WINPEC Working Paper Series No.E1812
January 2019

Stable Coalition Structures: Characterizations and Applications of Hart and Kurz's Four Stability Concepts

Takaaki Abe

Waseda INstitute of Political EConomy
Waseda University
Tokyo, Japan

Stable Coalition Structures: Characterizations and Applications of Hart and Kurz's Four Stability Concepts

Takaaki Abe*

January 11, 2019

Abstract

Hart and Kurz (1983) introduced four stability concepts, called α-, β-, γ-, and δ-stability. In contrast to the intensive studies on their conceptual aspects, these notions have rarely been adopted to analyze stable coalition structures in an application because the definitions consist of multiple intermediate steps. The purpose of this paper is to solve these practical difficulties. We provide an explicit form for each of the four stability concepts and reformulate each concept without using any intermediate steps. Moreover, we offer some sufficient conditions that guarantee the existence of stable coalition structures and the inclusion relation among the four stability notions. In addition, we propose a new approach to characterize the notions of stability. An application of our results to Cournot oligopoly is also provided.

Keywords: coalition formation; coalition structure; Cournot oligopoly; stability

JEL Classification: C71

1 Introduction

Hart and Kurz wrote two path-breaking papers on the stability of coalition structures: Hart and Kurz (1983, 1984). In these works, they combined two game-theoretic concepts, namely, value and stability. A value concept means a payoff distribution, which describes "who receives how much payoff in which coalition structure". Note that the term "coalition" means a group of players who jointly make a decision. A "coalition structure" means a collection of disjoint coalitions, namely, a partition of the set of players. Although most of traditional models had assumed that the coalition structure is given and fixed exogenously, Hart and Kurz regarded it as an endogenous outcome that players reach after iteratively splitting and merging coalitions. As this approach suggests, the purpose of their theory was to *predict* which coalition structure would be endogenously formed and be stable in the sense that no group of players attempts to deviate from the coalition structure to improve their payoffs. To analyze stable coalition structures, they introduced four stability concepts, called α-, β-, γ-, and δ-stability. These notions were widely accepted from a conceptual perspective. However, the connection they proposed, namely, the connection between the four stability notions, and the exploration of stable

* JSPS Research Fellow. Graduate School of Economics, Waseda University. 1-6-1, Nishi-waseda, Shinjuku-ku, Tokyo 169-8050, Japan. E-mail: takatomo3639@asagi.waseda.jp

The author gratefully acknowledges financial support from the Japan Society for the Promotion of Science (JSPS).

coalition structures, have gradually diverged in two directions.

On one hand, in the context of strategic form games, studies of the core concepts known as α-, β-, γ-, and δ-cores have incorporated the conceptual aspects of the four stability notions. These core notions are defined in strategic form games that admit joint strategies and binding commitments. However, the main purpose of the four cores is to specify stable *strategy profiles* instead of stable coalition structures.[*1] On the other hand, in the context of coalition formation, the framework of hedonic games (Banerjee *et al.*, 2001; and Bogomolnaia and Jackson, 2002) has played a central role in the analysis of stable coalition structures. For hedonic games, other stability notions, *e.g.*, Nash stability, individual stability, and the farsighted vNM stable set, have been proposed.[*2] Although the core is also defined, since most hedonic games assume the absence of externalities among coalitions (elaborated in Section 2), the four stability notions reduce to the unique core.

As summarized above, the four notions of stability have rarely been adopted to analyze stable coalition structures. The reason is that the definitions of the four stability notions consist of multiple steps and are not suitable for applications. As we will elaborate in Section 2, to check whether a coalition structure is α-stable in a game, we have to construct a coalition function form game with non-transferable utility (an NTU-game) based on the game and check whether the payoff distribution in the coalition structure belongs to the core of the NTU-game. We need another NTU-game for β-stability. Moreover, for γ- and δ-stability, we also have to construct a strategic form game and compute a strong Nash equilibrium. The steps of forming intermediate games, such as NTU-games and strategic form games, makes the use of the stability notions in application more difficult.

The purpose of this paper is to resolve these practical difficulties. We first provide an explicit form for each of the stability notions and reformulate each notion without using any intermediate games. In other words, we show that we can "remove" such intermediate steps from the definitions. Moreover, utilizing the explicit forms, we offer some sufficient conditions that are useful to examine the relationship among the four stability notions and the existence of stable coalition structures in application. In addition, we introduce the notion of a consistent partition to describe what partition consistently connects the partition from which the players deviate and the partition to which they deviate. We propose and demonstrate a constructive procedure to obtain such consistent partitions. This notion is a new approach to characterize the stability notions. We provide an application of our results to Cournot oligopoly.

The remainder of this paper is organized as follows. In section 2, we introduce basic definitions and the four stability concepts. In section 3, we provide an equivalent expression for each of the four stability concepts without using any intermediate steps. In section 4, we offer some conditions for externalities and internalities that are useful to analyze stable coalition structures in applications. In section 5, we introduce the notions of δ-consistent partitions and γ-consistent partitions. By using them, we offer another approach to characterize the stability concepts. In section 6, we analyze stable coalition structures under Cournot oligopoly. Section 7 provides a summary and possibilities for future research.

[*1] Therefore, we distinguish between the four cores and the four stability concepts. In this paper, we discuss the four stability concepts.

[*2] Nash stability is not Nash equilibrium. The former is the stability of a coalition structure, while the latter is a stability concept (or an equilibrium concept) of a strategy profile.

Most proofs are provided in the appendix. Some short proofs are offered in the main body of the paper.

2 Preliminaries

2.1 Coalitions and partitions

Let $N = \{1, ..., n\}$ be a finite set of players. A subset $S \subseteq N$ is a *coalition* of players. Let $\mathcal{P}$ be a *partition* (or a *coalition structure*). We typically use $\mathcal{P}$, $\mathcal{Q}$, or $\mathcal{R}$ to denote a partition. For any $S \subseteq N$, let $\Pi(S)$ be the set of all partitions of S. Let $\overline{\Pi}(S) = \{\mathcal{P} | \mathcal{P} \in \Pi(T), T \subseteq S\} = \bigcup_{T \subseteq S} \Pi(T)$. For any $i \in N$ and any $\mathcal{P} \in \Pi(N)$, $\mathcal{P}(i)$ is the coalition to which player i belongs. This is uniquely determined for any i and $\mathcal{P}$.

We now introduce some useful notions. Examples are provided after the definitions. For any nonempty coalition $S \subseteq N$ and any partition $\mathcal{P} \in \overline{\Pi}(N)$, the *projection of* $\mathcal{P}$ *on* S is given as $\mathcal{P}|_S = \{S \cap C | C \in \mathcal{P}, S \cap C \neq \emptyset\}$. For any nonempty coalition $S \subseteq N$ and any partition $\mathcal{P} \in \overline{\Pi}(N)$, the *covering of* $\mathcal{P}$ *on* S is given as $\overline{\mathcal{P}_S} = \{C | C \in \mathcal{P}, S \cap C \neq \emptyset\}$. Let $\widehat{\mathcal{P}_S} = \bigcup_{C \in \overline{\mathcal{P}_S}} C$. For any $\mathcal{P} \in \overline{\Pi}(N)$, let $\widehat{\mathcal{P}} = \bigcup_{C \in \mathcal{P}} C$. Hence, for any $\mathcal{P} \in \overline{\Pi}(N)$ and $S \subseteq \widehat{\mathcal{P}}$, we have $\overline{\mathcal{P}_S} \subseteq \mathcal{P}$ and $\widehat{\mathcal{P}_S} \subseteq \widehat{\mathcal{P}}$. For example, let $\mathcal{P} = \{\{1, 2\}, \{3, 4\}, \{5\}\}$ and $S = \{2, 3\}$. Then, $\mathcal{P}|_S = \{\{2\}, \{3\}\}$ and $\overline{\mathcal{P}_S} = \{\{1, 2\}, \{3, 4\}\}$. Moreover, $\widehat{\mathcal{P}_S} = \{1, 2, 3, 4\}$ and $\widehat{\mathcal{P}} = \{1, 2, 3, 4, 5\}$. Note that $\mathcal{P}|_S$ and $\overline{\mathcal{P}_S}$ are partitions and $\widehat{\mathcal{P}_S}$ and $\widehat{\mathcal{P}}$ are coalitions.

In addition, the following equality holds. Let $S \subseteq N$. For any $\mathcal{P} \in \Pi(S)$ and $T \subseteq S$, we have

$$\mathcal{P} \setminus \overline{\mathcal{P}_T} = \mathcal{P}|_{N \setminus \widehat{\mathcal{P}_T}}. \tag{2.1}$$

For example, let $S = \{1, 2, 3, 4, 5\}$. For $\mathcal{P} = \{\{1, 2\}, \{3, 4\}, \{5\}\}$ and $T = \{2, 3\}$, we have $\overline{\mathcal{P}_T} = \{\{1, 2\}, \{3, 4\}\}$ and $N \setminus \widehat{\mathcal{P}_T} = \{5\}$. Hence, the equality holds with partition $\{\{5\}\}$. This holds for any $\mathcal{P} \in \Pi(S)$ and $T \subseteq S$. Although this relationship is straightforward, since we sometimes use this equality, we provide its proof below.

Proof of (2.1): Let $C \in \mathcal{P} \setminus \overline{\mathcal{P}_T}$. For this C, since $C \in \mathcal{P}$, there exists a coalition $C' \in \mathcal{P}$ such that $C' \cap C = C$ (namely, C' is C). Hence, $C \in \mathcal{P}|_{N \setminus \widehat{\mathcal{P}_T}}$. Now, let $C \in \mathcal{P}|_{N \setminus \widehat{\mathcal{P}_T}}$. In view of the definition of $\widehat{\mathcal{P}_T}$ and the fact $\overline{\mathcal{P}_T} \subseteq \mathcal{P}$, we have $\mathcal{P}|_{N \setminus \widehat{\mathcal{P}_T}} \subseteq \mathcal{P}$. Hence, $C \in \mathcal{P}$. Moreover, since $\mathcal{P}|_{N \setminus \widehat{\mathcal{P}_T}}$ is the projection on $N \setminus \widehat{\mathcal{P}_T}$, we have $C \subseteq N \setminus \widehat{\mathcal{P}_T}$. Hence, $C \notin \overline{\mathcal{P}_T}$. Thus, together with $C \in \mathcal{P}$, we have $C \in \mathcal{P} \setminus \overline{\mathcal{P}_T}$. This completes the proof.

For any $S \subseteq N$, let $[S]$ denote the partition of S into singletons, $[S] := \{\{i\} | i \in S\}$. Similarly, for any $\mathcal{P} \in \overline{\Pi}(N)$, let $[\mathcal{P}] := \bigcup_{S \in \mathcal{P}} [S] = [\widehat{\mathcal{P}}]$.

2.2 A partition function form game and externalities

We introduce the notions of an embedded coalition and a partition function. These are basic concepts to describe externalities among coalitions. An *embedded coalition* of N is a pair $(S, \mathcal{P})$ satisfying $S \in \mathcal{P}$. The set of all embedded coalitions of N is given by

$$EC(N) = \{(S, \mathcal{P}) \mid \emptyset \neq S \subseteq N, \ \mathcal{P} \in \Pi(N), \text{ and } S \in \mathcal{P}\}.$$

A *partition function* v assigns a real number to each embedded coalition, namely, $v : EC(N) \to \mathbb{R}$. A *partition function form game* (also known as a *game with externalities*) is a pair (N, v). Omitting N, we often use v to denote a game with externalities. A game with externalities v has *positive externalities* if for any mutually disjoint nonempty coalitions $S, T_1, T_2 \subseteq N$ and any partition $\mathcal{P}' \in \Pi(N \setminus (S \cup T_1 \cup T_2))$,

$$v(S, \{S, T_1 \cup T_2\} \cup \mathcal{P}') > v(S, \{S, T_1, T_2\} \cup \mathcal{P}'). \tag{2.2}$$

Replacing the inequality with $<$, $\geq$, or $\leq$, we define *negative, nonnegative, or nonpositive externalities,* respectively. If equality $=$ holds, the game is said to have *no* externalities.

2.3 Hart and Kurz's four stability concepts

We now introduce the stability notions known as α-, β-, γ-, and δ-stability. As mentioned in Section 1, certain "intermediate" games are needed to define these concepts: Hart and Kurz (1983) use a non-transferable utility game in coalitional form (an NTU-game) to define α- and β-stability and a strategic form game to define γ- and δ-stability. As elaborated below, since the construction of the NTU-game for α- and β-stability depends on the strategic form game for γ- and δ-stability, we first introduce the strategic form game.

For every player $i \in N$, $\mathcal{A}_i$ is the set of coalitions containing player i, namely, $\mathcal{A}_i = \{S \subseteq N | i \in S\}$. We denote by $\mathcal{A}_i$ the set of strategies of player i: each player chooses a coalition to join. We use $\sigma_i \in \mathcal{A}_i$ to denote player i's strategy. For any $S \subseteq N$, let $\mathcal{A}_S = \times_{j \in S} \mathcal{A}_j$ and $\sigma_S = (\sigma_i)_{i \in S} \in \mathcal{A}_S$. We simply use σ to denote a strategy profile in $\mathcal{A}_N$, namely, $\sigma = (\sigma_i)_{i \in N} \in \mathcal{A}_N$. For any $\mathcal{P} \in \Pi(N)$, $\sigma^{\mathcal{P}}$ is the strategy profile such that $\sigma_i^{\mathcal{P}} = \mathcal{P}(i)$ for every $i \in N$, namely, the strategy profile in which every player chooses the coalition to which she belongs in the given partition $\mathcal{P}$. In the same manner, for any $S \subseteq N$, let $\sigma_S^{\mathcal{P}} = (\sigma_i^{\mathcal{P}})_{i \in S}$. For example, let $\mathcal{P} = \{\{1,2\},\{3,4\},\{5\}\}$. We have $\sigma_1^{\mathcal{P}} = \sigma_2^{\mathcal{P}} = \{1,2\}$, $\sigma_3^{\mathcal{P}} = \sigma_4^{\mathcal{P}} = \{3,4\}$, and $\sigma_5^{\mathcal{P}} = \{5\}$.

A strategic form game needs a function that assigns a payoff distribution to each strategy profile. We now establish such a function. In this framework, such a function consists of two functions. The first function assigns a partition to each strategy profile. The second function assigns a payoff distribution to each partition. For the first function, Hart and Kurz (1983) propose two types of functions. These functions generate the two types of strategic form games called *model* γ and *model* δ. The names of the stability concepts stem from these two models.

We first define model γ. Let $\mathcal{B}^\gamma : \mathcal{A}_N \to \Pi(N)$ be a function that assigns a partition to a strategy profile. The function $\mathcal{B}^\gamma$ is given as $\mathcal{B}^\gamma(\sigma) = \{T_\sigma^i | i \in N\}$, where

$$T_\sigma^i = \begin{cases} \sigma_i & \text{if } \sigma_i = \sigma_j \text{ for all } j \in \sigma_i, \\ \{i\} & \text{otherwise.} \end{cases}$$

To understand this function, let T be a coalition proposed to form. This formulation states that to form coalition T, *every* member of T has to choose coalition T as his strategy. If someone chooses a different coalition, the other members in T are partitioned into singletons. In other words, the players are required to submit a *unanimous* agreement on their coalition to form it. Let ϕ be a function that assigns a payoff profile to a partition. For any $\mathcal{P} \in \Pi(N)$, $\phi_i(\mathcal{P})$ is player i's payoff in partition $\mathcal{P}$.*3 As a result, the

*3 No assumption has to be imposed on the form of ϕ in this step. The role of function ϕ is elaborated in Section 3.

composite function $\phi \circ \mathcal{B}^\gamma$ assigns a payoff profile to every strategy profile. Let $G^\gamma_\phi := \phi \circ \mathcal{B}^\gamma$. Together with the player set N and the strategy sets $(\mathcal{A}_i)_{i \in N}$, tuple $(N, (\mathcal{A}_i)_{i\in N}, G^\gamma_\phi)$ is a strategic form game.

Model δ is defined in a similar manner. The only difference lies in the function $\mathcal{B}$. Define

$$\mathcal{B}^\delta(\sigma) = \{T \subseteq N | i, j \in T \iff \sigma_i = \sigma_j\}.$$

Let $G^\delta_\phi := \phi \circ \mathcal{B}^\delta$, where ϕ is the same as the function used in model γ. This formulation indicates that the players, say T, choosing the same coalition, say S, are assigned to the coalition T. Unlike model γ, only a *partial* agreement is needed to form a coalition. To see the difference, consider the following strategy profile: $(\sigma_1, \sigma_2, \sigma_3) = (\{1, 2, 3\}, \{1, 2, 3\}, \{2, 3\})$. The resulting partition is $\mathcal{B}^\gamma = \{\{1\}, \{2\}, \{3\}\}$ because of the different choice of player 3. In contrast, $\mathcal{B}^\delta = \{\{1, 2\}, \{3\}\}$, because players 1 and 2 choose the same strategy.

A coalition structure is γ- or δ-stable if no group of players has an incentive to change their coalition. We use the notion of a strong equilibrium: a strategy profile $\sigma \in \mathcal{A}_N$ is a *strong equilibrium* in a strategic form game $(N, (\mathcal{A}_i)_{i\in N}, \mathcal{G})$ if there exists no $T \subseteq N$ and no $\sigma'_T \in \mathcal{A}_T$ such that $\mathcal{G}_i(\sigma'_T, \sigma_{N\setminus T}) > \mathcal{G}_i(\sigma)$ for all $i \in T$.

Definition 2.1. Let $\phi : \Pi(N) \to \mathbb{R}^N$ and $\mathcal{P} \in \Pi(N)$. A partition $\mathcal{P}$ is *γ-stable* if strategy profile $\sigma^{\mathcal{P}}$ is a strong equilibrium in G^γ_ϕ. A partition $\mathcal{P}$ is *δ-stable* if strategy profile $\sigma^{\mathcal{P}}$ is a strong equilibrium in G^δ_ϕ.

Now, we define α-stability and β-stability. These notions are probably simpler and more popular than γ-stability and δ-stability. To define these notions, we use either $\mathcal{B}^\gamma$ or $\mathcal{B}^\delta$. As described below, whichever we choose, the definition is the same. Let $\mathcal{B}^*$ denote $\mathcal{B}^\gamma$ or $\mathcal{B}^\delta$. Let $\phi : \Pi(N) \to \mathbb{R}^N$. For any $S \subseteq N$, Hart and Kurz (1983) define

$$\begin{aligned} V^\alpha_\phi(S) = \{(x_i)_{i\in S} \in \mathbb{R}^S | &\text{ there is } \sigma_S \in \mathcal{A}_S \text{ such that for all } \sigma_{N\setminus S} \in \mathcal{A}_{N\setminus S}, \\ &\phi_i(\mathcal{B}^*(\sigma)) \geq x_i \text{ for all } i \in S\}. \end{aligned}$$

For any $S \subseteq N$, define

$$\begin{aligned} V^\beta_\phi(S) = \{(x_i)_{i\in S} \in \mathbb{R}^S | &\text{ for all } \sigma_{N\setminus S} \in \mathcal{A}_{N\setminus S}, \text{ there is } \sigma_S \in \mathcal{A}_S \text{ such that} \\ &\phi_i(\mathcal{B}^*(\sigma)) \geq x_i \text{ for all } i \in S\}. \end{aligned}$$

Together with the player set N, (N, V^α_ϕ) and (N, V^β_ϕ) are NTU-games.[*4]

We define the *core* of an NTU-game. Let V denote an NTU-game. The core of V is defined as follows: $C(V) = \{x \in V(N) |$ there is no $T \subseteq N$ and no $y \in V(T)$ such that $y_i > x_i$ for every $i \in T\}$.

Definition 2.2. Let $\phi : \Pi(N) \to \mathbb{R}^N$ and $\mathcal{P} \in \Pi(N)$. A partition $\mathcal{P}$ is *α-stable* if $\phi(\mathcal{P})$ is in the core of V^α_ϕ. A partition $\mathcal{P}$ is *β-stable* if $\phi(\mathcal{P})$ is in the core of V^β_ϕ.

For any ϕ, let $C^\alpha(\phi)$ be the set of α-stable partitions. We similarly define C^β, C^γ, and C^δ. The following relationship readily follows: for any ϕ,

$$C^\alpha(\phi) \supseteq C^\beta(\phi) \supseteq (C^\gamma(\phi) \cup C^\delta(\phi)). \tag{2.3}$$

[*4] For the definitions of V^α_ϕ and V^β_ϕ, we can consider $V^{\alpha\gamma}_\phi$ and $V^{\alpha\delta}_\phi$ and, similarly, $V^{\beta\gamma}_\phi$ and $V^{\beta\delta}_\phi$. Hart and Kurz (1983)'s definition states that $V^{\alpha\gamma}_\phi = V^{\alpha\delta}_\phi =: V^\alpha_\phi$ and $V^{\beta\gamma}_\phi = V^{\beta\delta}_\phi =: V^\beta_\phi$.

3 Equivalent expressions without intermediate games

In the previous section, the four stability concepts are defined by using intermediate steps: strategic form games are used to define γ- and δ-stability and NTU-games are used to define α- and β-stability. As stated in Section 1, although these stability notions are known as a theoretical approach to study stable coalition structures, it is difficult to apply them to economic/social situations. This is because the intermediate steps of the definitions make it difficult to compute the concepts in application. Our purpose is to solve this practical difficulty. Our first attempt is to "remove" the intermediate steps from these definitions and offer a simple equivalent expression.

The following lemma is a technical result.

Lemma 3.1. Let $T \subseteq N$ and $\sigma_T \in \mathcal{A}_T$. For any $\mathcal{P}, \mathcal{P}' \in \Pi(N \setminus T)$,

$$\mathcal{B}^*(\sigma_T, \sigma^{\mathcal{P}}_{N\setminus T})|_T = \mathcal{B}^*(\sigma_T, \sigma^{\mathcal{P}'}_{N\setminus T})|_T =: \mathcal{Q}.$$

Moreover, for any $\mathcal{P} \in \Pi(N \setminus T)$,

$$\mathcal{Q} \subseteq \mathcal{B}^*(\sigma_T, \sigma^{\mathcal{P}}_{N\setminus T})$$

and

$$\mathcal{Q} \cup \mathcal{P} = \mathcal{B}^*(\sigma_T, \sigma^{\mathcal{P}}_{N\setminus T}).$$

The following proposition is our first result.

