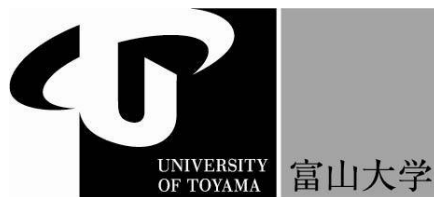**Working Paper No. 301**

Single-payoff farsighted stable sets in strategic games with dominant punishment strategies

Toshiyuki Hirai

First version: September 30, 2016;
Revised version: October 24, 2016

UNIVERSITY OF TOYAMA　富山大学

FACULTY OF ECONOMICS
UNIVERSITY OF TOYAMA

Single-payoff farsighted stable sets in strategic games with dominant punishment strategies[*]

Toshiyuki Hirai[†]

First version: September 30, 2016;
Revised version: October 24, 2016

Abstract

We investigate the farsighted stable set in a class of strategic games with dominant punishment strategies. In this class of games, each player has a strategy that uniformly minimizes the other players' payoffs for any given strategies of other players. We particularly investigate a special class of the farsighted stable sets each of which consists of strategy profiles yielding a single payoff vector. We call such farsighted stable sets as the single-payoff farsighted stable sets. We propose a concept called the inclusive set that completely characterizes the single-payoff farsighted stable sets in the strategic games with dominant punishment strategies. We also show that the set of payoff vectors yielded by the single-payoff farsighted stable sets is closely related to the strict α-core in strategic games. Further, we apply the results to the strategic games where each player has two strategies and strategic games associated with some market models.

JEL Classification: C70, C71
Keywords: farsighted stable set; dominant punishment strategy; inclusive set;
            strict α-core.

1 Introduction

The first solution concept in games with coalitional behavior is the stable set (von Neumann and Morgenstern, 1944), the definition of which is based on the dominance relation among the outcomes. Harsanyi (1974) argued the importance of taking the farsightedness of players into account in view of the underlying bargaining process for the stable set. He proposed the indirect dominance relation that incorporates the farsightedness of the players. Later, Chwe (1994) formulated the indirect dominance relation and the farsighted stable set in abstract games. His farsighted stable set can be applied to strategic games by employing the coalitional or individual contingent threats situation by Greenberg (1990).

Thereafter, the farsighted stable set has been investigated in various strategic games associated with economic situations. For example, Masuda (2002) for the average return game of a production economy; Suzuki and Muto (2005) and Nakanishi (2009) for the prisoners' dilemma game with and without coalitional deviation, respectively; Kawasaki and Muto (2009) for the indivisible public good provision game, Kamijo and Muto (2010) for cartel formation game; Shino and Kawasaki (2012) for the Hotelling's location game; Kawasaki, *et al.* (2015) for the tariff game, among others.[1] In the most of these literature, the farsighted stable sets each of which yields a single payoff vector were well-investigated and shown to have nice properties. We will refer such a farsighted stable set as the single-payoff farsighted stable set. The single-payoff farsighted stable set also has a nice property from the theoretical aspect. Dutta and Vohra (2016) defined the farsighted stable set with rational expectations and showed the equivalence to the farsighted stable set if both are single-payoff sets[2].

This paper characterizes the single-payoff farsighted stable set in a general class of strategic games, called the class of strategic games with dominant punishment strategies.

---

[1]The farsighted stable set is also applied to the problems in other forms. For example, Diamantoudi and Xue (2003) for the hedonic coalition formation problem; Diamantoudi (2005) for the cartel formation problem; Diamantoudi and Sartzetakis (2015) and Benchekroun and Chaudhuri (2015) for the international environmental agreement formation problems; Page and Wooders (2009) for the network formation problem; Kawasaki (2010) and Klaus, *et al.* (2010) for the exchange economy with indivisible goods under the weak and strong dominance relations, respectively; Klaus, *et al.* (2011) for the roommate problem; Mouleon, *et al.* (2011) for the two-sided matching markets, and it is extended to a more general setting by Roketskiy (2012), among others.

[2]Dutta and Vohra (2016) defined the farsighted stable set with rational expectations for more abstract games that includes not only the strategic games but also coalitional games and other forms. We can apply their result to the strategic games and the farsighted stable set due to Chwe (1994).

The dominant punishment strategy was formulated by Nakayama (1998). In words, a strategy of a player is called the dominant punishment strategy of that player (against the other players) if it uniformly minimizes the other players' payoffs with irrespective to the strategy choices of other players. The class of strategic games with dominant punishment strategies includes various games. Indeed, some of the strategic games and those appropriately associated with the problems in the previously introduced literature are included. The strategic games with monotone externalities[3] and the strategic games with punishment dominance relation (Masuzawa, 2003) are also the strategic games with dominant punishment strategies if the strategy sets are compact. See Nakayama (1998) for more examples.

The main result of this paper is summarized as follows. In a strategic game with dominant punishment strategies, the set of strategy profiles yielding a single payoff vector is a single-payoff farsighted stable set if and only if the set satisfies a condition that will be called the inclusiveness. The dominant punishment strategies are exploited to obtain the sufficiency of this equivalence, while the necessity can be shown without the dominant punishment strategies. Also, we will show that the set of the yielded payoff vectors by the inclusive sets is closely related to the strict α-core. Further, we will apply these results for the strategic games where each player has two strategies and strategic games associated with a pure exchange economy and a many-to-many matching problem with substitutable preferences, respectively.

The single-payoff farsighted stable set is also investigated in coalitional games. Béal, et al. (2008) showed the following interesting, but problematic result under a very mild condition[4]. In the TU coalitional games, any singleton consisting of an imputation not in the interior of the core is a farsighted stable set. Ray and Vohra (2015a) pointed out that this result is due to the unreasonable coalitional sovereignty in the definition of the indirect dominance relation. To resolve this problem, Ray and Vohra (2015a) introduced the effectivity of the coalitions to describe explicitly what outcomes a coalition can induce from a given outcome in coalitional games, which is a missed argument from Harsanyi's (1974) indirect dominance relation. Recently, Chander (2015) introduced a similar indirect dominance relation to the partition function games.

---

[3]For example, Milgrom and Roberts (1990, 1996) considered the subclass of the games with strategic complementarities that also has this property.

[4]For the NTU coalitional game, Bhattacharya and Brosi (2011) showed the existence of the farsighted stable set but did not argue its characteristics explicitly.

Under the new indirect dominance relation, Ray and Vohra (2015a) showed that the single-payoff farsighted stable set is equivalent to the separable payoff vector, which is closely related to the coalition structure core of the coalitional game. Also, Chander (2015) showed that the union of the single-payoff farsighted stable sets is equivalent to the strong-core. Further, he showed that they lie between the γ-core and the δ-core in the partition function game with positive or negative externalities. The main result of this paper is a counterpart of these results in the strategic game. In particular, our result is closely related to that of Ray and Vohra (2015a) as the inclusive set is a close, but slightly different concept from the separable payoff, which will be discussed in the later section.

Another related result is Corollary 1 by Kawasaki (2015). He showed a similar characterization of the single-payoff farsighted stable set in two-player strategic games where the minimax payoff coincides with the maxmin payoff for both players. He also showed that this condition is not sufficient when there are three players. In the strategic game with dominant punishment strategies, the coincidence of the minimax payoff and the maximin payoff of each player is guaranteed since the α- and β-coalitional games coincide with each other. (Nakayama, 1998.) Our result requires a stronger condition than that of Kawasaki (2015) but applies to n-player games.

The remaining of this paper is organized as follows. In the next section, we introduce some basic definitions. In section 3, we define two important concepts, the single-payoff farsighted stable set and the inclusive set. The main results are stated and proved in section 4. We apply our main results for the strategic game where each player has two strategies and strategic games associated with certain market models in section 5 and 6, respectively. In the final section, we conclude with a remark.

2   Preliminaries

The following notations are used throughout this paper. For any pair of real vectors $x, y \in \mathbb{R}^\ell$, we denote $x = y$ iff $x_i = y_i$ for all $i = 1, ..., \ell$; $x \geq y$ iff $x_i \geq y_i$ for all $i = 1, ..., \ell$; and $x > y$ iff $x_i > y_i$ for all $i = 1, ..., \ell$, respectively. For each finite set A, let $|A|$ denote the cardinality of A and 2^A denote the power set of A, respectively.

A strategic game is defined by a tuple $G = (N, (X_i)_{i \in N}, (v_i)_{i \in N})$, where $N = \{1, ..., n\}$ is the set of players, X_i is the set of strategies for each $i \in N$, and $v_i : \prod_{i \in N} X_i \to \mathbb{R}$ is

the payoff function of each $i \in N$. A nonempty subset of N is called a coalition. Let $\mathcal{N} = 2^N \setminus \{\emptyset\}$ denote the set of coalitions. For each $S \in \mathcal{N}$, we denote $X_S = \prod_{i \in S} X_i$. For the simplicity, we denote $X = X_N$ and $X_{-i} = X_{N \setminus \{i\}}$ for each $i \in N$. Similarly, for each $S \in \mathcal{N}$, $i \in N$, and $x \in X$, we denote $v_S(x) = (v_i(x))_{i \in S}$, $v(x) = v_N(x)$, and $v_{-i}(x) = v_{N \setminus \{i\}}(x)$. Further, for each $S \in \mathcal{N} \setminus \{N\}$, $i \in N$, and $x \in X$, we denote $x_S = (x_i)_{i \in S}$ and $x_{-i} = x_{N \setminus \{i\}}$.

In a strategic game G, $x \in X$ is said to be (Pareto) efficient iff there exists no $y \in X$ such that $v(y) \geq v(x)$ and $v(y) \neq v(x)$; $x \in X$ is said to be weakly (Pareto) efficient iff there exists no $y \in X$ such that $v(y) > v(x)$.

The following concept formulated by Nakayama (1998) and its strict modification will play important roles in this paper.

Definition 1 *Let G be a strategic game and $i \in N$.*

- *A strategy $d_i \in X_i$ is said to be a dominant punishment strategy of i (against $N \setminus \{i\}$) iff for any $x_i \in X_i$ and any $x_{-i} \in X_{-i}$, $v_{-i}(x) \geq v_{-i}(d_i, x_{-i})$.*

- *We say the dominant punishment strategy d_i of i is strict if for any $x_i \in X_i \setminus \{d_i\}$ and any $x_{-i} \in X_{-i}$, $v_{-i}(x) > v_{-i}(d_i, x_{-i})$.*

We say a strategic game G is a strategic game with dominant punishment strategies iff there exists the dominant punishment strategy of each $i \in N$. Similarly, G is a strategic game with strictly dominant punishment strategies iff there exists the strictly dominant punishment strategy of each $i \in N$.

