Big[\sum_{C\subsetneq I\setminus\{i,j\}} a_{(\theta_i,\theta_j,\theta_C)} + \sum_{C\subsetneq I\setminus\{i,j\}} a_{(\theta_i,\theta_C)}\Big] - \Big[\sum_{C\subsetneq I\setminus\{i,j\}} a_{(\theta_i,\theta'_j,\theta_C)} + \sum_{C\subsetneq I\setminus\{i,j\}} a_{(\theta_i,\theta_C)}\Big] \\ =& \sum_{C\subsetneq I\setminus\{i,j\}} a_{(\theta_i,\theta_j,\theta_C)} - \sum_{C\subsetneq I\setminus\{i,j\}} a_{(\theta_i,\theta'_j,\theta_C)}.\end{aligned}$$

Similarly, we can cancel the common terms in the other three square brackets. We thus have

$$\begin{aligned}\Delta =& \sum_{C\subsetneq I\setminus\{i,j\}} a_{(\theta_i,\theta_j,\theta_C)} - \sum_{C\subsetneq I\setminus\{i,j\}} a_{(\theta_i,\theta'_j,\theta_C)} + \sum_{C\subsetneq I\setminus\{i,j\}} a_{(\theta'_i,\theta'_j,\theta_C)} - \sum_{C\subsetneq I\setminus\{i,j\}} a_{(\theta'_i,\theta_j,\theta_C)} \\ &+ \sum_{C\subsetneq I\setminus\{i,j\}} a_{(\theta'_i,\theta_j,\theta_C)} - \sum_{C\subsetneq I\setminus\{i,j\}} a_{(\theta_i,\theta_j,\theta_C)} + \sum_{C\subsetneq I\setminus\{i,j\}} a_{(\theta_i,\theta'_j,\theta_C)} - \sum_{C\subsetneq I\setminus\{i,j\}} a_{(\theta'_i,\theta'_j,\theta_C)} = 0.\end{aligned}$$

On the other hand, we can reorder the eight terms in the definition of Δ, so that

$$0 = \Delta = [A_i(\theta) - A_j(\theta)] - [A_i(\theta'_j,\theta_{-j}) - A_j(\theta'_j,\theta_{-j})]$$

$$- [A_i(\theta_i', \theta_{-i}) - A_j(\theta_i', \theta_{-i})] + [A_i(\theta_i', \theta_j', \theta_{-i-j}) - A_j(\theta_i', \theta_j', \theta_{-i-j})]. \quad (7)$$

Both sides of expression (5) are $n|\Theta|$-dimensional vectors where each dimension corresponds to an agent and a type profile. Focus on the elements corresponding to type profile $\tilde{\theta}$ and agent $i \in S$ on both sides of expression (5) and the ones corresponding to type profile $\tilde{\theta}$ and agent $j \notin S$. We have $p(\tilde{\theta})A_i(\tilde{\theta}) + b_{\tilde{\theta}} = c_{\tilde{\theta}_S} p(\bar{\theta}_S, \tilde{\theta}_{-S})$ and $p(\tilde{\theta})A_j(\tilde{\theta}) + b_{\tilde{\theta}} = 0$. Thus,

$$A_i(\tilde{\theta}) - A_j(\tilde{\theta}) = c_{\tilde{\theta}_S} \frac{p(\bar{\theta}_S, \tilde{\theta}_{-S})}{p(\tilde{\theta})} = c_{\tilde{\theta}_S} \frac{p(\bar{\theta}_S) p(\tilde{\theta}_{-S}|\bar{\theta}_S)}{p(\tilde{\theta}_S) p(\tilde{\theta}_{-S}|\tilde{\theta}_S)}, \forall \tilde{\theta} \in \Theta.$$

