# Coalitional Strategic Games

Advanced Game Theory

CONTENTS

# 1 Strategic Equilibrium and Cores

Let $G = (N, \{X^i\}_{i \in N}, \{u_i\}_{i \in N})$ be a game in *strategic form*, where $N$ is a finite set of players, $X^i$ is a compact convex set of pure strategies of player $i$ and $u_i$ is a continuous payoff function of player $i$. A nonempty subset $S \subseteq N$ is called a *coalition*. For each coalition $S$, define $X^S = \Pi_{i \in S} X^i$, and $X = X^N$. Any singleton coalition $\{i\}$ is often identified with $i$.

## 1.1 Strong Nash, and Coalition-Proof Nash Equilibrium

For any $x, y \in X$, we define $x = (x^T, y^{N-T})$ if $T = N$, and $y = (x^T, y^{N-T})$ if $T = \emptyset$. For any given vector $f \in \Re^n$, we will denote $(f_i)_{i \in S}$ by $f_S$, and the relation $\geq$ is understood in coordinate-wise.

**Definition 1.** *Let $x \in X$ be given. Then, for each nonempty $S \subseteq N$, we say $S$ has a deviation $y^S \in X^S$ at $x$ if and only if*

$$u_S(y^S, x^{N \setminus S}) > u_S(x).$$

This definition seems to presume that coalition $S$ can make a binding agreement as to the joint choice of strategies in $S$. Later, we will modify this definition into one that may avoid the assumption of binding agreements.

**Definition 2.** *An n-tuple of strategies $x \in X$ is a* **strong Nash equilibrium** *if and only if there exists no coalition $S \subseteq N$ that has a deviation at $x$.*

This definition is due to Aumann [1]. A strong Nash equilibrium is *weakly Pareto efficient*, and generally hard to obtain. Exceptions are Aumann [1], Kalai, Postlewaite and Roberts [13], Peleg [20], Greenberg and Weber [7], Holzman and Law-Yone [10], Nishihara [19], and Hirai et al. [9].

#### 1.1.1 Voluntary Contribution to Public Goods

As a straightforward example of a strategic game, we consider a model of a voluntary provision of public goods.

Let $G$ be a strategic game with $X^i = [0, m_i]$, where $m_i > 0$ and $x^i \in X^i$ is $i$'s voluntary contribution to the public expenditure. The utility to each $i$ is given by

$$u_i(x) = v_i(\sum_{j \in N} x^j,\ m_i - x^i),$$

where $v_i$ is a continuous, quasiconcave and monotone increasing utility function. Then:

**Proposition 1.** *Let $x^*$ be a Nash equilibrium. If $x^{*i} = m_i$ for all $i \in N$, then $x^*$ is a strong Nash equilibrium.*

*Proof.* This is a consequence of the following lemma.

*Lemma. Let $x^* \in X$ be a Nash equilibrium, and let there be an $x \in X$ such that for some $i \in N$, $u_i(x) > u_i(x^*)$. Then, $\sum_{j \neq i} x^j > \sum_{j \neq i} x^{*j}$.*

Suppose that $x^*$ is not a strong Nash equilibrium. Then, there must exist an $S \subseteq N$ with $|S| > 1$ and $x^S \in X^S$ such that $u_S(x^S, x^{*N \setminus S}) > u_S(x^*)$. Then, letting $\overline{x} = (x^S, x^{*N \setminus S})$ and choosing $i \in S$ arbitrary, it follows from the lemma that

$$\begin{aligned}
\sum_{j \neq i} \overline{x}^j &= \sum_{j \notin S} x^{*j} + \sum_{j \in S \setminus \{i\}} x^j \\
&> \sum_{j \neq i} x^{*j} \\
&= \sum_{j \notin S} x^{*j} + \sum_{j \in S \setminus \{i\}} x^{*j}
\end{aligned}$$

Hence,

$$\sum_{j \in S \setminus \{i\}} x^j > \sum_{j \in S \setminus \{i\}} x^{*j} = \sum_{j \in S \setminus \{i\}} m_j,$$

which is a contradiction to the fact that $x^S \in X^S$. $\square$

When $x^*$ is not weakly Pareto optimal, we may take a weakly Pareto optimal $x \in X$ that satisfies the inequality in the lemma. Then, adding both sides for all $i \in N$, we see that the level of public expenditure at the Nash equilibrium is less than the one that is weakly Pareto optimal.

■**Problem 0** Try to prove the lemma. The monotonicity of $v_i$ is all we need.

#### 1.1.2 Coalition-Proof Nash Equilibrium

In an attempt to weaken the definition of strong Nash equilibrium, Peleg [4] formulated the notion of *coalition-proof Nash equilibrium* (CPNE for short). For any $x \in X$ and coalition $S$, denote by $G|x^{N\setminus S}$ the subgame $G'=(S, \{X^i\}_{i\in S}, \{u_i\}_{i\in S})$ induced by fixing $x^{N\setminus S}$. We assume that $G' = G$ when $S = N$.

**Definition 3.** *Let $x \in X$ be given. Then:*

1. *For each $i \in N$, we say $x^i$ is* **coalition-proof** *in $G|x^{N\setminus i}$ iff*

$$\forall y^i \in X^i,\ u_i(x) \geq u_i(y^i, x^{N\setminus i}).$$

2. *For each $S \subseteq N$ with $1 \leq |S| < m$, assume that the definition of the coalition-proofness of $x^S$ in $G|x^{N\setminus S}$ is completed. Then, for each $S \subseteq N$ with $|S| = m$, we say*

   (*a*) *$x^S$ is* **self-enforcing** *in $G|x^{N\setminus S}$ iff for all $T \subsetneq S$, $x^T$ is* **coalition-proof** *in $G|x^{N\setminus T}$; and*

   (*b*) *$x^S$ is coalition-proof in $G|x^{N\setminus S}$ iff it is* **self-enforcing** *in $G|x^{N\setminus S}$ and there does not exist $y^S \in X^S$ such that $y^S$ is* **self-enforcing** *in $G|x^{N\setminus S}$ and $\forall i \in S,\ u_i(y^S, x^{N\setminus S}) > u_i(x)$.*

*We say $x \in X$ is* **coalition-proof** *iff $x$ is coalition-proof in $G$.*

**Definition 4.** *Let $x \in X$ be given. Then, for each nonempty $S \subseteq N$, we say $S$ has a* **credible deviation** *at $x$ if and only if $S$ has a deviation $y^S \in X^S$ at $x$ such that there exists no $T \subsetneq S$ which has a* **credible deviation** *at $(y^S, x^{N\setminus S})$.*

When $S = \{i\}$, $i$ has a credible deviation $y^i$ at $x$ iff $u_i(y^i, x^{N\setminus i}) > u_i(x)$. Hence, for $S$ with $|S| > 1$, the definition of a credible deviation follows inductively.