Following Aumann and Peleg (1960), we define the α-coalitional game associated with a strategic game. In what follows, we denote $u_S = (u_i)_{i \in S}$ for each $S \in \mathcal{N}$ and $u \in \mathbb{R}^n$. Given a strategic game $G = (N, (X_i)_{i \in N}, (v_i)_{i \in N})$, the α-coalitional game associated with G is defined by $\Gamma = (N, V_\alpha^G)$, where $V_\alpha^G : \mathcal{N} \twoheadrightarrow \mathbb{R}^n$ is a characteristic correspondence such that for each $S \in \mathcal{N}$,

$$V_\alpha^G(S) = \begin{cases} \bigcup_{x_S \in X_S} \bigcap_{z_{N \setminus S} \in X_{N \setminus S}} \left\{ u \in \mathbb{R}^n | u_S \leq v_S\left(x_S, z_{N \setminus S}\right) \right\} & \text{if } S \neq N; \\ \bigcup_{x \in X} \left\{ u \in \mathbb{R}^n | u \leq v(x) \right\} & \text{if } S = N. \end{cases}$$

Let G be a strategic game, Γ be the α-coalitional game associated with G, and $u \in \mathbb{R}^n$ be a payoff vector. In Γ, u is said to be (Pareto) efficient iff $u \in V_\alpha^G(N)$ and there exists no $u' \in V_\alpha^G(N)$ such that $u' \geq u$ and $u \neq u'$; u is said to be weakly (Pareto) efficient iff

5

$u \in V_\alpha^G(N)$ and there exists no $u' \in V_\alpha^G(N)$ such that $u' > u$. Further, u is said to be in the strict α-core of G iff $u \in V_\alpha^G(N)$ and there exists no $S \in \mathcal{N}$ and $u' \in V_\alpha^G(S)$ such that $u'_S \geq u_S$ and $u'_S \neq u_S$; u is said to be in the relative interior of the strict α-core of G iff u is in the strict α-core and $u \notin V_\alpha^G(S)$ for all $S \in \mathcal{N} \setminus \{N\}$. Note that any payoff vector in the strict α-core is efficient by definition.

Remark 1 We summarize the following well-known properties of the α-coalitional game Γ associated with a strategic game G.

(a) Γ satisfies the superadditivity, that is, $V_\alpha^G(S) \cap V_\alpha^G(T) \subset V_\alpha^G(S \cup T)$ for any $S, T \in \mathcal{N}$ with $S \cap T = \emptyset$. To see this, let $S, T \in \mathcal{N}$ with $S \cap T = \emptyset$ and $u \in V_\alpha^G(S) \cap V_\alpha^G(T)$. Then, there exist $x_S \in X_S$ and $x_T \in X_T$ such that $u_S \leq v_S(x_S, z_{N\setminus S})$ for any $z_{N\setminus S} \in X_{N\setminus S}$ and $u_T \leq v_T(x_T, z_{N\setminus T})$ for any $z_{N\setminus T} \in X_{N\setminus T}$. By $S \cap T = \emptyset$, $u_S \leq v_S(x_{S\cup T}, z_{N\setminus(S\cup T)})$ and $u_T \leq v_T(x_{S\cup T}, z_{N\setminus(S\cup T)})$ for any $z_{N\setminus(S\cup T)} \in X_{N\setminus(S\cup T)}$. Hence, $u \in V_\alpha^G(S \cup T)$. In particular, $\bigcap_{S\in\mathcal{P}} V_\alpha^G(S) \subset V_\alpha^G(N)$ for all partition $\mathcal{P}$ of N.

(b) If $u \in \mathbb{R}^n$ is efficient in Γ, then there exists some $x \in X$ such that $u = v(x)$. To see this, let $u \in \mathbb{R}^n$ be an efficient payoff vector in Γ. By $u \in V_\alpha^G(N)$, there exists some $\hat{x} \in X$ such that $u \leq v(\hat{x})$. The equality must hold by the efficiency of u and $v(\hat{x}) \in V_\alpha^G(N)$. On the other hand, this statement may fail for the weakly efficient payoff vector.

3   The farsighted stable set and the inclusive set

We introduce the concept of (coalitionally) farsighted stable set in strategic games due to Chwe (1994). We begin with defining the inducibility and the indirect dominance relation between two strategy profiles. Let $x, y \in X$ and $S \in \mathcal{N}$. We say y is inducible from x via S, denoted by $x \rightarrow_S y$, iff $x_{N\setminus S} = y_{N\setminus S}$. We say y indirectly dominates x, denoted by $y \succ x$, iff there exist a sequence of strategy profiles $z^0, ..., z^m$ with $z^0 = x$ and $z^m = y$, and a sequence of coalitions $S^1, ..., S^m$ such that for all $h = 1, ..., m$, $z^{h-1} \rightarrow_{S^h} z^h$ and $v_{S^h}(y) > v_{S^h}(z^{h-1})$.

In this paper, we will concentrate on a special class of the farsighted stable set, called the single-payoff farsighted stable set. Define $[x] = \{y \in X | v(y) = v(x)\}$ for any $x \in X$ that will be sometimes called single-payoff strategy set.

Definition 2 *Let G be a strategic game.*

- *We say $K \subset X$ is a farsighted stable set (FSS) in G iff K satisfies the following two stabilities:*

  Internal stability: *for any $x, y \in K$, $x \succ y$ does not hold;*

  External stability: *for any $x \in X \setminus K$, there exists some $y \in K$ such that $y \succ x$.*

- *We say K is a single-payoff farsighted stable set (SP-FSS) in G iff K is a FSS and $K = [x]$ for some $x \in X$.*

In coalitional games, Ray and Vohra (2015a) showed that the set of outcomes yielding a single payoff vector is a SP-FSS defined by their indirect dominance relation if and only if the payoff vector is separable. Here, we give the definition of the separable payoff vector in terms of the α-coalitional game. Let $\Gamma = (N, V_\alpha^G)$ be an α-coalitional game associated with a strategic game G. A payoff vector $u \in \mathbb{R}^n$ is said to be separable in Γ iff (i) u is efficient in Γ;[5] (ii) for any $T \in \mathcal{N} \setminus \{N\}$ and any partition $\mathcal{P}(T)$ of T, if $u \in V_\alpha^G(T')$ for all $T' \in \mathcal{P}(T)$, there exists some nonempty $S \subset N \setminus T$ such that $u \in V_\alpha^G(S)$. We introduce a corresponding concept for the single-payoff strategy set, called the inclusive set.

Definition 3 *Let G be a strategic game and $x \in X$. We say $[x]$ is an inclusive set in G iff for any $S \in \mathcal{N}$ and any $y_S \in X_S$, $v_S(y_S, z_{N \setminus S}) \geq v_S(x)$ for any $z_{N \setminus S} \in X_{N \setminus S}$ implies $y_S = x'_S$ for some $x' \in [x]$.*

Given $x \in X$, a single-payoff strategy set $[x]$ is called an inclusive set if it includes any strategy profile such that a coalition can guarantee for themselves the payoffs at x by taking the part of the strategy profile. The definition of the inclusive set can be restated as follows in a strategic game with dominant punishment strategies.

Remark 2 *Let G be a strategic game with dominant punishment strategies and $x \in X$. For each $i \in N$, let d_i denote the dominant punishment strategy of i. We say $[x]$ is an inclusive set in G iff for any $S \in \mathcal{N}$ and any $y_S \in X_S$, $v_S(y_S, d_{N \setminus S}) \geq v_S(x)$ implies $y_S = x'_S$ for some $x' \in [x]$.*

---

[5]In Ray and Vohra (2015a), the definition of the efficiency of a payoff vector in a (possibly not superadditive) coalitional game (N, V) is slightly different from ours as follows: u is efficient if $u \in \bigcap_{S \in \mathcal{P}} V(S)$ for some partition $\mathcal{P}$ of N and there exists no partition $\mathcal{P}'$ of N and $u' \in \bigcap_{S \in \mathcal{P}'} V(S)$ such that $u' \geq u$ and $u' \neq u$. This definition is equivalent to that in the present paper for the α-coalitional games by its generic superadditivity as mentioned in Remark 1(a).

Ray and Vohra (2015a) argued the close relationship between the separable payoff vectors and the coalition structure core in coalitional games. We show that the inclusive set also has a close relationship with the strict α-core in strategic games.

Proposition 1 *Let G be a strategic game and $x \in X$.*

(a) *If $v(x)$ is in the relative interior of the strict α-core of G, then $[x]$ is an inclusive set in G.*

(b) *If $[x]$ is an inclusive set in G, then $v(x)$ is in the strict α-core of G.*

Proof. Let $\Gamma = \left(N, V_\alpha^G\right)$ be the α-coalitional game associated with a strategic game G and $x \in X$.

(a) Assume that $v(x)$ is in the relative interior of the strict α-core of G. By the efficiency of $v(x)$, $v(y) \geq v(x)$ implies $v(y) = v(x)$ and $y \in [x]$ for any $y \in X$. By $v(x) \notin V_\alpha^G(S)$ for any $S \in \mathcal{N} \setminus \{N\}$, there exists no $T \in \mathcal{N} \setminus \{N\}$ and $y_T \in X_T$ such that $v_T(y_T, z_{N \setminus T}) \geq v_T(x)$ for any $z_{N \setminus T} \in X_{N \setminus T}$ at all. Thus, $[x]$ is an inclusive set in G.

(b) Assume that $[x]$ is an inclusive set in G. Suppose that $v(x)$ is not in the strict α-core of G. Then, there exist some $S \in \mathcal{N}$ and $u' \in V_\alpha^G(S)$ such that $u'_S \geq v_S(x)$ and $u'_S \neq v_S(x)$. Let $j \in S$ such that $u'_j > v_j(x)$. By $u' \in V_\alpha^G(S)$, there exists some $y_S \in X_S$ such that $v_S\left(y_S, z_{N \setminus S}\right) \geq u'_S$ for any $z_{N \setminus S} \in X_{N \setminus S}$. Thus, for any $z_{N \setminus S} \in X_{N \setminus S}$, $v_j\left(y_S, z_{N \setminus S}\right) \geq u'_j > v_j(x)$. Hence, $\left(y_S, z_{N \setminus S}\right) \notin [x]$ for any $z_{N \setminus S} \in X_{N \setminus S}$, contradicting that $[x]$ is an inclusive set. ∎

Therefore, the union of the payoff vectors yielded by the inclusive sets is very close to the strict α-core. We also point out that the inclusive sets in a strategic game is related to the separable payoff vectors by Ray and Vohra (2015a) in the associated α-coalitional game.

Proposition 2 *Let G be a strategic game, Γ be the α-coalitional game associated with G, and $x \in X$. If $u = v(x)$ is separable in Γ, then $[x]$ is an inclusive set in G.*

Proof. Let G be a strategic game, Γ be the α-coalitional game associated with G, and $x \in X$. Assume that $u = v(x)$ is separable in Γ.