Proposition 3.2. For any partition $\mathcal{P} \in \Pi(N)$, the following four statements hold.

(i) $\mathcal{P}$ is α-stable if and only if there exist no $T \subseteq N$ and no $\mathcal{Q} \in \Pi(T)$ such that for any $\mathcal{P}' \in \Pi(N\setminus T)$, $\phi_i(\mathcal{Q} \cup \mathcal{P}') > \phi_i(\mathcal{P})$ for every $i \in T$.

(ii) $\mathcal{P}$ is β-stable if and only if there exist no $T \subseteq N$ such that for any $\mathcal{P}' \in \Pi(N \setminus T)$, there exists $\mathcal{Q} \in \Pi(T)$ such that $\phi_i(\mathcal{Q} \cup \mathcal{P}') > \phi_i(\mathcal{P})$ for every $i \in T$.

(iii) $\mathcal{P}$ is γ-stable if and only if there exist no $T \subseteq N$ and no $\mathcal{Q} \in \Pi(T)$ such that $\phi_i(\mathcal{Q} \cup [\widehat{\mathcal{P}_T} \setminus T] \cup (\mathcal{P} \setminus \overline{\mathcal{P}_T})) > \phi_i(\mathcal{P})$ for every $i \in T$.

(iv) $\mathcal{P}$ is δ-stable if and only if there exist no $T \subseteq N$ and no $\mathcal{Q} \in \Pi(T)$ such that $\phi_i(\mathcal{Q} \cup (\mathcal{P}|_{N\setminus T})) > \phi_i(\mathcal{P})$ for every $i \in T$.

In the necessary and sufficient conditions above, each of the stability notions is defined in terms of partition, and any strategic form games and NTU-games are no longer used. It is easy to interpret these conditions. For each condition, coalition $T \subseteq N$ is a deviating coalition. Partition $\mathcal{P}$ is the partition from which the coalition T deviates. The players in coalition T choose an internal partition $\mathcal{Q} \in \Pi(T)$ that they form after their deviation. The difference among the stability notions lies in the reaction from the other players $N \setminus T$. The concept of α-stability exhibits the deviating players' careful attitude: for *all* the coalition structures that the other players can organize as their reaction, namely, $\mathcal{P}' \in \Pi(N \setminus T)$, the deviating players' coalition structure $\mathcal{Q}$ must be beneficial. The notion of β-stability may be closer to the concept of best response: the deviating players obtain their benefits by choosing an appropriate coalition structure $\mathcal{Q}$ for each reaction $\mathcal{P}'$ by the other players. In other words, α-stability requires the deviating players to have a certain coalition structure that *guarantees* the benefits, while β-stability requires that they are *not prevented* from obtaining the benefits.[*5]

A similarity between α-stability and β-stability is to focus on a reaction from *all* the non-deviating players. In contrast, γ-stability and δ-stability only take into account a reaction from the players who shared the same coalition in the original partition. We first consider γ-stability. This notion indicates that the players who shared the same coalition in the original partition are partitioned into singletons. For example, let $\mathcal{P} = \{\{1,2,3\},\{4,5,6\},\{7,8\}\}$, $T = \{3,4\}$ and $\mathcal{Q} = \{\{3,4\}\}$. In view of (iii) of Proposition 3.2, the coalition $\widehat{\mathcal{P}_T}$ is $\{1,2,3,4,5,6\}$, and $\overline{\mathcal{P}_T}$ is partition $\{\{1,2,3\},\{4,5,6\}\}$. In other words, $\widehat{\mathcal{P}_T}$ is the set of players who share the coalitions with the members of T in the original partition $\mathcal{P}$. Therefore, such players are partitioned into singletons after the deviation $\mathcal{Q}$ of T, and the resulting coalition is $\mathcal{Q} \cup [\widehat{\mathcal{P}_T} \setminus T] \cup (\mathcal{P} \setminus \overline{\mathcal{P}_T}) = \{\{1\},\{2\},\{3,4\},\{5\},\{6\},\{7,8\}\}$. In this sense, γ-stability exhibits a disintegrative reaction. The concept of δ-stability can be seen as an integrative reaction: the remaining players maintain their (sub)coalitions. In the same example, the resulting partition is now $\mathcal{Q} \cup (\mathcal{P}|_{N\setminus T}) = \{\{1,2\},\{3,4\},\{5,6\},\{7,8\}\}$.

We have established the equivalent expressions of the stability concepts without using any intermediate game such as a strategic form game and an NTU-game (Proposition 3.2). Henceforth, we will regard $\phi : \Pi(N) \to \mathbb{R}^N$ as a game. Function ϕ describes who obtains how much payoff under which coalition structure: $\phi_i(\mathcal{P})$ is player i's payoff in partition $\mathcal{P}$. To avoid ambiguity, we call ϕ a *coalition structure game* (a *CS*-game). A CS-game can be thought of as a variation of the hedonic games introduced by Banerjee *et al.* (2001) and Bogomolnaia and Jackson (2002), in which each player has ordinal preferences over coalitions. Our game is its cardinal version when including externalities.[*6] This setting is also used by, for example, Hart and Kurz (1984), Casajus (2009), and Abe (2018).

4 Externalities, internalities, and the stability concepts

4.1 Externalities and the stability concepts

We begin by analyzing the externalities of a CS-game. We show that some basic properties of externalities bring about a coincidence and an inclusion relationship between the stability notions. The summary of the results is provided at the end of this section.

A CS-game ϕ is said to have *nonnegative externalities* if for any mutually disjoint nonempty coalitions $S, T_1, T_2 \subseteq N$ and any partition $\mathcal{P}' \in \Pi(N \setminus (S \cup T_1 \cup T_2))$,

$$\phi_i(\{S, T_1 \cup T_2\} \cup \mathcal{P}') \geq \phi_i(\{S, T_1, T_2\} \cup \mathcal{P}') \text{ for any } i \in S.$$

[*5] As Hart and Kurz (1983) state, these notions correspond to the α-core and the β-core introduced by Aumann (1967). As mentioned in Section 1, the α-core and the β-core are defined to analyze the stability of strategy profiles in a strategic form game, while α-stability and β-stability are introduced to examine the stability of coalition structures.

[*6] Bogomolnaia and Jackson (2002) call their model a hedonic game, and Banerjee *et al.* (2001) call theirs a coalition formation game. Although the names are different, these two models are the same. Since our game is slightly different from their model in the sense of its cardinal setting, we refer to our model ϕ as a CS-game. This is consistent with the terminology of Hart and Kurz (1984), Casajus (2009), and Abe (2018), as each contribution uses a *coalition structure value* (a *CS-value*) to construct a game.

Replacing the inequality with $\leq$, $>$, or $<$, we define *nonpositive, positive, or negative externalities*, respectively. If $=$ holds, the game has no externalities. Now, we define its slight variations. A CS-game ϕ has *weak nonnegative externalities* if for any nonempty coalition $S \subseteq N$ and any partition $\mathcal{P}' \in \Pi(N \backslash S)$,

$$\phi_i(\{S\} \cup \mathcal{P}') \geq \phi_i(\{S\} \cup [N \setminus S]) \text{ for any } i \in S.$$

Similarly, a CS-game ϕ has *weak nonpositive externalities* if for any nonempty coalition $S \subseteq N$ and any partition $\mathcal{P}' \in \Pi(N \setminus S)$,

$$\phi_i(\{S\} \cup \mathcal{P}') \geq \phi_i(\{S, N \setminus S\}) \text{ for any } i \in S.$$

The partition $[N \setminus S]$ is the finest partition of $N \setminus S$. The condition of weak nonnegative externalities compares the finest partition and each partition of $N \setminus S$. Regarding the condition of weak nonpositive externalities, in contrast, the coarsest partition $\{N \setminus S\}$ is compared with each partition of $N \setminus S$.

Combining externalities and Proposition 3.2 reveals the relationship among the stability notions. The following proposition shows that even the weaker conditions are sufficient for the notions of α-stability and β-stability to coincide.

Proposition 4.1. If a CS-game ϕ has weak nonnegative externalities or weak nonpositive externalities, then $C^\alpha(\phi) = C^\beta(\phi)$.

Note that we readily have the following relationship ("ext." means externalities): positive ext. $\Rightarrow$ nonnegative ext. $\Rightarrow$ weak nonnegative ext.; negative ext. $\Rightarrow$ nonpositive ext. $\Rightarrow$ weak nonpositive ext. Many economic applications have either weak nonnegative externalities or weak nonpositive externalities. For example, as discussed in 6, Cournot oligopoly is known as a model that has positive externalities.[*7]

Externalities also determine the relationship between γ-stability and δ-stability.

Proposition 4.2. If a CS-game ϕ has nonnegative externalities, then $C^\gamma(\phi) \supseteq C^\delta(\phi)$. If a CS-game ϕ has nonpositive externalities, then $C^\delta(\phi) \supseteq C^\gamma(\phi)$.

An intuition behind this proposition can be derived from Proposition 3.2. In model γ, the non-deviating players are partitioned into singletons. This disintegrative reaction makes it more difficult for the deviating players to deviate under nonnegative externalities, which makes more coalition structures γ-stable. In contrast, the model δ exhibits the integrative reactions from the non-deviating players. Therefore, the non-deviating players form larger coalitions, which makes it easier for the deviating players to deviate in the presence of nonnegative externalities. As a result, fewer coalition structures are δ-stable

4.2 Internalities and the stability concepts

In this subsection, we focus on internalities. Internalities describe the relationship between a merger of coalitions and the benefit of the coalitions. In the same manner as externalities, internalities specify

[*7] Moreover, Yi (1997) formulates a wide range of economic and social problems as games with positive or negative externalities, *e.g.*, the provision of public goods and customs unions in international trade.

integrative/disintegrative tendencies of coalition formation. The most typical concepts of internalities are probably superadditivity and subadditivity as defined below.

Definition 4.3. A CS-game ϕ is *superadditive* (SUPA) if for any $\mathcal{P} \in \Pi(N)$, $S, T \in \mathcal{P}$, and $i \in S \cup T$, $\phi_i(\{S \cup T\} \cup (\mathcal{P} \setminus \{S, T\})) \geq \phi_i(\mathcal{P})$. Similarly, a CS-game ϕ is *subadditive* (SUBA) if $\phi_i(\{S \cup T\} \cup (\mathcal{P} \setminus \{S, T\})) \leq \phi_i(\mathcal{P})$.

Superadditivity is an integrative condition that facilitates a merger of coalitions, while subadditivity is a disintegrative condition. By using integrative/disintegrative internalities and externalities, we divide the class of all CS-games into four subclasses, namely, {integrative internalities, disintegrative internalities} × {integrative externalities, disintegrative externalities}. To avoid ambiguity, let integrative externalities include positive, nonnegative, and weak nonnegative externalities. Similarly, disintegrative externalities include negative, nonpositive, and weak nonpositive externalities. Integrative internalities include superadditivity and its variations defined below; disintegrative internalities include subadditivity and its variations. Table 2 and the list provided at the end of this section summarize these conditions and results.

One might conjecture that if both internalities and externalities are integrative, the grand coalition is stable, and similarly, if they are both disintegrative, then the partition into singletons is stable. Propositions 4.4 and 4.5 show that this conjecture is correct. Moreover, Proposition 4.4 shows that integrative internalities guarantee that the grand coalition is stable even in the presence of disintegrative externalities. In contrast, Example 4.6 shows that the existence of stable coalition structures is not straightforwardly guaranteed under the combination of disintegrative internalities and integrative externalities. To see this, we define a weak form of superadditivity: a CS-game ϕ satisfies *weak partition-wise superadditivity* (PSUPA^-) if for any $\mathcal{P} \in \Pi(N)$, $\mathcal{Q} \subseteq \mathcal{P}$, $T \in \mathcal{Q}$, there exists $i \in T$ such that

$$\phi_i(\{\widehat{\mathcal{Q}}\} \cup (\mathcal{P} \setminus \mathcal{Q})) \geq \phi_i(\mathcal{P}). \tag{4.1}$$

We note that SUPA implies PSUPA^-.

Proposition 4.4. If a CS-game ϕ satisfies PSUPA^-, then the grand coalition $\{N\}$ satisfies all four stability concepts.

Proof. Let $S \subseteq N$ and $\mathcal{Q} \in \Pi(S)$. Assume that there exists $\mathcal{P}' \in \Pi(N \setminus S)$ such that $\phi_i(\mathcal{Q} \cup \mathcal{P}') > \phi_i(\{N\})$ for any $i \in S$. Since $\{N\} = \{\widehat{\mathcal{Q} \cup \mathcal{P}'}\}$, PSUPA^- implies that there exists $i^* \in S$ such that $\phi_{i^*}(\{N\}) \geq \phi_{i^*}(\mathcal{Q} \cup \mathcal{P}') > \phi_{i^*}(\{N\})$. This is a contradiction. $\square$

Although Proposition 4.4 holds independent of externalities, if the game has nonnegative (nonpositive) externalities, then $C^\gamma(\phi) \supseteq C^\delta(\phi)$ ($C^\delta(\phi) \supseteq C^\gamma(\phi)$). Moreover, if the game has nonnegative externalities, the following weaker condition is also sufficient for the grand coalition to satisfy all four stability concepts: for any $S \subseteq N$ there exists $i \in S$ such that

$$\phi_i(\{N\}) \geq \phi_i(\{S, N \setminus S\}). \tag{4.2}$$

Proposition 4.4 shows that integrative internalities allow us to provide a simple sufficient condition in the presence of both integrative and disintegrative externalities. In contrast, we need to perform a careful investigation for disintegrative internalities. If externalities are also disintegrative, then an analog of (4.2) guarantees the nonemptiness of stable coalition structures as follows: a CS-game ϕ satisfies *weak singleton subadditivity* (SSUBA$^-$) if for any $S \subseteq N$ there exists $i \in S$ such that

$$\phi_i([N]) \geq \phi_i(\{S\} \cup [N \setminus S]). \tag{4.3}$$

Note that SUBA implies SSUBA$^-$.

Proposition 4.5. If a CS-game ϕ has nonpositive externalities and satisfies SSUBA$^-$, then $[N]$ satisfies all four stability concepts.

Proof. Let $S \subseteq N$ and $\mathcal{Q} \in \Pi(S)$. Assume that there exists $\mathcal{P}' \in \Pi(N \setminus S)$ such that $\phi_i(\mathcal{Q} \cup \mathcal{P}') > \phi_i([N])$ for any $i \in S$. Fix a coalition $T \in \mathcal{Q}$. Note that $T \subseteq S$. Since ϕ has nonpositive externalities, we have $\phi_i(\{T\} \cup [N \setminus T]) \geq \phi_i(\mathcal{Q} \cup \mathcal{P}')$ for any $i \in T$. SSUBA$^-$ implies that there exists $i^* \in T$ such that $\phi_{i^*}([N]) \geq \phi_{i^*}(\{T\} \cup [N \setminus T])$. This is a contradiction. Hence, $[N]$ is γ-stable. Since ϕ has nonpositive externalities, $[N]$ satisfies all four stability notions. $\square$

For a game with disintegrative internalities and integrative externalities, even the following stronger version of subadditivity is no longer sufficient: a CS-game ϕ satisfies *partition-wise subadditivity* (PSUBA) if for any $\mathcal{P} \in \Pi(N)$, $\mathcal{Q} \subseteq \mathcal{P}$, and $i \in \widehat{\mathcal{Q}}$,

$$\phi_i(\mathcal{P}) \geq \phi_i(\{\widehat{\mathcal{Q}}\} \cup (\mathcal{P} \setminus \mathcal{Q})). \tag{4.4}$$

Example 4.6. Let $N = \{1, 2, 3, 4\}$. Consider the game given in Table 1. In this example, we write, for example, $12, 3, 4$ instead of $\{\{1, 2\}, \{3\}, \{4\}\}$ for simplicity. This game is symmetric: for example, $\phi_1(34, 1, 2) = \phi_2(34, 1, 2) = 6$ and $\phi_3(34, 1, 2) = \phi_4(34, 1, 2) = 3$. This game has positive externalities

Table 1 All coalition structures are neither γ- nor δ-stable

$\mathcal{P}$	ϕ_1	ϕ_2	ϕ_3	ϕ_4
1234	1	1	1	1
$123, 4$	2	2	2	7
$12, 34$	5	5	5	5
$12, 3, 4$	3	3	6	6
$1, 2, 3, 4$	4	4	4	4

and satisfies PSUBA, while both C^γ and C^δ are empty.

Example 4.6 shows that we need another perspective to guarantee the existence of stable partitions in this class. We introduce the following two conditions. A CS-game ϕ satisfies *weak partition-wise singleton subadditivity* (PSSUBA$^-$) if for any $\mathcal{P} \in \Pi(N)$ and $\mathcal{Q} \subseteq \mathcal{P}$, there exists $i \in \widehat{\mathcal{Q}}$ such that

$$\phi_i([\widehat{\mathcal{Q}}] \cup (\mathcal{P} \setminus \mathcal{Q})) \geq \phi_i(\mathcal{P}). \tag{4.5}$$

Let $\mathcal{P} \in \Pi(N)$. Partition $\mathcal{P}$ satisfies *individual rationality* (IR) if for any $i \in N$,

$$\phi_i(\mathcal{P}) \geq \phi_i(\{i\} \cup \{\mathcal{P}(i) \setminus \{i\}\} \cup (\mathcal{P} \setminus \{\mathcal{P}(i)\})) = \phi_i(\{i\} \cup (\mathcal{P}|_{N \setminus \{i\}})). \tag{4.6}$$

Proposition 4.7. Let ϕ have nonnegative externalities and satisfy PSSUBA$^-$. For any $\mathcal{P} \in \Pi(N)$, if $\mathcal{P}$ satisfies IR, then $\mathcal{P}$ satisfies all four stability concepts.

Proof. Let partition $\mathcal{P}^*$ satisfy $\phi_i(\mathcal{P}^*) \geq \phi_i(\{i\} \cup (\mathcal{P}^*|_{N\setminus\{i\}}))$ for any $i \in N$. Consider $S \subseteq N$ and $\mathcal{Q} \in \Pi(S)$, and set $\mathcal{P} := \mathcal{Q} \cup (\mathcal{P}^*|_{N\setminus S})$. Assume that $\phi_i(\mathcal{P}) > \phi_i(\mathcal{P}^*)$ for any $i \in S$. It follows from PSSUBA$^-$ that there exists $i \in S$ such that $\phi_i([S] \cup (\mathcal{P} \setminus \mathcal{Q})) \geq \phi_i(\mathcal{P})$. Fix player i. We have $\phi_i([S] \cup (\mathcal{P} \setminus \mathcal{Q})) > \phi_i(\{i\} \cup (\mathcal{P}^*|_{N\setminus\{i\}}))$.

Since $\mathcal{P} \setminus \mathcal{Q} = \mathcal{P}^*|_{N\setminus S}$, for any $C \in \mathcal{P} \setminus \mathcal{Q}$, there exists $C' \in \mathcal{P}^*$ such that $C \subseteq C'$. Moreover, $[S]$ is the finest partition of S. Hence, in view of $N \setminus S \subseteq N \setminus \{i\}$, partition $[S] \cup (\mathcal{P} \setminus \mathcal{Q})$ is finer than $\{i\} \cup (\mathcal{P}^*|_{N\setminus\{i\}})$. From nonnegative externalities, it follows that $\phi_i(\{i\} \cup (\mathcal{P}^*|_{N\setminus\{i\}})) \geq \phi_i([S] \cup (\mathcal{P} \setminus \mathcal{Q}))$. This is a contradiction. Hence, $\mathcal{P}$ is δ-stable and satisfies all four stability concepts in the presence of nonnegative externalities. □

Proposition 4.7 shows that under PSSUBA$^-$, we can regard individual rationality as a condition for the partition to be stable in all four senses. The proposition also states that the combination of PSSUBA$^-$ and nonnegative externalities prevents all coalitions consisting of two or more players from deviating.

The class of games with disintegrative internalities and integrative externalities contains some important economic applications. For example, Cournot oligopoly belongs to this class. Cournot oligopoly does not satisfy the conditions of Proposition 4.7. However, this is consistent because Cournot oligopoly has no δ-stable coalition structure. This is further investigated in Section 6.

The following list and table summarize the results discussed in this section. In Table 2, NNE (NPE) means nonnegative (nonpositive) externalities.

- Integrative internalities: SUPA(Def.4.3) $\Rightarrow$ PSUPA$^-$(4.1)
- Disintegrative internalities: PSUBA(4.4) $\Rightarrow$ SUBA(Def.4.3) $\Rightarrow$ PSSUBA$^-$(4.5) $\Rightarrow$ SSUBA$^-$(4.3)
- Integrative externalities: Positive ext.$\Rightarrow$Nonnegative ext.$\Rightarrow$Weak nonnegative ext.
- Disintegrative externalities: Negative ext.$\Rightarrow$Nonpositive ext.$\Rightarrow$Weak nonpositive ext.
- (2.3): $C^\alpha \supseteq C^\beta \supseteq (C^\gamma \cup C^\delta)$
- Prop.4.1: weak nonnegative ext. or weak nonpositive ext. $\Rightarrow C^\alpha = C^\beta$
- Prop.4.2: nonnegative ext. $\Rightarrow C^\gamma \supseteq C^\delta$; nonpositive ext. $\Rightarrow C^\delta \supseteq C^\gamma$
- Ex.4.6: Positive ext., PSUBA, $C^\delta = \emptyset$, $C^\gamma = \emptyset$

5 Necessary and sufficient conditions

As we have already seen in Definitions 2.1 and 2.2, there are some structural differences between the pair of α- and β-stability and that of γ- and δ-stability. One is the difference in the intermediate game needed to define the stability concepts: an NTU-game and a strategic game. Another lies in how to determine the resulting partition. In this section, we focus on this second difference.