This definition avoids at least the bindingness of the coalition itself: deviations may trigger further deviations by subcoalitions. Hence, only credible deviations deserve consideration.

**Theorem 1.** *Let $x \in X$ be given. Then, for each nonempty $S \subseteq N$, $x^S$ is coalition-proof in $G|x^{N\setminus S}$ iff there exists no $T \subseteq S$ which has a credible deviation at $x$ in $G|x^{N\setminus S}$.*

*Proof.* When $|S| = 1$, the assertion follows directly from the definitions. Assume that $|S| = m > 1$ and that the assertion is true for any $S$ with $|S| < m$.

**necessity:**. Suppose that there exists $T \subseteq S$ which has a credible deviation $y^T \in X^T$. Then, for any $R \subsetneq T$, there exists no credible deviation at $(y^T, x^{N\setminus T})$. By the induction hypothesis, this means that for any fixed $R \subsetneq T$, $y^R$ is coalition-proof in $G|(y^{T\setminus R}, x^{N\setminus T})$ since no subset of $R$ has a credible deviation. Hence, by Definition 3, $y^T$ is self-enforcing in $G|x^{N\setminus T}$. But, since $y^T$ was a credible deviation at $x$, it must be true that $\forall i \in T,\ u_i(y^T, x^{N\setminus T}) > u_i(x)$. Hence, $x^T$ is not coalition-proof in $G|x^{N\setminus T}$; so that, $x^S$ is not self-enforcing in $G|x^{N\setminus S}$. Hence, $x^S$ cannot be coalition-proof in $G|x^{N\setminus S}$.

**sufficiency:**. Suppose that $x^S$ is not coalition-proof in $G|x^{N\setminus S}$. If $x^S$ is not self-enforcing in $G|x^{N\setminus S}$, then for some $T \subsetneq S$, $x^T$ is not coalition-proof in $G|x^{N\setminus T}$. It then follows from the induction hypothesis that there exists $R \subseteq T$ which has a credible deviation at $x$. On the other hand, if $x^S$ is self-enforcing, then, by Definition 3, there exists another self-enforcing $y^S \in X^S$ such that

$\forall i \in S,\ u_i(y^S, x^{N\setminus S}) > u_i(x)$. Then, for any $T \subsetneq S$, $y^T$ must be coalition-proof in $G|(y^{S\setminus T}, x^{N\setminus S})$. By the induction hypothesis, there must not exist any $T' \subseteq T \subsetneq S$ which has a credible deviation at $(y^S, x^{N\setminus S})$. Hence, by Definition 4, $y^S$ must be a credible deviation at $x$. □

**Corollary 1.** *A strong Nash equilibrium is coalition-proof.*

Thus, if a strong Nash equilibrium exists, it can be reached without binding agreements among players. If not, at any strategy combination, some coalition can deviate from it. But, this deviation in general will require a binding agreement as to the choice of strategies.

**Problem 1.** *Assume that $S \subseteq N$ has a credible deviation at $x \in X$. Then, there exists a strategy profile $x^{*S} \in X^S$ that is coalition-proof in the subgame $G|\, x^{N\setminus S}$.*

**Remark 1.** *The CPNE in the above problem depends on the given $x$ except for $S = N$. Conditions for general existence of CPNE are not known.*

**Theorem 2.** *Let $x \in X$ be a unique Nash equilibrium of a game $G$, and for any nonempty $S \subseteq N$, the subgame $G|x^{N\setminus S}$ has also a unique Nash equilibrium. Then, $x$ is a coalition-proof Nash equilibrium of $G$.*

**Problem 2.** Prove this theorem.

**Example 1.** *A Three Person Game*

$$
\begin{array}{ccc}
 & C_1^{\circ} & \\
 & B_1 & B_2^{\circ} \\
A_1 & 1, 1, -5 & -5, -5, 0 \\
A_2^{\circ} & x, -5, 0 & 0, 0, 10
\end{array}
\qquad
\begin{array}{ccc}
 & C_2^{*} & \\
 & B_1^{*} & B_2 \\
A_1^{*} & -1, -1, 5 & -5, -5, 0 \\
A_2 & -5, -5, 0 & -2, -2, 0
\end{array}
$$

**case x=-5:** $(A_2, B_2, C_1)$ is a Nash equilibrium (Check it!); but not coalition-proof. Fix $C_1$ and consider the subgame $G|C_1$. $(A_2, B_2)$ is a Nash equilibrium, and hence, self-enforcing in $G|C_1$. But $(A_1, B_1)$ is also a Nash equilibrium (self-enforcing) in the same subgame which is better for both players A and B. Hence $(A_2, B_2)$ is not coalition-proof in $G|C_1$. Hence, $(A_2, B_2, C_1)$ is not self-enforcing, which shows that $(A_2, B_2, C_1)$ is not coalition-proof.

$(A_1, B_1, C_2)$ is a coalition-proof Nash equilibrium.

**case x=2:** $(A_2, B_2, C_1)$ is now a coalition-proof Nash equilibrium; but, $(A_1, B_1, C_2)$ is not. In the subgame $G|C_1$, $(A_2, B_2)$ is self-enforcing as before. But now, $(A_1, B_1)$ is not self-enforcing in $G|C_1$, because $(A_1, B_1)$ is no longer a Nash equilibrium in this subgame ($A$ will deviate). Hence $(A_2, B_2)$ is coalition-proof, because it is undominated by any self-enforcing strategy pair in $G|C_1$. Coalition-proofness for every other coalition is easily checked. Hence $(A_2, B_2, C_1)$ is self-enforcing and undominated by any other self-enforcing strategy triple.