Fix arbitrary $S_1 \in \mathcal{N}$ and $x^*_{S_1} \in X_{S_1}$ such that $v_{S_1}(x^*_{S_1}, z_{N \setminus S_1}) \geq u_{S_1}$ for any $z_{N \setminus S_1} \in X_{N \setminus S_1}$. Then, $u \in V_\alpha^G(S_1)$. If $S_1 = N$, then $v(x^*) = u$ follows from the efficiency of u, and thus $x^* \in [x]$. Therefore, assume that $S_1 \neq N$.

Since u is separable in Γ, there exist some $S_2 \in \mathcal{N}$ such that $S_2 \subset N \setminus S_1$ and $u \in V_\alpha^G(S_2)$, and some $S_3 \in \mathcal{N}$ such that $S_3 \subset N \setminus (S_1 \cup S_2)$ and $u \in V_\alpha^G(S_3)$, and so on. Eventually, there exists a partition $\{S_1, ..., S_k\}$ of N such that $u \in V_\alpha^G(S_h)$ for all $h = 1, ..., k$. Then, for each $h = 2, ..., k$, there exists some $x_{S_h}^* \in X_{S_h}$ such that $v_{S_h}(x_{S_h}^*, z_{N \setminus S_h}) \geq u_{S_h}$ for any $z_{N \setminus S_h} \in X_{N \setminus S_h}$. Therefore, $v(x^*) \geq u$. By the efficiency of u, $v(x^*) = u$. Thus, $x^* \in [x]$, and $[x]$ is an inclusive set in G. ∎

Note that neither Proposition 1 nor 2 requires the dominant punishment strategies. On the other hand, the converse of Proposition 2 does not hold even in a strategic game with dominant punishment strategies. We show this fact by employing the lumpy public good game due to Taylor (1987), the FSS of which was investigated by Kawasaki and Muto (2009).

The lumpy public good game is a strategic game, where $X_i = \{C, D\}$ for all $i \in N$ and the payoff functions are defined as follows. Let $g : \{C, D\} \times \{0, ..., n-1\} \to \mathbb{R}$ be a function such that

$$g(C, h) = \begin{cases} B - K & \text{if } h \geq r^* - 1; \\ -K & \text{if } h < r^* - 1; \end{cases}$$

$$g(D, h) = \begin{cases} B & \text{if } h \geq r^*; \\ 0 & \text{if } h < r^*, \end{cases}$$

where $0 < K < B$ and $r^* = 1, ..., n-1$ is a constant. Then,

$$v_i(x) = \begin{cases} g(C, |C(x)| - 1) & \text{if } x_i = C; \\ g(D, |C(x)|) & \text{if } x_i = D, \end{cases}$$

where $C(x) = \{i \in N | x_i = C\}$. Note that D is the dominant punishment strategy of each player.

Define $x^* \in X$ be a strategy profile such that $x_i^* = C$ for all $i = 1, ..., r^*$ and $x_i^* = D$ for all $i = r^* + 1, ..., n$. It can be easily confirmed that $[x^*] = \{x^*\}$ is an inclusive set. Note that $v_i(x^*) = B - K$ for all $i = 1, ..., r^*$ and $v_i(x^*) = B$ for all $i = r^* + 1, ..., n$. Let $R^* = \{1, ..., r^*\}$. Then, $v(x^*) \in V_\alpha^G(R^*)$, but $v(x^*) \notin V_\alpha^G(S)$ for all nonempty $S \subset N \setminus R^*$ since either the payoffs of at least r^* members in S are at most $B - K$ (when $|S| \geq r^*$) or any member in S obtains at most the payoff 0 (when $|S| < r^*$), while $v_i(x^*) = B$ for all $i \in N \setminus R^*$. Hence, $v(x^*)$ is not separable.

In some strategic games, on the other hand, any inclusive set turns out to be a separable payoff vectors in the associated α-coalitional games. We will discuss this

point later in Section 6.

4   Main results

In this section, we state and prove the main results.

Theorem 1 *Let G be a strategic game with dominant punishment strategies and $x \in X$. In G, if $[x]$ is an inclusive set, then $[x]$ is a SP-FSS.*

Proof. Let G be a strategic game with dominant punishment strategies and d_i denote the dominant punishment strategy of each $i \in N$. Fix an arbitrary $x \in X$. Assume that $[x]$ is an inclusive set. Note that x is efficient in G by Proposition 1(b).

Claim 1 $[x]$ *satisfies the internal stability.*

Proof of Claim 1. Fix arbitrary $y, y' \in [x]$. Suppose that $y \succ y'$. Then, there exist sequences of strategy profiles $z^0 (= y'), ..., z^m (= y)$ and coalitions $S^1, ..., S^m$ such that for each $h = 1, ..., m$, $z^{h-1} \rightarrow_{S^h} z^h$ and $v_{S^h}\left(z^{h-1}\right) < v_{S^h}(y)$. However, constructing such sequences is impossible by $v(y') = v(z^0) = v(y)$. Hence, $y \succ y'$ is impossible. $\qquad \square$

We turn to the proof of the external stability of $[x]$. For any $x' \in X$, define

$$L(x', x) = \{i \in N | v_i(x') < v_i(x)\} \, .$$

Claim 2 *For any $x' \in X$, $L(x', x) = \emptyset$ if and only if $x' \in [x]$.*

Proof of Claim 2. Fix an arbitrary $x' \in X$. First, assume that $L(x', x) = \emptyset$. Then, $v(x') \geq v(x)$. By the efficiency of x, $v(x') = v(x)$. Thus, $x' \in [x]$.

Next, assume that $L(x', x) \neq \emptyset$. Then, there exists some $j \in N$ such that $v_j(x') < v_j(x)$. Thus, $x' \notin [x]$. $\qquad \square$

Fix an arbitrary $y \in X \setminus [x]$. We construct a sequence of strategy profiles and a sequence of coalitions that consist an indirect dominance relation $x' \succ y$ for some $x' \in [x]$. Define $z^0 = y$ and

$$S^h = \left\{i \in L(z^{h-1}, x) | z_i^{h-1} \neq d_i\right\};$$
$$z^h = \left(d_{S^h}, z_{N \setminus S^h}^{h-1}\right)$$

10

for each $h = 1, ...m$ iteratively as long as $S^h \neq \emptyset$. Thus, m is the maximum natural number such that $S^m \neq \emptyset$. Note that if $S^1 = \emptyset$, then $m = 0$.

We confirm that such a finite m exists. Suppose that $S^h \neq \emptyset$ for any natural number h. For any h, if $i \in S^h$, then $z_i^{h'-1} = d_i$ and $i \notin S^{h'}$ for all $h' > h$. Thus, $S^h \cap S^{h'} = \emptyset$ for any different h and h'. This means that the number of players taking the dominant punishment strategies at z^h monotonically increases as h increases. This contradicts the finiteness of N. Thus, there exists the maximum m such that $S^m \neq \emptyset$.

Claim 3 (i) *If $z^m \in [x]$, then $z^m \succ y$.*

(ii) *If $z^m \notin [x]$ and there exist some $\hat{S} \in \mathcal{N}$ and $\hat{z} \in [x]$ such that $z^m \to_{\hat{S}} \hat{z}$ and $v_{\hat{S}}(z^m) < v_{\hat{S}}(\hat{z})$, then $\hat{z} \succ y$.*

Proof of Claim 3. For each $h = 1, ..., m$, $z^{h-1} \to_{S^h} z^h$ since $z_i^h \neq z_i^{h-1}$ implies $i \in S^h$ for each $i \in N$. For each $h = 1, ..., m$, we have $v_{S^h}(x) > v_{S^h}(z^{h-1})$ by $S^h \subset L(z^{h-1}, x)$. By these observations, (i) if $z^m \in [x]$, then the sequences $z^0(= y), ..., z^m(\in [x])$ and $S^1, ..., S^m$ consist an indirect dominance relation $z^m \succ y$; (ii) if $z^m \notin [x]$ and there exist some $\hat{S} \in \mathcal{N}$ and $\hat{z} \in [x]$ such that $z^m \to_{\hat{S}} \hat{z}$ and $v_{\hat{S}}(z^m) < v_{\hat{S}}(\hat{z})$, then the sequences $z^0(= y), ..., z^m, \hat{z}(\in [x])$ and $S^1, ..., S^m, \hat{S}$ consist an indirect dominance relation $\hat{z} \succ y$. $\square$

Assume that $L(z^m, x) = \emptyset$. Then $z^m \in [x]$ by Claim 2. Thus, $z^m \succ y$ by Claim 3(i).

Assume that $L(z^m, x) \neq \emptyset$. Then, $z^m \notin [x]$ by Claim 2. By the construction of m, we have $z_{L(z^m,x)}^m = d_{L(z^m,x)}$. First, consider the case where $L(z^m, x) = N$. In this case, $z^m \to_N x$ and $v(z^m) < v(x)$. Thus, $x \succ y$ by Claim 3(ii). Next, consider the case where $L(z^m, x) \neq N$. In this case, $v_{N \setminus L(z^m,x)}(z_{N \setminus L(z^m,x)}^m, d_{L(z^m,x)}) \geq v_{N \setminus L(z^m,x)}(x)$ by $z_{L(z^m,x)}^m = d_{L(z^m,x)}$. Since $[x]$ is an inclusive set, there exists some $x^* \in [x]$ such that $x_{N \setminus L(z^m,x)}^* = z_{N \setminus L(z^m,x)}^m$. Then, $z^m \to_{L(z^m,x)} x^*$ and $v_{L(z^m,x)}(x^*) > v_{L(z^m,x)}(z^m)$. Thus, $x^* \succ y$ by Claim 3(ii), where $x^* \in [x]$. Hence, $[x]$ is a SP-FSS. $\blacksquare$

Theorem 2 *Let G be a strategic game and $x \in X$. In G, if $[x]$ is a SP-FSS, then $[x]$ is an inclusive set.*

Proof. Let G be a strategic game and $x \in X$. We show the contraposition. Assume that $[x]$ is not an inclusive set in G. Then, there exist some $S \in \mathcal{N}$ and some $y_S \in X_S$

such that $v_S(y_S, z_{N \setminus S}) \geq v_S(x)$ for any $z_{N \setminus S} \in X_{N \setminus S}$ and $y_S \neq x'_S$ for all $x' \in [x]$. Fix an arbitrary $y_{N \setminus S} \in X_{N \setminus S}$. Note that $v_S(y) \geq v_S(x)$ and $y \notin [x]$ by the choice of y_S.