Since $c_{\theta_S} = 0$, $\theta_S' = (\theta_i', \theta_{S\setminus\{i\}})$, and $\theta_{-S}' = (\theta_j', \theta_{I\setminus(S\cup\{j\})})$, expression (7) implies that

$$0 = \Delta = 0 - 0 - c_{\theta_S'} \frac{p(\bar{\theta}_S) p(\theta_{-S}|\bar{\theta}_S)}{p(\theta_S') p(\theta_{-S}|\theta_S')} + c_{\theta_S'} \frac{p(\bar{\theta}_S) p(\theta_{-S}'|\bar{\theta}_S)}{p(\theta_S') p(\theta_{-S}'|\theta_S')}.$$

Thus, expression (6) holds. □

Claim 4. *If a type profile $\theta_S \in \Theta_S$ is such that c_{θ_S}=0, then for each type profile θ_S' that differs from θ_S for only one agent, $c_{\theta_S'} = 0$.*

Proof of Claim 4. Fix any type profile $\theta_S \in \Theta_S$ such that $c_{\theta_S} = 0$ and type profile θ_S' that differs from θ_S for only one agent.

For type profiles θ_{-S} and θ_{-S}' that only differ from each other for one agent, we know from Claim 3 that expression (6) holds. For θ_{-S} and θ_{-S}' that differ for at least two agents, there exists a finite sequence $(\theta_{-S}^k)_{k=1,\ldots,K}$ such that $\theta_{-S}^1 = \theta_{-S}$, $\theta_{-S}^K = \theta_{-S}'$, and any two consecutive elements in the sequence only differ from each other for one agent. Applying Claim 3 on consecutive elements recursively, we know that expression (6) holds for all $\theta_{-S}, \theta_{-S}' \in \Theta_{\_S}$.

Then we discuss two cases. Case 1: when $\theta_S' = \bar{\theta}_S$. We have $c_{\theta_S'} = 0$ because $c_{\bar{\theta}_S}$ is defined to be zero. Case 2: when $\theta_S' \neq \bar{\theta}_S$. Recall that we have concluded from the previous paragraph that expression (6) holds for all $\theta_{-S}, \theta_{-S}' \in \Theta_{\_S}$. By the CBDP property, there must exist a pair of type profiles $\theta_{-S} \neq \theta_{-S}'$ such that the term in the square bracket in expression (6) is non-zero. Also, recall that $p(\bar{\theta}_S), p(\hat{\theta}_S) > 0$. For expression (6) to hold, it must be that $c_{\theta_S'} = 0$. Hence, we obtain that $c_{\theta_S'} = 0$ in both cases. □

Claim 5. *For each type profile $\theta_S' \in \Theta_S$, it holds that $c_{\theta_S'} = 0$.*

Proof of Claim 5. Recall that by definition $c_{\bar{\theta}_S} = 0$. For any type profile $\theta_S' \neq \bar{\theta}_S$, we can find a finite sequence $(\theta_S^k)_{k=1,\ldots,K}$ such that $\theta_S^1 = \bar{\theta}_S$, $\theta_S^K = \theta_S'$, and every two consecutive elements differ from each other for only one agent. By applying Claim 4 on consecutive elements recursively, we also know that $c_{\theta_S'} = 0$. Hence, we have established Claim 5. □

Claim 5 contradicts the fact that $c_{\theta_S} > 0$ for some $\theta_S \in \Theta_S \backslash \{\bar{\theta}_S\}$. Hence, we have established Lemma 4. $\square$

Proof of Theorem 1. Statement 4 ⇒ Statement 3 ⇒ Statement 2 holds trivially. Statement 2 ⇒ Statement 1 has been proved in Lemma 3. It remains to establish Statement 1 ⇒ Statement 4.

Given any profile of $(u_i)_{i \in I}$ and efficient allocation rule q, define a budget balanced transfer rule $\eta = (\eta_i : \Theta \to \mathbb{R})_{i \in I}$ by $\eta_i(\theta) \equiv \frac{1}{n}\sum_{j \in I} u_j(q(\theta), \theta) - u_i(q(\theta), \theta)$ for all $\theta \in \Theta$ and $i \in I$. As the CBDP property holds, by Lemma 4, for each $S \in 2^I \backslash \{\varnothing, I\}$ and $\bar{\theta}_S \in \Theta_S$, there exists a budget balanced transfer rule $\phi^{\bar{\theta}_S}$ satisfying the three conditions in Lemma 4.