There are several economic applications of CPNE. Wako [27] considers an economy with multiple indivisible goods, and shows that a strategic game describing the transactions in this economy has a unique CPNE. Nishihara [19] presents an analysis of N-person prisoners' dilemma, in which a Nash equilibrium is coalition-proof, and if the Nash equilibrium is Pareto efficient, then it is a strong equilibrium. See also Hirai et al. [9].

## 1.2 The $\alpha$– and $\beta$– Cores

### 1.2.1 NTU games

The cooperative solution concept of cores have been extended to strategic games with coalitions by Aumann and Peleg [3].

Let $G = (N, \{X^i\}_{i\in N}, \{u_i\}_{i\in N})$ be a strategic game with coalitions, and let $S$ be a nonempty subset of $N$.

**Definition 5.** *Coalition $S$ is said to be $\alpha$-effective for a payoff vector $v_S \in \Re^S$ iff*

$$\exists x^S \in X^S,\ \forall x^{N \setminus S} \in X^{N \setminus S},\ \ u_S(x^S, x^{N \setminus S}) \geq v_S.$$

*The set of all payoff vectors for which $S$ is $\alpha$-effective is denoted by $v_\alpha(S)$.*

Thus, $S$ is $\alpha$-effective for $v_S$ if $S$ *can assure itself of getting at least* $v_S$.

**Definition 6.** *Coalition $S$ is said to be $\beta$-effective for a payoff vector $v_S \in \Re^S$ iff*

$$\forall x^{N \setminus S} \in X^{N \setminus S},\ \exists x^S \in X^S,\ \ u_S(x^S, x^{N \setminus S}) \geq v_S.$$

*The set of all payoff vectors for which $S$ is $\beta$-effective is denoted by $v_\beta(S)$.*

Thus, $S$ is $\beta$-effective for $v_S$ if $N \setminus S$ *cannot prevent $S$ from getting at least* $v_S$.

Let $v_\alpha(\emptyset)$=$v_\beta(\emptyset) = \emptyset$; and

$$v_\alpha(N) = v_\beta(N) = \{v \in \Re^N | u_N(x) \geq v,\ \exists x \in X\}.$$

**Definition 7.** *An NTU game* $(N, v)$ *derived from the strategic form game* $G$ *is called the* $\alpha$*-NTU game if* $\forall S \subseteq N$, $v(S)$= $v_\alpha(S)$. *Similarly, the NTU game* $(N, v)$ *is called the* $\beta$*-NTU game if* $\forall S \subseteq N$, $v(S) = v_\beta(S)$.

**Remark 2.** By definition, $v_\alpha(S) \subset v_\beta(S)$. (see also the problem below).

**Definition 8.** *The core of a NTU game* $(N, v)$ *is the set of undominated payoff vectors* $v \in v(N)$*, i.e., a subset of* $v(N)$ *for which there is no coalition* $S \subseteq N$ *such that*

$$\exists u_S \in v(S),\ u_S > v_S.$$

**Definition 9.** *The core of an* $\alpha$ $(\beta)$*-NTU game* $(N, v)$ *is called the* $\alpha$ $(\beta)$*-core.*

**Problem 3.** Show that, given a game $G$, $uSE \subset \beta$-core $\subset \alpha$-core, where $uSE$ is the set of utility vectors corresponding to the strong Nash equilibria in $G$.

### 1.2.2 Scarf's Core Existence Theorem

Scarf [23] proved a beautiful theorem on the existence of an $\alpha$-core.

**Theorem 3. (Scarf.)** *Assume that for each $i \in N$, $X^i$ is a compact convex set, and $u_i$ is quasi-concave in $x \in X$. Then the $\alpha$-core is nonempty.*

**Example 2. (Scarf)***:( empty $\beta$-core with nonempty $\alpha$-core)*

$N = \{1, 2, 3\}$*; one good A and one bad B;*

$$X^i = \{x^i = (x_A^{i1}, x_A^{i2}, x_A^{i3}; x_B^{i1}, x_B^{i2}, x_B^{i3}) \\ \mid \sum_{j \in N} x_A^{ij} \leq 1, \sum_{j \in N} x_B^{ij} = 1, x_A^{ij} \geq 0 \text{ and } x_B^{ij} \geq 0\},$$

$$u_i(x) = u_i(\sum_{j \in N} x^{ji}) = \sum_{j \in N} x_A^{ji} - \sum_{j \in N} x_B^{ji}.$$

**Claim:** $\alpha$**-core** $\neq \emptyset$ **and** $u_0 = (0, 0, 0) \in \alpha - core$**.**

*Proof.* Every condition of Scarf is satisfied; so that the $\alpha$-core $\neq \emptyset$.

The payoff vector $u_0$ cannot be dominated via singleton set $\{3\}$ because for any $x^3$, $\{1, 2\}$ can dump 2 units of bads onto $\{3\}$ so that the utility of 3 is negative.

The payoff vector $u_0$ cannot be dominated via 2-person set $\{1, 2\}$ either, since for any $x^{12}$, $\{3\}$ can dump one unit of bad onto one of $\{1, 2\}$, thereby making the utility of one of them non-positive. $u_0$ is Pareto efficient. $\square$

**Claim: $\beta$-core is empty.**

*Proof.* Let $u = (u_1, u_2, u_3), u_1 \geq u_2 \geq u_3$ be any feasible utility allocation. Then, $\sum_{j \in N} u_j \leq 0$. Every 2-person coalition $\{i, j\}$ can attain any $(u_i, u_j)$ with $u_i + u_j \leq 1$, because for any $x^k$,$\{i, j\}$ can dump 2-units of bads onto $\{k\}$, thereby re-allocating 2-units of goods between them to compensate the disutility caused by $x^k$ (i.e., by the bad dumped by $k$ onto $\{i, j\}$).

If $u \in \beta$-core, then $u_3 < 0$; otherwise $u_1 = u_2 = u_3 = 0$, which is $\beta$-dominated via any 2-person coalition. But then $u_2 \geq 1$, because $u_2 < 1$ implies that $u$ is $\beta$-dominated via $\{2, 3\}$. Since $u_1 \geq u_2$, it follows that $u_3 \leq -2$, which is $\beta$-dominated via $\{3\}$. (Note that $v_\beta(\{i\}) = (-\infty, -1]$, for all $i \in N$). Contradiction. $\square$

We now state a sufficient condition for the existence of a $\beta$-core appeared in Nakayama [16].