Suppose that there exists some $x' \in [x]$ such that $x' \succ y$. Then, there exist a sequence of strategy profiles $z^0, ..., z^m$ and a sequence of coalitions $S^1, ..., S^m$ such that $z^0 = y$, $z^m = x'$, and for all $h = 1, ..., m$, $z^{h-1} \rightarrow_{S^h} z^h$ and $v_{S^h}(x') > v_{S^h}(z^{h-1})$. By $v_S(y) \geq v_S(x) = v_S(x')$, $S \cap S^1 = \emptyset$.

Fix an arbitrary $h = 2, ..., m$. We claim that if $S \cap S^\ell = \emptyset$ for all $\ell = 1, ..., h-1$, then $S \cap S^h = \emptyset$. Since $S \cap S^\ell = \emptyset$ for all $\ell = 1, ..., h-1$, $z_S^{h-1} = y_S$. By the choice of y_S, $v_S(z^{h-1}) = v_S(y_S, z_{N \setminus S}^{h-1}) \geq v_S(x')$. By $v_{S^h}(x') > v_{S^h}(z^{h-1})$, $S \cap S^h = \emptyset$.

By this mathematical induction, $S \cap \left(\bigcup_{h=1}^m S^h \right) = \emptyset$. Thus, $z_S^0 = z_S^m$. This contradicts that $y_S \neq x'_S$. Hence, $\hat{x} \succ y$ does not hold for all $\hat{x} \in [x]$, and $[x]$ does not satisfy the external stability. ∎

From Theorem 1 and 2, we obtain the following corollary that completely characterizes the SP-FSS in strategic games with dominant punishment strategies.

Corollary 1 *Let G be a strategic game with dominant punishment strategies and $x \in X$. In G, $[x]$ is a SP-FSS if and only if $[x]$ is an inclusive set.*

Corollary 1 is a counterpart of Theorem 2 by Ray and Vohra (2015a), where the SP-FSS of the coalitional game defined by their indirect dominance relation is characterized by the separable payoff vectors. Since the inclusive set is closely related to the strict α-core, Corollary 1 is also related to the result by Chander (2015) who showed the equivalence between the SP-FSS and the strong-core in the partition function game.

In the proof of Ray and Vohra (2015a), singleton coalitions play important roles to construct the indirect dominance relation. Note that the construction of an indirect dominance relation in the proof of Theorem 1 in the present paper is similar with that in the proof of Theorem 2 in Ray and Vohra (2015a), though the inclusive set is different from the separable payoff vector. The dominant punishment strategies play the similar role with the singleton coalitions in their proof. Further, a similar manner of constructing the indirect dominance relation in a specific model is also found in the literature, for example Suzuki and Muto (2005).

Note that the dominant punishment strategies are not required in Theorem 2. On the other hand, the following example cited from Kawasaki (2015) demonstrates that the dominant punishment strategies are essential to obtain Theorem 1.

Example 1 (Kawasaki, 2015, Example 1) Let G be a strategic game where $N = \{1, 2, 3\}$ and $X_i = \{C, D\}$ for each $i = 1, 2, 3$. The payoffs are defined as follows, where player 1 chooses rows, player 2 chooses columns, and player 3 chooses matrices.

	C	D				C	D	
C	1,1,1	0,5,0	C		C	0,0,5	0,5,0	D
D	5,0,0	5,0,0			D	0,0,5	5,0,0	

Player 2 is the only player who does not have the dominant punishment strategy since $v_1(D, C, D) = 0 < 5 = v_1(D, D, D)$ and $v_3(D, C, D) = 5 > 0 = v_3(D, D, D)$, while D is the dominant punishment strategy of both player 1 and 3. It is easy to see that (C, C, C) is efficient. It is also easy to see that $v \in V_\alpha^G(\{i\})$ implies $v_i \leq 0$ for all $i = 1, 2, 3$ and $v \in V_\alpha^G(\{i, j\})$ implies either $v_i \leq 0$ or $v_j \leq 0$ for all $i, j = 1, 2, 3$ with $i \neq j$. Thus, $[(C, C, C)] = \{(C, C, C)\}$ is an inclusive set by Proposition 1(a). However, Kawasaki (2015) showed that $(C, C, C) \succ (D, C, C)$ does not hold, and $[(C, C, C)]$ is not a FSS. Hence, Theorem 1 does not hold even if there is only one player without the dominant punishment strategy.

The following corollary is obtained from Remark 1(b), Proposition 2, and Corollary 1.

Corollary 2 *Let G be a strategic game with dominant punishment strategies and Γ be the α-coalitional game associated with G. If there exists a separable payoff vector in Γ, then the SP-FSS exists in G.*

In other words, a SP-FSS in the sense of Chwe (1994) exists in a strategic game with dominant punishment strategies if a SP-FSS in the sense of Ray and Vohra (2015a) exists in the α-coalitional game associated with the strategic game. However, the SP-FSS may fail to exist even if the strict α-core is nonempty. To see this, the following result will be helpful, which strengthens Proposition 1(b) with the strictly dominant punishment strategies.

Proposition 3 *Let G be a strategic game with strictly dominant punishment strategies and $x \in X$. For each $i \in N$, let d_i denote the strictly dominant punishment strategy of i. If $[x]$ is a SP-FSS in G, then (i) $v(x)$ is in the strict α-core of G, and (ii) either $v(x) > v(d_N)$ or $v(x) = v(d_N)$.*

13

Proof. Let G be a strategic game with strictly dominant punishment strategies and $x \in X$. Let d_i denote the strictly dominant punishment strategy of each $i \in N$. Assume that $[x]$ is a SP-FSS in G. Then, $[x]$ is an inclusive set in G by Theorem 2. Proposition 1(b) guarantees (i). For (ii), suppose that $v(x) \neq v(d_N)$ and there exists some $i \in N$ such that $v_i(x) \leq v_i(d_N)$. Since $[x]$ is an inclusive set, there exists some $x' \in [x]$ such that $x'_i = d_i$. By $v(x') = v(x) \neq v(d_N)$, $x'_j \neq d_j$ for some $j \neq i$. Then, $v_i(x') > v_i(x'_{-j}, d_j) \geq v_i(d_N) \geq v_i(x) = v_i(x')$, a contradiction. Hence, x satisfies both (i) and (ii). $\blacksquare$

Proposition 3 says that the inclusive set does not exist if neither $v(x) > v(d_N)$ nor $v(x) = v(d_N)$ for any $x \in X$ such that $v(x)$ is in the strict α-core when each player has the strictly dominant punishment strategy. We can easily construct such an example as follows.

Example 2 Let G be a strategic game where $N = \{1, 2, 3\}$ and $X_i = \{C, D\}$ for each $i = 1, 2, 3$. The payoffs are defined as follows, where player 1 chooses rows, player 2 chooses columns, and player 3 chooses matrices.

	C	D			C	D	
C	2,2,2	1,3,1	C	C	1,1,5	0,2,4	D
D	3,1,1	2,2,0		D	2,0,4	1,1,3	

Note that D is the strictly dominant punishment strategy of each player. It is easy to see that (C, C, D) is the unique strategy profile that yields the payoff vector in the strict α-core because $v(C, C, D) \geq v(D, D, D)$ with $v_3(C, C, D) > v_3(D, D, D)$, and for any $x \in X \setminus \{(C, C, D), (D, D, D)\}$, $v_j(x) < v_j(D, D, D)$ for at least one $j \in N$. Thus, for any strategy profile $x \in X \setminus \{(C, C, D)\}$, $[x]$ is not an inclusive set by Proposition 1(b). However, (C, C, D) does not satisfy (ii) in Proposition 3 by $v_{\{1,2\}}(C, C, D) = v_{\{1,2\}}(D, D, D)$. Thus, $[(C, C, D)]$ is not an inclusive set. Hence, the inclusive set does not exist in this example.

The converse of Proposition 3 does not hold in general. We show it by another example.

Example 3 Let G be a strategic game such that $N = \{1, 2\}$, $X_i = [0, 1 - \varepsilon]$ for each $i = 1, 2$, where $\varepsilon > 0$ is a sufficiently small real number, and $v_i(x) = (1 - x_i)(x_1 + x_2)$ for each $i = 1, 2$. It is easy to see that $d_i = 0$ is the strictly dominant punishment strategy of

each $i = 1, 2$. Note that this is the most simple case of the voluntary contribution game of a public good with a slight modification to make 0 the strictly dominant punishment strategy.

By a straightforward calculation, it is easy to see that payoff vector $(1/4, 3/4) > (0, 0) = v(d_N)$ is in the strict α-core that is solely yielded by $x^* = (3/4, 1/4)$. Thus, x^* satisfies (i) and (ii) in Proposition 3 and $[x^*] = \{x^*\}$. However, player 1 can guarantee the payoff $1/4$ for herself by taking strategy $y_1 = 1/2$ even if player 2 takes 0. Thus, $[x^*]$ is not an inclusive set in G. By Theorem 2, $[x^*]$ is neither a SP-FSS in G.

On the other hand, we can obtain the converse of Proposition 3 in a special subclass of the strategic games with strictly dominant punishment strategy. We show this fact in the next section.

5   Binary games

We say a strategic game is binary when every player has two strategies. Such a game often describes a situation where the players are facing to decide whether they participate or not to an institution, an agreement, or a mechanism. A typical example is the cartel formation problem due to d'Aspremont, *et al.* (1983)[6]. In this situation, the players decide their strategies with foreseeing the outcome of the subsequent market stage that varies according to the formed cartel in the current cartel formation stage, as pointed out by Diamantoudi (2005). Therefore, the players are at least implicitly assumed farsighted, and it seems consistent that the players also foresee the reaction by the other players after changing their own decision in the cartel formation stage. In this sense, the binary game is a nice application to consider the FSS.

Formally, a strategic game $G = (N, (X_i)_{i \in N}, (v_i)_{i \in N})$ is a binary game if $X_i = \{C, D\}$ for all $i \in N$. Without loss of generality, we may assume that D is the (strictly) dominant punishment strategy of each $i \in N$ if a binary game G is a strategic game with (strictly) dominant punishment strategies. Throughout this section, denote $C(x) = \{i \in N | x_i = C\}$ for all $x \in X$; for all $S \in \mathcal{N}$, denote $C_S = x_S$ such that $x_i = C$ for all $i \in S$ and $D_S = x_S$ such that $x_i = D$ for all $i \in S$. The following result strengthens Proposition 3 for binary games.

---

[6]d'Aspremont, *et al.* (1983) formulated the problem in an abstract model. Later, Kamijo and Muto (2010) reformulated it to the strategic game and investigated the FSS in the sense of Chwe (1994) in it.

Proposition 4 *Let G be a binary game with strictly dominant punishment strategies and $x \in X$. In G, $[x]$ is a SP-FSS if and only if (i) x is efficient, and (ii) either $v(x) > v(D_N)$ or $v(x) = v(D_N)$.*

Proof. Let G be a binary game with strictly dominant punishment strategies and $x \in X$. The necessity follows from Proposition 3 and the efficiency of the strict α-core. For the sufficiency, it suffices to show that $[x]$ is an inclusive set if x satisfies (i) and (ii) by Theorem 1.