Pick a sufficiently large constant $M > 0$ that is weakly larger than

$$\frac{\sum_{\theta_{-S} \in \Theta_{-S}} \left[\frac{|S|}{n}\sum_{j \in I} u_j(q(\bar{\theta}_S, \theta_{-S}), (\bar{\theta}_S, \theta_{-S})) - \sum_{i \in S}[u_i(q(\hat{\theta}_S, \theta_{-S}), (\bar{\theta}_S, \theta_{-S})) + \eta_i(\hat{\theta}_S, \theta_{-S})]\right] p(\theta_{-S}|\bar{\theta}_S)}{\sum_{i \in S}\sum_{\theta_{-S} \in \Theta_{-S}} \phi_i^{\bar{\theta}_S}(\hat{\theta}_S, \theta_{-S}) p(\theta_{-S}|\bar{\theta}_S)}$$

for any non-grand coalition $S \in 2^I \backslash \{\varnothing, I\}$ and type profiles $\bar{\theta}_S, \hat{\theta}_S \in \Theta_S$ with $\bar{\theta}_S \neq \hat{\theta}_S$.

Let $T = \{\eta + M\phi^{\tilde{\theta}_C} : C \in 2^I \backslash \{\varnothing, I\}, \tilde{\theta}_C \in \Theta_C\}$. It is easy to see that (q, T) satisfies the ex-post BB condition, because η and each $\phi^{\tilde{\theta}_C}$ are budget balanced. We verify below that (q, T) satisfies the conditions of interim CIC, interim CR, and interim AR.

To establish the interim CR condition and the interim AI condition, recall that each $t \in T$ can be expressed as $t = \eta + M\phi^{\tilde{\theta}_C}$ for some $C \in 2^I \backslash \{\varnothing, I\}$ and $\tilde{\theta}_C \in \Theta_C$. The on-path aggregate utility of the grand coalition with type profile θ under (q, t) is equal to

$$V_I(t, \theta) = \sum_{i \in I}[u_i(q(\theta), \theta) + t_i(\theta)] = \sum_{i \in I} u_i(q(\theta), \theta) \geqslant 0.$$

The on-path aggregate utility of type $\bar{\theta}_S$ non-grand coalition S under (q, t) is

$$\begin{aligned} V_S(t, \bar{\theta}_S) =& \left[\frac{|S|}{n}\sum_{j \in I}\sum_{\theta_{-S} \in \Theta_{-S}} u_j(q(\bar{\theta}_S, \theta_{-S}), (\bar{\theta}_S, \theta_{-S})) + M\sum_{i \in S}\sum_{\theta_{-S} \in \Theta_{-S}} \phi_i^{\tilde{\theta}_C}(\bar{\theta}_S, \theta_{-S})\right] p(\theta_{-S}|\bar{\theta}_S) \\ =& \frac{|S|}{n}\sum_{j \in I}\sum_{\theta_{-S} \in \Theta_{-S}} u_j(q(\bar{\theta}_S, \theta_{-S}), (\bar{\theta}_S, \theta_{-S})) p(\theta_{-S}|\bar{\theta}_S) \geqslant 0, \end{aligned} \quad (8)$$

where the first equality follows from the definition of η and the construction of t, the second equality follows from Condition 2 of $\phi^{\tilde{\theta}_C}$ in Lemma 4, and the inequality follows from the non-negative ex-post social surplus of q. Since in both cases, aggregate utility within the coalition is positive and independent of $t \in T$, (q, T) satisfies the interim CR and AI conditions.