**Definition 10.** *Let $\emptyset \neq S \subsetneq N$. Then, $N \setminus S$ is said to have a* **dominant punishment strategy** $d^{N\setminus S} \in X^{N\setminus S}$ *against $S$ iff*

$$\forall z^S \in X^S, \forall z^{N\setminus S} \in X^{N\setminus S},\ u_S(z^S, z^{N\setminus S}) \geq u_S(z^S, d^{N\setminus S}).$$

**Definition 10+.** Let $\emptyset \neq S \subsetneq N$. Then $S$ has a **dominant strategy ($S$−dominant strategy)** $x^S \in X^S$ iff for all $z \in X$, $u_S(x^S, z^{N\setminus S}) \geq u_S(z)$.

**Theorem 4.** *Assume that for each nonempty $S \subsetneq N$, either $N \setminus S$ has a dominant punishment strategy against $S$, or $S$ has an $S-$ dominant strategy. Then, under the Scarf's condition the $\beta$-core is nonempty and identical to the $\alpha$-core.*

*Proof.* It will be enough to show that for any given payoff vector $v_S$, $S$ is $\alpha$-effective for $v_S$ iff $S$ is $\beta$-effective for $v_S$.

Let $S$ be $\beta$-effective for $v_S$. Then, for all $x^{N\setminus S} \in X^{N\setminus S}$ there is a strategy $x^S \in X^S$ such that $\forall i \in S$, $u_i(x^S, x^{N\setminus S}) \geq v_i$. Thus, for the dominant punishment strategy $d^{N\setminus S} \in X^{N\setminus S}$, there is a strategy $x(d^{N\setminus S}) \in X^S$ such that $\forall i \in S$, $u_i(x(d^{N\setminus S}), d^{N\setminus S}) \geq v_i$. But, since $d^{N\setminus S}$ is a dominant punishment strategy, it follows that

$$u_S(x(d^{N\setminus S}), z^{N\setminus S}) \geq u_S(x(d^{N\setminus S}), d^{N\setminus S}) \geq v_S,$$

which implies that by choosing the strategy $x(d^{N\setminus S})$, $S$ can assure itself the payoff vector $v_S$. The converse follows by definition. (*Problem*: Complete the proof.)

$\square$

Thus, what $S$ can assure itself is precisely those payoff vectors which $N \setminus S$ cannot prevent $S$ from getting. The set of payoff vectors that $S$ can assure itself is determined by the dominant punishment strategy of $N \setminus S$. More than thirty years ago, Jentzsch [12] called such a strategy *optimal* and the payoffs structure with the optimal strategy *classical* paying attention to the zero-sum-like situation between $S$ and $N \setminus S$. Thus, the existence of a dominant punishment strategy will be limited only to a narrow class of games - the *classical games*. Nevertheless, there is a natural important economic example of a game with this payoffs structure; namely, the pure exchange game (see Scarf [23], and also Mas-Colell [14]).

**Problem 4.** Let $G$ be the pure exchange game with nonnegative strategies where $X^i = \{x^i = (x^{i1}, \ldots, x^{in}) | \quad \sum_{j \in N} x^{ij} \leq \omega^i$,and $\forall j \in N,\ x^{ij} \in \Re^m_+\}$ $u_i(x) = f_i(\sum_{j \in N} x^{ji})$, where $f_i$ is continuous, quasi-concave and monotone nondecreasing in $\sum_{j \in N} x^{ji}$. Show that this game has a nonempty $\beta$-core.

The public good game mentioned in Mas-Colell [14] is also an example of a game with a nonempty $\beta-$core.

Recently, a progress on the existence of $\beta$-cores is made also in the class of **TU games** by Zhao [29], in which a slightly weaker condition is shown to be sufficient for the existence of the $\beta$-core that is identical to the $\alpha$-core. He also shows a similar condition is sufficient for the existence of the $\alpha$-core in TU games.

## 2 Self-Binding Coalitions

Implicit in the theory of cooperative games is an assumption that players can make a *binding agreement*. Due to this assumption, any coalition can be formed once players agree to do so. In this section, we review an attempt in Nakayama [16] of a cooperative game with *self-binding coalitions* that may enable to dispense with the assumption of binding agreements.

Before doing so, we list some comments on the binding agreements appeared in the literature.

## 2.1 Comments on the Binding Agreements in the Literature

### 2.1.1 von Neumann and Morgenstern [25, pp.223–224]

Two players who wish to collaborate must get together on this subject before the play, i.e., outside game. The player who lives up to his agreement must possess the conviction that the partner too will do likewise. As long as we are concerned only with the rules of the game, we are in no position to judge what the basis for such a conviction may be. In other words **what, if anything, enforces the ”sanctity” of such agreements?** …On a later occasion we propose to investigate what theoretical structures are required in order to eliminate these concepts. (I.e., **auxiliary concepts such as ”agreements”, ”understandings”**,etc.) We shall …make use of the possibility of the establishment of coalitions outside the game; this will include the hypothesis that they are respected by the contracting parties.

#### 2.1.2 Nash [17, pp.286–295]

By a cooperative game we mean a situation . . . with the assumption that the players can and will collaborate as they do in the von-Neumann and Morgenstern theory. This means the players may communicate and **form coalitions which will be enforced by an umpire.**

. . . The problem of analyzing a cooperative game becomes the problem of obtaining a suitable, and convincing, non-cooperative model for the negotiation.