Assume that x satisfies (i) and (ii). Suppose that $[x]$ is not an inclusive set. Then, there exist some $S \in \mathcal{N}$ and $y_S \in X_S$ such that $v_S\left(y_S, D_{N \setminus S}\right) \geq v_S(x)$ and $y_S \neq x'_S$ for all $x' \in [x]$. By (i), $S \neq N$.

First, consider the case where $v(x) = v(D_N)$. Then, $D_N \in [x]$. By $y_S \neq D_S$, $v_{N \setminus S}\left(y_S, D_{N \setminus S}\right) > v_{N \setminus S}(D_N) = v_{N \setminus S}(x)$. This contradicts (i) by $v_S\left(y_S, D_{N \setminus S}\right) \geq v_S(x)$ and $S \neq N$.

Next, consider the case where $v(x) > v(D_N)$. Note that $C(x) \neq \emptyset$ in this case. By $v_S(y_S, D_{N \setminus S}) \geq v_S(x) > v_S(D_N)$, $y_i = C$ for some $i \in S$. Denote $T = C\left(y_S, D_{N \setminus S}\right) \neq \emptyset$. Note that $T \subset S$, and thus $v_T\left(y_S, D_{N \setminus S}\right) \geq v_T(x)$. If $T = C(x)$, then $y_S = x_S$, contradicting the choice of y_S. Thus, $T \neq C(x)$. If $T \subset C(x)$, then $v_T\left(y_S, D_{N \setminus S}\right) < v_T(x)$, contradicting $v_T(y_S, D_{N \setminus S}) \geq v_T(x)$. Therefore, assume that $T \setminus C(x) \neq \emptyset$. Then,

$$v_{C(x)}\left(C_{T \cup C(x)}, D_{N \setminus (T \cup C(x))}\right) > v_{C(x)}(x);$$
$$v_{N \setminus (T \cup C(x))}\left(C_{T \cup C(x)}, D_{N \setminus (T \cup C(x))}\right) > v_{N \setminus (T \cup C(x))}(x).$$

Further, by $T \setminus C(x) \subset T \subset T \cup C(x)$ and $v_T\left(y_S, D_{N \setminus S}\right) \geq v_T(x)$,

$$v_{T \setminus C(x)}\left(C_{T \cup C(x)}, D_{N \setminus (T \cup C(x))}\right) \geq v_{T \setminus C(x)}\left(y_S, D_{N \setminus S}\right) \geq v_{T \setminus C(x)}(x).$$

These three inequalities contradict (i) by $C(x) \neq \emptyset$. Hence, $[x]$ is an inclusive set in G. ∎

By the finiteness of X, we immediately obtain the following corollary from Proposition 4.

Corollary 3 *In a binary game with strictly dominant punishment strategies G, a SP-FSS exists if and only if D_N is efficient or not weakly efficient.*

Therefore, it suffices to check the (weak) efficiency of D_N for the existence of SP-FSS in a binary game with strictly dominant punishment strategies. Indeed, Example 2 appeared in Section 4 is an example such that D_N is weakly efficient, but not efficient. The following two examples show that the strictness of the dominant punishment strategies is essential for both the sufficiency and the necessity in Proposition 4.

Example 4 Let G be a strategic game where $N = \{1, 2, 3\}$ and $X_i = \{C, D\}$ for each $i = 1, 2, 3$. The payoffs are defined as follows, where player 1 chooses rows, player 2 chooses columns, and player 3 chooses matrices.

	C	D				C	D	
C	3,3,0	3,3,0	C		C	3,3,1	3,3,0	D
D	3,3,0	1,1,0			D	3,3,0	0,0,0	

Note that D is the dominant punishment strategy of each player, but none of which is strict. In this game, $v(C, C, D) = (3, 3, 1) > (0, 0, 0) = v(D, D, D)$ and (C, C, D) is efficient. Also, $[(C, C, D)] = \{(C, C, D)\}$. However $[(C, C, D)]$ is not an inclusive set in G because $\{1, 2\}$ can guarantee payoffs $(3, 3)$ for themselves by taking (C, D) and (D, C) even if player 3 takes D. Hence, the sufficiency of Proposition 4 may not hold without the strictly dominant punishment strategies.

Example 5 Let G be a strategic game where $N = \{1, 2, 3\}$ and $X_i = \{C, D\}$ for each $i = 1, 2, 3$. The payoffs are defined as follows, where player 1 chooses rows, player 2 chooses columns, and player 3 chooses matrices.

	C	D				C	D	
C	3,3,1	0,4,1	C		C	3,3,1	0,4,1	D
D	4,0,1	1,1,0			D	4,0,1	0,0,1	

Note that D is the dominant punishment strategy of each player, but none of which is strict. We show that $[(C, C, C)] = \{(C, C, C), (C, C, D)\}$ is an inclusive set. For each $i = 1, 2$ and $x_i \in \{C, D\}$, $v_i(x_i, D_{-i}) = 0$. For player 3, both C and D take a part of some strategy profile in $[(C, C, C)]$. For any $x_{\{1,2\}} \neq (C, C)$, $v_i(x_{\{1,2\}}, D_3) = 0$ for at least either one of $i = 1, 2$. For each $i = 1, 2$ and $x_{\{i,3\}} \in X_{\{i,3\}}$, $v_i(x_{\{i,3\}}, D_j) \leq 1$, where $j \neq i, 3$. Thus, $[(C, C, C)]$ is an inclusive set in G. However, $v_3(C, C, C) = v_3(D, D, D)$. Hence,

the necessity of Proposition 4 may not hold without the strictly dominant punishment strategies.

The n-player prisoners' dilemma (Okada, 1993) defined as follows is a strategic game with strictly dominant punishment strategies, in which all the FSS were specified by Suzuki and Muto (2005).

Let $f : \{C, D\} \times \{0, ..., n-1\} \to \mathbb{R}$ be a function such that (i) both $f(C, k)$ and $f(D, k)$ are increasing in k; (ii) $f(D, k) > f(C, k)$ for all $k = 0, ..., k-1$; (iii) $f(C, n-1) > f(D, 0)$. A binary game G is called the prisoners' dilemma iff

$$v_i(x) = \begin{cases} f(C, |C(x)| - 1) & \text{if } x_i = C; \\ f(D, |C(x)|) & \text{if } x_i = D. \end{cases}$$

Suzuki and Muto (2005) showed that for any $x \in X$, $[x] = \{x\}$ is a SP-FSS if and only if x is efficient and $v(x) > v(D_N)$. Proposition 4 in the present paper extends this result because D_N is never efficient in the prisoners' dilemma.

In the prisoners' dilemma, Suzuki and Muto (2005) showed that there exists at most one FSS with multiple payoffs, and it exists only for the degenerate case. On the other hand, the lumpy public good game appeared in section 3 has a quite different type of the FSS with multiple payoffs as shown by Kawasaki and Muto (2009). They showed that there exist an inefficient FSS except for the case where $r^* = n - 1$: any $x \in X$ with $r^* \leq |C(x)| < n$ is contained in at least one FSS, while x is efficient if and only if $|C(x)| = r^*$. This difference shows that we need to exploit the structures of the strategic games other than the dominant punishment strategies for characterizing the multiple-payoff FSS, even though the strategic game is binary.

Kamijo and Muto (2010) formulated the cartel formation problem by d'Aspremont, *et al.* (1983) in a binary strategic game, where each firm chooses to participate or not to participate to the cartel, and studied the FSS where coalitional behavior is allowed. As pointed out by d'Aspremont, *et al.* (1983), each player may not have the dominant punishment strategy in such a game. On the other hand, by Donsimoni, *et al.* (1986), we can confirm that not to participate becomes the strictly dominant punishment strategy of each firm in the strategic cartel formation game when the demand function and the marginal cost function are linear. Similarly, we can derive the binary game with strictly dominant punishment strategy of each player from the cartel formation problem with heterogeneous firms by Donsimoni (1985) when the demand and marginal cost functions

are linear. The heterogeneity of the players may make the problem more complex, but Proposition 4 says that we only need to check the efficiency and to compare the payoffs with $v(D_N)$ for each strategy profile.

6 Market games

This section considers primitive strategic games naturally associated with some market models. Some kinds of the stable sets with farsighted players have been well investigated in market models such as Greenberg, *et al.* (2002), Kawasaki (2010), Klaus, *et al.* (2010), and Mouleon, *et al.* (2011), among others. The FSS by Ray and Vohra (2015a) also applies to the market economy models since they are frequently formulated in coalitional games.

This section particularly considers two types of market models, the pure exchange economy and the (many-to-many) matching problem. The strategic games associated with these markets are shown to be those with dominant punishment strategies under certain conditions, and the results in Section 4 apply. In each strategic game, the "status quo" strategy becomes the dominant punishment strategies.

6.1 The exchange economy

This subsection considers the strategic game associated with the classical pure exchange economy (Scarf, 1971). Consider an exchange economy with m commodities. Let $N = \{1, ..., n\}$ denote the set of agents (players). The preference relation of each $i \in N$ is represented by a utility function $u_i : \mathbb{R}^m_+ \to \mathbb{R}$. Each $i \in N$ is endowed with $\omega_i \in \mathbb{R}^m_+ \backslash \{0\}$. Let $E = (N, (u_i)_{i \in N}, (\omega)_{i \in N})$ denote an exchange economy.

Throughout this subsection, we assume that u_i is continuous on $\mathbb{R}^m_+$; monotonically increasing, *i.e.* for any $x, x' \in \mathbb{R}^m_+$ with $x \geq x'$ and $x \neq x'$, $u_i(x) > u_i(x')$; strictly quasi-concave, *i.e.* for any $x, x' \in \mathbb{R}^m_+$ with $u_i(x) \geq u_i(x')$ and $x \neq x'$, and any $\theta \in \mathbb{R}$ with $0 < \theta < 1$, $u_i(\theta x + (1 - \theta)x') > u_i(x')$ for each $i \in N$.

In an exchange economy E, an n-tuple of consumption bundles $a = (a_1, ..., a_n) \in \mathbb{R}^{nm}_+$ is called an allocation, and a positive m-vector $p = (p_1, ..., p_m) \in \mathbb{R}^m_{++}$ is called a price. A pair of a price and an allocation $(p^*, a^*) \in \mathbb{R}^m_{++} \times \mathbb{R}^{nm}_+$ is said to be a Walras equilibrium iff $a_i^* \in \arg\max\{u_i(a_i) \mid a_i \geq 0, \ p^* \cdot a_i \leq p^* \cdot \omega_i\}$ for all $i \in N$ and $\sum_{i \in N} a_i^* = \sum_{i \in N} \omega_i$.