Proposition 3.2 shows that a deviation in models γ and δ specifies the partition that results from the deviation: $\mathcal{Q} \cup [\widehat{\mathcal{P}_T} \setminus T] \cup (\mathcal{P} \setminus \overline{\mathcal{P}_T})$ for γ and $\mathcal{Q} \cup (\mathcal{P}|_{N\setminus T})$ for δ. In other words, given a partition $\mathcal{P} \in \Pi(N)$ and a pair $(T, \mathcal{Q})$ of a deviating coalition T and its partition $\mathcal{Q} \in \Pi(T)$, we can specify the resulting

Table 2 Summary of internalities and externalities

		Externalities: Integrative	Externalities: Disintegrative
Internalities	Integrative	Prop.4.4 PSUPA$^-$ $\Rightarrow \{N\} : \alpha\beta\gamma\delta$	
	Disintegrative	Prop.4.7 NNE, PSSUBA$^-$ $\Rightarrow \mathcal{P}$ with IR $: \alpha\beta\gamma\delta$	Prop.4.5 NPE, SSUBA$^-$ $\Rightarrow [N] : \alpha\beta\gamma\delta$
Note:		$C^\alpha = C^\beta \supseteq C^\gamma \supseteq C^\delta$	$C^\alpha = C^\beta \supseteq C^\delta \supseteq C^\gamma$

partition $\mathcal{P}'$. Now, our new question is as follows: given two partitions $\mathcal{P}$ and $\mathcal{P}'$, what pair $(T, \mathcal{Q})$ consistently connects $\mathcal{P}$ and $\mathcal{P}'$ for the pair $(T, \mathcal{Q})$ to deviate from $\mathcal{P}$ to $\mathcal{P}'$? This can be illustrated as follows.

$$\begin{array}{lll} \text{(i)} & \text{Proposition 3.2:} & \mathcal{P} \overset{(T,\mathcal{Q})}{\longrightarrow} ? \\ \text{(ii)} & \text{The new question:} & \mathcal{P} \overset{?}{\longrightarrow} \mathcal{P}' \end{array} \tag{5.1}$$

The answer to question (i) is provided in Proposition 3.2, namely, as mentioned above, $\mathcal{P}' \overset{\gamma}{=} \mathcal{Q} \cup [\widehat{\mathcal{P}_T} \setminus T] \cup (\mathcal{P} \setminus \overline{\mathcal{P}_T})$ and $\mathcal{P}' \overset{\delta}{=} \mathcal{Q} \cup (\mathcal{P}|_{N\setminus T})$. This offers a necessary and sufficient condition for a partition to be γ- or δ-stable. Below, by answering question (ii), we provide a new approach that leads us to another characterization. To this end, we need some notions. Fix $\mathcal{P} \in \Pi(N)$. For any $\mathcal{P}' \in \overline{\Pi}(N)$ with $\mathcal{P} \neq \mathcal{P}'$, we define

- $\mathcal{L}^{\mathcal{P}}(\mathcal{P}') = \{C' \in \mathcal{P}' | C' \subseteq C \text{ for some } C \in \mathcal{P}\} \subseteq \mathcal{P}'$,
- $\mathcal{Z}^{\mathcal{P}}(\mathcal{P}') = \mathcal{P}' \setminus \mathcal{L}^{\mathcal{P}}(\mathcal{P}') \subseteq \mathcal{P}'$.

For example, setting $\mathcal{P} = \{\{1,2,3\},\{4,5,6\},\{7,8\}\}$ and $\mathcal{P}' = \{\{1,2\},\{3,4\}\}$, we have $\mathcal{L}^{\mathcal{P}}(\mathcal{P}') = \{\{1,2\}\}$ and $\mathcal{Z}^{\mathcal{P}}(\mathcal{P}') = \{\{3,4\}\}$. We provide a technical lemma used in the proofs.

Lemma 5.1. For any $\mathcal{P}, \mathcal{P}' \in \Pi(N)$ and any $T \subseteq N$,

$$\mathcal{P}|_T \subseteq \mathcal{P}' \Rightarrow \mathcal{P}|_T \subseteq \mathcal{L}^{\mathcal{P}}(\mathcal{P}').$$

Proof. Let $S \in \mathcal{P}|_T$. Hence, $S \in \mathcal{P}'$. There exists $C \in \mathcal{P}$ such that $S = T \cap C$. Hence, $S \subseteq C$, $C \in \mathcal{P}$. Since $S \in \mathcal{P}'$, we have $S \in \mathcal{L}^{\mathcal{P}}(\mathcal{P}')$. □

5.1 δ-consistent partitions

Since model δ is simpler than model γ, we begin with δ. We reformulate the question above as follows. Given $\mathcal{P}, \mathcal{P}' \in \Pi(N)$, what partition $\mathcal{Q} \in \overline{\Pi}(N)$ satisfies $\mathcal{P}' = \mathcal{Q} \cup (\mathcal{P}|_{N\setminus\widehat{\mathcal{Q}}})$? How can we construct such $\mathcal{Q}$ by using the two given partitions $\mathcal{P}$ and $\mathcal{P}'$? We refer to such a partition $\mathcal{Q}$ as a δ-consistent partition from $\mathcal{P}$ to $\mathcal{P}'$.

Definition 5.2. Let $\mathcal{P}, \mathcal{P}' \in \Pi(N)$ with $\mathcal{P} \neq \mathcal{P}'$. A partition $\mathcal{Q} \in \overline{\Pi}(N)$ is a *δ-consistent partition from $\mathcal{P}$ to $\mathcal{P}'$* if $\mathcal{Q}$ satisfies $\mathcal{P}' = \mathcal{Q} \cup (\mathcal{P}|_{N \setminus \widehat{\mathcal{Q}}})$.

We first offer some examples.

Example 5.3. Let

$$\begin{aligned} \mathcal{P} &= \{\{1,2,3\},\{4,5,6\}\}, \\ \mathcal{P}' &= \{\{1,2\},\{3,4\},\{5,6\}\}. \end{aligned}$$

What partition is a δ-consistent partition from $\mathcal{P}$ to $\mathcal{P}'$? The most intuitive answer might be $\mathcal{Q} = \{\{3,4\}\}$, which actually satisfies $\mathcal{P}' = \mathcal{Q} \cup (\mathcal{P}|_{N \setminus \widehat{\mathcal{Q}}})$. However, this is not the only instance. The list of all δ-consistent partitions is

$$\{\{3,4\}\}; \{\{1,2\},\{3,4\}\}; \{\{3,4\},\{5,6\}\}; \{\{1,2\},\{3,4\},\{5,6\}\}.$$

We now briefly analyze these partitions. We first notice that every partition contains coalition $\{3,4\}$. This coalition is in $\mathcal{Z}^{\mathcal{P}}(\mathcal{P}')$. Moreover, each of the partitions consists of the coalition $\{3,4\}$ and a collection of the other coalitions in $\mathcal{P}'$, namely $\{1,2\}$ and $\{5,6\}$. These coalitions are in $\mathcal{L}^{\mathcal{P}}(\mathcal{P}')$. Therefore, one might conjecture that a δ-consistent partition is composed of the partitions in $\mathcal{Z}^{\mathcal{P}}(\mathcal{P}')$ and some partitions in $\mathcal{L}^{\mathcal{P}}(\mathcal{P}')$; formally, $\mathcal{Q} = \mathcal{Z}^{\mathcal{P}}(\mathcal{P}') \cup \mathcal{R}$ for some $\mathcal{R} \subseteq \mathcal{L}^{\mathcal{P}}(\mathcal{P}')$, where $\mathcal{R}$ can be empty. This conjecture is correct even in the case with empty $\mathcal{Z}$ as the next example shows.

Example 5.4. Now, let

$$\begin{aligned} \mathcal{P} &= \{\{1,2,3\}\}, \\ \mathcal{P}' &= \{\{1\},\{2\},\{3\}\}. \end{aligned}$$

The list of all δ-consistent partitions is

$$\{\{1\},\{2\}\}; \{\{1\},\{3\}\}; \{\{2\},\{3\}\}; \{\{1\},\{2\},\{3\}\}.$$

The points of this example are as follows: (i) $\mathcal{Z}^{\mathcal{P}}(\mathcal{P}')$ is empty, and (ii) $\mathcal{P}'|_{N \setminus \widehat{\mathcal{Q}}} = \mathcal{P}|_{N \setminus \widehat{\mathcal{Q}}}$ for any δ-consistent partition $\mathcal{Q}$. Although $\mathcal{Z}^{\mathcal{P}}(\mathcal{P}')$ is empty, we can obtain a δ-consistent partition in the same manner as the previous example. Moreover, every partition listed above clearly satisfies the second condition.*8 In addition, as a remark, we note that even if a partition is a δ-consistent partition from $\mathcal{P}$ to $\mathcal{P}'$, it is not that from $\mathcal{P}'$ to $\mathcal{P}$, as shown in the above examples.

From these observations, we can derive a characterization of a δ-consistent partition.

Proposition 5.5. Let $\mathcal{P}, \mathcal{P}' \in \Pi(N)$ with $\mathcal{P} \neq \mathcal{P}'$. Partition $\mathcal{Q}$ is a δ-consistent partition from $\mathcal{P}$ to $\mathcal{P}'$ if and only if $\mathcal{Q} \in \overline{\Delta}^{\mathcal{P}}(\mathcal{P}')$, where

$$\begin{aligned} \overline{\Delta}^{\mathcal{P}}(\mathcal{P}') = \{\mathcal{Q} \in \overline{\Pi}(N) | \ & \text{there exists } \mathcal{R} \subseteq \mathcal{L}^{\mathcal{P}}(\mathcal{P}') \text{ such that} \\ & \mathcal{Q} = \mathcal{Z}^{\mathcal{P}}(\mathcal{P}') \cup \mathcal{R} \text{ and } \mathcal{P}'|_{N \setminus \widehat{\mathcal{Q}}} = \mathcal{P}|_{N \setminus \widehat{\mathcal{Q}}}\}. \end{aligned}$$

Note that $\mathcal{R}$ can be empty.

*8 The same observation holds in Example 5.3.

Given this result and the examples above, one might consider that some δ-consistent partitions are "minimal." Formally, a partition $\mathcal{Q} \in \overline{\Delta}^{\mathcal{P}}(\mathcal{P}')$ is a *minimal δ-consistent partition from $\mathcal{P}$ to $\mathcal{P}'$* if for any $T \in \mathcal{Q}$, $\mathcal{Q} \setminus \{T\} \notin \overline{\Delta}^{\mathcal{P}}(\mathcal{P}')$. Let $\Delta^{\mathcal{P}}(\mathcal{P}')$ be the set of minimal δ-consistent partitions from $\mathcal{P}$ to $\mathcal{P}'$.

The advantage of defining a minimal δ-consistent partition lies in the fact that we can straightforwardly construct it by removing some coalitions from $\mathcal{P}'$. Below is the procedure. We call this procedure *procedure* λ. Let $\mathcal{P}, \mathcal{P}' \in \Pi(N)$ with $\mathcal{P} \neq \mathcal{P}'$.

> <u>Procedure λ:</u> For every $S \in \mathcal{P}$, if there is a nonempty coalition $T \in \mathcal{P}'$ such that $T \subseteq S$, then we remove the coalition T from $\mathcal{P}'$. Note that if there are two or more such partitions T in $\mathcal{P}'$, choose one and remove it from $\mathcal{P}'$.

The partition consisting of the remaining coalitions is a minimal δ-consistent partition (Proposition 5.6). Now we formulate a procedure λ as a function. For $\mathcal{P}$ and $\mathcal{P}'$ in $\Pi(N)$, let λ be a function $\lambda : \mathcal{P} \to \mathcal{P}' \cup \{\emptyset\}$ given as follows: for any $S \in \mathcal{P}$,

$$\lambda(S) = \begin{cases} T & \text{if there exists } T \in \mathcal{P}' \text{ such that } T \subseteq S, \\ \emptyset & \text{otherwise.} \end{cases} \tag{5.2}$$

For a function λ, we define $\Lambda^{\mathcal{P}}(\mathcal{P}') = \{\lambda(S) | S \in \mathcal{P}, \lambda(S) \neq \emptyset\}$. Therefore, $\Lambda^{\mathcal{P}}(\mathcal{P}')$ is the set of coalitions that we remove through procedure λ.

> <u>Demonstration of procedure λ:</u> Following procedure λ, we demonstrate the construction of a minimal δ-consistent partition. We consider the partitions in Example 5.3: $\mathcal{P} = \{\{1,2,3\},\{4,5,6\}\}$, $\mathcal{P}' = \{\{1,2\},\{3,4\},\{5,6\}\}$. First, for coalition $\{1,2,3\} \in \mathcal{P}$, coalition $\{1,2\} \in \mathcal{P}'$ satisfies $\{1,2\} \subseteq \{1,2,3\}$. Hence, we remove this coalition from $\mathcal{P}'$ and obtain partition $\{\{3,4\},\{5,6\}\}$. Next, for coalition $\{4,5,6\} \in \mathcal{P}$, coalition $\{5,6\} \in \mathcal{P}'$ satisfies $\{5,6\} \subseteq \{4,5,6\}$. Hence, we remove this coalition from $\{\{3,4\},\{5,6\}\}$ and obtain partition $\{\{3,4\}\}$. The procedure stops here. The remaining partition $\{\{3,4\}\}$ is δ-consistent from $\mathcal{P}$ to $\mathcal{P}'$ and clearly minimal. This is the only minimal δ-consistent partition in this example.
>
> Now, we consider the partitions in Example 5.4: $\mathcal{P} = \{\{1,2,3\}\}$, $\mathcal{P}' = \{\{1\},\{2\},\{3\}\}$. The coalition $\{1,2,3\}$ is the only coalition in $\mathcal{P}$. For this coalition, there are three coalitions T in $\mathcal{P}'$ such that $T \subseteq \{1,2,3\}$, namely, the three one-person coalitions. We choose one of them, say $\{1\}$, and remove this from $\mathcal{P}'$. We obtain partition $\{\{2\},\{3\}\}$. Since $\mathcal{P}$ consists of one coalition, the procedure stops here. The resulting partition $\{\{2\},\{3\}\}$ is one of the minimal δ-consistent partitions from $\mathcal{P}$ to $\mathcal{P}'$. If we choose $\{2\}$ or $\{3\}$, the resulting partition is $\{\{1\},\{3\}\}$ or $\{\{1\},\{2\}\}$, respectively. These three partitions are all of the minimal δ-consistent partitions of this example.

The following proposition formally states that procedure λ generates a minimal δ-consistent partition of $\mathcal{P}'$.

Proposition 5.6. For any $\mathcal{P}$ and $\mathcal{P}' \in \Pi(N)$ with $\mathcal{P} \neq \mathcal{P}'$, $\mathcal{P}' \setminus \Lambda^{\mathcal{P}}(\mathcal{P}')$ is a minimal δ-consistent partition from $\mathcal{P}$ to $\mathcal{P}'$.

Remark 5.7. For any $\mathcal{P}, \mathcal{P}' \in \Pi(N)$ with $\mathcal{P} \neq \mathcal{P}'$, there must exist a δ-consistent partition from

$\mathcal{P}$ to $\mathcal{P}'$, because the partition $\mathcal{P}'$ is an obvious δ-consistent partition. In other words, $\overline{\Delta}^{\mathcal{P}}(\mathcal{P}') \neq \emptyset$. Moreover, $\Delta^{\mathcal{P}}(\mathcal{P}')$ is also nonempty. This is because $\mathcal{P}'$ is a finite set of coalitions. To be more specific, let $\mathcal{Q} \in \overline{\Delta}^{\mathcal{P}}(\mathcal{P}')$. If $\mathcal{Q} \setminus \{T\} \notin \overline{\Delta}^{\mathcal{P}}(\mathcal{P}')$ for any $T \in \mathcal{Q}$, then by definition, the partition $\mathcal{Q}$ is minimal. If $\mathcal{Q}$ is not minimal, then there is $T_1 \in \mathcal{Q}$ such that $\mathcal{Q} \setminus \{T_1\} \in \overline{\Delta}^{\mathcal{P}}(\mathcal{P}')$. If $\mathcal{Q} \setminus \{T_1\}$ is minimal, then this process ends; if not, then there is $T_2 \in \mathcal{Q} \setminus \{T_1\}$ such that $\mathcal{Q} \setminus \{T_1, T_2\} \in \overline{\Delta}^{\mathcal{P}}(\mathcal{P}')$. However, since $\mathcal{Q}$ is a *finite* set of coalitions and $\mathcal{P} \neq \mathcal{P}'$, some partition $\mathcal{Q} \setminus \{T_1, , ..., T_k\}$ must be minimal for some k.

By using the notion of δ-consistent partitions, we can obtain a necessary and sufficient condition for a partition to be δ-stable.

Proposition 5.8. For any $\mathcal{P} \in \Pi(N)$, $\mathcal{P}$ is δ-stable if and only if for any $\mathcal{P}' \in \Pi(N) \setminus \{\mathcal{P}\}$ and any $\mathcal{Q} \in \Delta^{\mathcal{P}}(\mathcal{P}')$, there exists $i \in \widehat{\mathcal{Q}}$ such that $\phi_i(\mathcal{P}) \geq \phi_i(\mathcal{P}')$.

Proposition 5.8 shows that by comparing a partition $\mathcal{P} \in \Pi(N)$ with the other partitions $\mathcal{P}' \in \Pi(N)$, we can check whether the partition $\mathcal{P}$ is δ-stable. Moreover, we only have to examine all *minimal* δ-consistent partitions.

5.2 γ-consistent partitions

Now we analyze γ-stability. The approach is the same as that for δ-stability. The corresponding consistent partition is now defined as follows.

Definition 5.9. Let $\mathcal{P}, \mathcal{P}' \in \Pi(N)$ with $\mathcal{P} \neq \mathcal{P}'$. A partition $\mathcal{Q} \in \overline{\Pi}(N)$ is a *γ-consistent partition from $\mathcal{P}$ to $\mathcal{P}'$* if $\mathcal{Q}$ satisfies $\mathcal{P}' = \mathcal{Q} \cup (\mathcal{P} \setminus \overline{\mathcal{P}_{\widehat{\mathcal{Q}}}}) \cup [\widehat{\mathcal{P}_{\widehat{\mathcal{Q}}}} \setminus \widehat{\mathcal{Q}}]$.

We apply this notion to the partitions in Examples 5.3 and 5.4. What partitions are γ-consistent? For the partitions $\mathcal{P} = \{\{1,2,3\},\{4,5,6\}\}$ and $\mathcal{P}' = \{\{1,2\},\{3,4\},\{5,6\}\}$ (Example 5.3), partition

$$\{\{1,2\},\{3,4\},\{5,6\}\}$$

is the only γ-consistent partition. For the partitions $\mathcal{P} = \{\{1,2,3\}\}$ and $\mathcal{P}' = \{\{1\},\{2\},\{3\}\}$ (Example 5.4), the list of all γ-consistent partitions is

$$\{\{1\}\}; \{\{2\}\}; \{\{3\}\}; \{\{1\},\{2\}\}; \{\{1\},\{3\}\}; \{\{2\},\{3\}\}; \{\{1\},\{2\},\{3\}\}.$$

From these examples, one might derive the fact that the partition of the players outside $\widehat{\mathcal{Q}}$ consists of singletons. This conjecture is partly true. To see this more precisely, consider the following variation of the first example: $\mathcal{P} = \{\{1,2,3\},\{4,5,6\},\{7,8\}\}$ and $\mathcal{P}' = \{\{1,2\},\{3,4\},\{5,6\},\{7,8\}\}$. Partition $\{\{1,2\},\{3,4\},\{5,6\}\}$ is still γ-consistent even in this variation. Similarly, for the partitions $\mathcal{P} = \{\{1,2,3\},\{4,5\}\}$ and $\mathcal{P}' = \{\{1\},\{2\},\{3\},\{4,5\}\}$, that is a modification of the second example, all seven γ-consistent partitions listed above are still γ-consistent. As these examples describe,[*9] the coalitions in $\mathcal{P}$ that do *not* have an intersection with $\widehat{\mathcal{Q}}$ can be straightforwardly projected onto $\mathcal{P}'$ as

[*9] To make the symbols in this sentence easier to read, we provide the instances below. For the last example $\mathcal{P} = \{\{1,2,3\},\{4,5\}\}$ and $\mathcal{P}' = \{\{1\},\{2\},\{3\},\{4,5\}\}$, consider $\mathcal{Q} = \{\{1\}\}$. Hence, $\widehat{\mathcal{Q}} = \{1\}$. We have $\overline{\mathcal{P}_{\widehat{\mathcal{Q}}}} = \{\{1,2,3\}\}$, and hence, $\mathcal{P} \setminus \overline{\mathcal{P}_{\widehat{\mathcal{Q}}}} = \{\{4,5\}\}$. Moreover, $\widehat{\mathcal{P}_{\widehat{\mathcal{Q}}}} = \{1,2,3\}$.

$\mathcal{P}\setminus\overline{\mathcal{P}_{\widehat{\mathcal{Q}}}}$. Note that $\overline{\mathcal{P}_{\widehat{\mathcal{Q}}}}$ refers to the coalitions in $\mathcal{P}$, each of which has an intersection with $\widehat{\mathcal{Q}}$. Therefore, every player who is in the set $\widehat{\mathcal{P}_{\widehat{\mathcal{Q}}}}$ but is not a member of $\widehat{\mathcal{Q}}$, namely the player in $\widehat{\mathcal{P}_{\widehat{\mathcal{Q}}}}\setminus\widehat{\mathcal{Q}}$, is supposed to form a one-person coalition in $\mathcal{P}'$.

Given the observation above, the following proposition, which corresponds to Proposition 5.5, characterizes a γ-consistent partition.