#### 2.1.3 Nash [18, pp.128–140]

The word cooperative is used because the two individuals are supposed to be able to discuss the situation and agree on a rational joint plan of action, **an agreement that should be assumed to be enforceable** . . . . A game is non-cooperative if it is impossible for the players to communicate or collaborate in any way. Supposing $A$ and $B$ to be rational beings, it is essential for the success of the threat that $A$ be compelled to carry out his threat if $B$ fails to comply ... The point of discussion is that we must assume there is an adequate mechanism for forcing the players to stick to their threats and demands once made; and one to enforce the bargain, once agreed. Thus **we need a sort of umpire, who will enforce contracts or commitments.**

#### 2.1.4 Aumann [2, pp.67–96]

.... both the non-cooperative and the cooperative theory involve agreement anong the players, the difference being only in that in one case the agreement is self-enforcing, **whereas in the other case it must be externally enforced.**

### 2.1.5 Harsanyi and Selten [8]

A non-cooperative game is a game modeled by making the assumption that the players are unable to make enforceable agreements except insofar as the extensive form of the game explicitly gives them an ability to do so. In contrast, a cooperative game is **a game modeled by making the assumption that the players are able to make enforceable agreements** even if their ability to do so is not shown explicitly by the extensive form of the game. . . . . (By our solution theory), the problem of defining a solution for a cooperative game $G$ can always be reduced to the problem of defining a solution for a non-cooperative bargaining game $B(G)$.

## 2.2 Self-Binding Strategies

**Definition 11.** *For all nonempty $T \subseteq N$, $T$ has an $\alpha$-deviation $y^T \in X^T$ at $x \in X$ if and only if*

$$\forall z \in X,\ u_T(y^T, z^{N\setminus T}) > u_T(x).$$

**Definition 12.** *For all nonempty $T \subseteq N$, $T$ has a* **credible $\alpha$ - deviation** *at $x \in X$ if and only if $T$ has an $\alpha$-deviation $y^T \in X^T$ at $x$ such that for all $z \in X$ there exists no $R \subsetneq T$ which has a credible $\alpha$-deviation at $(y^T, z^{N \setminus T})$.*

Note that the definition is not circular. Note also that every deviating subset $R$ of $T$ must confront the same strategic environment as that of $T$, i.e., $R$ must take all the reactions of $N \setminus R$ into consideration.

**Definition 13.** *For all $S \subseteq N$, $x^S \in X^S$ is a* **self-binding strategy** *for $S$ if and only if for all $z \in X$, no $T \subseteq S$ has a credible $\alpha$-deviation at $(x^S, z^{N \setminus S})$.*

A *self-binding coalition* is one that has a self-binding strategy. In characteristic function form, Ray [21] defined a *credible coalition* to be one that can sustain itself by assuring each of the members a certain level of utility. Note that 1-person coalition is always self-binding by the maximin strategy.

**Problem 5.** Let $x \in X$, and $T \subseteq N$. Then, if $T$ has an $\alpha$-deviation at $x$, some $R \subseteq T$ has a credible $\alpha$-deviation at x. Prove this fact.

The self-binding coalition can be related to the $\alpha-$core as follows.

**Theorem 5.**

1. *$N$ is a self-binding coalition iff the $\alpha$-core is nonempty.*
2. *Let $S \subsetneq N$ have a credible $\alpha-$ deviation at $x \in X$. Then $S$ is self-binding.*

*Proof.* **1**. Suppose that the $\alpha - core$ is empty. Then, at any $x$ there is a subset $S \subseteq N$ that has an $\alpha$-deviation . Then, Problem 5 implies that there is a subset $T \subseteq S$ that has a credible $\alpha$-deviation at $x$. Hence, $x$ cannot be a self-binding strategy of $N$, which contradicts the hypothesis that $N$ is self-binding.

Conversely, let $x \in X$ be in the $\alpha$-core. Then, no subset $T$ has an $\alpha$-deviation at $x$, which implies that no subset $S$ has a credible $\alpha$- deviation at $x$.

**2**. Let the credible $\alpha$-deviation at $x \in X$ be $y^S$. Then, for any $z \in X$, no proper subset $T \subsetneq S$ has a credible $\alpha-$ deviation at $(y_S, z_{N\setminus S})$. Thus, if $S$ itself does not have a credible $\alpha-$ deviation at $(y_S, z_{N\setminus S})$, then $y_S$ is a self-binding strategy of $S$. Indeed, simply taking $y_S$ to be one undominated by any other credible $\alpha-$deviation at $x$ by $S$, $y_S$ satisfies the requirement for the self-binding strategy. □

The next theorem provides a sufficient condition for a given coalition to be self-binding.

**Theorem 6.** *Assume that for all $i \in N$, $u_i$ is continuous and quasi-concave in $x \in X$; and that $S \subseteq N$ is nonempty and proper. Then, $S$ is a self-binding coalition if $N \setminus S$ has a dominant punishment strategy against $S$.*

*Proof*. Let $d^{N\backslash S}$ be the dominant punishment strategy. Then, since $u_i(\cdot\,, d^{N\backslash S})$ is quasi-concave for all $i \in S$, it follows from Theorem 5 and the Scarf's theorem [23] that there exists a self-binding strategy $x^S \in X^S$ for $S$ in the subgame induced by holding $x^{N\backslash S}$ fixed to $d^{N\backslash S}$. Then, for any $T \subseteq S$ and any $y^T \in X^T$, there must exist $z \in X^S$ such that $u_i(y^T, z^{S-T}, d^{N\backslash S}) \leq u_i(x^S, d^{N\backslash S})$ for some $i \in T$. Hence, there exists $w \in X$ such that $u_i(y^T, w^{N-T}) \leq u_i(x^S, d^{N\backslash S})$ for some $i \in T$. Since $d^{N\backslash S}$ is a dominant punishment strategy, it follows that for all $x^{N\backslash S} \in X^{N\backslash S}$,

$$\exists i \in T,\ u_i(y^T, w^{N-T}) \leq u_i(x^S, d^{N\backslash S}) \leq u_i(x^S, x^{N\backslash S}),$$

which implies that no $T \subseteq S$ has an $\alpha$-deviation at $(x^S, x^{N\backslash S})$. Hence, for all $x^{N\backslash S}$, there exists no $T \subseteq S$ that has a credible $\alpha$-deviation at $(x^S, x^{N\backslash S})$, so that $x^S$ is a self-binding strategy for $S$.

□

**Problem 6.** Show that in the pure exchange game with nonnegative strategies (see Problem 4), every nonempty subset of $N$ is a self-binding coalition.