We will employ the following well-known facts, the proofs of which can be found in standard textbooks such as Mas-colell, *et al.* (1995).

Remark 3 A Walras equilibrium (p^*, a^*) exists and is Pareto efficient under our conditions, that is, there exists no allocation a such that $\sum_{i \in N} a_i \leq \sum_{i \in N} \omega_i$, $u_i(a_i^*) \geq u_i(a_i)$ for all $i \in N$, and $u_j(a_j^*) > u_j(a_j)$ for some $j \in N$.

We introduce a strategic exchange game $G^E = (N, (X_i)_{i \in N}, (u_i)_{i \in N})$ associated with E due to Scarf (1971). The set of players N is identical to the set of agents in E. For each $i \in N$, define

$$X_i = \left\{ x_i = (x_{i1}, ..., x_{in}) \in \mathbb{R}_+^{mn} \,\middle|\, \begin{array}{l} x_{ij} = (x_{ij}^1, ..., x_{ij}^m), \forall j \in N; \\ \sum_{j \in N} x_{ij}^h = \omega_i^h, \forall h = 1, ..., m. \end{array} \right\}.$$

For each $i, j \in N$ and $h = 1, ..., m$, x_{ij}^h denotes the amount of h-th commodity that i gives to j. To define the payoff function, we define an allocation function $a : X \to \mathbb{R}_+^{mn}$ such that $a(x) = (a_1(x), ..., a_n(x))$ and $a_i(x) = \omega_i - \sum_{j \in N} x_{ij} + \sum_{j \in N} x_{ji}$ for each $i \in N$. Then, the payoff function $v_i : X \to \mathbb{R}$ is defined by $v_i(x) = u_i(a_i(x))$ for all $i \in N$ and any $x \in X$.

The following fact follows from the monotonicity of the utility functions, which was also pointed out by Hirai, *et al.* (2004).

Remark 4 *For each $i \in N$, $d_i = (d_{i1}, ..., d_{in})$ such that*

$$d_{ij} = \begin{cases} \omega_i & \text{if } i = j; \\ 0 & \text{otherwise} \end{cases}$$

is the dominant punishment strategy in G^E.

Throughout this subsection, we denote d_i the dominant punishment strategy of each i described in Remark 4. The strategic exchange game is an example where the inclusive sets are equivalent to the separable payoffs (Ray and Vohra, 2015) as pointed out in section 3.

Proposition 5 *Let E be an exchange economy, G^E be a strategic exchange game associated with E, and Γ^E be an α-coalitional game associated with G^E.*

(a) *Let (p^*, a^*) be a Walras equilibrium in E and $x^* \in X$ be such that $a(x^*) = a^*$. Then, $[x^*]$ is an inclusive set, and thus, $[x^*]$ is a SP-FSS.*

(b) *Let $x \in X$. Then, $[x]$ is an inclusive set if and only if $v(x)$ is separable in Γ^E.*

Proof. Let E be an exchange economy, G^E be a strategic exchange game associated with E, and Γ^E be an α-coalitional game associated with G^E.

(a) Let (p^*, a^*) be a Walras equilibrium and $x^* \in X$ be such that $a(x^*) = a^*$. Assume that there exist some $S \in \mathcal{N}$ and $y_S \in X_S$ such that $v_S(y_S, d_{N \setminus S}) \geq v_S(x^*)$. If $S = N$, then $v(y) = v(x^*)$ follows from the efficiency of the Walras equilibrium, and $y \in [x^*]$. Thus, assume that $S \neq N$.

Denote $a' = a(y_S, d_{N \setminus S})$. Note that $\sum_{i \in S} a'_i \leq \sum_{i \in S} \omega_i$ since $d_{ji} = 0$ for all $j \in N \setminus S$ and all $i \in S$. We claim that the equality holds. Suppose that $\sum_{i \in S} a'_i \neq \sum_{i \in S} \omega_i$. Then, there exists some $\hat{y}_S \in X_S$ such that $a_i(\hat{y}_S, d_{N \setminus S}) \geq a'_i$ for all $i \in S$, and $a_j(\hat{y}_S, d_{N \setminus S}) \neq a'_j$ for some $j \in S$. Denote $\hat{a} = a(\hat{y}_S, d_{N \setminus S})$. Then, $u_i(\hat{a}_i) \geq u_i(a'_i) \geq u_i(a^*_i)$ for all $i \in S$, and $u_j(\hat{a}_j) > u_j(a'_j) \geq u_j(a^*_i)$ for some $j \in S$. Since $a^*_i \in \arg\max\{u_i(a_i)| a_i \geq 0, p^* \cdot a_i \leq p^* \cdot \omega_i\}$ for all $i \in N$, $p^* \cdot \hat{a}_i \geq p^* \cdot \omega_i$ for all $i \in S$, and $p^* \cdot \hat{a}_j > p^* \cdot \omega_j$ for some $j \in S$. Thus, $p^* \cdot \sum_{i \in S} \hat{a}_i > p^* \cdot \sum_{i \in S} \omega_i$. By $p^* > 0$, $\sum_{i \in S} \hat{a}_i \leq \sum_{i \in S} \omega_i$ does not hold. This contradicts that $d_{ji} = 0$ for all $j \in N \setminus S$ and $i \in S$. Thus, $\sum_{i \in S} a'_i = \sum_{i \in S} \omega_i$. Note that this implies that $y_{ij} = 0$ for all $i \in S$ and all $j \in N \setminus S$ since $a'_i = a_i(y_S, d_{N \setminus S})$ for all $i \in S$.

Suppose that $a'_j \neq a^*_j$ for some $j \in S$. By $u_i(a'_i) = v_i(y_S, d_{N \setminus S}) \geq v_i(x^*) = u_i(a^*_i)$ for all $i \in S$ and $a^*_i \in \arg\max\{u_i(a_i)| a_i \geq 0, p^* \cdot a_i \leq p^* \cdot \omega_i\}$ for all $i \in N$, $p^* \cdot a'_i \geq p^* \cdot \omega_i$ for all $i \in S$. By $a'_j \neq a^*_j$ for some $j \in S$ and the strict quasi-concavity of u_j, $p^* \cdot a'_j > p^* \cdot \omega_j$ for some $j \in S$. Thus, $p^* \cdot \sum_{i \in S} a'_i > p^* \cdot \sum_{i \in S} \omega_i$. This contradicts that $\sum_{i \in S} a'_i = \sum_{i \in S} \omega_i$. Thus, $a'_i = a^*_i$ for all $i \in S$.

Since $\sum_{i \in S} a'_i = \sum_{i \in S} \omega_i$ and $a'_i = a^*_i$ for all $i \in S$, $\sum_{i \in S} \omega_i = \sum_{i \in S} a'_i = \sum_{i \in S} a^*_i$ and

$$\sum_{i \in N \setminus S} a^*_i = \sum_{i \in N} a^*_i - \sum_{i \in S} a^*_i = \sum_{i \in N} \omega_i - \sum_{i \in S} \omega_i = \sum_{i \in N \setminus S} \omega_i.$$

Thus, there exists some $y_{N \setminus S} \in X_{N \setminus S}$ such that $a_i(d_S, y_{N \setminus S}) = a^*_i$ for all $i \in N \setminus S$. By $y_{ij} = d_{ij} = 0$ for all $i \in S$ and $j \in N \setminus S$, $a_i(y) = a_i(d_S, y_{N \setminus S}) = a^*_i$ for all $i \in N \setminus S$. Thus, $v_{N \setminus S}(y) = v_{N \setminus S}(x^*)$. Then, by $v_S(y) \geq v_S(y_S, d_{N \setminus S}) \geq v_S(x^*)$ and the efficiency of the Walras equilibrium, $v(y) = v(x^*)$. Hence, $y \in [x^*]$ and $[x^*]$ is an inclusive set. By Theorem 1, $[x^*]$ is a SP-FSS.

(b) The sufficiency has already been proved by Proposition 2. For the necessity, fix an arbitrary $x \in X$ such that $[x]$ is an inclusive set. Let $a = a(x)$. By Proposition 1(b)

and the efficiency of the strict α-core, $v(x)$ is efficient.

Assume that there exist some T and its partition $\mathcal{P}(T)$ such that $v(x) \in V_\alpha^E(S)$ for all $S \in \mathcal{P}(T)$. Then, for each $S \in \mathcal{P}(T)$, there exists some $y_S \in X_S$ such that $v_S(y_S, d_{N \setminus S}) \geq v_S(x)$. Let $\hat{a}^S = (y_S, d_{N \setminus S})$ for each $S \in \mathcal{P}(T)$.

Fix an arbitrary $Q \in \mathcal{P}(T)$. By Proposition 1(b) and the monotonicity of u_i, $v_Q(y_Q, d_{N \setminus Q}) = v_Q(x)$ and $\sum_{i \in Q} \hat{a}_i^Q = \sum_{i \in Q} \omega_i$. Thus, $y_{ij} = 0$ for all $i \in Q$ and $j \in N \setminus Q$. Since this holds for all $S \in \mathcal{P}(T)$, $y_{ij} = 0$ for all $i \in T$ and $j \in N \setminus T$.

Since $[x]$ is an inclusive set, there exists some $y_{N \setminus T} \in X_{N \setminus T}$ such that $v(y) = v(x)$. By $y_{ij} = 0$ for all $i \in T$ and $j \in N \setminus T$, $a_i(y) = a_i(d_T, y_{N \setminus T})$ for all $i \in N \setminus T$, and thus, $v_{N \setminus T}(d_T, y_{N \setminus T}) = v_{N \setminus T}(x)$. Hence, $v(x) \in V_\alpha^E(N \setminus T)$, and $v(x)$ is separable. $\blacksquare$

6.2   The many-to-many matching

This subsection considers a primitive strategic game with dominant punishment strategies associated with a many-to-many matching problem with substitutable preferences. The (SP-)FSS of the one-to-one and many-to-one matching problem with substitutable preferences was well-investigated by Mouleon, *et al.* (2011) who considered a reasonable definition of the indirect dominance relation. They showed that a singleton of a matching is a FSS in their sense if and only if the matching is a (strong) core matching. Ray and Vohra (2015a) as well as their Supplementary Note (Ray and Vohra, 2015b) argued the close relationship between their FSS and the FSS by Mouleon, *et al.* (2011). Further, Roketskiy (2012) considered a more general class of matching problems by dropping the assumption of the substitutable preferences and adding contracts, while he retained the feature of the indirect dominance relation of Mouleon, *et al.* (2011).