We now establish the interim CIC condition. The grand coalition cannot benefit from jointly misreporting, because of the ex-post BB condition of (q, T) and the ex-post efficiency of q. For a non-grand coalition S with type profile $\bar\theta_S$, as $\eta + M\phi^{\bar\theta_S} \in T$, the MEU for type-$\bar\theta_S$ coalition S to adopt strategy $\delta_S : \Theta_S \to \Delta(\Theta_S)$, $\min_{t\in T} V_S[\delta_S](t, \bar\theta_S)$, is no higher than

$$\sum_{i\in S}\sum_{\hat\theta_S\in\Theta_S}\sum_{\theta_{-S}\in\Theta_{-S}}[u_i(q(\hat\theta_S,\theta_{-S}),(\bar\theta_S,\theta_{-S})) + \eta_i(\hat\theta_S,\theta_{-S}) + M\phi_i^{\bar\theta_S}(\hat\theta_S,\theta_{-S})]p(\theta_{-S}|\bar\theta_S)\delta_S[\bar\theta_S](\hat\theta_S)$$

$$\leqslant \frac{|S|}{n}\sum_{\theta_{-S}\in\Theta_{-S}}\sum_{j\in I}u_j(q(\bar\theta_S,\theta_{-S}),(\bar\theta_S,\theta_{-S}))p(\theta_{-S}|\bar\theta_S) = \min_{t\in T} V_S(t,\bar\theta_S),$$

where the inequality follows from the choice of M, and the equality follows from the equivalent form of $V_S(t, \bar\theta_S)$ given by expression (8). Hence, the interim CIC condition holds. □

A.6 Proofs in Section 6

Proof of Proposition 3. Statement 3 ⇒ Statement 2 is trivial and Statement 2 ⇒ Statement 1 follows from Lemma 3. It remains to establish Statement 1 ⇒ Statement 3.

Fix any ex-post efficient allocation rule $q : \Theta \to A$. For each $i \in I$ and $\theta_i \in \Theta_i$, define $w_i(\theta_i) \equiv \sum_{\theta_{-i}\in\Theta_{-i}} u_i(q(\theta),\theta)p(\theta_{-i}|\theta_i) - \frac{1}{n}SS$, where $SS \equiv \sum_{j\in I}\sum_{\tilde\theta\in\Theta} u_j(q(\tilde\theta),\tilde\theta)p(\tilde\theta)$ is the ex-ante social surplus. It is clear that $\sum_{i\in I}\sum_{\theta_i\in\Theta_i} w_i(\theta_i)p(\theta_i) = 0$. By Lemma A.3 of Kosenok and Severinov (2008), there exists a transfer rule $\tau \equiv (\tau_i : \Theta \to \mathbb{R})_{i\in I}$ satisfying the ex-post BB condition such that $\sum_{\theta_{-i}\in\Theta_{-i}} \tau_i(\theta)p(\theta_{-i}|\theta_i) = w_i(\theta_i)$ for all $i \in I$ and $\theta_i \in \Theta_i$. For each $i \in I$ and $\theta \in \Theta$, define $\eta_i(\theta) \equiv \tau_i(\theta) + \frac{1}{n}SS$. Apparently, $\sum_{i\in I}\eta_i(\theta) = SS$ for all $\theta \in \Theta$. Also,

$$\sum_{\theta_{-i}\in\Theta_{-i}}\eta_i(\theta)p(\theta_{-i}|\theta_i) = w_i(\theta_i) + \frac{1}{n}SS = \sum_{\theta_{-i}\in\Theta_{-i}}u_i(q(\theta),\theta)p(\theta_{-i}|\theta_i). \tag{9}$$