### 2.3 Derivation of NTU Market Games

In the pure exchange game, every coalition $S$ acquires the set $v(S)$ of payoff vectors that S can assure itself:

$$v(S) = \{v \in \Re^N | \; \exists x^S \in X^S \; \forall z \in X, \; u_S(x^S, z^{N \setminus S}) \geq v_S := (v_i)_{i \in S}\}$$

On the other hand, defining for each coalition $S$ an $S$-*allocation* to be $y^S = \{y^i\}_{i \in S}$ satisfying $\forall i \in S, \; y^i \in \Re^m_+$ and $\sum_{i \in S} y^i = \sum_{i \in S} w^i$, the following set $V(S)$ of payoff vectors can be associated to each coalition $S$:

$$V(S) = \{v \in \Re^N | \; \exists S - allocation \; y^S \; \forall i \in S, \; f_i(y^i) \geq v_i\},$$

where the utility function $f_i$ is the same as in **Problem 4**.

Then, it is easy to show that $v(S)$ is identical to $V(S)$ by considering for any given $S$-allocation $y^S$ the strategy $x^S$ defined by

$$x^{ij,k} = \frac{\omega^{i,k} y^{j,k}}{\sum_{i \in S} \omega^{i,k}}, \quad i, j \in S, \quad k = 1 \ldots m.$$

The collection $\{V(S) | S \subseteq N\}$ is called the **NTU market game** (see Scarf [22]).

**Problem 7.** Show that $v(S)$ is identical to $V(S)$ for all $S \subseteq N$.

In this way, the NTU market game is derived from the pure exchange game, so that every coalition of the market game is also self-binding. It is this result that may explain why a market game, whether it is of TU or NTU, has been a central economic application of cooperative game theory.

# 3 NTU Cores and Related Topics

## 3.1 Core Equivalent Strong Nash Equilibria of the Pure Exchange Game

### 3.1.1 Scarf's Pure Exchange Game

A typical economic example of strategic games with coalitions is the *pure exchange game* due to Scarf [23].

Let $N = \{1, \ldots, n\}$ be the set of players, and let $w = (w^1, \ldots, w^n) \in \mathfrak{R}^{nm}_{+}$, where $w^i = (w^i_1, \ldots, w^i_m) \in \mathfrak{R}^m_+$, be a vector of initial endowments. For each $S \subseteq N$, a vector $y = (y^1, \ldots, y^n) \in \mathfrak{R}^{nm}_+$ with $y^i = (y^i_1, \ldots, y^i_m) \in \mathfrak{R}^m_+$ is an *$S$-allocation* if it satisfies $\sum_{i \in S} y^i = \sum_{i \in S} w^i$, i.e.,

$$\sum_{i \in S} y^i_h = \sum_{i \in S} w^i_h, \; h = 1, \ldots, m.$$

The utility function $v_i(y^i)$ of player $i$ is assumed to be continuous, quasi-concave and monotone nondecreasing in $y^i$.

The *core* of a pure exchange economy is the set of $N-$allocations $y^*$ that are not dominated, i.e., the set of $N-$allocations $y^*$ such that for any $S \subseteq N$ there is no $S-$allocation $y$ satisfying: $v_i(y^i) > v_i(y^{*i}),\ \forall i \in S$.

Let $X^i$ be the set of strategies of player $i$ given by

$$X^i = \{x^i \in \Re^{nm} \,|\, x^i = (x^{i1}, \ldots, x^{in}),\quad x^{ij} = (x_1^{ij}, \ldots, x_m^{ij});\ \sum_{j \in N} x_h^{ij} = w_h^i,\ h = 1, \ldots m\}.$$

Then, the *pure exchange game* is a strategic game $G = (N, \{X^i\}, \{u_i\})$ with an *outcome function* $g : X \to \Re_+^{nm}$ such that

$$g(x) = \begin{cases} (\sum_{j \in N} x^{j1}, \ldots, \sum_{j \in N} x^{jn}) & if\ (\cdot) \in \Re_+^{nm}, \\ w & otherwise \end{cases}$$

The payoff $u_i(x)$ to player $i$ in the game $G$ is defined to be

$$u_i(x) = v_i(g(x)_i)$$

where $g(x)_i$ is the $i$-th component of $g(x)$.

Note that the outcome function $g$ generates an $N-$allocation, since

$$\sum_{i \in N} \sum_{j \in N} x^{ji} = \sum_{j \in N} \sum_{i \in N} x^{ji} = \sum_{j \in N} w^j.$$

**Remark 3.** Here, the strategies are not assumed to be nonnegative. The original Scarf's pure exchange game restricts the strategies to be nonnegative. Later, we also assume the nonnegativity.

**Proposition 2.** *Any $N-$allocation at a strong Nash equilibrium of $G = (N, \{X^i\}, \{u_i\})$ is contained in the core of the pure exchange economy.*

*Proof.* Let $y^*$ be an $N-$allocation that is not in the core. Then, there is an $S-$allocation $y$ that dominates $y^*$. Let $x^* \in X$ be any strategy combination that attains $y^*$, and define:

$$x_h^{ij} = \frac{w_h^i}{\sum_{i \in S} w_h^i} (y_h^j - \sum_{k \in N \setminus S} x_h^{*kj}).$$

Then,

$$\sum_{j\in N} x_h^{ij} = \frac{w_h^i}{\sum_{i\in S} w_h^i}\Big(\sum_{i\in N} y_h^i - \sum_{j\in N}\sum_{k\in N\setminus S} x_h^{*kj}\Big)$$
$$= \frac{w_h^i}{\sum_{i\in S} w_h^i}\Big(\sum_{i\in N} w_h^i - \sum_{i\in N\setminus S} w_h^i\Big) \quad = \quad w_h^i.$$

Hence,

$$x^i \in X^i \ \text{ for all } \ i \in N.$$

Moreover,

$$\sum_{i\in S} x_h^{ij} = y_h^j - \sum_{k\in N\setminus S} x_h^{*kj}.$$

Therefore, the $S-$allocation $y$ is attained by the deviation $x^S$ at $x^*$. By assumption, then, we have

$$u_i(x^S, x^{*N\setminus S}) > u_i(x^*), \text{ for all } i \in S,$$

which shows that $x^*$ is not a strong Nash equilibrium. □

Note that the strategy $x^S$ defined in the proof by

$$x_h^{ij} = \frac{w_h^i}{\sum_{i\in S} w_h^i}(y_h^j - \sum_{k\in N\setminus S} x_h^{*kj})$$

satisfies, as shown below, the equality that

$$\sum_{i\in S}\sum_{j\in N\setminus S} x_h^{ij} = \sum_{i\in S}\sum_{j\in N\setminus S} x_h^{*ji}, \quad h = 1, \ldots, m.$$

This means that when deviating from $x^*$, coalition $S$ returns the amount of goods to $N \setminus S$ exactly what it obtains from $N \setminus S$. In other words, $S$ is deviating with only resources available within $S$. In this sense, the deviation $x^S$ may be called the *self-supporting* deviation. Thus, the above proposition still holds if deviations are restricted to the self-supporting ones.