We investigate the SP-FSS in a primitive strategic game associated with a many-to-many matching problem. If the preference of each player satisfies the substitutability assumption, the associated strategic game turns out to be that with dominant punishment strategies. On the other hand, our strategic game excludes the matching that is not individually rational. By this feature, the SP-FSS in the sense of Chwe (1994) becomes equivalent to the individually rational core matching, rather than the core matching. Echenique and Oviedo (2006) showed that the individually rational core matching has nice properties under a stronger condition than the present paper. They showed that the individually rational core matching exists and can be found by a certain fixed-points

algorithm called the T-algorithm if the preferences of the players in one side satisfy the substitutability and the preferences of the players in the other side satisfy the strong substitutability.[7]

We introduce the many-to-many matching model by following Echenique and Oviedo (2006) with slight modifications. A (many-to-many) matching problem is defined by a tuple $M = (F, W, P)$, where $F = \{f_1, ..., f_r\}$ is the set of firms, $W = \{w_1, ..., w_s\}$ is the set of workers, and $P = (P(f_1), ..., P(f_r), P(w_1), ..., P(w_s))$ is the preference profile of $F \cup W$.

For each $f \in F$, $S \in 2^W$ is called a set of f's partners, and for each $w \in W$, $S \in 2^F$ is called a set of w's partners. The preference relation of each $i \in F \cup W$ is a strict and linear ordering on the set of all sets of i's partners. For each $i \in F \cup W$ and two sets of i's partners S, S', we denote $S \ P(i) \ S'$ when i prefers S to S', and $S \ R(i) \ S'$ when either $S \ P(i) \ S'$ or $S = S'$. For each $i \in F \cup W$ and a set of i's partners S, let $Ch_i(S)$ denote the most preferable subset of S for i that is called the choice set of S for i. Therefore, $Ch_i(S) \ R(i) \ S'$ for all $S' \subset S$. For each $i \in F \cup W$, the preference relation $P(i)$ is said to be substitutable iff $j \in Ch_i(S' \cup \{j\})$ implies $j \in Ch_i(S \cup \{j\})$ for any i's partner j and two sets of i's partners S, S' such that $S \subset S'$. A matching problem M is said to be a matching problem with substitutability iff the preference relation of any $i \in F \cup W$ is substitutable.

A (many-to-many) matching in M is a mapping $\mu : F \cup W \to 2^F \cup 2^W$, where for any $f \in F$ and $w \in W$, (i) $\mu(f) \in 2^W$; (ii) $\mu(w) \in 2^F$; (iii) $f \in \mu(w)$ if and only if $w \in \mu(f)$. In M, a matching μ is said to be individually rational iff $\mu(i) = Ch_i(\mu(i))$ for all $i \in N$. Given a matching μ, a block of μ is a tuple (F', W', μ') where $F' \subset F$, $W' \subset W$, and μ' is a matching such that (i) $F' \cup W' \neq \emptyset$; (ii) $\mu'(i) \subset F' \cup W'$ for all $i \in F' \cup W'$; (iii) $\mu'(i) \ R(i) \ \mu(i)$ for all $i \in F' \cup W'$; (iv) $\mu'(i) \ P(i) \ \mu(i)$ for some $i \in F' \cup W'$. A block (F', W', μ') of a matching μ is said to be individually rational iff $\mu'(i) = Ch_i(\mu'(i))$ for all $i \in F' \cup W'$. A matching μ is said to be a core matching iff there exists no block of μ.[8] A matching μ is said to be an individually rational core matching iff μ is individually rational and there exists no individually rational block of μ.

Now, we associate a primitive strategic game $G^M = (N, (X_i)_{i \in N}, (v_i)_{i \in N})$ with a

---

[7]We omit the definitions of the strong substitutability and the T-algorithm because we do not employ them in this paper. See Echenique and Oviedo (2006) for the detail.

[8]We follow the terminology of Echenique and Oviedo (2006), while the core matching should probably be called the *strong* core matching, which is more familiar.

matching problem with substitutability $M = (F, W, P)$. Let $N = F \cup W$. Let $X_f = \{S \in 2^W | Ch_f(S) = S\}$ for each $f \in F$, and $X_w = \{S \in 2^F | Ch_w(S) = S\}$. Note that $X_i \neq \emptyset$ for all $i \in F \cup W$ since $\emptyset \in X_i$ for all $i \in F \cup W$. A function ν from X to the set of matchings is said to be a matching rule iff for any $x \in X$,

$$\nu(x)(f) = \{w \in x_f | f \in x_w\} \text{ for each } f \in F;$$
$$\nu(x)(w) = \{f \in x_w | w \in x_f\} \text{ for each } w \in W.$$

The payoff function of $i \in N$ is defined consistently with $P(i)$ as follows: $v_i(x) > v_i(x')$ if and only if $\nu(x)(i) P(i) \nu(x')(i)$ for all $i \in N$ and all $x, x' \in X$.[9]

We can easily confirm that $\nu(x)$ is a matching for any $x \in X$ as the following way. Let $x \in X$, $f \in F$, and $w \in W$. It is obvious that (i) and (ii) of the definition of the matching is satisfied in $\nu(x)$. For (iii), if $f \in \nu(x)(w)$, then $w \in x_f$ and $f \in x_w$. Thus, $w \in \nu(x)(f)$, and *vice versa*. Further, the following lemma points out that our strategic game completely captures the individually rational matchings.

Lemma 1 *Let M be a matching problem with substitutability and G^M be the strategic game associated with M. A matching μ is individually rational in M if and only if $\nu(x) = \mu$ for some $x \in X$ in G^M.*

Proof. Let M be a matching problem with substitutability and G^M be the strategic game associated with M. We first show the necessity. Let μ be an individually rational matching in M. By the individual rationality, $Ch_i(\mu(i)) = \mu(i)$ for all $i \in N$. Thus, $\mu(i) \in X_i$ for all $i \in N$. Define $x_i = \mu(i)$ for all $i \in N$. Then, $\nu(x)(i) = \mu(i)$ for all $i \in N$ by the definition of ν.

Next, we show the sufficiency. Let μ be a matching such that $\mu = \nu(x)$ for some $x \in X$. Fix an arbitrary $i \in N$. We assume that $i \in F$. The case where $i \in W$ is omitted because it can be proved in the same way as the following proof. It suffices to show that $\nu(x)(i) \subset Ch_f(\nu(x)(i))$. If $\nu(x)(i) = \emptyset$, then the proof is done. Therefore, assume that $\nu(x)(i) \neq \emptyset$.

Fix an arbitrary $j \in \nu(x)(i)$. Note that $j \in \nu(x)(i) \subset x_i$, and thus, $\nu(x)(i) \setminus \{j\} \subset x_i \setminus \{j\}$. By the definition of X_i, $j \in x_i = Ch_i(x_i) = Ch_i((x_i \setminus \{j\}) \cup \{j\})$. By the

[9]Our strategic game is very similar with that by Konishi and Ünver (1999). The only difference is that we restrict the strategy sets to the individually rational partners.

substitutability of the preference relation and $\nu(x)(i) \setminus \{j\} \subset x_i \setminus \{j\}$, $j \in Ch_i((\nu(x)(i) \setminus \{j\}) \cup \{j\}) = Ch_i(\nu(x)(i))$. Hence, $\nu(x)(i) \subset Ch_i(\nu(x)(i))$. ∎

Now, we state and prove the result of this subsection.

Proposition 6 *Let M be a matching problem with substitutability and G^M be the strategic game associated with M.*

(a) *In G^M, $d_i = \emptyset$ is the dominant punishment strategy of i for each $i \in N$.*

(b) *For any $x \in X$, $[x]$ is a SP-FSS in G^M if and only if $\nu(x)$ is the individually rational core matching in M.*

Proof. Let M be a matching problem with substitutability and G^M be a strategic game associated with M.

(a) Fix an arbitrary $i \in N$, an arbitrary $x_i \in X_i \setminus \{\emptyset\}$, and an arbitrary $z_{-i} \in X_{-i}$. We assume $i \in F$. The case where $i \in W$ is omitted because it can be proved in the same way as the following proof. For each $f \in F \setminus \{i\}$, $\nu(x_i, z_{-i})(f) = \nu(\emptyset, z_{-i})(f)$, and thus, $v_f(x_i, z_{-i}) = v_f(\emptyset, z_{-i})$.

Fix an arbitrary $w \in W$. First, assume that either $i \notin z_w$ or $w \notin x_i$. Then, $i \notin \nu(x_i, z_{-i})(w)$. Thus, $\nu(x_i, z_{-i})(w) = \nu(\emptyset, z_{-i})(w)$ and $v_w(x_i, z_{-i}) = v_w(\emptyset, z_{-i})$. Next, assume that both $i \in z_w$ and $w \in x_i$. Then, $i \in \nu(x_i, z_{-i})(w)$ and $\nu(\emptyset, z_{-i})(w) = \nu(x_i, z_{-i})(w) \setminus \{i\}$. By Lemma 1, $Ch_w(\nu(x_i, z_{-i})(w)) = \nu(x_i, z_{-i})(w)$. It follows that $\nu(x_i, z_{-i})(w)$ $P(w)$ $\nu(x_i, z_{-i})(w) \setminus \{i\} = \nu(\emptyset, z_{-i})(w)$. Thus, $v_w(x_i, z_{-i}) > v_w(\emptyset, z_{-i})$. Hence, $\emptyset$ is the dominant punishment strategy of i.

(b) Let $x \in X$. First, assume that $[x]$ is a SP-FSS. Then, $[x]$ is an inclusive set in G^M by Theorem 2. Suppose that $\nu(x)$ is not an individually rational core matching in M. By Lemma 1, $\nu(x)$ is individually rational. Then, there exists an individually rational block (F', W', μ') of μ. Thus, (i) $F' \cup W' \neq \emptyset$, (ii) $\mu'(i) \subset F' \cup W'$ for all $i \in F' \cup W'$, (iii) $\mu'(i)$ $R(i)$ $\mu(i)$ for all $i \in F' \cup W'$, (iv) $\mu'(i)$ $P(i)$ $\mu(i)$ for some $i \in F' \cup W'$, and (v) $Ch_i(\mu'(i)) = \mu'(i)$ for all $F' \cup W'$. By (v), $\mu'(i) \in X_i$ for all $i \in F' \cup W'$. Define $x'_i = \mu'(i)$ for all $i \in F' \cup W'$. Then, $\nu(x'_{F' \cup W'}, \emptyset_{N \setminus (F' \cup W')})(i) = \mu'(i)$ for all $i \in F' \cup W'$ by (ii). This contradicts Proposition 1(b) by (iii) and (iv). Hence, $\nu(x)$ is an individually rational core matching in M.

Next, assume that $\nu(x)$ is an individually rational core matching in M. By Theorem 1 and (a) of this proposition, it suffices to show that $[x]$ is an inclusive set.