Fix any constant $M > 0$ that is weakly larger than

$$\frac{\sum_{\theta_{-S}\in\Theta_{-S}}\sum_{i\in S}[u_i(q(\bar\theta_S,\theta_{-S}),(\bar\theta_S,\theta_{-S})) - \tau_i(\bar\theta_S,\theta_{-S}) - u_i(q(\hat\theta_S,\theta_{-S}),\theta) + \tau_i(\hat\theta_S,\theta_{-S})]p(\theta_{-S}|\bar\theta_S)}{\sum_{i\in S}\sum_{\theta_{-S}\in\Theta_{-S}}\phi_i^{\bar\theta_S}(\hat\theta_S,\theta_{-S})p(\theta_{-S}|\bar\theta_S)}$$

for any coalition $S \in 2^I\backslash\{\varnothing, I\}$ and type profiles $\bar\theta_S, \hat\theta_S \in \Theta_S$ with $\bar\theta_S \neq \hat\theta_S$. Define $T = \{-\eta + M\phi^{\tilde\theta_C} : C \in 2^I\backslash\{\varnothing, I\}, \tilde\theta_C \in \Theta_C\}$, where each $\phi^{\tilde\theta_C}$ satisfies the conditions in Lemma 4.

Since each $\phi^{\tilde\theta_C}$ is ex-post budget balanced, it is easy to see that for each $t \in T$,

$$-\sum_{\theta\in\Theta}\sum_{i\in I}t_i(\theta)p(\theta) = \sum_{\theta\in\Theta}\sum_{i\in I}\eta_i(\theta)p(\theta) = SS.$$

For each $t \in T$, there exists $C \in 2^I\backslash\{\varnothing, I\}$ and $\tilde\theta_C \in \Theta_C$ such that $t = -\eta + M\phi^{\tilde\theta_C}$. By expression (9) and Condition 2 of $\phi^{\tilde\theta_C}$ in Lemma 4, we have $V_i(t, \theta_i) = \sum_{\theta_{-i}\in\Theta_{-i}}[u_i(q(\theta),\theta) - \eta_i(\theta) + M\phi_i^{\tilde\theta_C}(\theta)]p(\theta_{-i}|\theta_i) = 0$. One can thus verify the interim IR condition.

To demonstrate the interim CIC condition, one can follow an argument similar to Theorem 1 and verify that non-grand coalitions cannot benefit from misreporting. Moreover, since each $\phi^{\tilde{\theta}_C}$ is budget balanced, by following strategy δ_I, type-θ coalition I earns a utility

$$\sum_{\theta' \in \Theta} \sum_{i \in I} [u_i(q(\theta'), \theta) - \eta_i(\theta')] \delta_I[\theta](\theta') = \sum_{\theta' \in \Theta} \sum_{i \in I} u_i(q(\theta'), \theta) \delta_I[\theta](\theta') - SS.$$

Since q is ex-post efficient and SS is constant, misreporting is not profitable.

Hence, (q, T) extracts the full surplus and satisfies the interim CIC condition. □

Proof of Corollary 1. We modify the ambiguous mechanism designed in Theorem 1 so that $T \equiv \{\eta + M\phi^{\theta_C} : C \in 2^I \backslash \{\varnothing, I\}, \theta_C \in \Theta_C\} \cup \{\eta - M\phi^{\theta_C} : C \in 2^I \backslash \{\varnothing, I\}, \theta_C \in \Theta_C\}$. It is easy to see that (q, T) satisfies the interim CR, interim AI, and ex-post BB conditions, and that the grand coalition has no incentive to misreport. It remains to show that for M sufficiently large, the interim CIC constraints hold for non-grand coalitions.

Fix a coalition S, type profile $\bar{\theta}_S$, and collusive deviating strategy $\delta_S : \Theta_S \to \Delta(\Theta_S)$. The interim utility for this coalition to adopt δ_S is $\int_{\pi \in \Pi^\epsilon} \psi\Big(\int_{t \in T} V_S[\delta_S](t, \bar{\theta}_S) d\pi(t) \Big) d\mu(\pi)$. For each $\pi \in \Pi^\epsilon$, the argument in function ψ has two components. The first component,

$$\int_{t \in T} \sum_{i \in S} \sum_{\hat{\theta}_S \in \Theta_S} \sum_{\theta_{-S} \in \Theta_{-S}} [u_i(q(\hat{\theta}_S, \theta_{-S}), (\bar{\theta}_S, \theta_{-S})) + \eta_i(\hat{\theta}_S, \theta_{-S})] p(\theta_{-S}|\bar{\theta}_S) \delta_S[\bar{\theta}_S](\hat{\theta}_S) d\pi(t),$$