Proposition 5.10. Let $\mathcal{P},\mathcal{P}'\in\Pi(N)$ with $\mathcal{P}\neq\mathcal{P}'$. Partition $\mathcal{Q}$ is a γ-consistent partition from $\mathcal{P}$ to $\mathcal{P}'$ if and only if $\mathcal{Q}\in\overline{\Gamma}^{\mathcal{P}}(\mathcal{P}')$, where

$$\overline{\Gamma}^{\mathcal{P}}(\mathcal{P}')=\{\mathcal{Q}\in\overline{\Pi}(N)|\text{ there exists }\mathcal{R}\subseteq\mathcal{L}^{\mathcal{P}}(\mathcal{P}')\text{ such that }\\ \mathcal{Q}=\mathcal{Z}^{\mathcal{P}}(\mathcal{P}')\cup\mathcal{R}\text{ and }\mathcal{P}'|_{N\setminus\widehat{\mathcal{Q}}}=(\mathcal{P}\setminus\overline{\mathcal{P}_{\widehat{\mathcal{Q}}}})\cup[\widehat{\mathcal{P}_{\widehat{\mathcal{Q}}}}\setminus\widehat{\mathcal{Q}}]\}.$$

Note that $\mathcal{R}$ can be empty.

A minimal γ-consistent partition is defined in the same manner. A partition $\mathcal{Q}\in\overline{\Gamma}^{\mathcal{P}}(\mathcal{P}')$ is a *minimal γ-consistent partition from $\mathcal{P}$ to $\mathcal{P}'$* if for any $T\in\mathcal{Q}$, $\mathcal{Q}\setminus\{T\}\notin\overline{\Gamma}^{\mathcal{P}}(\mathcal{P}')$. Let $\Gamma^{\mathcal{P}}(\mathcal{P}')$ be the set of minimal γ-consistent partitions from $\mathcal{P}$ to $\mathcal{P}'$. We now offer the procedure to obtain a minimal γ-consistent partition. Compared with procedure λ, this procedure, say procedure κ, is slightly complicated. Let $\mathcal{P}$ and $\mathcal{P}'\in\Pi(N)$ with $\mathcal{P}\neq\mathcal{P}'$.

> <u>Procedure κ:</u> For every $S\in\mathcal{P}$, if there is the same coalition in $\mathcal{P}'$, then we remove the coalition from $\mathcal{P}'$. If there is a strict subset $T\subsetneq S$ satisfying $[T]\subseteq\mathcal{P}'$, we remove the partition $[T]$ from $\mathcal{P}'$. Note that if some strict subsets $T_1,...,T_m$ of S satisfy the condition, we choose the largest coalition T from the subsets $T_1,...,T_m$ and remove $[T]$ from $\mathcal{P}'$. If there are two or more such largest coalitions, then we choose one, say T^*, and remove $[T^*]$ from $\mathcal{P}'$.

Let $\mathcal{K}^{\mathcal{P}}(\mathcal{P}')$ be the set of coalitions that we choose to remove through the procedure. Below, we formally define $\mathcal{K}^{\mathcal{P}}(\mathcal{P}')$. For $\mathcal{P}$ and $\mathcal{P}'$ in $\Pi(N)$, let κ be a function $\kappa:\mathcal{P}\to 2^{\mathcal{P}'}\cup\{\emptyset\}$ given as follows: for any $S\in\mathcal{P}$,

$$\kappa(S)=\begin{cases}\{T\} & \text{if there exists } T\in\mathcal{P}' \text{ such that } T=S,\\ [T] & \text{if there exists } T\subsetneq S \text{ such that } [T]\subseteq\mathcal{P}' \text{ and}\\ & \quad\text{no } T' \text{ with } T\subsetneq T'\subsetneq S \text{ satisfies } [T']\subseteq\mathcal{P}',\\ \emptyset & \text{otherwise.}\end{cases}\tag{5.3}$$

For a function κ, we define $\mathcal{K}^{\mathcal{P}}(\mathcal{P}')=\bigcup_{S\in\mathcal{P},\kappa(S)\neq\emptyset}\kappa(S)$. The following demonstration shows how this procedure works.

> <u>Demonstration of procedure κ:</u> We begin with the partitions in Example 5.3: $\mathcal{P}=\{\{1,2,3\},\{4,5,6\}\}$, $\mathcal{P}'=\{\{1,2\},\{3,4\},\{5,6\}\}$. For coalition $\{1,2,3\}\in\mathcal{P}$, $\mathcal{P}'$ has neither this coalition nor the partition of any subset of this coalition into singletons. Therefore, we remove no coalition from $\mathcal{P}'$. The same applies to coalition $\{4,5,6\}\in\mathcal{P}$. Consequently, the remaining partition is $\{\{1,2\},\{3,4\},\{5,6\}\}$. This is the only minimal γ-consistent partition in this example.
>
> We now consider the partitions in Example 5.4: $\mathcal{P}=\{\{1,2,3\}\}$, $\mathcal{P}'=\{\{1\},\{2\},\{3\}\}$. For the coalition $\{1,2,3\}\in\mathcal{P}$, there are three strict subsets such that $\mathcal{P}'$ contains its partition into

singletons: $\{1,2\}$, $\{1,3\}$, and $\{2,3\}$. We exclude the other strict subsets $\{1\}$, $\{2\}$, and $\{3\}$ because each has some strict supersets. We choose, for example, $\{1,2\}$ and remove $[\{1,2\}]$ from $\mathcal{P}'$. The resulting partition is $\{\{3\}\}$. This is a minimal γ-consistent partition. The others are $\{\{1\}\}$ and $\{\{2\}\}$.

In general, procedure κ yields a minimal γ-consistent partition. As stated in Remark 5.7, for any $\mathcal{P}$ and $\mathcal{P}' \in \Pi(N)$ with $\mathcal{P} \neq \mathcal{P}'$, there is at least one minimal γ-consistent partition from $\mathcal{P}$ to $\mathcal{P}'$.

Proposition 5.11. For any $\mathcal{P}$ and $\mathcal{P}' \in \Pi(N)$ with $\mathcal{P} \neq \mathcal{P}'$, $\mathcal{P}' \setminus \mathcal{K}^{\mathcal{P}}(\mathcal{P}')$ is a minimal γ-consistent partition from $\mathcal{P}$ to $\mathcal{P}'$.

We characterize a γ-stable coalition structure by using the notion of minimal γ-consistent partitions.

Proposition 5.12. For any $\mathcal{P} \in \Pi(N)$, $\mathcal{P}$ is γ-stable if and only if for any $\mathcal{P}' \in \Pi(N) \setminus \{\mathcal{P}\}$ and any $\mathcal{Q} \in \Gamma^{\mathcal{P}}(\mathcal{P}')$, there exists $i \in \widehat{\mathcal{Q}}$ such that $\phi_i(\mathcal{P}) \geq \phi_i(\mathcal{P}')$.

In view of Propositions 5.8 and 5.12, the difference between γ-stability and δ-stability boils down to the difference between $\Gamma^{\mathcal{P}}(\mathcal{P}')$ and $\Delta^{\mathcal{P}}(\mathcal{P}')$. Note that these two collections consist only of the terms of partitions: neither $\Gamma^{\mathcal{P}}(\mathcal{P}')$ nor $\Delta^{\mathcal{P}}(\mathcal{P}')$ includes any information about payoffs and incentives. Independent of $\Gamma^{\mathcal{P}}(\mathcal{P}')$ and $\Delta^{\mathcal{P}}(\mathcal{P}')$, the two stability notions share the same system of inequalities on payoffs.

5.3 Another characterization approach

The consistent partition approach in the previous subsection does not apply to α-stability and β-stability because, unlike γ-stability and δ-stability, a deviating coalition does not specify its resulting partition. Can we fill such a gap and compare all four concepts in a simple term? We conclude this section by offering an attempt to achieve this.

The root of this gap can be ascribed to the fact that no *expectation formation rule* applies to α-stability and β-stability. The concept of an expectation formation rule is introduced by Bloch and van den Nouweland (2014) as a function f that assigns a partition of the player set N to a tuple $(S, \mathcal{Q}, \mathcal{P}, v)$, where $S \subseteq N$ is a coalition of deviating players, $\mathcal{Q} \in \Pi(S)$ is its partition, $\mathcal{P} \in \Pi(N)$ is the initial partition from which the players deviate, and v is a partition function (see Subsection 2.2).

The reason that no expectation formation rule applies to α-stability is that a CS-game ϕ does not necessarily have transferable utility. An illustrating example is the *pessimistic rule* formulated by Bloch and van den Nouweland (2014), which is often referred to as an analog of α-stability:

$$f(S, \mathcal{Q}, \mathcal{P}, v) = \underset{\mathcal{P}' \in \Pi(N \setminus S)}{\arg\min} \sum_{T \in \mathcal{Q}} v(T, \mathcal{Q} \cup \mathcal{P}').$$

The pessimistic rule indicates that the non-deviating players $N \setminus S$ reorganize their coalition structure $\mathcal{P}' \in \Pi(N \setminus S)$ to minimize the *total* worth of the deviating coalition S, namely, $\sum_{T \in \mathcal{Q}} v(T, \mathcal{Q} \cup \mathcal{P}')$. However, this summation is clearly based on the premise that the underlying game v has transferable utility.[*10] Since we do not assume that utility is transferable in a CS-game ϕ, we cannot adopt this type of minimization.

In addition to the NTU-setting of a CS-game ϕ, β-stability has another reason: it consists of two steps of reactions. First, the deviating players form S. Next, as the first reaction, the players in $N \setminus S$ choose a coalition structure $\mathcal{P}'$ of $N \setminus S$. Finally, as the second reaction, the players in S choose a coalition structure $\mathcal{Q}$ of S. Since an expectation formation rule is defined to formulate a one-step reaction, it does not cover such multi-step reactions as β-stability.

We provide a framework to extend the notion of an expectation formation rule. As mentioned above, the purpose is to incorporate all four stability notions into this framework. The structure of the framework is similar to that of two-sided matching. For any $S \subseteq N$ and any $\mathcal{P} \in \Pi(N)$, define

$$\mathscr{G}^{+}_{S,\mathcal{P}} := \{(\mathcal{Q}, \mathcal{P}') \in \Pi(S) \times \Pi(N \setminus S) | \text{ for any } i \in S,\ \phi_i(\mathcal{Q} \cup \mathcal{P}') > \phi_i(\mathcal{P})\},$$
$$\mathscr{G}^{-}_{S,\mathcal{P}} := \{(\mathcal{Q}, \mathcal{P}') \in \Pi(S) \times \Pi(N \setminus S) | \text{ there exists } i \in S \text{ such that } \phi_i(\mathcal{Q} \cup \mathcal{P}') \leq \phi_i(\mathcal{P})\}.$$

We regard $\mathscr{G}^{+}_{S,\mathcal{P}}$ and $\mathscr{G}^{-}_{S,\mathcal{P}}$ as graphs, both of which are sets of links that connect partitions $\mathcal{Q} \in \Pi(S)$ and $\mathcal{P}' \in \Pi(N \setminus S)$. For each partition $\mathcal{P} \in \Pi(N)$ and coalition $S \subseteq N$, an expectation formation rule associates exactly one partition of $N \setminus S$ with a partition of S, while each graph may assign some partitions of $N \setminus S$ to a partition of S. In that sense, these notions can also be thought of as multi-valued functions. If pair $(\mathcal{Q}, \mathcal{P}')$ is in $\mathscr{G}^{+}_{S,\mathcal{P}}$, then the partition $\mathcal{P}'$ is beneficial for all players in S. If it is in $\mathscr{G}^{-}_{S,\mathcal{P}}$, then some player in S does not agree with the partition $\mathcal{Q}$ because of the partition $\mathcal{P}'$ of $N \setminus S$.

Below, let $\mathscr{G}^{*}_{S,\mathcal{P}}$ denote $\mathscr{G}^{+}_{S,\mathcal{P}}$ or $\mathscr{G}^{-}_{S,\mathcal{P}}$. A partition $\mathcal{Q} \in \Pi(S)$ is *single* in $\mathscr{G}^{*}_{S,\mathcal{P}}$ if there is no $\mathcal{P}' \in \Pi(N \setminus S)$ such that $(\mathcal{Q}, \mathcal{P}') \in \mathscr{G}^{*}_{S,\mathcal{P}}$. Similarly, a partition $\mathcal{P}' \in \Pi(N \setminus S)$ is *single* in $\mathscr{G}^{*}_{S,\mathcal{P}}$ if there is no $\mathcal{Q} \in \Pi(S)$ such that $(\mathcal{Q}, \mathcal{P}') \in \mathscr{G}^{*}_{S,\mathcal{P}}$. By using the notion of a single partition, Proposition 3.2 can be written as follows.

Corollary 5.13. Let $\mathcal{P} \in \Pi(N)$.

$\mathcal{P}$ is α-stable if and only if no $\mathcal{Q} \in \Pi(S)$ is single in $\mathscr{G}^{-}_{S,\mathcal{P}}$ for any $S \subseteq N$.
$\mathcal{P}$ is β-stable if and only if some $\mathcal{P}' \in \Pi(N \setminus S)$ is single in $\mathscr{G}^{+}_{S,\mathcal{P}}$ for any $S \subseteq N$.
$\mathcal{P}$ is γ-stable if and only if $(\mathcal{P} \setminus \overline{\mathcal{P}_S}) \cup [\widehat{\mathcal{P}_S} \setminus S]$ is single in $\mathscr{G}^{+}_{S,\mathcal{P}}$ for any $S \subseteq N$.
$\mathcal{P}$ is δ-stable if and only if $\mathcal{P}|_{N \setminus S}$ is single in $\mathscr{G}^{+}_{S,\mathcal{P}}$ for any $S \subseteq N$.

6 Cournot oligopoly

In this section, we analyze stable coalition structures in oligopolistic competition. Specifically, we focus on the CS-game based on Cournot oligopoly formulated by Ray and Vohra (1997). Following their discussion, we first define the game.

[*10] This does not mean that they place a special assumption on a partition function form game. A partition function form game is often assumed to have transferable utility.

6.1 The model

We assume that $n \geq 3$. Let c be the constant marginal cost of producing one unit of identical divisible goods. Let $p(\bar{q}) = a - \bar{q}$ be an inverse demand function, where a is a parameter, and $\bar{q}$ is the (total) quantity. Fix a partition $\mathcal{P} \in \Pi(N)$. In this partition, every coalition $S \in \mathcal{P}$ simultaneously determines its quantity q_S. Let $q = (q_S)_{S \in \mathcal{P}}$ be a quantity profile. For each quantity profile q and each coalition $S \in \mathcal{P}$, coalition S's profit is given as $b_S(q) = q_S \cdot (p(\sum_{S \in \mathcal{P}} q_S) - c)$. For the partition $\mathcal{P}$, solving for the (unique) Nash equilibrium, we obtain coalition S's equilibrium profit $\frac{M}{(|\mathcal{P}|+1)^2}$, where $M = (a-c)^2$. For simplicity, let $M = 1$ and $\frac{1}{(|\mathcal{P}|+1)^2}$. Since all players are symmetric, we assume that the members of each coalition S equally share their profit: in the partition $\mathcal{P} \in \Pi(N)$, player i receives

$$\phi_i(\mathcal{P}) = \frac{1}{(|\mathcal{P}|+1)^2 \cdot |\mathcal{P}(i)|}.$$

Recall that $\mathcal{P}(i)$ is the coalition in $\mathcal{P}$ that contains player i. Applying this computation to all partitions in $\Pi(N)$, we obtain function $\phi : \Pi(N) \to \mathbb{R}^N$.*11 We treat ϕ as a CS-game of the Cournot oligopoly. We call ϕ a *Cournot CS-game*. For example, Table 3 is the Cournot CS-game with five players. A check mark means that the coalition structure satisfies the corresponding stability, and a "-" means that it violates the stability concept.

Table 3 The five-player Cournot CS-game

$\mathcal{P}$	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	α	β	γ	δ
12345	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	✓	✓	✓	-
1234, 5	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{9}$	-	-	-	-
123, 45	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{18}$	$\frac{1}{18}$	✓	✓	-	-
123, 4, 5	$\frac{1}{48}$	$\frac{1}{48}$	$\frac{1}{48}$	$\frac{1}{16}$	$\frac{1}{16}$	-	-	-	-
12, 34, 5	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{16}$	✓	✓	-	-
12, 3, 4, 5	$\frac{1}{50}$	$\frac{1}{50}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	-	-	-	-
1, 2, 3, 4, 5	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	-	-	-	-

6.2 Stable coalition structures

First, for any number of players n, a Cournot CS-game has positive externalities. Therefore, Proposition 4.1 implies that $C^\alpha(\phi) = C^\beta(\phi)$. Moreover, Proposition 4.2 implies that $C^\gamma(\phi) \supseteq C^\delta(\phi)$. Hence, for any n, it immediately holds that

$$C^\alpha(\phi) = C^\beta(\phi) \supseteq C^\gamma(\phi) \supseteq C^\delta(\phi).$$

The following result shows that no partition meets δ-stability.

*11 To be more precise, ϕ depends on the number of players n. Therefore, ϕ^n is the more formal notation. For simplicity, we omit n and write ϕ.

Proposition 6.1. For any n, no coalition structure is δ-stable.

Although C^{δ} is empty, some partitions satisfy the other three stability notions. The following proposition states that the grand coalition satisfies the other three stability concepts for any n.

Proposition 6.2. For any n, partition $\{N\}$ is α-, β-, and γ-stable.

Note that the stability of a coalition structure means that of a cartel in the context of oligopoly. As Proposition 6.1 shows, some coalition, say S, has an incentive to deviate form the grand coalition when the non-deviating players "cling" to the remaining coalition, namely, $N \setminus S$. This implies that a monopoly, namely, the grand coalition, may endogenously split into a duopoly. However, as Proposition 6.2 shows, it is difficult for a monopoly to directly split into multiple small firms in one shot. Therefore, a process of splits, if any, should be a sequence of small splits from the grand coalition into some fine coalition structure, *e.g.*, the partition into singletons.

Is there a coalition structure that stops such a process of splits? Proposition 6.3 states that some coalition structures can be stable and stop the process.

Proposition 6.3. Let $i \in N$. Partition $\{N \setminus \{i\}, \{i\}\}$ is γ-stable if and only if $n = 6, 8$.

Note that, since $C^{\alpha} = C^{\beta} \supseteq C^{\gamma}$, if a partition is γ-stable then it is α- and β-stable. One may interpret this result to mean that *few* partitions are stable, and the process of splits is likely to continue since it is stable only if $n = 6, 8$. However, in a game with more players, more coalition structures can be stable. For example, in the game with $n = 12$, the partitions consisting of a seven-player coalition and a five-player coalition are γ-stable. Moreover, one might derive the irregularity of γ-stability (and, hence, that of α-stability and β-stability) from Proposition 6.3 because of the exceptions of $n = 6$ and 8. According to our computations, for all $n \leq 8$, all partitions except $\{N\}$ and $\{N \setminus \{i\}, \{i\}\}$ are not γ-stable. However, as the example with $n = 12$ shows, other partitions can be γ-stable in a game with more players.

7 Concluding remarks

The main topic of this paper is the four stability notions proposed by Hart and Kurz (1983), called α-, β-, γ-, and δ-stability. These notions were introduced to study stable coalition structures in economic, political, or game-theoretic models. The conceptual aspects of these notions has been widely accepted. However, they have rarely been adopted to analyze stable coalition structures in applications. The reason lies in the fact that these definitions contain some intermediate games such as NTU-games and strategic form games. To make these notions more useful and practical, in Section 3, we provide an explicit form with each of the stability notions without using any intermediate game. In Section 4, we utilize this result and provide some conditions for externalities and internalities that determine the relationship among the four stability notions. We offer a weak condition for externalities for α-stability to coincide with β-stability and a condition for the set of γ-stable coalition structures to be a subset of δ-stable coalition structures (and the condition for its converse inclusion). In addition, we provide conditions for internalities for a game to have a coalition structure that satisfies all four stability concepts. In Section 5, another characterization is provided. We introduce the notions of a δ-consistent partition and a γ-consistent partition to describe what partition consistently connects the partition from which the players deviate and the partition to which they deviate. We also propose a constructive procedure to specify such consistent partitions. In Section 6, we apply our results to Cournot oligopoly and analyze its stable coalition structures.

In this paper, except in Section 6, we consider a general CS-game ϕ. We can derive a CS-game ϕ from a coalition structure value (a CS-value), namely, a solution concept for a TU-game with coalition structures. For example, Hart and Kurz (1984) and Abe (2018) focus on the Owen value (Owen, 1977) and analyze coalition structures in a symmetric majority game and an apex game. Casajus (2009) proposes his value concept, called χ-value, as a new solution concept and analyzes stable coalition structures in a gloves game. These value concepts are closely related to efficiency properties: the class of CS-values can be divided into two classes. One is the class of N-efficient values: the Owen value and the Kamijo (2009) value belong to this class. N-efficiency is also known as overall efficiency since the summation of the distributed payoffs is equal to the worth of the grand coalition for any partition. The other is the class of values satisfying coalitional efficiency: the Aumann-Drèze (1974) value, the Casajus value, and the Wiese (2007) value belong to this class. This efficiency notion requires the payoff distribution to be efficient within each coalition in a partition. Our future research topic is to study the relationship between the efficiency notions and the four stability concepts. This research will reveal (i) the general relationship between CS-games and CS-values and (ii) that between stable coalition structures and efficiency properties.