Suppose that there exist some $S \in \mathcal{N}$ and x_S^* such that $v_S(x_S^*, \emptyset_{N \setminus S}) \geq v_S(x)$. Note that $Ch_i(\nu(x_S^*, \emptyset_{N \setminus S})(i)) = \nu(x_S^*, \emptyset_{N \setminus S})(i)$ for all $i \in S$ by Lemma 1. Note also that $\nu(x_S^*, \emptyset_{N \setminus S})(i) \subset S$ for all $i \in S$ by the definition of ν. Since $\nu(x)$ is an individually rational core matching and the preferences are strict, $\nu(x_S^*, \emptyset_{N \setminus S})(i) = \nu(x)(i)$ for all $i \in S$. Then, $\nu(x)(j) \subset N \setminus S$ for all $j \in N \setminus S$ by $\nu(x)(i) = \nu(x_S^*, \emptyset_{N \setminus S})(i) \subset S$ for all $i \in S$.

Since $\nu(x)$ is an individually rational core matching, $\nu(x)(j) \in X_j$ for all $j \in N \setminus S$. Denote $x_j^* = \nu(x)(j)$ for each $j \in N \setminus S$. For all $j \in N \setminus S$, $\nu(\emptyset_S, x_{N \setminus S}^*)(j) = \nu(x)(j)$ by $\nu(x)(j) \subset N \setminus S$. Since $i \notin x_j^*$ for all $i \in S$ and $j \in N \setminus S$, $\nu(x^*)(i) = \nu(x_S^*, \emptyset_{N \setminus S})(i) = \nu(x)(i)$ for all $i \in S$ and $\nu(x^*)(j) = \nu(\emptyset_S, x_{N \setminus S}^*)(j) = \nu(x)(j)$ for all $j \in N \setminus S$. Thus, $x^* \in [x]$, and $[x]$ is an inclusive set in G^M. ∎

Proposition 6(a) heavily relies on the substitutability of the preferences. Consider the following simple example. Let $F = \{f_1\}$ and $W = \{w_1, w_2\}$. The preference relation of f_1 is $\{w_1, w_2\} \ P(f_1) \ \{w_1\} \ P(f_1) \ \emptyset \ P(f_1) \ \{w_2\}$, which is not substitutable, and that of w_i is $\{f_1\} \ P(w_i) \ \emptyset$ for each $i = 1, 2$. Then, neither $\emptyset$ nor $\{f_1\}$ is the dominant punishment strategy of w_2 since $v_{f_1}(\{w_1, w_2\}, \emptyset, \emptyset) > v_{f_1}(\{w_1, w_2\}, \emptyset, \{f_1\})$ and $v_{f_1}(\{w_1, w_2\}, \{f_1\}, \{f_1\}) > v_{f_1}(\{w_1, w_2\}, \{f_1\}, \emptyset)$. Therefore, our result cannot be extended to the model by Roketskiy (2012) who allowed general preferences by dropping the substitutability condition.

7   Concluding remarks

This paper characterized the SP-FSS in the strategic games with dominant punishment strategies by the notion of the inclusiveness. The union of the inclusive sets has a very close connection to the strict α-core. A further characterization is given to the SP-FSS in the binary game with strictly dominant punishment strategies. Some applications are also investigated. Our results may simplify the investigation on the SP-FSS for a variety of strategic games as we mentioned in section 1. We conclude with a remark.

This paper did not consider the FSS whose elements yield multiple payoffs. In coalitional and partition function games, Béal, *et al.* (2008), Ray and Vohra (2015a), and Chander (2015) showed that there is no FSS including multiple payoffs at least under a certain condition. On the other hand, a FSS yielding multiple payoffs may exist in a number of strategic games. See for example, Suzuki and Muto (2005) and

Kawasaki and Muto (2009). Although we did not consider such a FSS, investigating it is also important since it may have novel implications. For example, Kawasaki and Muto (2009) showed that the multiple-payoff FSS supports the over-contribution to a public good in a voluntary contributing situation as we reviewed in Section 3, which has been rarely discussed in the literature so far. As we pointed out in Section 5, investigating the FSS with multiple payoffs should exploit the structure of the specific game. The results in the present paper allow us to concentrate for investigating such a FSS for various strategic games. We remain this problem for future research.

References

Aumann, R.J. and Peleg, B. (1960) "Von Neumann-Morgenstern solutions to cooperative games without side payments," *Bulletin of American Mathematical Society* **66**, 173-179.

Béal, S., Durieu, J., and Solal, P. (2008) "Farsighted coalitional stability in TU-games," *Mathematical Social Sciences* **56**, 303-313.

Benchekroun, H. and Chaudhuri, A.R. (2015) "Cleaner technologies and the stability of international environmental agreements," *Journal of Public Economic Theory* **17**, 887-915.

Bhattacharya, A. and Brosi, B. (2011) "An Existence result for farsighted stable sets of games in characteristic function form," *International Journal of Game Theory* **40**, 393-401.

Chander, P. (2015) "A farsighted stable set for partition function games," *Preliminary Draft* http://www.parkashchander.com/pdf/stability.pdf

Chwe, M. S.-Y. (1994) "Farsighted coalitional stability," *Journal of Economic Theory* **63**, 299-325.

d'Aspremont, C., Jacquemin, A., Gabszewicz, J.J., and Weymark, J.A. (1983) "On the stability of collusive price leadership," *Canadian Journal of Economics* **16**, 17-25.

Diamantoudi, E. (2005) "Stable cartels revisited," *Economic Theory*, **26**, 907-921.

Diamantoudi, E. and Sartzetakis, E.S. (2015) "International environmental agreements: Coordinated action under foresight," *Economic Theory* **59**, 527-546.

Diamantoudi, E. and Xue, L. (2003) "Farsighted stability in hedonic games," *Social Choice and Welfare* **21**, 39-61.

Donsimoni, M.-P. (1985) "Stable heterogeneous cartels," *International Journal of Industrial Organization*, **3**, 451-467.

Donsimoni, M.-P., Economides, N.S., and Polemarchakis, H.M. (1986) "Stable cartels," *International Economic Review* **27**, 317-327.

Dutta, B. and Vohra, R. (2016) "Rational expectations and farsighted stability," *Theoretical Economics*, forthcoming.

Echenique, F. and Oviedo, J. (2006) "A theory of stability in many-to-many matching markets," *Theoretical Economics* **1**, 233-273.

Greenberg, J. (1990) "*The theory of social situations: An alternative game-theoretic approach*," Cambridge University Press.

Greenberg, J., Luo, X., Oladi, R. and Shitovitz, B. (2002) "(Sophisticated) stable sets in exchange economies" *Games and Economic Behavior* **39**, 54-70.

Harsanyi, J. (1974) "An equilibrium-point interpretation of stable sets and a proposed alternative definition," *Management Science* **20**, 1472-1495.

Hirai, T., Masuzawa, T., and Nakayama, M. (2004) "Credible deviations and retaliations in a class of strategic games," *KUMQRP Discussion Paper Series* **DP2004-012**.

Kamijo, Y. and Muto, S. (2010) "Farsighted coalitional stability of a price leardership cartel," *Japanese Economic Review* **61**, 455-465.

Kawasaki, R. (2010) "Farsighted stability of the competitive allocations in an exchange economy with indivisible goods," *Mathematical Social Sciences* **59**, 46-52.

Kawasaki, R. (2015) "Maximin, minimax, and von Neumann-Morgenstern farsighted stable sets," *Mathematical Social Sciences* **74**, 8-12.

Kawasaki, R. and Muto, S. (2009) "Farsighted stability in provision of perfectly "lumpy" public goods," *Mathematical Social Sciences* **58**, 98-109.

Kawasaki, R., Sato, T., and Muto, S. (2015) "Farsightedly stable tariffs," *Mathematical Social Sciences* **76**, 118-124.

Klaus, B., Klijn, F., and Walzl, M. (2010) "Farsighted house allocation," *Journal of Mathematical Economics* **47**, 84-98.

Klaus, B., Klijn, F., and Walzl, M. (2011) "Farsighted stability for roommate markets," *Journal of Public Economic Theory* **13**, 921-931.

Konishi, H. and Ünver, M.U. (1999) "Credible group-stability in many-to-many matching problems," *Journal of Economic Theory* **129**, 57-80.

Mas-colell, A., Whinston, M.D., and Green, J.R. (1995) "*Microeconomic Theory*," Oxford University Press.

Masuda, T. (2002) "Farsighted stability in average return games," *Mathematical Social Sciences* **44**, 169-181.

Masuzawa, T. (2003) "Punishment strategies make the α-coalitional game ordinally convex and balanced," *International Journal of Game Theory* **32**, 479-483.

Milgrom, P. and Roberts, J. (1990) "Rationalizability, learning and equilibrium in games with strategic complementarities," *Econometrica* **58**, 1255-1278.

Milgrom, P. and Roberts, J. (1996) "Coalition-proofness and correlation with arbitrary communication possibilities," *Games and Economic Behavior* **17**, 113-128.

Mouleon, A., Vannetelbosch, V.J., and Vergote, W. (2011) "von Neumann-Morgenstern farsightedly stable sets in two-sided matching," *Theoretical Economics* **6**, 499-521.

Nakanishi, N. (2009) "Noncooperative farsighted stable set in an n-player prisoners' dilemma," *International Journal of Game Theory* **38**, 249-261.

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

Okada, A. (1993) "The possibility of cooperation in an n-person prisoners' dilemma with institutional arrangements," *Public Choice* **77**, 629-656.

Page, F.H. and Wooders, M. (2009) "Strategic basins of attraction, the path dominance core, and network formation games," *Games and Economic Behavior* **66**, 462-287.

Ray, D. and Vohra, R. (2015a) "The farsighted stable set," *Econometrica* **83**, 977-1011.

Ray, D. and Vohra, R. (2015b) "Supplement to 'The farsighted stable set' (Econometrica, Vol.83, No.3, May 2015, 977-1011)."
http://www.econ.nyu.edu/user/debraj/Papers/RayVohraFarSightedSuppl.pdf

Roketskiy, N. (2012) "Farsightedly stable matchings," *NET Institute Working Paper* **No. 12-26**.

Shino, J. and Kawasaki, R. (2012) "Farsighted stable sets in Hotelling's location games," *Mathematical Social Sciences* **63**,23-30.

Scarf, H.E. (1971) "On the existence of a cooperative solution for a general class of N-person games," *Journal of Economic Theory* **3**, 169-181.

Suzuki, A. and Muto, S. (2005) "Farsighted stability in an n-person prisoner's dilemma," *International Journal of Game Theory*, 431-445.

Taylor, M. "*The Possibility of Cooperation*," Cambridge University Press.

von Neumann, J., Morgenstern, O. (1944) "*Theory of Games and Economic Behavior*," Princeton University Press.