We will show that if only self-supporting deviations are allowed, the converse of the above proposition also holds. Before this, we shall verify that the deviation $x^S$ is self-supporting.

Note that since $y$ is an $S$-allocation, it is an $N \setminus S$-allocation whenever $y$ is an $N$-allocation. Hence, for this $y$ we have

$$\begin{aligned}
\sum_{i\in S} \sum_{j\in N\setminus S} x_h^{ij} &= \sum_{i\in S} \frac{w_h^i}{\sum_{i\in S} w_h^i} \left( \sum_{j\in N\setminus S} y_h^j - \sum_{j\in N\setminus S} \sum_{k\in N\setminus S} x_h^{*kj} \right) \\
&= \frac{\sum_{i\in S} w_h^i}{\sum_{i\in S} w_h^i} \left( \sum_{j\in N\setminus S} w_h^j - ( \sum_{k\in N\setminus S} w_h^k - \sum_{j\in S} \sum_{k\in N\setminus S} x_h^{*kj}) \right) \\
&= \sum_{j\in S} \sum_{k\in N\setminus S} x_h^{*kj}, \quad h = 1, \ldots, m.
\end{aligned}$$

**Proposition 3.** *The core of the pure exchange economy coincides with the set of $N$-allocations attained by the strong Nash euilibrium with only self-supporting deviations being permissible.*

*Proof.* Suppose that $x^* \in X$ admits a self-supporting deviation $x^S \in X^S$ by a coalition $S$. Then, the allocation $g(x^S, x^{*N\setminus S})$ is an $S$-allocation as shown below. Since $x^S$ is in $X^S$,

$$\sum_{i\in S}\sum_{j\in S} x_h^{ij} + \sum_{i\in S}\sum_{k\in N\setminus S} x_h^{ik} = \sum_{i\in S} w_h^i, \quad h = 1, ..., m.$$

But then, the definition of self-supporting deviations implies

$$\sum_{i\in S}\sum_{j\in S} x_h^{ij} + \sum_{j\in N\setminus S}\sum_{i\in S} x_h^{*ji} = \sum_{i\in S} w_h^i, \quad h = 1, ..., m.$$

Hence, $g(x^S, x^{*N\setminus S})$ is an $S$-allocation. Then, the allocation $g(x^*)$ is improved by the $S$-allocation $g(x^S, x^{*N\setminus S})$, which implies that when the allocation $g(x^*)$ is in the core, the strategy profile $x^*$ is a strong Nash equilibrium with deviations being restricted to be self-supporting.

The converse follows from the above proposition and the remark that the deviation there is self-supporting. □

Recall that an NTU game $(N, V)$, or simply $V$ is *convex* iff

$$V(S) \cap V(T) \subseteq V(S \cup T) \cup V(S \cap T) \quad \forall S, T \subseteq N.$$

Let $V$ satisfy the condition

$$S, T \subseteq N \Rightarrow V(S) \cap V(T) \subseteq V(S \cup T).$$

Then, this game $V$ is convex, and moreover, balanced. Convexity is straightforward. Balancedness follows from the repeated use of the above condition to obtain

$$\bigcap_{S \in \mathcal{B}} V(S) \subseteq V(\bigcup_{S \in \mathcal{B}} S) \subseteq V(N).$$

Such a game is of course exceptional. But Masuzawa[15] presented an interesting condition for an NTU game to have this property. There are many examples of the TU convex game, but a strategic environment generating

NTU convex game is not known. We show below that the condition called the **punishment dominance** makes it possible to obtain NTU convex games.

**Definition 14.** *Let $S \subseteq N$, $x_S, y_S \in X_S$. Then, we say $x_S$ is punishment dominant over $y_S$ against $N \setminus S$, and write $x_S P_S y_S$, if for all $z \in X$ we have*

$$u_i(z_{N \setminus S}, y_S) \geq u_i(z_{N \setminus S}, x_S) \;\; \forall i \in N \setminus S.$$

**Assumption 1.** *For all $i \in N$, and all $x_i, y_i \in X_i$, $x_i$ is punishment dominant over $y_i$ against $N \setminus \{i\}$, or $y_i$ is punishment dominant over $x_i$ against $N \setminus \{i\}$.*

This assumption says that a strategy of a player can make all of the other players' payoffs larger or smaller. Typical examples of this situations would be the n-person Prisoner's Dilemma, the greenhouse effects, public goods.

**Theorem 7.** *Under Assumption (1) ,*

$$V(S) \cap V(T) \subseteq V(S \cup T) \;\; \forall S, T \subseteq N,$$

*where $V = V_\alpha$.*

*Proof.* For each player $i$, let $x_i, y_i \in X_i$ be two arbitrary strategies, and define

$$p(x_i, y_i) = x_i \iff y_i P_i x_i, where P_i := P_{\{i\}}.$$

That is, $p(x_i, y_i)$ is the one that is 'less' punishment dominant among the two $x_i$ and $y_i$.

Let $v \in V(S) \cap V(T)$. By the definition of $V_\alpha$, the $alpha-$effectiveness, there exist $a_S \in X_S$ and $b_T \in X_T$ so that

$$\forall i \in S,\ \forall d_{N \setminus S} \in X_{N \setminus S},\ u_i(a_S, d_{N \setminus S}) \geq v_i$$
$$\forall j \in T,\ \forall d_{N \setminus T} \in X_{N \setminus T},\ u_j(b_T, d_{N \setminus T}) \geq v_j.$$

Define the strategy $z_{S \cup T} \in X_{S \cup T}$ of $S \cup T$ as follows.