Appendix

Proof of Lemma 3.1

Let $T \subseteq N$ and $\sigma_T \in \mathcal{A}_T$. For any $\mathcal{P}, \mathcal{P}' \in \Pi(N \setminus T)$,

$$\mathcal{B}^*(\sigma_T, \sigma^{\mathcal{P}}_{N\setminus T})|_T = \mathcal{B}^*(\sigma_T, \sigma^{\mathcal{P}'}_{N\setminus T})|_T =: \mathcal{Q}.$$

Moreover, for any $\mathcal{P} \in \Pi(N \setminus T)$,

$$\mathcal{Q} \subseteq \mathcal{B}^*(\sigma_T, \sigma^{\mathcal{P}}_{N\setminus T})$$

and

$$\mathcal{Q} \cup \mathcal{P} = \mathcal{B}^*(\sigma_T, \sigma^{\mathcal{P}}_{N\setminus T}).$$

Proof. We begin with $\mathcal{B}^\gamma$. We first show that $\mathcal{B}^\gamma(\sigma_T, \sigma^{\mathcal{P}}_{N\setminus T})|_T = \mathcal{B}^\gamma(\sigma_T, \sigma^{\mathcal{P}'}_{N\setminus T})|_T$. Let $S \in \mathcal{B}^\gamma(\sigma_T, \sigma^{\mathcal{P}}_{N\setminus T})|_T$. There exists $C \in \mathcal{B}^\gamma(\sigma_T, \sigma^{\mathcal{P}}_{N\setminus T})$ such that $T \cap C = S$. By the definition of $\sigma^{\mathcal{P}}$, $\mathcal{P} \subseteq \mathcal{B}^\gamma(\sigma_T, \sigma^{\mathcal{P}}_{N\setminus T})$. We have $C \subseteq T$ because if there is $i \in C\setminus T$, then $C \overset{i\in C \in \mathcal{B}^\gamma}{=} \mathcal{B}^\gamma(\sigma_T, \sigma^{\mathcal{P}}_{N\setminus T})(i) \overset{i \in \mathcal{P}(i) \in \mathcal{B}^\gamma}{=} \mathcal{P}(i)$ holds, while $C \cap T \neq \emptyset$ and $\mathcal{P}(i) \cap T \overset{\mathcal{P}\in\Pi(N\setminus T)}{=} \emptyset$. Now, since $C \subseteq T$, the fact $T \cap C = S$ implies $C = S$. Hence, $S \in \mathcal{B}^\gamma(\sigma_T, \sigma^{\mathcal{P}}_{N\setminus T})$ and $S \subseteq T$ hold. In view of the definition of $\mathcal{B}^\gamma$,

$$\sigma_i = S \text{ for any } i \in S. \tag{A.1}$$

Hence, $S \in \mathcal{B}^\gamma(\sigma_T, \sigma^{\mathcal{P}'}_{N\setminus T})$ and $S \subseteq T$ follow. Thus, $S \in \mathcal{B}^\gamma(\sigma_T, \sigma^{\mathcal{P}'}_{N\setminus T})|_T$. The opposite inclusion relation is also proved in the same manner. As for $\mathcal{B}^\delta$, replace (A.1) by "$\sigma_i = \sigma_{i'}$ for any $i, i' \in S$." The fact that for any $j \in N \setminus T$, $\sigma^{\mathcal{P}}(j) \cap T = \emptyset$ implies $\sigma^{\mathcal{P}}(j) \cap S = \emptyset$. Hence, $S \in \mathcal{B}^\delta(\sigma_T, \sigma^{\mathcal{P}'}_{N\setminus T})$ also follows.

Now, we prove the second statement. Since $S \in \mathcal{B}^\gamma(\sigma_T, \sigma^{\mathcal{P}}_{N\setminus T})$ for any $S \in \mathcal{B}^\gamma(\sigma_T, \sigma^{\mathcal{P}}_{N\setminus T})|_T$, we have $\mathcal{B}^\gamma(\sigma_T, \sigma^{\mathcal{P}}_{N\setminus T})|_T = \mathcal{B}^\gamma(\sigma_T, \sigma^{\mathcal{P}}_{N\setminus T}) \setminus \mathcal{P}$. The same holds for $\mathcal{B}^\delta$. This completes the second statement.

The third statement follows from $\mathcal{P}, \mathcal{Q} \subseteq \mathcal{B}^\gamma(\sigma_T, \sigma^{\mathcal{P}}_{N\setminus T})$, $\mathcal{P} \in \Pi(N \setminus T)$, and $\mathcal{Q} \in \Pi(T)$. □

Proof of Proposition 3.2

For any partition $\mathcal{P} \in \Pi(N)$, the following four statements hold.

i. $\mathcal{P}$ is α-stable if and only if there exist no $T \subseteq N$ and no $\mathcal{Q} \in \Pi(T)$ such that for any $\mathcal{P}' \in \Pi(N\setminus T)$, $\phi_i(\mathcal{Q} \cup \mathcal{P}') > \phi_i(\mathcal{P})$ for every $i \in T$.
ii. $\mathcal{P}$ is β-stable if and only if there exist no $T \subseteq N$ such that for any $\mathcal{P}' \in \Pi(N \setminus T)$, there exists $\mathcal{Q} \in \Pi(T)$ such that $\phi_i(\mathcal{Q} \cup \mathcal{P}') > \phi_i(\mathcal{P})$ for every $i \in T$.
iii. $\mathcal{P}$ is γ-stable if and only if there exist no $T \subseteq N$ and no $\mathcal{Q} \in \Pi(T)$ such that $\phi_i(\mathcal{Q} \cup [\widehat{\mathcal{P}_T} \setminus T] \cup (\mathcal{P} \setminus \overline{\mathcal{P}_T})) > \phi_i(\mathcal{P})$ for every $i \in T$.
iv. $\mathcal{P}$ is δ-stable if and only if there exist no $T \subseteq N$ and no $\mathcal{Q} \in \Pi(T)$ such that $\phi_i(\mathcal{Q} \cup (\mathcal{P}|_{N\setminus T})) > \phi_i(\mathcal{P})$ for every $i \in T$.

Proof. The if-part of α-stability: Assume that $\mathcal{P}$ is not α-stable, namely, $\phi(\mathcal{P}) \notin C(V^\alpha)$. There are $T \subseteq N$ and $y \in V^\alpha(T)$ such that $y_i > \phi_i(\mathcal{P})$ for any $i \in T$. For the T and y, by the definition of V^α, there exists $\tilde{\sigma}_T \in \mathcal{A}_T$ such that for any $\sigma_{N\setminus T} \in \mathcal{A}_{N\setminus T}$, $\phi_i(\mathcal{B}^*(\tilde{\sigma}_T, \sigma_{N\setminus T})) \geq y_i > \phi_i(\mathcal{P})$ for any $i \in T$. Fix the $\tilde{\sigma}_T$. Lemma 3.1 implies that for the fixed $\tilde{\sigma}_T$ and any $\mathcal{P}' \in \Pi(N \setminus T)$, $\mathcal{Q} := \mathcal{B}^*(\tilde{\sigma}_T, \sigma^{\mathcal{P}'}_{N\setminus T})|_T$ satisfies $\mathcal{Q} \cup \mathcal{P}' = \mathcal{B}^*(\tilde{\sigma}_T, \sigma^{\mathcal{P}'}_{N\setminus T})$. Thus, for the T and $\mathcal{Q}$, we have $\phi_i(\mathcal{Q} \cup \mathcal{P}') > \phi_i(\mathcal{P})$ for any $\mathcal{P}' \in \Pi(N \setminus T)$ and any $i \in T$. This is a contradiction.

The only-if-part of α-stability: Assume that there exist $T \subseteq N$ and $\mathcal{Q} \in \Pi(T)$ such that for any $\mathcal{P}' \in \Pi(N \setminus T)$, $\phi_i(\mathcal{Q} \cup \mathcal{P}') > \phi_i(\mathcal{P})$ for every $i \in T$. For any $\sigma_{N\setminus T} \in \mathcal{A}_{N\setminus T}$, since $\mathcal{Q} \subseteq \mathcal{B}^*(\sigma^{\mathcal{Q}}_T, \sigma_{N\setminus T})$, we have $\phi_i(\mathcal{B}^*(\sigma^{\mathcal{Q}}_T, \sigma_{N\setminus T})) > \phi_i(\mathcal{P})$ for every $i \in T$. Now, for any $y \in \mathbb{R}^T$, let $F^T(y) = \{x_T \in \mathbb{R}^T | x_j \leq y_j \text{ for all } j \in T\}$. For simplicity, write $\phi_T(\sigma_{N\setminus T}) := (\phi_i(\mathcal{B}^*(\sigma^{\mathcal{Q}}_T, \sigma_{N\setminus T})))_{i\in T}$. We define

$$z := \sup \bigcap_{\sigma_{N\setminus T} \in \mathcal{A}_{N\setminus T}} F^T(\phi_T(\sigma_{N\setminus T})).$$

Note that z is unique because of the construction of F^T. The payoff vector z is in $V^\alpha(T)$, while $z_i > \phi_i(\mathcal{P})$ for any $i \in T$. Hence, $\mathcal{P}$ is not α-stable.

The if-part of β-stability: Assume that $\mathcal{P}$ is not β-stable, namely, $\phi(\mathcal{P}) \notin C(V^\beta)$. There are $T \subseteq N$ and $y \in V^\beta(T)$ such that $y_i > \phi_i(\mathcal{P})$ for any $i \in T$. For the T and y, by the definition of V^β, for any $\sigma_{N\setminus T} \in \mathcal{A}_{N\setminus T}$, there exists $\tilde{\sigma}_T \in \mathcal{A}_T$ such that $\phi_i(\mathcal{B}^*(\tilde{\sigma}_T, \sigma_{N\setminus T})) \geq y_i > \phi_i(\mathcal{P})$ for any $i \in T$. Hence, for any $\mathcal{P}' \in \Pi(N \setminus T)$, there exists $\tilde{\sigma}_T \in \mathcal{A}_T$ such that $\phi_i(\mathcal{B}^*(\tilde{\sigma}_T, \sigma^{\mathcal{P}'}_{N\setminus T})) > \phi_i(\mathcal{P})$ for any $i \in T$. For any $\mathcal{P}' \in \Pi(N \setminus T)$, since $\mathcal{B}^*(\tilde{\sigma}_T, \sigma^{\mathcal{P}'}_{N\setminus T}) \supseteq \mathcal{P}'$, set $\mathcal{Q}(\mathcal{P}') := \mathcal{B}^*(\tilde{\sigma}_T, \sigma^{\mathcal{P}'}_{N\setminus T}) \setminus \mathcal{P}'$. For any $\mathcal{P}' \in \Pi(N \setminus T)$, such

$\mathcal{Q}(\mathcal{P}')$ establishes that $\phi_i(\mathcal{Q}(\mathcal{P}') \cup \mathcal{P}') > \phi_i(\mathcal{P})$ for any $i \in T$.

The only-if-part of β-stability: Assume that there exists $T \subseteq N$ such that for any $\mathcal{P}' \in \Pi(N \setminus T)$ there is $\mathcal{Q} \in \Pi(T)$ such that $\phi_i(\mathcal{Q} \cup \mathcal{P}') > \phi_i(\mathcal{P})$ for every $i \in T$. Fix the coalition T. For any $\mathcal{P}' \in \Pi(N \setminus T)$, let $\mathcal{Q}(\mathcal{P}')$ be a partition satisfying the inequality. Define

$$z_T := \sup \bigcap_{\mathcal{P}' \in \Pi(N \setminus T)} F^T(\phi_T(\mathcal{Q}(\mathcal{P}') \cup \mathcal{P}')).$$

For any $\mathcal{P}' \in \Pi(N \setminus T)$,

$$\phi_i(\mathcal{Q}(\mathcal{P}') \cup \mathcal{P}') \geq z_i > \phi_i(\mathcal{P}) \text{ for any } i \in T. \tag{A.2}$$

Now, we show that $z_T \in V^\beta(T)$. For any $\sigma_{N\setminus T} \in \mathcal{A}_{N\setminus T}$, set $\dot{\mathcal{B}}^\gamma(\sigma_{N\setminus T}) := \{\dot{T}^i_{\sigma_{N\setminus T}} | i \in N \setminus T\}$, where

$$\dot{T}^i_{\sigma_{N\setminus T}} = \begin{cases} \sigma_i & \text{if } \sigma_i = \sigma_j \text{ for all } j \in \sigma_i \text{ and } \sigma_i \subseteq N \setminus T, \\ \{i\} & \text{otherwise.} \end{cases}$$

Note that $\dot{\mathcal{B}}^\gamma$ is defined so as to satisfy $\mathcal{B}^\gamma(\sigma_T^{\mathcal{Q}(\dot{\mathcal{B}}^\gamma(\sigma_{N\setminus T}))}, \sigma_{N\setminus T}) = \mathcal{Q}(\dot{\mathcal{B}}^\gamma(\sigma_{N\setminus T})) \cup \dot{\mathcal{B}}^\gamma(\sigma_{N\setminus T})$ for any $\sigma_{N\setminus T} \in \mathcal{A}_{N\setminus T}$.[*12] Hence, in view of (A.2), for any $\sigma_{N\setminus T} \in \mathcal{A}_{N\setminus T}$, we have

$$\phi_i(\mathcal{B}^\gamma(\sigma_T^{\mathcal{Q}(\dot{\mathcal{B}}^\gamma(\sigma_{N\setminus T}))}, \sigma_{N\setminus T})) = \phi_i(\mathcal{Q}(\dot{\mathcal{B}}^\gamma(\sigma_{N\setminus T})) \cup \dot{\mathcal{B}}^\gamma(\sigma_{N\setminus T})) \overset{(A.2)}{\geq} z_i$$

for any $i \in T$. Hence, $z_T \in V^\beta(T)$, and $\mathcal{P}$ is not β-stable because of such T and z_T.

The if-part of γ-stability: Let $\mathcal{P} \in \Pi(N)$. There is no $T \subseteq N$ and no $\mathcal{Q} \in \Pi(T)$ such that $\phi_i(\mathcal{Q} \cup [\widehat{\mathcal{P}_T} \setminus T] \cup (\mathcal{P} \setminus \overline{\mathcal{P}_T})) > \phi_i(\mathcal{P})$ for every $i \in T$. Hence, for any $T \subseteq N$ and any $\mathcal{Q} \in \Pi(T)$, there exists $i^* \in T$ such that

$$\phi_{i^*}(\mathcal{Q} \cup [\widehat{\mathcal{P}_T} \setminus T] \cup (\mathcal{P} \setminus \overline{\mathcal{P}_T})) \leq \phi_{i^*}(\mathcal{P}). \tag{A.3}$$

Now, assume that the strategy profile $\sigma^{\mathcal{P}}$ is not a strong equilibrium in G^γ_ϕ: there exists $(T^*, \hat{\sigma}_{T^*})$ such that for any $i \in T^*$, $\hat{\sigma}_i \neq \sigma_i^{\mathcal{P}}$ and

$$\phi_i(\mathcal{B}^\gamma(\hat{\sigma}_{T^*}, \sigma^{\mathcal{P}}_{N\setminus T^*})) > \phi_i(\mathcal{B}^\gamma(\sigma^{\mathcal{P}})) = \phi_i(\mathcal{P}), \tag{A.4}$$

where the equality holds because $\mathcal{B}^\gamma(\sigma^{\mathcal{P}}) = \mathcal{P}$. We fix T^*. From (A.3) and (A.4), it follows that for any $\mathcal{Q} \in \Pi(T^*)$, there exists $i^* \in T^*$ such that

$$\phi_{i^*}(\mathcal{B}^\gamma(\hat{\sigma}_{T^*}, \sigma^{\mathcal{P}}_{N\setminus T^*})) > \phi_{i^*}(\mathcal{P}) \geq \phi_{i^*}(\mathcal{Q} \cup [\widehat{\mathcal{P}_{T^*}} \setminus T^*] \cup (\mathcal{P} \setminus \overline{\mathcal{P}_{T^*}})). \tag{A.5}$$

We now focus on the partition $\mathcal{B}^\gamma(\hat{\sigma}_{T^*}, \sigma^{\mathcal{P}}_{N\setminus T^*})$. Note that $T^* \subseteq \widehat{\mathcal{P}_{T^*}} \subseteq N$. In view of the strategy profile $(\hat{\sigma}_{T^*}, \sigma^{\mathcal{P}}_{N\setminus T^*})$, every player $i \in N \setminus T^*$ chooses $\sigma_i^{\mathcal{P}} = \mathcal{P}(i)$. Hence, for every $i \in N \setminus \widehat{\mathcal{P}_{T^*}}$, $\sigma_i^{\mathcal{P}} = \mathcal{P}(i)$. Note that $\mathcal{P} \setminus \overline{\mathcal{P}_{T^*}}$ is a partition of $N \setminus \widehat{\mathcal{P}_{T^*}}$, which means $\mathcal{P}(i) \in \mathcal{P} \setminus \overline{\mathcal{P}_{T^*}}$ for every $i \in N \setminus \widehat{\mathcal{P}_{T^*}}$. In other words, for each coalition C in $\mathcal{P} \setminus \overline{\mathcal{P}_{T^*}}$, every member in C chooses the same C as his/her strategy. Hence, $\mathcal{B}^\gamma$ assigns $\mathcal{P}(i)$ to each player $i \in N \setminus \widehat{\mathcal{P}_{T^*}}$. We next arbitrarily fix a player i in $\widehat{\mathcal{P}_{T^*}} \setminus T^*$. Since $\hat{\sigma}_j \neq \sigma_j^{\mathcal{P}}$ for every $j \in T^*$, the new strategy of player $j \in \mathcal{P}(i) \cap T^*$, namely $\hat{\sigma}_j$, is different from the player i's strategy $\sigma_i^{\mathcal{P}} = \mathcal{P}(i)$. Namely, we have $\sigma_i^{\mathcal{P}} = \mathcal{P}(i) = \sigma_j^{\mathcal{P}} \neq \hat{\sigma}_j$. Hence, for the strategy profile $(\hat{\sigma}_{T^*}, \sigma^{\mathcal{P}}_{N\setminus T^*})$, $\mathcal{B}^\gamma$

[*12] For $\mathcal{B}^\delta$, we set $\dot{\mathcal{B}}^\delta(\sigma_{N\setminus T}) := \{\dot{T} \subseteq N \setminus T | i.j \in \dot{T} \iff \sigma_i = \sigma_j\}$.

assigns $\{i\}$ to each player $i \in \widehat{\mathcal{P}_{T^*}} \setminus T^*$. As a result, $\mathcal{B}^\gamma(\hat{\sigma}_{T^*}, \sigma^{\mathcal{P}}_{N\setminus T^*}) = [\widehat{\mathcal{P}_{T^*}} \setminus T^*] \cup (\mathcal{P} \setminus \overline{\mathcal{P}_{T^*}}) \cup \mathcal{Q}'$ for some $\mathcal{Q}' \in \Pi(T^*)$, where $\mathcal{Q}'$ is a partition of T^* induced by the strategy profile $\hat{\sigma}_{T^*}$. However, (A.5) must hold for any $\mathcal{Q} \in \Pi(T^*)$. This is a contradiction.

The only-if-part of γ-stability: Let $\mathcal{P}$ be γ-stable. The strategy profile $\sigma^{\mathcal{P}}$ is a strong equilibrium in G^γ_ϕ, which implies that for any $T \subseteq N$ and any $\hat{\sigma}_T \in \mathcal{A}_T$, there exists $i^* \in T$ such that

$$\phi_{i^*}(\mathcal{B}^\gamma(\hat{\sigma}_T, \sigma^{\mathcal{P}}_{N\setminus T})) \leq \phi_{i^*}(\mathcal{P}). \tag{A.6}$$

Now assume that there exist $T^* \subseteq N$ and $\mathcal{Q}^* \in \Pi(T^*)$ such that for any $i \in T^*$

$$\phi_i(\mathcal{Q}^* \cup [\widehat{\mathcal{P}_{T^*}} \setminus T^*] \cup (\mathcal{P} \setminus \overline{\mathcal{P}_{T^*}})) > \phi_i(\mathcal{P}). \tag{A.7}$$

For the fixed T^* and $\mathcal{Q}^*$, (A.6) and (A.7) result in

$$\phi_{i^*}(\mathcal{Q}^* \cup [\widehat{\mathcal{P}_{T^*}} \setminus T^*] \cup (\mathcal{P} \setminus \overline{\mathcal{P}_{T^*}})) > \phi_{i^*}(\mathcal{B}^\gamma(\hat{\sigma}_{T^*}, \sigma^{\mathcal{P}}_{N\setminus T^*})) \tag{A.8}$$

for any $\hat{\sigma}_{T^*} \in \mathcal{A}_{T^*}$. We set $\hat{\sigma}_i := \mathcal{Q}^*(i)$ for every $i \in T^*$. Since $\mathcal{Q}^*$ is a partition of T^*, we obtain $\mathcal{B}^\gamma(\hat{\sigma}_{T^*}, \sigma^{\mathcal{P}}_{N\setminus T^*}) = [\widehat{\mathcal{P}_{T^*}} \setminus T^*] \cup (\mathcal{P} \setminus \overline{\mathcal{P}_{T^*}}) \cup \mathcal{Q}^*$ in the same manner with the if-part.[*13] However, this contradicts (A.8).

The if-part of δ-stability: This is similar to the if-part of γ-stability. Replacing $\phi_i(\mathcal{Q} \cup [\widehat{\mathcal{P}_T} \setminus T] \cup (\mathcal{P}|_{N\setminus\widehat{\mathcal{P}_T}})) > \phi_i(\mathcal{P})$ by $\mathcal{Q} \cup (\mathcal{P}|_{N\setminus T})$, we have the following inequality that corresponds to (A.5): for any $\mathcal{Q} \in \Pi(T^*)$, there exists $i^* \in T^*$ such that $\phi_{i^*}(\mathcal{B}^\delta(\hat{\sigma}_{T^*}, \sigma^{\mathcal{P}}_{N\setminus T^*})) > \phi_{i^*}(\mathcal{Q} \cup (\mathcal{P}|_{N\setminus T^*}))$. For the given T^* and $\hat{\sigma}_{T^*}$, we set $\mathcal{Q}^* := \{R \subseteq T^* | i, j \in R \iff \hat{\sigma}_i = \hat{\sigma}_j\}$. Note that, as in the case of γ-stability, $\hat{\sigma}_i \neq \sigma^{\mathcal{P}}_i$ for every $i \in T^*$. In view of the definition of $\mathcal{B}^\delta$, we obtain $\mathcal{Q}^* \cup (\mathcal{P}|_{N\setminus T^*}) = \mathcal{B}^\delta(\hat{\sigma}_{T^*}, \sigma^{\mathcal{P}}_{N\setminus T^*})$. This is a contradiction.