is independent of $t \in T$ and equal to $V_S[\delta_S](\eta, \bar{\theta}_S)$. The second component is M times

$$\begin{aligned}\Phi[S, \bar{\theta}_S, \delta_S](\pi) \equiv & \sum_{C \in 2^I \backslash \{\varnothing, I\}, \theta_C \in \Theta_C} \sum_{\hat{\theta}_S \in \Theta_S} \sum_{\theta_{-S} \in \Theta_{-S}} \phi_i^{\theta_C}(\hat{\theta}_S, \theta_{-S}) p(\theta_{-S}|\bar{\theta}_S) \delta_S[\bar{\theta}_S](\hat{\theta}_S) \pi(\eta + M\phi^{\theta_C}) \\ & - \sum_{C \in 2^I \backslash \{\varnothing, I\}, \theta_C \in \Theta_C} \sum_{\hat{\theta}_S \in \Theta_S} \sum_{\theta_{-S} \in \Theta_{-S}} \phi_i^{\theta_C}(\hat{\theta}_S, \theta_{-S}) p(\theta_{-S}|\bar{\theta}_S) \delta_S[\bar{\theta}_S](\hat{\theta}_S) \pi(\eta - M\phi^{\theta_C}),\end{aligned}$$

which is linear in π. Notice that $\Phi[S, \bar{\theta}_S, \delta_S]$ is as a random variable over Π^ϵ. By linearity of $\Phi[S, \bar{\theta}_S, \delta_S](\pi)$ in π and the fact that $\mathbb{E}_{\pi \in \Pi^\epsilon}[\pi] = \bar{\pi}$ ($\bar{\pi}$ is the uniform distribution), we have $\mathbb{E}_{\pi \in \Pi^\epsilon}[\Phi[S, \bar{\theta}_S, \delta_S]] = \Phi[S, \bar{\theta}_S, \delta_S](\mathbb{E}_{\pi \in \Pi^\epsilon}[\pi]) = \Phi[S, \bar{\theta}_S, \delta_S](\bar{\pi}) = 0$. It is also easy to see that $\Phi[S, \bar{\theta}_S, \delta_S]$ has a positive measure of positive realizations, a postive measure of negative realizations, and a zero measure of zero realizations. We let $\mathbb{E}$ denote $\mathbb{E}_{\pi \in \Pi^\epsilon}$ for simplicity.

Consider a new random variable that has realizations $\mathbb{E}[\Phi[S, \bar{\theta}_S, \delta_S] | \Phi[S, \bar{\theta}_S, \delta_S] < 0] < 0$ and $\mathbb{E}[\Phi[S, \bar{\theta}_S, \delta_S] | \Phi[S, \bar{\theta}_S, \delta_S] > 0] > 0$ with probabilities $\bar{\mu}(\Phi[S, \bar{\theta}_S, \delta_S] < 0) > 0$ and $\bar{\mu}(\Phi[S, \bar{\theta}_S, \delta_S] > 0) > 0$ respectively. The new random variable is a mean preserving contraction of $\Phi[S, \bar{\theta}_S, \delta_S]$ and has zero mean. From the zero mean feature of the new random variable and the strict concavity of ψ, one can verify that function $\beta : \mathbb{R}_+ \to \mathbb{R}$ defined by

$$\beta(M) \equiv \bar{\mu}(\Phi[S, \bar{\theta}_S, \delta_S] > 0) \psi\big(V_S[\delta_S](\eta, \bar{\theta}_S) + M\mathbb{E}[\Phi[S, \bar{\theta}_S, \delta_S] | \Phi[S, \bar{\theta}_S, \delta_S] > 0]\big)$$

$$+ \bar{\mu}(\Phi[S, \bar{\theta}_S, \delta_S] < 0)\psi\big(V_S[\delta_S](\eta, \bar{\theta}_S) + M\mathbb{E}[\Phi[S, \bar{\theta}_S, \delta_S] | \Phi[S, \bar{\theta}_S, \delta_S] < 0]\big)$$

for each $M \in \mathbb{R}_+$ has $\beta'(\cdot) < 0$ and $\beta''(\cdot) < 0$. Hence, $\beta(\cdot)$ decreases unboundedly.