1. $z_i = a_i\ \ \forall i \in S \setminus T$,
2. $z_i = b_i\ \ \forall i \in T \setminus S$,
3. $z_i = p(a_i, b_i)\ \ \forall i \in S \cap T$.

That is, players in $S \cap T$ are choosing less punishment dominant strategies among $a$ and $b$. Then, members $i \in S$ with $z_i = a_i$ can assure the amount of payoffs obtainable when any member in $j \in S \cap T$ takes the strategy $a_j$. Indeed, for any player $i \in S$ with $z_i = a_i$, and any $d_{N \setminus (S \cup T)} \in X_{N \setminus (S \cup T)}$, we have

$$\begin{aligned} v_i &\leq u_i(a_i, a_{S \setminus \{i\}}, b_{T \setminus S}, d_{N \setminus (S \cup T)}) \\ &\leq u_i(a_i, a_{S \setminus (T \cup \{i\})}, z_{S \cap T \setminus \{i\}}, b_{T \setminus S}, d_{N \setminus (S \cup T)}) \\ &= u_i(z_{S \cup T}, d_{N \setminus (S \cup T)}). \end{aligned}$$

Similarly, players $i \in T$ can guarantee the amount of payoffs obtained when any member $j \in S \cap T$ takes $b_j$:

$$v_i \leq u_i(z_{S \cup T}, d_{N \setminus (S \cup T)}) \ \ \forall i \in S \cup T$$

Hence, we have that $v \in V(S \cup T)$. $\square$

Note that this result does not need the quasiconcavity of payoff functions. Therefore, the class of games in this theorem can be different from that of

Scarf [23].

## 参考文献


[1] R.J.Aumann, “Acceptable points in general cooperative n-person games,” *Annals of Mathematics Studies* **40**(1959), 287–324.

[2] R.J.Aumann, “Subjectivity and correlation in randomized strategies,” *Journal of Mathematical Economics*, **1**(1973),67–96.

[3] R.J.Aumann and B.Peleg, “Von-Neumann-Morgenstern solutions to co-operative games without side payments,” *Bulletin of the American Mathematical Society* **66**(1960),173–179.

[4] B.D.Bernheim, B.Peleg and M.D.Whinston, “Coalition-proof equilibria I. concepts,” *Journal of Economic Theory*, **42**(1987), 1–12.

[5] O.N.Bondareva, “Nekotrye Primeneniia Metodov Linejnogo Program-mirovaniia k Teorii Kooperativnykh Igr” (Some Applications of Linear Programing Methods to the Theory of Games), **10**(1963), 119–139.

[6] T.Ichiishi, *Game Theory for Economic Analysis*, Academic Press, New York, 1983.

[7] J.Greenberg and S.Weber, “Stable coalition structures with a unidimensional set of alternatives,” *Journal of Economic Theory* **60**(1993),62–82.

[8] J.C.Harsanyi and R.Selten, *A General Theory of Equilibrium Selection in Games*, MIT press, 1988.

[9] T.Hirai, T.Masuzawa and M.Nakayama, “Coalition-proof Nash equilibria and cores in a strategic pure exchange game of bads,” *Mathematical Social Sciences* **51** (2006), 162-170.

[10] R.Holzman and N. Law-Yone, “Strong Equilibrium in Congestion Games,” *Games and Economic Behavior* **21** (1997), 85-101.

[11] T.Ichiishi, *Game Theory for Economic Analysis*, Academic Press, New York, 1983.

[12] G.Jentzsch, “Some thoughts on the theory of cooperative games,” in *Advances in Game Theory, Ann. of Math. Study*, **52** M.Dresher,

L.S.Shapley and A.W.Tucker, eds., Princeton University Press, 1964.

[13] E.Kalai, A.Postlewaite and J.Roberts, "A Group Incentive Compatible Mechanism Yielding Core Allocations," *Journal of Economic Theory* **20**(1979), 13–22.

[14] A.Mas-Colell, "Cooperative equilibrium," in *The New Palgrave: Game Theory*, J.Eatwell, M.Milgate and P.Newman eds., The Macmillan Press, 1987.

[15] T.Masuzawa, "Punishment Strategies Make the $\alpha-$Coalitional Game Ordinally Convex and Balanced," *International Journal of Game Theory*, **32** (2003), 479–483.

[16] M.Nakayama, "Self-binding coalitions," *Keio Economic Studies* **35**(1998),1-8.

[17] J.F.Nash, "Non-cooperative games," *Annals of Mathematics*, **54** (1951), 286–295.

[18] J.F.Nash, "Two-person cooperative games," *Econometrica*,

**21**(1953),128–140]

[19] K.Nishihara, “Stability of the cooperative equilibrium in N-person prisoners’ dilemma with sequential moves,” *Economic Theory*, **13**(1999),483–494.

[20] B.Peleg, *Game Theoretic Analysis of Voting in Committees*, Cambridge Univ.Press, London, 1984.

[21] D.Ray, “Credible coalitions and the core,” *International Journal of Game Theory*, **18**(1989),185–187.

[22] H.Scarf, “The core of an n person game,” *Econometrica*, **35**(1967), 50–69.

[23] H.Scarf, “On the existence of a cooperative solution for a general class of n-person games,” *Journal of Economic Theory*, **3**(1971),169–181.

[24] L.S.Shapley, “On balanced sets and cores,” *Naval Research Logistics Quarterly* **14**(1967),453–460.

[25] J. von Neumann and O.Morgenstern, *Theory of Games and Economic*

*Behavior,3rd edition*, John Wiley & Sons, 1953.

[26] Utsumi,Y. and M.Nakayama, "Strategic Cores in a Public Goods Economy," *International Game Theory Review* **6** (2004), 509-524.

[27] J.Wako, "Coalition-proof Nash allocation in a barter game with multiple indivisible goods," mimeo., 1998.

[28] J.Zhao, "The Existence of TU $\alpha$-core in normal form games," *International Journal of Game Theory*, **28** (1999), 25–34.

[29] J.Zhao, "A $\beta$-core Existence Result and its Application to Oligopoly Markets," *Games and Economic Behavior*, **27** (1999), 153–168.