The only-if-part of δ-stability: In the same manner with the only-if-part of γ-stability, there exist $T^* \subseteq N$, $\mathcal{Q}^* \in \Pi(T^*)$, and $i^* \in T^*$ such that for any $\hat{\sigma}_{T^*} \in \mathcal{A}_{T^*}$, $\phi_{i^*}(\mathcal{Q}^* \cup (\mathcal{P}|_{N\setminus T^*})) > \phi_{i^*}(\mathcal{B}^\delta(\hat{\sigma}_{T^*}, \sigma^{\mathcal{P}}_{N\setminus T^*}))$. Setting $\hat{\sigma}_i := \mathcal{Q}^*(i)$ for any $i \in T^*$, since $\mathcal{Q}^*$ is a partition of T^*, we have $\mathcal{B}^\delta(\hat{\sigma}_{T^*}, \sigma^{\mathcal{P}}_{N\setminus T^*}) = \mathcal{Q}^* \cup (\mathcal{P}|_{N\setminus T^*})$. This is a contradiction. □

Proof of Proposition 4.1

If a CS-game ϕ has weak nonnegative externalities or weak nonpositive externalities, then $C^\alpha(\phi) = C^\beta(\phi)$.

Proof. Let ϕ be a CS-game that has weak nonnegative externalities. In view of (2.3), it holds that $C^\alpha(\phi) \supseteq C^\beta(\phi)$. We show that $C^\alpha(\phi) \subseteq C^\beta(\phi)$. We prove that if $\mathcal{P}$ is not β-stable, then it is not α-stable. In view of Proposition 3.2, let $T \subseteq N$ be a coalition such that for any $\mathcal{P}' \in \Pi(N \setminus T)$, there exists $\mathcal{Q} \in \Pi(T)$ such that $\phi_i(\mathcal{Q} \cup \mathcal{P}') > \phi_i(\mathcal{P})$ for every $i \in T$. Hence, for partition $[N \setminus T]$ ($\in \Pi(N \setminus T)$),

[*13] Note that for any $i \in \widehat{\mathcal{P}_{T^*}} \setminus T^*$ and any $j \in \mathcal{P}(i) \cap T^*$, we have $\sigma^{\mathcal{P}}_i \neq \mathcal{Q}^*(j)$ because if this equality holds for some $i \in \widehat{\mathcal{P}_{T^*}} \setminus T^*$ and $j \in \mathcal{P}(i) \cap T^*$, then, in view of the definition of $\widehat{\mathcal{P}_{T^*}}$, $T^* \subsetneq \mathcal{P}(i) = \sigma^{\mathcal{P}}_i = \mathcal{Q}^*(j)$, which contradicts the fact that $\mathcal{Q}^*$ is a partition of T^*.

there exists $\mathcal{Q}^* \in \Pi(T)$ such that $\phi_i(\mathcal{Q}^* \cup [N \setminus T]) > \phi_i(\mathcal{P})$ for every $i \in T$. Since ϕ has weak nonnegative externalities, for any $\mathcal{P}' \in \Pi(N \setminus T)$ we have

$$\phi_i(\mathcal{Q}^* \cup \mathcal{P}') \geq \phi_i(\mathcal{Q}^* \cup [N \setminus T]) > \phi_i(\mathcal{P}) \text{ for every } i \in T.$$

Because of this pair $(T, \mathcal{Q}^*)$, partition $\mathcal{P}$ is not α-stable. If ϕ has weak nonpositive externalities, replacing $[N \setminus T]$ by $\{N \setminus T\}$ completes the poof. □

Proof of Proposition 4.2

If a CS-game ϕ has nonnegative externalities, then $C^\gamma(\phi) \supseteq C^\delta(\phi)$. If a CS-game ϕ has nonpositive externalities, then $C^\delta(\phi) \supseteq C^\gamma(\phi)$.

Proof. Let ϕ have nonnegative externalities. We prove that if $\mathcal{P}$ is not γ-stable, then it is not δ-stable. Since $\mathcal{P}$ is not γ-stable, in view of Proposition 3.2, there are $T \subseteq N$ and $\mathcal{Q} \in \Pi(T)$ such that $\phi_i(\mathcal{Q} \cup [\widehat{\mathcal{P}_T} \setminus T] \cup (\mathcal{P} \setminus \overline{\mathcal{P}_T})) > \phi_i(\mathcal{P})$ for every $i \in T$. We focus on the subpartition $[\widehat{\mathcal{P}_T} \setminus T] \cup (\mathcal{P} \setminus \overline{\mathcal{P}_T})$. For any $S \in \mathcal{P} \setminus \overline{\mathcal{P}_T}$, there exists $S' \in \mathcal{P}|_{N \setminus T}$ such that $S \subseteq S'$ because

$$\mathcal{P} \setminus \overline{\mathcal{P}_T} \overset{(2.1)}{=} \mathcal{P}|_{N \setminus \widehat{\mathcal{P}_T}} \overset{T \subseteq \widehat{\mathcal{P}_T}}{\subseteq} \mathcal{P}|_{N \setminus T}.$$

Moreover, $[\widehat{\mathcal{P}_T} \setminus T]$ is the finest partition of $\widehat{\mathcal{P}_T} \setminus T$. Hence, $[\widehat{\mathcal{P}_T} \setminus T] \cup (\mathcal{P} \setminus \overline{\mathcal{P}_T})$ is finer than $\mathcal{P}|_{N \setminus T}$. In view of nonnegative externalities, we have $\phi_i(\mathcal{Q} \cup (\mathcal{P}|_{N \setminus T})) \geq \phi_i(\mathcal{Q} \cup [\widehat{\mathcal{P}_T} \setminus T] \cup (\mathcal{P} \setminus \overline{\mathcal{P}_T})) > \phi_i(\mathcal{P})$ for any $i \in T$. Hence, $\mathcal{P}$ is not δ-stable.

If ϕ has nonpositive externalities, then we have $\phi_i(\mathcal{Q} \cup [\widehat{\mathcal{P}_T} \setminus T] \cup (\mathcal{P} \setminus \overline{\mathcal{P}_T})) \geq \phi_i(\mathcal{Q} \cup (\mathcal{P}|_{N \setminus T})) > \phi_i(\mathcal{P})$ for any $i \in T$ in the same manner. □

Proof of Proposition 5.5

Let $\mathcal{P}, \mathcal{P}' \in \Pi(N)$ with $\mathcal{P} \neq \mathcal{P}'$. Partition $\mathcal{Q}$ is a δ-consistent partition from $\mathcal{P}$ to $\mathcal{P}'$ if and only if $\mathcal{Q} \in \overline{\Delta}^{\mathcal{P}}(\mathcal{P}')$, where

$$\begin{aligned}\overline{\Delta}^{\mathcal{P}}(\mathcal{P}') = \{\mathcal{Q} \in \overline{\Pi}(N) |\ & \text{there exists } \mathcal{R} \subseteq \mathcal{L}^{\mathcal{P}}(\mathcal{P}') \text{ such that} \\ & \mathcal{Q} = \mathcal{Z}^{\mathcal{P}}(\mathcal{P}') \cup \mathcal{R} \text{ and } \mathcal{P}'|_{N \setminus \widehat{\mathcal{Q}}} = \mathcal{P}|_{N \setminus \widehat{\mathcal{Q}}}\}\end{aligned}$$

Note that $\mathcal{R}$ can be empty.

Proof. We first show that a δ-consistent partition from $\mathcal{P}$ to $\mathcal{P}'$ is in $\overline{\Delta}^{\mathcal{P}}(\mathcal{P}')$. Let $\mathcal{Q}$ be a δ-consistent partition from $\mathcal{P}$ to $\mathcal{P}'$. The equality $\mathcal{P}'|_{N \setminus \widehat{\mathcal{Q}}} = \mathcal{P}|_{N \setminus \widehat{\mathcal{Q}}}$ readily follows from $\mathcal{P}' = \mathcal{Q} \cup (\mathcal{P}|_{N \setminus \widehat{\mathcal{Q}}})$. We assume that $\mathcal{Z}^{\mathcal{P}}(\mathcal{P}') \not\subseteq \mathcal{Q}$. Then there exists $C' \in \mathcal{Z}^{\mathcal{P}}(\mathcal{P}')$ such that $C' \notin \mathcal{Q}$. By $C' \in \mathcal{Z}^{\mathcal{P}}(\mathcal{P}')$, we have $C' \in \mathcal{P}'$. The three facts $C' \in \mathcal{P}'$, $C' \notin \mathcal{Q}$, and $\mathcal{P}' = \mathcal{Q} \cup (\mathcal{P}|_{N \setminus \widehat{\mathcal{Q}}})$ imply that $C' \in \mathcal{P}|_{N \setminus \widehat{\mathcal{Q}}}$. Moreover, $\mathcal{P}' = \mathcal{Q} \cup (\mathcal{P}|_{N \setminus \widehat{\mathcal{Q}}})$ implies that $\mathcal{P}|_{N \setminus \widehat{\mathcal{Q}}} \subseteq \mathcal{P}'$. Hence, in view of Lemma 5.1, $\mathcal{P}|_{N \setminus \widehat{\mathcal{Q}}} \subseteq \mathcal{L}^{\mathcal{P}}(\mathcal{P}')$. Since $C' \in \mathcal{P}|_{N \setminus \widehat{\mathcal{Q}}}$, we obtain $C' \in \mathcal{L}^{\mathcal{P}}(\mathcal{P}')$. This contradicts $C' \in \mathcal{Z}^{\mathcal{P}}(\mathcal{P}')$ as $\mathcal{L}^{\mathcal{P}}(\mathcal{P}') = \mathcal{P}' \setminus \mathcal{Z}^{\mathcal{P}}(\mathcal{P}')$. Hence, $\mathcal{Z}^{\mathcal{P}}(\mathcal{P}') \subseteq \mathcal{Q}$. From $\mathcal{L}^{\mathcal{P}}(\mathcal{P}') = \mathcal{P}' \setminus \mathcal{Z}^{\mathcal{P}}(\mathcal{P}')$, it follows that $\mathcal{R} := \mathcal{Q} \setminus \mathcal{Z}^{\mathcal{P}}(\mathcal{P}') \subseteq \mathcal{L}^{\mathcal{P}}(\mathcal{P}')$. Thus, $\mathcal{Q} = \mathcal{Z}^{\mathcal{P}}(\mathcal{P}') \cup \mathcal{R}$.

Now, we show that $\mathcal{Q} \in \overline{\Delta}^{\mathcal{P}}(\mathcal{P}')$ is δ-consistent. In view of $\mathcal{P}'|_{N\setminus\widehat{\mathcal{Q}}} = \mathcal{P}|_{N\setminus\widehat{\mathcal{Q}}}$, we have $\mathcal{Q} \cup (\mathcal{P}|_{N\setminus\widehat{\mathcal{Q}}}) = \mathcal{Q} \cup (\mathcal{P}'|_{N\setminus\widehat{\mathcal{Q}}})$. Since $\mathcal{Z}^{\mathcal{P}}(\mathcal{P}') \cup \mathcal{R} \subseteq \mathcal{P}'$, we have $\mathcal{Q} = \mathcal{Z}^{\mathcal{P}}(\mathcal{P}') \cup \mathcal{R} \subseteq \mathcal{P}'$. Hence, we have $\mathcal{Q} \cup (\mathcal{P}'|_{N\setminus\widehat{\mathcal{Q}}}) = \mathcal{Q} \cup (\mathcal{P}' \setminus \mathcal{Q}) = \mathcal{P}'$. Thus, we obtain $\mathcal{Q} \cup (\mathcal{P}|_{N\setminus\widehat{\mathcal{Q}}}) = \mathcal{P}'$. $\square$

Proof of Proposition 5.6

For any $\mathcal{P}$ and $\mathcal{P}' \in \Pi(N)$ with $\mathcal{P} \neq \mathcal{P}'$, $\mathcal{P}' \setminus \Lambda^{\mathcal{P}}(\mathcal{P}')$ is a minimal δ-consistent partition.

Proof. We fix $\mathcal{P}$ and $\mathcal{P}' \in \Pi(N)$ with $\mathcal{P} \neq \mathcal{P}'$ through the proof. For simplicity, omitting $\mathcal{P}$ and $\mathcal{P}'$, we write Λ, $\mathcal{Z}$, and $\mathcal{L}$ instead of $\Lambda^{P}(\mathcal{P}')$, $\mathcal{Z}^{P}(\mathcal{P}')$, and $\mathcal{L}^{P}(\mathcal{P}')$. Define $\mathcal{R}^{\mathcal{P}}_{\mathcal{P}'} := \mathcal{P}' \setminus (\Lambda \cup \mathcal{Z})$. Similarly, we write $\mathcal{R}$ instead of $\mathcal{R}^{P}_{\mathcal{P}'}$. The proof consists of two parts. The first part shows $\mathcal{P}' \setminus \Lambda$ is a δ-consistent partition. In the second part, we prove $\mathcal{P}' \setminus \Lambda$ is minimal.

Part 1: $\mathcal{P}' \setminus \Lambda \in \overline{\Delta}^{\mathcal{P}}(\mathcal{P}')$. First we have $\Lambda \subseteq \mathcal{L}$ because in view of (5.2), for any $T \in \Lambda$, there exists $S \in \mathcal{P}$ such that $\lambda(S) = T \subseteq S$. Since $\mathcal{Z} = \mathcal{P}' \setminus \mathcal{L}$, we have $\Lambda \cap \mathcal{Z} = \emptyset$. Hence, $\mathcal{R}$, Λ, and $\mathcal{Z}$ are disjoint, and

$$\mathcal{P}' = \mathcal{R} \cup \Lambda \cup \mathcal{Z}. \tag{A.9}$$

We have

$$\begin{aligned}\mathcal{R} &= \mathcal{P}' \setminus (\Lambda \cup \mathcal{Z}) = (\mathcal{P}' \setminus \Lambda) \cup (\mathcal{P}' \setminus \mathcal{Z}) \\ &= (\mathcal{P}' \setminus \Lambda) \cup \mathcal{L} \subseteq \mathcal{L}.\end{aligned} \tag{A.10}$$

It readily follows from (A.9) that

$$\mathcal{P}' \setminus \Lambda = \mathcal{R} \cup \mathcal{Z}. \tag{A.11}$$

Now, let $\mathcal{Q} := \mathcal{R} \cup \mathcal{Z}(= \mathcal{P}' \setminus \Lambda)$. In view of the definition of $\overline{\Delta}^{\mathcal{P}}(\mathcal{P}')$, since we have (A.10) and (A.11), it suffices to show $\mathcal{P}'|_{N\setminus\widehat{\mathcal{Q}}} = \mathcal{P}|_{N\setminus\widehat{\mathcal{Q}}}$. Note that by $\mathcal{Q} = \mathcal{Z} \cup \mathcal{R} \subseteq \mathcal{P}'$, we have

$$\mathcal{P}'|_{N\setminus\widehat{\mathcal{Q}}} \overset{(2.1)}{=} \mathcal{P}' \setminus (\mathcal{Z} \cup \mathcal{R}) \overset{(A.9)}{=} \Lambda. \tag{A.12}$$

We first show that $\mathcal{P}'|_{N\setminus\widehat{\mathcal{Q}}} \subseteq \mathcal{P}|_{N\setminus\widehat{\mathcal{Q}}}$. Let $T \in \mathcal{P}'|_{N\setminus\widehat{\mathcal{Q}}}$, namely, $T \in \Lambda$. By (5.2), there exists a coalition $S \in \mathcal{P}$ such that $T \subseteq S$. For this coalition $S \in \mathcal{P}$, we have

$$T \subseteq (N \setminus \widehat{\mathcal{Q}}) \cap S,$$

because $T \subseteq S$ and $T \in \mathcal{P}'|_{N\setminus\widehat{\mathcal{Q}}}$, which implies $T \subseteq N \setminus \widehat{\mathcal{Q}}$. Now assume that there exists a player i in $((N \setminus \widehat{\mathcal{Q}}) \cap S) \setminus T$. Since $i \in S$ and $S \in \mathcal{P}$, we have $\mathcal{P}(i) = S$. Since $i \in N \setminus \widehat{\mathcal{Q}} = N \setminus (\widehat{\mathcal{Z} \cup \mathcal{R}})$, we have $i \in \widehat{\Lambda}$, which implies that $\mathcal{P}'(i) \in \Lambda$ since $\Lambda \subseteq \mathcal{P}'$. Since $\mathcal{P}'(i) \in \Lambda$ and $\mathcal{P}(i) = S$, we obtain $\mathcal{P}(i) = \mathcal{P}'(i)$. However, the facts $T \in \mathcal{P}'$ and $i \notin T$ imply $\mathcal{P}'(i) \cap T = \emptyset$. By $\emptyset \neq T \subseteq S = \mathcal{P}(i)$, we have $T \subseteq \mathcal{P}(i) \setminus \mathcal{P}'(i)$, which contradicts $\mathcal{P}(i) = \mathcal{P}'(i)$. Hence, we obtain

$$T = (N \setminus \widehat{\mathcal{Q}}) \cap S. \tag{A.13}$$

Since $S \in \mathcal{P}$ and $\mathcal{P}|_{N\setminus\widehat{\mathcal{Q}}} = \{(N \setminus \widehat{\mathcal{Q}}) \cap C | C \in \mathcal{P}, (N \setminus \widehat{\mathcal{Q}}) \cap C \neq \emptyset\}$, setting $C := S$, from (A.13) it follows that $T \in \mathcal{P}|_{N\setminus\widehat{\mathcal{Q}}}$.

We now prove $\underline{\mathcal{P}'|_{N\setminus\widehat{\mathcal{Q}}} \supseteq \mathcal{P}|_{N\setminus\widehat{\mathcal{Q}}}}$. Let $T \in \mathcal{P}|_{N\setminus\widehat{\mathcal{Q}}}$. We have $T \subseteq N \setminus \widehat{\mathcal{Q}} \overset{\text{(A.9)}}{=} \widehat{\Lambda}$. Assume that there exist $T_1, T_2 \in \Lambda$ such that $T_1 \neq T_2$, $T_1 \cap T \neq \emptyset$, and $T_2 \cap T \neq \emptyset$. From $T_1 \in \Lambda$, it follows that there exists $S_1 \in \mathcal{P}$ such that $\lambda(S_1) = T_1 \subseteq S_1$. Similarly, there exists $S_2 \in \mathcal{P}$ such that $\lambda(S_2) = T_2 \subseteq S_2$. If $S_1 = S_2$, then $T_1 = \lambda(S_1) = \lambda(S_2) = T_2$, which contradicts $T_1 \neq T_2$. If $S_1 \neq S_2$, then let $i_1 \in T_1 \cap T$ and $i_2 \in T_2 \cap T$. Since $i_1 \in S_1 \in \mathcal{P}$ and $i_1 \in T \in \mathcal{P}$, we have $S_1 = \mathcal{P}(i_1) = T$. In the same manner, $S_2 = \mathcal{P}(i_2) = T$. However, this contradicts $S_1 \neq S_2$. Hence, one or no $T' \in \Lambda$ satisfies $T' \cap T \neq \emptyset$. If there is no such T', it readily contradicts $T \subseteq \widehat{\Lambda}$. Hence, exactly one $T' \in \Lambda$ satisfies $T' \cap T \neq \emptyset$. In view of $T \subseteq \widehat{\Lambda}$, $T \subseteq T'$ follows. Since $T' \in \Lambda$, there exists $S' \in \mathcal{P}$ such that $T' \subseteq S'$. Since $T \in \mathcal{P}$ and $T' \cap T \neq \emptyset$, we have $S' = T$ and $T' \subseteq T$. Hence, $T' = T$. By $T' \in \Lambda$, $T \in \Lambda$. In view of (A.12), we obtain $T \in \mathcal{P}'|_{N\setminus\widehat{\mathcal{Q}}}$.

Thus, $\mathcal{P}'|_{N\setminus\widehat{\mathcal{Q}}} = \mathcal{P}|_{N\setminus\widehat{\mathcal{Q}}}$ holds. Together with (A.10) and (A.11), we obtain $\mathcal{P}' \setminus \Lambda \in \overline{\Delta}^{\mathcal{P}}(\mathcal{P}')$.

Part 2: Minimality. Consider $C \in \mathcal{P}' \setminus \Lambda$. Note that $\Lambda \subsetneq \mathcal{P}'$ because if $\Lambda = \mathcal{P}'$, $\mathcal{P} = \mathcal{P}'$, which contradicts $\mathcal{P} \neq \mathcal{P}'$. Now, define

$$\mathcal{Q}' := (\mathcal{P}' \setminus \Lambda) \setminus \{C\} = \mathcal{P}' \setminus (\Lambda \cup \{C\}). \tag{A.14}$$

Our purpose is to show that $\mathcal{P}'|_{N\setminus\widehat{\mathcal{Q}'}} \neq \mathcal{P}|_{N\setminus\widehat{\mathcal{Q}'}}$. We have $\mathcal{P}'|_{N\setminus\widehat{\mathcal{Q}'}} \overset{(2.1)}{=} \mathcal{P}' \setminus \mathcal{Q}' \overset{\text{(A.14)}}{=} \Lambda \cup \{C\}$. Since $N \setminus \widehat{\mathcal{Q}'} \overset{\text{(A.14)}}{=} \widehat{\Lambda} \cup C$, we have $\mathcal{P}|_{N\setminus\widehat{\mathcal{Q}'}} = \mathcal{P}|_{\widehat{\Lambda}\cup C}$. Hence, our purpose is equivalent to

$$\Lambda \cup \{C\} \neq \mathcal{P}|_{\widehat{\Lambda}\cup C}.$$

Now, since $C \notin \Lambda$, $C \in \mathcal{P}'$ and (A.9), we have $C \in \mathcal{Z}$ or $C \in \mathcal{R}$. <u>If $C \in \mathcal{Z}$</u>, then there is no $S \in \mathcal{P}$ such that $C \subseteq S$. Assume that $\Lambda \cup \{C\} = \mathcal{P}|_{\widehat{\Lambda}\cup C}$. Since $C \in \mathcal{P}|_{\widehat{\Lambda}\cup C}$, there exists $S \in \mathcal{P}$ such that $S \cap (\widehat{\Lambda} \cup C) = C$, which implies that, in view of $C \notin \Lambda$, $C \subseteq S$. This is a contradiction. Hence, $\Lambda \cup \{C\} \neq \mathcal{P}|_{\widehat{\Lambda}\cup C}$. <u>If $C \in \mathcal{R}$</u>, then $C \in \mathcal{L}$ since $\mathcal{R} \subseteq \mathcal{L}$. Hence, there exists $S \in \mathcal{P}$ such that $C \subseteq S$. We now consider $\lambda(S)$. In view of (5.2), $\lambda(S)$ satisfies $\lambda(S) \in \mathcal{P}'$ and $\lambda(S) \subseteq S$. Since $C \notin \Lambda$ and $\lambda(S) \in \Lambda$, we have $\lambda(S) \cap C = \emptyset$. Since $C \subseteq S$ and $\lambda(S) \subseteq S$, we have $(\lambda(S) \cup C) \subseteq S$. Moreover, $\lambda(S) \cup C = (\widehat{\Lambda} \cup C) \cap S$. By setting $C' := S$ below, we have

$$(\lambda(S) \cup C) \in \{(\widehat{\Lambda} \cup C) \cap C' | C' \in \mathcal{P}, (\widehat{\Lambda} \cup C) \cap C' \neq \emptyset\} = \mathcal{P}|_{\widehat{\Lambda}\cup C}.$$

However, partition $\mathcal{P}'$ (more specifically, its subpartition $\Lambda \cup \{C\}$) contains disjoint coalitions $\lambda(S)$ and C separately. Thus, $\Lambda \cup \{C\} \neq \mathcal{P}|_{\widehat{\Lambda}\cup C}$. □

Proof of Proposition 5.8

For any $\mathcal{P} \in \Pi(N)$, $\mathcal{P}$ is δ-stable if and only if for any $\mathcal{P}' \in \Pi(N) \setminus \{\mathcal{P}\}$ and any $\mathcal{Q} \in \Delta^{\mathcal{P}}(\mathcal{P}')$, there exists $i \in \widehat{\mathcal{Q}}$ such that $\phi_i(\mathcal{P}) \geq \phi_i(\mathcal{P}')$.