Due to the construction of T and the Condition 2 satisfied by each ϕ^{θ_C} in the statement of Lemma 4, type-$\bar{\theta}_S$ coalition S has an on-path interim utility level $\psi\big(V_S(\eta, \bar{\theta}_S)\big)$.

We claim that there exists a sufficiently large constant $M > 0$ such that

$$\psi\big(V_S(\eta, \bar{\theta}_S)\big) > \beta(M) > \int_{\pi \in \Pi^\epsilon} \psi\big(\int_{t \in T} V_S[\delta_S](t, \bar{\theta}_S) d\pi(t)\big) d\mu(\pi).$$

The first inequality comes from the fact that β decreases unboundedly. The second inequality follows from the strict concavity of ψ and the fact that the new random variable is a mean preserving contraction of $\Phi[S, \bar{\theta}_S, \delta_S]$. Hence, it is not profitable for type-$\bar{\theta}_S$ coalition S to misreport by following δ_S.

One can enlarge M such that the above sequence of inequalities holds for every coalition S (there is a finite set of such coalitions), every mapping $\delta_S : \Theta_S \to \Delta(\Theta_S)$ (there is a compact set of such mappings), and every $\bar{\theta}_S \in \Theta_S$ (there is a finite set of such elements). When M is sufficiently large, the ambiguous mechanism (q, T) satisfies the interim CIC condition. □

References

Bennett, E. and Conn, D. (1977). The group incentive properties of mechanisms for the provision of public goods. *Public Choice*, 29:95–102.

Bergemann, D. and Morris, S. (2005). Robust mechanism design. *Econometrica*, 73(6):1771–1813.

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

Bodoh-Creed, A. L. (2012). Ambiguous beliefs and mechanism design. *Games and Economic Behavior*, 75(2):518–537.

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

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

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

Bose, S. and Daripa, A. (2017b). Eliciting second-order beliefs. Working Paper.

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

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

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

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

Chen, Y.-C. and Xiong, S. (2011). The genericity of beliefs-determine-preferences models revisited. *Journal of Economic Theory*, 146(2):751–761.

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

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

d'Aspremont, C., 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*, 177:678–707.

De Castro, L. I., Liu, Z., and Yannelis, N. C. (2017). Implementation under ambiguity. *Games and Economic Behavior*, 101:20–33.

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

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *The Quarterly Journal of Economics*, 75(4):643–669.

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

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

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

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

Guo, H. and Yannelis, N. C. (2020). Incentive compatibility under ambiguity. *Economic Theory*. Forthcoming.

Guo, H. and Yannelis, N. C. (2021). Full implementation under ambiguity. *American Economic Journal: Microeconomics*, 13(1):148–78.

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

Klibanoff, P., Marinacci, M., and Mukerji, S. (2005). A smooth model of decision making under ambiguity. *Econometrica*, 73(6):1849–1892.

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

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

Krasa, S. and Yannelis, N. C. (1994). The value allocation of an economy with differential information. *Econometrica*, 62(4):881–900.

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

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

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

Lopomo, G., Rigotti, L., and Shannon, C. (2021). Uncertainty in mechanism design. Working Paper.

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

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

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

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

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

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

Ollár, M. and Penta, A. (2017). Full implementation and belief restrictions. *American Economic Review*, 107(8):2243–77.

Pintér, M. (2021). How to make ambiguous strategies. Working Paper.

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

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

Song, Y. (2020). Efficient design with small ambiguity. Working Paper.

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

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