Proof. <u>If-part:</u> Suppose that $\mathcal{P}$ is not δ-stable. There exist $S \subseteq N$ and $\mathcal{Q} \in \Pi(S)$ such that for any $i \in S$, $\phi_i(\mathcal{Q} \cup (\mathcal{P}|_{N\setminus S})) > \phi_i(\mathcal{P})$. Set $\mathcal{P}' := \mathcal{Q} \cup (\mathcal{P}|_{N\setminus S})$. In view of Proposition 5.5, $\mathcal{Q} \in \overline{\Delta}^{\mathcal{P}}(\mathcal{P}')$. Hence, for the given $\mathcal{P}' = \mathcal{Q} \cup (\mathcal{P}|_{N\setminus S})$ and $\mathcal{Q} \in \overline{\Delta}^{\mathcal{P}}(\mathcal{P}')$, we have $\phi_i(\mathcal{P}') > \phi_i(\mathcal{P})$ for any $i \in S = \widehat{\mathcal{Q}}$. In view of

Remark 5.7, there exists $\mathcal{Q}^* \in \Delta^{\mathcal{P}}(\mathcal{P}')$ such that $\mathcal{Q}^* \subseteq \mathcal{Q}$, $\mathcal{Q}^* \cup (\mathcal{P}|_{N\setminus\widehat{\mathcal{Q}^*}}) = \mathcal{P}'$, and $\phi_i(\mathcal{P}') > \phi_i(\mathcal{P})$ for any $i \in \widehat{\mathcal{Q}^*}$.

Only-if-part: Let $\mathcal{P}$ be δ-stable. If there exist $\mathcal{P}' \in \Pi(N) \setminus \{\mathcal{P}\}$ and $\mathcal{Q} \in \Delta^{\mathcal{P}}(\mathcal{P}')$ such that $\phi_i(\mathcal{P}') > \phi_i(\mathcal{P})$ for any $i \in \widehat{\mathcal{Q}}$, then $\mathcal{P}' = \mathcal{Q} \cup (\mathcal{P}|_{N\setminus S})$ follows from Proposition 5.5. Hence, $\phi_i(\mathcal{Q} \cup (\mathcal{P}|_{N\setminus S})) > \phi_i(\mathcal{P})$ for any $i \in \widehat{\mathcal{Q}}$. This $\mathcal{Q}$ contradicts the fact that $\mathcal{P}$ is δ-stable. □

Proof of Proposition 5.10

Let $\mathcal{P}, \mathcal{P}' \in \Pi(N)$ with $\mathcal{P} \neq \mathcal{P}'$. Partition $\mathcal{Q}$ is a γ-consistent partition from $\mathcal{P}$ to $\mathcal{P}'$ if and only if $\mathcal{Q} \in \overline{\Gamma}^{\mathcal{P}}(\mathcal{P}')$, where

$$\overline{\Gamma}^{\mathcal{P}}(\mathcal{P}') = \{\mathcal{Q} \in \overline{\Pi}(N) | \text{ there exists } \mathcal{R} \subseteq \mathcal{L}^{\mathcal{P}}(\mathcal{P}') \text{ such that } \mathcal{Q} = \mathcal{Z}^{\mathcal{P}}(\mathcal{P}') \cup \mathcal{R} \text{ and } \mathcal{P}'|_{N\setminus\widehat{\mathcal{Q}}} = (\mathcal{P} \setminus \overline{\mathcal{P}_{\widehat{\mathcal{Q}}}}) \cup [\widehat{\mathcal{P}_{\widehat{\mathcal{Q}}}} \setminus \widehat{\mathcal{Q}}]\}.$$

Note that $\mathcal{R}$ can be empty.

Proof. We first show that a γ-consistent partition from $\mathcal{P}$ to $\mathcal{P}'$ is in $\overline{\Gamma}^{\mathcal{P}}(\mathcal{P}')$. Let $\mathcal{Q}$ be a γ-consistent partition from $\mathcal{P}$ to $\mathcal{P}'$. The equality $\mathcal{P}'|_{N\setminus\widehat{\mathcal{Q}}} = (\mathcal{P} \setminus \overline{\mathcal{P}_{\widehat{\mathcal{Q}}}}) \cup [\widehat{\mathcal{P}_{\widehat{\mathcal{Q}}}} \setminus \widehat{\mathcal{Q}}]$ follows from the definition of a γ-consistent partition: $\mathcal{P}' = \mathcal{Q} \cup (\mathcal{P} \setminus \overline{\mathcal{P}_{\widehat{\mathcal{Q}}}}) \cup [\widehat{\mathcal{P}_{\widehat{\mathcal{Q}}}} \setminus \widehat{\mathcal{Q}}]$. We assume that $\mathcal{Z}^{\mathcal{P}}(\mathcal{P}') \not\subseteq \mathcal{Q}$. Then there exists $C' \in \mathcal{Z}^{\mathcal{P}}(\mathcal{P}')$ such that $C' \notin \mathcal{Q}$. By $C' \in \mathcal{Z}^{\mathcal{P}}(\mathcal{P}')$, we have $C' \in \mathcal{P}'$. The three facts $C' \in \mathcal{P}'$, $C' \notin \mathcal{Q}$, and $\mathcal{P}' = \mathcal{Q} \cup (\mathcal{P} \setminus \overline{\mathcal{P}_{\widehat{\mathcal{Q}}}}) \cup [\widehat{\mathcal{P}_{\widehat{\mathcal{Q}}}} \setminus \widehat{\mathcal{Q}}]$ imply that $C' \in (\mathcal{P} \setminus \overline{\mathcal{P}_{\widehat{\mathcal{Q}}}}) \cup [\widehat{\mathcal{P}_{\widehat{\mathcal{Q}}}} \setminus \widehat{\mathcal{Q}}]$. If $C' \in [\widehat{\mathcal{P}_{\widehat{\mathcal{Q}}}} \setminus \widehat{\mathcal{Q}}]$, then C' is one-person coalition. Hence, $C' \in \mathcal{L}^{\mathcal{P}}(\mathcal{P}')$. If $C' \in (\mathcal{P} \setminus \overline{\mathcal{P}_{\widehat{\mathcal{Q}}}})$, then $C' \in \mathcal{P}$. Hence, $C' \in \mathcal{L}^{\mathcal{P}}(\mathcal{P}')$. However, in any case, this contradicts $C' \in \mathcal{Z}^{\mathcal{P}}(\mathcal{P}')$. Hence, $\mathcal{Z}^{\mathcal{P}}(\mathcal{P}') \subseteq \mathcal{Q}$. From $\mathcal{L}^{\mathcal{P}}(\mathcal{P}') = \mathcal{P}' \setminus \mathcal{Z}^{\mathcal{P}}(\mathcal{P}')$, it follows that $\mathcal{R} := \mathcal{Q} \setminus \mathcal{Z}^{\mathcal{P}}(\mathcal{P}') \subseteq \mathcal{L}^{\mathcal{P}}(\mathcal{P}')$. Thus, $\mathcal{Q} = \mathcal{Z}^{\mathcal{P}}(\mathcal{P}') \cup \mathcal{R}$.

Now, we show that $\mathcal{Q} \in \overline{\Gamma}^{\mathcal{P}}(\mathcal{P}')$ is γ-consistent. We have the following two facts (i) $\mathcal{P}'|_{N\setminus\widehat{\mathcal{Q}}} = (\mathcal{P} \setminus \overline{\mathcal{P}_{\widehat{\mathcal{Q}}}}) \cup [\widehat{\mathcal{P}_{\widehat{\mathcal{Q}}}} \setminus \widehat{\mathcal{Q}}]$ and (ii) $\mathcal{Q} \subseteq \mathcal{P}'$. Hence, we obtain $\mathcal{Q} \cup (\mathcal{P} \setminus \overline{\mathcal{P}_{\widehat{\mathcal{Q}}}}) \cup [\widehat{\mathcal{P}_{\widehat{\mathcal{Q}}}} \setminus \widehat{\mathcal{Q}}] \overset{\text{(i)}}{=} \mathcal{Q} \cup (\mathcal{P}'|_{N\setminus\widehat{\mathcal{Q}}}) \overset{\text{(ii)}}{=} \mathcal{Q} \cup (\mathcal{P}' \setminus \mathcal{Q}) = \mathcal{P}'$. □

Proof of Proposition 5.11

For any $\mathcal{P}$ and $\mathcal{P}' \in \Pi(N)$ with $\mathcal{P} \neq \mathcal{P}'$, $\mathcal{P}' \setminus \mathcal{K}^{\mathcal{P}}(\mathcal{P}')$ is a minimal γ-consistent partition.

Proof. We fix $\mathcal{P}$ and $\mathcal{P}' \in \Pi(N)$ with $\mathcal{P} \neq \mathcal{P}'$ through the proof. Omitting $\mathcal{P}$ and $\mathcal{P}'$, we write $\mathcal{K}$, $\mathcal{Z}$, and $\mathcal{L}$. Define $\mathcal{R}^{\mathcal{P}}_{\mathcal{P}'} := \mathcal{P}' \setminus (\mathcal{K} \cup \mathcal{Z})$. Similarly, we write $\mathcal{R}$ instead of $\mathcal{R}^{\mathcal{P}}_{\mathcal{P}'}$. The approach of the proof is similar to that of Proposition 5.6: the proof consists of two parts.

Part 1: $\mathcal{P}' \setminus \mathcal{K} \in \overline{\Gamma}^{\mathcal{P}}(\mathcal{P}')$. For any $T \in \mathcal{K}$, the following (i) or (ii) holds:

(i) there exists $S \in \mathcal{P}$ such that $T = S$,

(ii) $|T| = 1$ and there exist $S \in \mathcal{P}$ and $T' \subsetneq S$ such that $[T'] \subseteq \mathcal{P}'$, $T \in [T']$, and no T' with $T \subsetneq T' \subsetneq S$ satisfies $[T'] \subseteq \mathcal{P}'$.

For every $T \in \mathcal{K}$, if (i) holds, $T \in \mathcal{L}$ readily follows. If (ii) holds, then $|T| = 1$ and, hence, $T \in \mathcal{L}$. Hence, $\mathcal{K} \subseteq \mathcal{L}$. Since, $\mathcal{Z} = \mathcal{P}' \setminus \mathcal{L}$, we have $\mathcal{Z} \cap \mathcal{K} = \emptyset$. Hence, $\mathcal{R}, \mathcal{K}$, and $\mathcal{Z}$ are disjoint, and

$$\mathcal{P}' = \mathcal{R} \cup \mathcal{K} \cup \mathcal{Z}. \tag{A.15}$$

In the same manner with (A.11) of δ-stability, we have

$$\mathcal{R} \subseteq \mathcal{L}. \tag{A.16}$$

In view of (A.15),

$$\mathcal{P}' \setminus \mathcal{K} = \mathcal{R} \cup \mathcal{Z}. \tag{A.17}$$

Now, let $\mathcal{Q} := \mathcal{R} \cup \mathcal{Z} (= \mathcal{P}' \setminus \mathcal{K})$. In view of the definition of $\overline{\Gamma}^{\mathcal{P}}(\mathcal{P}')$ and (A.16) and (A.17), it suffices to show $\mathcal{P}'|_{N \setminus \widehat{\mathcal{Q}}} = (\mathcal{P} \setminus \overline{\mathcal{P}_{\widehat{\mathcal{Q}}}}) \cup [\widehat{\mathcal{P}_{\widehat{\mathcal{Q}}}} \setminus \widehat{\mathcal{Q}}]$.

We first show that $\underline{\mathcal{P}'|_{N \setminus \widehat{\mathcal{Q}}} \subseteq (\mathcal{P} \setminus \overline{\mathcal{P}_{\widehat{\mathcal{Q}}}}) \cup [\widehat{\mathcal{P}_{\widehat{\mathcal{Q}}}} \setminus \widehat{\mathcal{Q}}]}$. Since $\mathcal{Q} = \mathcal{R} \cup \mathcal{Z} \subseteq \mathcal{P}'$, we have $\mathcal{P}'|_{N \setminus \widehat{\mathcal{Q}}} \overset{(2.1)}{=} \mathcal{P}' \setminus \mathcal{Q} = \mathcal{K}$. Let $T \in \mathcal{P}'|_{N \setminus \widehat{\mathcal{Q}}}$, namely $T \in \mathcal{K}$.

<u>If (i) holds</u>, $T \in \mathcal{P}$ follows. Assume that $T \in \overline{\mathcal{P}_{\widehat{\mathcal{Q}}}}$. Since $\overline{\mathcal{P}_{\widehat{\mathcal{Q}}}} = \{C \in \mathcal{P} | \widehat{\mathcal{Q}} \cap C \neq \emptyset\}$, we have $T \cap \widehat{\mathcal{Q}} \neq \emptyset$. However, since $T \in \mathcal{P}'|_{N \setminus \widehat{\mathcal{Q}}}$, we have $T \subseteq N \setminus \widehat{\mathcal{Q}}$, which implies that $T \cap \widehat{\mathcal{Q}} = \emptyset$. This is a contradiction. Hence, $T \notin \overline{\mathcal{P}_{\widehat{\mathcal{Q}}}}$. From $T \in \mathcal{P}$, it follows that $T \in \mathcal{P} \setminus \overline{\mathcal{P}_{\widehat{\mathcal{Q}}}}$.

<u>If (ii) holds</u>, $|T| = 1$ follows. Let t be the player in the one-person coalition T. Since $T \in \mathcal{P}'|_{N \setminus \widehat{\mathcal{Q}}}$, we have $T \subseteq N \setminus \widehat{\mathcal{Q}}$ and $t \notin \widehat{\mathcal{Q}}$. Assume that $t \notin \widehat{\mathcal{P}_{\widehat{\mathcal{Q}}}}$. Since $\widehat{\mathcal{P}_{\widehat{\mathcal{Q}}}} = \cup_{C \in \mathcal{P}: C \cap \widehat{\mathcal{Q}} \neq \emptyset} C$, we have $\mathcal{P}(t) \cap \widehat{\mathcal{Q}} = \emptyset$. Now in view of (ii), the coalition S satisfies $t \in S \in \mathcal{P}$. Hence, we set $S := \mathcal{P}(t)$, and the condition (ii) implies that there exists coalition $T' \subsetneq \mathcal{P}(t)$. Let $i \in \mathcal{P}(t) \setminus T'$. Note that the coalition T' satisfies $[T'] \subseteq \mathcal{P}'$, $\{t\} \in [T']$ (*i.e.*, $t \in T'$), and no T'' with $T' \subsetneq T'' \subsetneq \mathcal{P}(t)$ satisfies $[T''] \subseteq \mathcal{P}'$ because of (ii). Now, consider coalition $\mathcal{P}'(i)$. <u>If $|\mathcal{P}'(i)| = 1$</u>, then set $T^* := T' \cup \{i\}$. The coalition T^* satisfies $[T^*] \subseteq \mathcal{P}'$ and $T' \subsetneq T^* \subsetneq \mathcal{P}(t)$. However, this readily contradicts the constraint that no T'' with $T' \subsetneq T'' \subsetneq \mathcal{P}(t)$ satisfies $[T''] \subseteq \mathcal{P}'$. <u>If $|\mathcal{P}'(i)| \geq 2$</u>, then we have $\mathcal{P}'(i) \in \mathcal{K}$, because if $\mathcal{P}'(i) \notin \mathcal{K}$ then $\mathcal{P}'(i) \in \mathcal{Q}$ and $i \in \widehat{\mathcal{Q}}$, while the facts $i \in \mathcal{P}(t)$ and $\mathcal{P}(t) \cap \widehat{\mathcal{Q}} = \emptyset$ imply $i \notin \widehat{\mathcal{Q}}$, a contradiction. Since $|\mathcal{P}'(i)| \geq 2$ and $\mathcal{P}'(i) \in \mathcal{K}$, condition (i) implies $\mathcal{P}'(i) = \mathcal{P}(i)$. From $i \in \mathcal{P}(t)$, it follows that $\mathcal{P}(i) = \mathcal{P}(t)$. However, since $t \in T'$ and $[T'] \subseteq \mathcal{P}'$, we have $\mathcal{P}'(t) = \{t\}$, which means $t \notin \mathcal{P}'(i)$. Hence, it holds that $t \in \mathcal{P}(t) \setminus \mathcal{P}'(i)$ and $\mathcal{P}(t) \neq \mathcal{P}'(i)$. Given that $\mathcal{P}(i) = \mathcal{P}(t)$, this contradicts $\mathcal{P}'(i) = \mathcal{P}(i)$.

We now show that $\underline{\mathcal{P}'|_{N \setminus \widehat{\mathcal{Q}}} \supseteq (\mathcal{P} \setminus \overline{\mathcal{P}_{\widehat{\mathcal{Q}}}}) \cup [\widehat{\mathcal{P}_{\widehat{\mathcal{Q}}}} \setminus \widehat{\mathcal{Q}}]}$. Let $T \in (\mathcal{P} \setminus \overline{\mathcal{P}_{\widehat{\mathcal{Q}}}}) \cup [\widehat{\mathcal{P}_{\widehat{\mathcal{Q}}}} \setminus \widehat{\mathcal{Q}}]$. If $T \in [\widehat{\mathcal{P}_{\widehat{\mathcal{Q}}}} \setminus \widehat{\mathcal{Q}}]$, then $T \subseteq N \setminus \widehat{\mathcal{Q}}$. Similarly, if $T \in (\mathcal{P} \setminus \overline{\mathcal{P}_{\widehat{\mathcal{Q}}}})$, then $T \subseteq N \setminus \widehat{\mathcal{Q}}$ follows from $\widehat{\mathcal{Q}} \subseteq \widehat{\mathcal{P}_{\widehat{\mathcal{Q}}}}$. Hence,

$$T \subseteq N \setminus \widehat{\mathcal{Q}} = \widehat{\mathcal{K}}. \tag{A.18}$$

Note that this does not mean $T \in \mathcal{K}$. Now, we show $T \in \mathcal{P}'$. Assume $T \notin \mathcal{P}'$. One of the following two disjoint statements holds:

(a) there exist $i_1, i_2 \in T$ such that $i_1 \neq i_2$ and $\mathcal{P}'(i_1) \neq \mathcal{P}'(i_2)$,
(b) there exists $i \in N \setminus T$ such that for any $j \in T$, $\mathcal{P}'(i) = \mathcal{P}'(j)$.

<u>If $T \in [\widehat{\mathcal{P}_{\widehat{\mathcal{Q}}}} \setminus \widehat{\mathcal{Q}}]$</u>, T is a one-person coalition. Write $\{j\} = T$. Clearly (a) does not hold. Hence (b) holds. By (b), since $i \neq j$, we have $|\mathcal{P}'(j)| \geq 2$. Since $T \overset{(A.18)}{\subseteq} \widehat{\mathcal{K}}$, $j \in \widehat{\mathcal{K}}$. Hence, $\mathcal{K} \subseteq \mathcal{P}'$ implies $\mathcal{P}'(j) \in \mathcal{K}$. From $|\mathcal{P}'(j)| \geq 2$, $\mathcal{P}'(j) \in \mathcal{K}$, and condition (i), it follows that $\mathcal{P}'(j) = \mathcal{P}(j)$. Since $j \in \widehat{\mathcal{P}_{\widehat{\mathcal{Q}}}} \setminus \widehat{\mathcal{Q}} \subseteq \widehat{\mathcal{P}_{\widehat{\mathcal{Q}}}}$ and $\widehat{\mathcal{P}_{\widehat{\mathcal{Q}}}} = \cup_{C \in \mathcal{P}: C \cap \widehat{\mathcal{Q}} \neq \emptyset} C$, we have $\mathcal{P}(j) \cap \widehat{\mathcal{Q}} \neq \emptyset$. Hence, $\mathcal{P}'(j) \cap \widehat{\mathcal{Q}} \neq \emptyset$, which implies $\mathcal{P}'(j) \not\subseteq \widehat{\mathcal{K}}$. This contradicts $\mathcal{P}'(j) \in \mathcal{K}$.

If $T \in (\mathcal{P} \setminus \overline{\mathcal{P}_{\widehat{\mathcal{Q}}}})$, then $T \in \mathcal{P}$. If (a) holds, $i_1 \in T \overset{(A.18)}{\subseteq} \widehat{\mathcal{K}}$. Hence, in view of $\mathcal{K} \subseteq \mathcal{P}'$, $\mathcal{P}'(i_1) \in \mathcal{K}$. For $\mathcal{P}'(i_1) \in \mathcal{K}$, if (i) holds, then $\mathcal{P}'(i_1) = \mathcal{P}(i_1)$. Since $i_1 \in T \in \mathcal{P}$, $\mathcal{P}(i_1) = T$. Hence, $\mathcal{P}'(i_1) = T$. In view of (a), $\mathcal{P}'(i_1) \neq \mathcal{P}'(i_2)$ implies $\mathcal{P}'(i_1) \cap \mathcal{P}'(i_2) = \emptyset$. Hence, $T \cap \mathcal{P}'(i_2) = \emptyset$. However, both T and $\mathcal{P}'(i_2)$ contain i_2, a contradiction. Thus, for $\mathcal{P}'(i_1) \in \mathcal{K}$, (ii) holds. In the same manner, (ii) also holds for $\mathcal{P}'(i_2) \in \mathcal{K}$. Hence, we have $|\mathcal{P}'(i_1)| = |\mathcal{P}'(i_2)| = 1$, and there exist $S \in \mathcal{P}$ and $T' \subsetneq S$ such that $[T'] \subseteq \mathcal{P}'$, $[T'] = \kappa(S)$, $i_1 \in T'$, and $i_2 \in T'$.[*14] Let $j \in S \setminus T'$. Since $i_1 \in T \in \mathcal{P}$ and $i_1 \in S \in \mathcal{P}$, we have $S = T$. Hence, $j \in S \setminus T' = T \setminus T' \subseteq T \overset{(A.18)}{\subseteq} \widehat{\mathcal{K}}$, which implies $\mathcal{P}'(j) \in \mathcal{K}$. Now, assume that (i) holds for $\mathcal{P}'(j) \in \mathcal{K}$. We have $\mathcal{P}(j) = \mathcal{P}'(j)$. Since $|\mathcal{P}'(i_1)| = 1$, we have $j \notin \mathcal{P}'(i_1)$, equivalently, $i_1 \notin \mathcal{P}'(j)$. By $i_1 \in T$, $T \neq \mathcal{P}'(j)$. By $j \in T \in \mathcal{P}$, $\mathcal{P}(j) = T$. Hence, $\mathcal{P}(j) \neq \mathcal{P}'(j)$. This is a contradiction. Thus, (ii) holds for $\mathcal{P}'(j) \in \mathcal{K}$. We have $\{j\} = \mathcal{P}'(j)$, and by (ii) there exists $T'' \subsetneq T(= S = \mathcal{P}(j))$ such that $[T''] \subseteq \mathcal{P}'$, $j \in T''$, and $[T''] = \kappa(T)$. By $S = T$, we have $[T''] = \kappa(T) = \kappa(S) = [T']$. However, this contradicts $j \notin T'$ and $j \in T''$.

If (b) holds, there exists $i \in N \setminus T$ such that for any $j \in T$, $\mathcal{P}'(i) = \mathcal{P}'(j)$. Fix $j \in T \overset{(A.18)}{\subseteq} \widehat{\mathcal{K}}$. We have $\mathcal{P}'(j) \in \mathcal{K}$. Hence, $\mathcal{P}'(i) \in \mathcal{K}$, and $\mathcal{P}'(i)$ contains both i and j: $|\mathcal{P}'(i)| \geq 2$. Condition (i) implies $\mathcal{P}'(i) = \mathcal{P}(i)$. Hence, $\mathcal{P}'(j) = \mathcal{P}(i)$ follows. We have $j \in \mathcal{P}'(j) = \mathcal{P}(i)$, equivalently, $\mathcal{P}(i) = \mathcal{P}(j)$. By $j \in T \in \mathcal{P}$, we have $T = \mathcal{P}(j)$; by $i \in N \setminus T$, we have $\mathcal{P}(i) \neq T$. This is a contradiction.

Thus, $\mathcal{P}'|_{N \setminus \widehat{\mathcal{Q}}} = (\mathcal{P} \setminus \overline{\mathcal{P}_{\widehat{\mathcal{Q}}}}) \cup [\widehat{\mathcal{P}_{\widehat{\mathcal{Q}}}} \setminus \widehat{\mathcal{Q}}]$ holds. Together with (A.16) and (A.17), we obtain $\mathcal{P}' \setminus \mathcal{K} \in \overline{\Gamma}^{\mathcal{P}}(\mathcal{P}')$.

Part 2: Minimality. Let $C \in \mathcal{P}' \setminus \mathcal{K}$ and

$$\mathcal{Q}' := (\mathcal{P}' \setminus \mathcal{K}) \setminus \{C\} = \mathcal{P}' \setminus (\mathcal{K} \cup \{C\}). \tag{A.19}$$

Below, we show that $\mathcal{P}'|_{N \setminus \widehat{\mathcal{Q}'}} \neq (\mathcal{P} \setminus \overline{\mathcal{P}_{\widehat{\mathcal{Q}'}}}) \cup [\widehat{\mathcal{P}_{\widehat{\mathcal{Q}'}}} \setminus \widehat{\mathcal{Q}'}]$. First, we have $\mathcal{P}'|_{N \setminus \widehat{\mathcal{Q}'}} \overset{(2.1)}{=} \mathcal{P}' \setminus \mathcal{Q}' \overset{(A.19)}{=} \mathcal{K} \cup \{C\}$. Since $C \notin \mathcal{K}$, there is no $S \in \mathcal{P}$ such that $S = C$, which means $C \notin \mathcal{P}$. Hence, we obtain $C \notin (\mathcal{P} \setminus \overline{\mathcal{P}_{\widehat{\mathcal{Q}'}}})$. Next, we prove $C \notin [\widehat{\mathcal{P}_{\widehat{\mathcal{Q}'}}} \setminus \widehat{\mathcal{Q}'}]$. Assume that $C \in [\widehat{\mathcal{P}_{\widehat{\mathcal{Q}'}}} \setminus \widehat{\mathcal{Q}'}]$. Then $|C| = 1$. As C is a one-person coalition, some $S \in \mathcal{P}$ satisfies $C \subseteq S$. We consider the following two cases: $|S| = 1$ and $|S| \geq 2$. If $|S| = 1$, then $C = S$. However, $C \notin \mathcal{K}$ implies that there is no such S, a contradiction. If $|S| \geq 2$, then for the $S \in \mathcal{P}$, there exists $T' \subsetneq S$ such that $C \in [T']$ and $[T'] \subseteq \mathcal{P}'$, *e.g.*, $T' := C$. However, $C \notin \mathcal{K}$ implies that there is no such T', a contradiction. Thus, $\mathcal{P}'|_{N \setminus \widehat{\mathcal{Q}'}} = \mathcal{K} \cup \{C\} \neq (\mathcal{P} \setminus \overline{\mathcal{P}_{\widehat{\mathcal{Q}'}}}) \cup [\widehat{\mathcal{P}_{\widehat{\mathcal{Q}'}}} \setminus \widehat{\mathcal{Q}'}]$. □

Proof of Proposition 5.12

For any $\mathcal{P} \in \Pi(N)$, $\mathcal{P}$ is γ-stable if and only if for any $\mathcal{P}' \in \Pi(N) \setminus \{\mathcal{P}\}$ and any $\mathcal{Q} \in \Gamma^{\mathcal{P}}(\mathcal{P}')$, there exists $i \in \widehat{\mathcal{Q}}$ such that $\phi_i(\mathcal{P}) \geq \phi_i(\mathcal{P}')$.

Proof. This proof is the same as Proposition 5.8. Given Proposition 5.10, replacing $\mathcal{P}' = \mathcal{Q} \cup (\mathcal{P}|_{N \setminus S})$ in the proof of Proposition 5.8 by $\mathcal{P}' = \mathcal{Q} \cup [\widehat{\mathcal{P}_S} \setminus S] \cup (\mathcal{P}|_{N \setminus \widehat{\mathcal{P}_S}})$ completes the proof. □

[*14] If $i_1 \notin T'$ or $i_2 \notin T'$ holds, say $i_1 \notin T'$, then the fact that (ii) holds for $\mathcal{P}'(i_1) \in \mathcal{K}$ implies that there exist $\tilde{S} \in \mathcal{P}$ and $\tilde{T}' \subsetneq \tilde{S}$ such that $[\tilde{T}'] \subseteq \mathcal{P}'$, $[\tilde{T}'] = \kappa(\tilde{S})$, and $i_1 \in \tilde{T}'$. However, if $\tilde{S} \neq S$, then the fact that $i_1, i_2 \in T \in \mathcal{P}$ contradicts that $i_1 \in \tilde{S} \in \mathcal{P}$ and $i_2 \in S \in \mathcal{P}$. Hence, $\tilde{S} = S$. If $\tilde{T}' \neq T'$ ($\tilde{T}', T' \subsetneq S$), then $\kappa(S)$ has two different values $[\tilde{T}']$ and $[\tilde{T}]$, a contradiction.

Proof of Proposition 6.1

For any n, no coalition structure is δ-stable.

Proof. Let $\mathcal{P} \in \Pi(N) \setminus [N]$. Since $\mathcal{P} \neq [N]$, there exists a coalition $S \in \mathcal{P}$ such that $|S| \geq 2$. Fix such a coalition $S \in \mathcal{P}$ and a player $i \in S$. Set $\mathcal{P}' := \{S \setminus \{i\}, \{i\}\} \cup (\mathcal{P} \setminus \{S\})$. Let $k := |\mathcal{P}|$. For simplicity, we write $s = |S|$. We have $|\mathcal{P}'| = k + 1$. Hence,

$$\phi_i(\mathcal{P}) = \frac{1}{(k+1)^2 \cdot s},\ \phi_i(\mathcal{P}') = \frac{1}{(k+2)^2}$$

We have

$$\phi_i(\mathcal{P}') - \phi_i(\mathcal{P}) > 0 \iff s > \frac{(k+2)^2}{(k+1)^2}.$$

For $k \geq 2$, it holds that $s \geq 2 > \frac{(k+2)^2}{(k+1)^2}$. For $k = 1$, since $n \geq 3$, we have $s = n \geq 3 > \frac{9}{4} = \frac{(k+2)^2}{(k+1)^2}$. Hence, player i has an incentive to deviate from $\mathcal{P}$, and $\mathcal{P}$ is not δ-stable. □

Proof of Proposition 6.2

For any n, partition $\{N\}$ is γ-stable.

Proof. First, for any $\mathcal{P}' \in \Pi(N)$, we have

$$\sum_{j \in N} \phi_j(\{N\}) \geq \sum_{j \in N} \phi_j(\mathcal{P}'), \tag{A.20}$$

because $\sum_{j \in N} \phi_j(\{N\}) = \frac{1}{4} \geq \frac{|\mathcal{P}'|}{(|\mathcal{P}'|+1)^2} = \sum_{j \in N} \phi_j(\mathcal{P}')$ holds for any $1 \leq |\mathcal{P}'| \leq n$. Hence, for each n-person deviation, some player does not have an incentive to participate.

Now, let $S \subsetneq N$ and $\mathcal{Q} \in \Pi(S)$. Let $\mathcal{P}' := \mathcal{Q} \cup [N \setminus S]$. Assume that $\phi_i(\mathcal{P}') > \phi_i(\{N\})$ for every $i \in S$. Then, it follows that $\phi_j(\{N\}) \geq \phi_j(\mathcal{P}')$ for any $j \in N \setminus S$ because of (A.20) and the fact that $\phi_j(\mathcal{P}') = \phi_{j'}(\mathcal{P}')$ for any $j, j' \in N \setminus S$. Fix $i \in S$ and $j \in N \setminus S$. Since $|\mathcal{P}'(j)| = 1$ and $|\mathcal{P}'(i)| \geq 1$, we have $\phi_i(\mathcal{P}') \leq \phi_j(\mathcal{P}')$. Hence, we obtain $\phi_j(\{N\}) \geq \phi_j(\mathcal{P}') \geq \phi_i(\mathcal{P}') > \phi_i(\{N\})$, while $\phi_i(\{N\}) = \frac{1}{4n} = \phi_j(\{N\})$. This is a contradiction. □

Proof of Proposition 6.3

Let $i \in N$. Partition $\{N \setminus \{i\}, \{i\}\}$ is γ-stable if and only if $n = 6, 8$.

Proof. Fix $h \in N$ and $\mathcal{P} := \{N \setminus \{h\}, \{h\}\}$. Let $S \subsetneq N \setminus \{h\}$ and $\mathcal{Q} \in \Pi(S)$. Note that $|S| \leq n-2$. Let $\mathcal{P}' := \mathcal{Q} \cup [N \setminus S]$. Partition $\mathcal{P}$ ($\mathcal{P}'$) is the partition from (to) which players in S deviate. For simplicity, we write $s = |S|$ and $q = |\mathcal{Q}|$. For any $j \in N \setminus \{h\}$, we have $\phi_j(\mathcal{P}) = \frac{1}{9(n-1)}$. For any $j \in (N \setminus \{h\}) \setminus S$, $\phi_j(\mathcal{P}') = \frac{1}{(q+n-s+1)^2}$. Moreover, $\sum_{j \in N \setminus \{h\}} \phi_j(\mathcal{P}) = \frac{1}{9}$, and $\sum_{j \in N \setminus \{h\}} \phi_j(\mathcal{P}') = \frac{q+n-s-1}{(q+n-s+1)^2}$. Note that $q + n - s \geq 3$.

Step 1: Now, in general, for any natural number m, we have $\frac{1}{9} \geq \frac{m-1}{(m+1)^2} \iff m \geq 5$. Hence, if $q+n-s \geq 5$, then

$$\sum_{j \in N \setminus \{h\}} \phi_j(\mathcal{P}) \geq \sum_{j \in N \setminus \{h\}} \phi_j(\mathcal{P}'). \tag{A.21}$$

In the same manner with Proposition 6.2, assume that $\phi_i(\mathcal{P}') > \phi_i(\mathcal{P})$ for every $i \in S$. It follows that $\phi_j(\mathcal{P}) \geq \phi_j(\mathcal{P}')$ for any $j \in (N \setminus \{h\}) \setminus S$ because of (A.21) and the fact that $\phi_j(\mathcal{P}') = \phi_{j'}(\mathcal{P}')$ for any $j, j' \in (N \setminus \{h\}) \setminus S$. Fix $i \in S$ and $j \in (N \setminus \{h\}) \setminus S$. Since $|\mathcal{P}'(j)| = 1$ and $|\mathcal{P}'(i)| \geq 1$, we have $\phi_i(\mathcal{P}') \leq \phi_j(\mathcal{P}')$. Hence, we obtain $\phi_j(\mathcal{P}) \geq \phi_j(\mathcal{P}') \geq \phi_i(\mathcal{P}') > \phi_i(\mathcal{P})$, while $\phi_i(\mathcal{P}) = \frac{1}{9(n-1)} = \phi_j(\mathcal{P})$. This is a contradiction. Thus, if $q+n-s \geq 5$, $\phi_i(\mathcal{P}') \leq \phi_i(\mathcal{P})$ for some $i \in S$. In other words, no $S \subsetneq N \setminus \{h\}$ has an incentive to deviate from $\mathcal{P}$ as long as $q+n-s \geq 5$.

Step 2: If $q+n-s=3$, it follows from $s \leq n-2$ that $q=1$ and $s=n-2$. For any $i \in S$,

$$\phi_i(\mathcal{P}) - \phi_i(\mathcal{P}') = \frac{1}{9(n-1)} - \frac{1}{16(n-2)} \geq 0 \iff n \geq 4.$$

Hence, if $n=3$, coalition S with $|S|=1$ and $\mathcal{Q} = \{S\}$ deviates from $\mathcal{P}$. For $n=3$, $\mathcal{P}$ is not γ-stable. For $n \geq 4$, such S has no incentive to deviate from $\mathcal{P}$.

Step 3: If $q+n-s=4$, one of ($q=2$ and $s=n-2$) or ($q=1$ and $s=n-3$) holds. For the former case, let T be the larger coalition in $\mathcal{Q}$: $t=|T| \geq \frac{n-2}{2}$. In the same manner as Step 2, for any $i \in T$, we have

$$\phi_i(\mathcal{P}) - \phi_i(\mathcal{P}') = \frac{1}{9(n-1)} - \frac{1}{25t} \geq 0 \iff t \geq \frac{9}{25}(n-1).$$

In words, if $t \geq \frac{9}{25}(n-1)$, the coalition T has no incentive to deviate from $\mathcal{P}$. For any $n \geq 3$ and every natural number $t \geq \frac{n-2}{2}$, it holds that $t \geq \frac{9}{25}(n-1)$.

For the latter case, for any $i \in S$, we have

$$\phi_i(\mathcal{P}) - \phi_i(\mathcal{P}') = \frac{1}{9(n-1)} - \frac{1}{25(n-3)} \geq 0 \iff n \geq 5.$$

Hence, if $n=3$ or 4, $\mathcal{P}$ is not γ-stable. For $n \geq 5$, such S, namely $s=n-3$ and $\mathcal{Q}=\{S\}$, has no incentive to deviate from $\mathcal{P}$. This completes Step 3.

In view of Steps 1-3, if $n \geq 5$ then no coalition $S \subsetneq N \setminus \{h\}$ and its partition $\mathcal{Q} \in \Pi(S)$ have an incentive to deviate from $\mathcal{P} = \{N \setminus \{h\}, \{h\}\}$. Note that $\phi_h(\mathcal{P}) = \frac{1}{9} \geq \phi_j(\mathcal{P}'')$ for any $j \in N$ and $\mathcal{P}'' \in \Pi(N)$. Hence, player h does not participate in any deviation.

Now, let $S = N \setminus \{h\}$. For any $\mathcal{Q} \in \Pi(S)$, we have

$$\sum_{j \in N \setminus \{h\}} \phi_j(\mathcal{P}) - \sum_{j \in N \setminus \{h\}} \phi_j(\mathcal{P}') = \frac{1}{9} - \frac{q}{(q+1+1)^2} \geq 0 \iff q \geq 4.$$

Hence, for players in $S = N \setminus \{h\}$ to deviate from $\mathcal{P}$, their partition $\mathcal{Q}$ must satisfy $q=2$ or 3. If $q=2$, then let T be the larger coalition in $\mathcal{Q}$, namely $t \geq \frac{n-1}{2}$. For any $i \in T$,

$$\phi_i(\mathcal{P}) - \phi_i(\mathcal{P}') = \frac{1}{9(n-1)} - \frac{1}{16t} \geq 0 \iff t \geq \frac{9}{16}(n-1).$$

Hence, if T satisfies $t < \frac{9}{16}(n-1)$, then the T has an incentive to deviate. For $n = 3, 5, 7, 9$, and more than 9, there exists natural number t satisfying $\frac{n-1}{2} \leq t < \frac{9}{16}(n-1)$. Hence, for these n, partition $\mathcal{P} = \{N \setminus \{h\}, \{h\}\}$ is not γ-stable. For $n = 4$, it is not γ-stable because of Step 3.

Below, we show that $\mathcal{P}$ is γ-stable for $n = 6, 8$. By the discussion above, the remaining case is $S = N \setminus \{h\}$ and $q = 3$. For $n = 8$, since $n - 1 = 7$, Q can be $< 5, 1, 1 >$, $< 4, 2, 1 >$, or $< 3, 2, 2 >$, where $< x, y, z >$ is an integer partition of 7. For example, $< 5, 1, 1 >$ means one five-person coalition and two one-person coalitions. We focus on the largest coalition for each partition: 5,4,3. We have $\phi_i(\mathcal{P}') = \frac{1}{125}(= \frac{1}{5 \cdot (4+1)^2})$ for the five-person coalition in $\mathcal{P}'$; $\frac{1}{100}$ for the four-person coalition; and $\frac{1}{75}$ for the three-person coalition. Since $\phi_i(\mathcal{P}) = \frac{1}{9(n-1)} = \frac{1}{63} > \max\{\frac{1}{125}, \frac{1}{100}, \frac{1}{75}\}$ for $n = 8$, no partition $\mathcal{Q}$ deviates. Similarly, for $n = 6$ $(n - 1 = 5)$, consider $< 3, 1, 1 >$ and $< 2, 2, 1 >$. We have $\phi_i(\mathcal{P}) = \frac{1}{45} > \max\{\frac{1}{75}, \frac{1}{50}\}$. Thus, partition $\mathcal{P} = \{N \setminus \{h\}, \{h\}\}$ is γ-stable for $n = 6, 8$. □

References

[1] Abe T (2018) Stable coalition structures in symmetric majority games: A coincidence between myopia and farsightedness. *Theory and Decision* 85:353-374

[2] Aumann RJ, Drèze JH (1974) Cooperative games with coalition structures. *International Journal of Game Theory* 3:217-237

[3] Banerjee S, Konishi H, Sönmez T (2001) Core in a simple coalition structure game. *Social Choice and Welfare* 18:135-153

[4] Bogomolnaia A, Jackson MO (2002) The stability of hedonic coalition structures. *Games and Economic Behavior* 38:201-230

[5] Casajus, A (2009) Outside options, component efficiency, and stability. *Games and Economic Behavior* 65:49-61

[6] Hart S, Kurz M (1983) Endogenous formation of coalitions. *Econometrica* 51:1047-1064

[7] Hart S, Kurz M (1984) Stable coalition structures. Holler MJ (ed.), *Coalitions and Collective Action* 51:235-258

[8] Kamijo Y (2009) A two-step Shapley value for cooperative games with coalition structures. *International Game Theory Review* 11:207-214

[9] Owen G (1977) Values of games with a priori unions. Henn R, Moeschlin O (eds.), *Essays in Mathematical Economics and Game Theory*. Springer, Berlin, 7688

[10] Ray D, Vohra R (1997) Equilibrium binding agreements. *Journal of Economic Theory* 73:3078

[11] Wiese H (2007) Measuring the power of parties within government coalitions. *International Game Theory Review* 9:307-322

[12] Yi SS (1997) Stable coalition structures with externalities. *Games and Economic Behavior* 20:201-237