

Working Paper No. 279

**The Stable Set of the Social Conflict Game
with Delegations: Existence, Uniqueness, and
Efficiency**

Toshiyuki Hirai

March 5, 2013



FACULTY OF ECONOMICS
UNIVERSITY OF TOYAMA

The Stable Set of the Social Conflict Game with Delegations: Existence, Uniqueness, and Efficiency*

Toshiyuki Hirai[†]

March 5, 2013

Abstract

We investigate the stable sets of social conflict games by employing the framework of the (abstract) system by Greenberg (*Theory of Social Situations: An Alternative Game theoretic Approach*, Cambridge University Press, 1990). The social conflict game includes the prisoners' dilemma and the chicken game. We show that the stable set may fail to exist in the system directly derived from the social conflict game. The stable set exists if and only if the strong equilibrium exists in the underlying game. We consider another system where an agent is prepared and each player is allowed to delegate his decision to the agent. Then, the stable set always exists and consists of Pareto efficient outcomes with a certain property. We also discuss the relationship between the strong equilibrium and the stable set for the model with delegations.

Keywords: social conflict game, (abstract) system, theory of social situations, stable set, delegation

JEL Classification: C71, D74

*The author is grateful for the financial support from the JSPS Grant-in-Aid for Young Scientists (B) 22730155 and the JSPS Grant-in-Aid for Scientific Research (B) 24310110.

[†]Faculty of Economics, University of Toyama. thirai@eco.u-toyama.ac.jp

1 Introduction

Game theory well treats situations where individual interests conflict with the social interest completely (the prisoners' dilemma) or partially (the chicken game). In such a game, each player has two strategies, cooperation and defection, and there is usually an insufficient number of cooperating players for the efficiency at the Nash equilibrium. We will refer to the games describing such conflicting situations as social conflict games. Examples include the public good provision game with binary choices (Taylor and Ward, 1982) and the game of voluntary participation in a public good provision mechanism (Saijo and Yamato, 1999).

In line with the usual inefficiency of the Nash equilibrium, the Nash equilibrium is usually unstable against coalitional deviations; the grand coalition can deviate from a Nash equilibrium unless the Nash equilibrium is efficient. Indeed, a strong equilibrium (Aumann, 1959) exists if and only if the Nash equilibrium is efficient in the games with social conflicts. The coalition-proof equilibrium (Bernheim *et al.*, 1989) is a relaxed concept that requires the deviating coalition to be credible. The coalition-proof equilibria in the voluntary participation game was characterized by Shinohara (2010). However, the inefficiency of the equilibrium remains unchanged.

We consider the stable set for a strategic game by applying the theory of social situations by Greenberg (1990). Specifically, we employ the (abstract) system and its stable set. The relationship between the set of the strong equilibrium and the stable set is just like that between the core and the stable set for the coalitional game. In the theory of social situations, the stable set is regarded as a recommendation of an acceptable course of action. However, we will show that the stable set may not exist. Therefore, no recommendation is acceptable for the players in such a situation.

Several studies have approached the stable set for the social conflict game by applying the theory of social situations, especially in the prisoners' dilemma. Nakanishi (2001) considered the stable set when only individual deviations are allowed. He showed that the stable set exists and contains at least one efficient strategy profile. Suzuki and Muto (2005) considered the farsighted stable set by Chwe (1994), allowing coalitional deviations. They showed that any singleton consisting of an efficient and individually rational strategy profile is always a farsighted stable set. Further, Nakanishi (2009) considered the farsighted stable sets where only individual deviations are allowed. These studies imposed some requirements on the ability of the players. Nakanishi (2001, 2009) restricted the players from forming a coalition, and Suzuki and Muto (2005) and Nakanishi (2009) required the players to be fully farsighted.

This paper takes another approach to the stable set for social conflict games by exogenously giving the players the option to delegate. Specifically, we consider a model with a (non-player) agent, to whom each player can choose to delegate his decision or not, such as an international or intranational arbitration organization, a national or local government, or a lawyer. Each player has the option to delegate his decision to the agent or not, and the agent, who is common to all the players, acts on behalf of all the delegating players.

The idea of delegation is traced back to Schelling (1960): a player may be benefited by delegating his decision to the agent (who may or may not be a player), who has a different interest from the delegating player. In this paper, each player is, of course, interested in his own payoff only, while the agent is uniformly interested in the payoffs of all the delegating players. Then, delegation plays the role of a commitment device in the sense that delegating

players must give up their decision to change their strategies selfishly. The role of the agent here differs from that in Okada (1993), who also considered the prisoners' dilemma model with an agent. In his model, the agent plays as a punishment device to force the players to abide by the agreement in an endogenously arranged institution.

To make the delegation an effective commitment device, we will assume that each player cannot withdraw his delegation once he has delegated his decision. However, the players are completely free to decide whether or not to delegate; they are never forced. Therefore, the stable set recommends an acceptable course of action for the players that includes whether they should delegate their decision or not.

This paper shows that the stable set uniquely exists in the model with delegations and consists of efficient outcomes that include some outcomes where no player delegates his decision. In this sense, we can recommend an acceptable course of action in a social conflict by arranging an agent exogenously. This contrasts with the model without delegations where the stable set does not exist iff the Nash equilibria are not strong equilibria. On the other hand, some outcomes with full cooperation, (these are sometimes regarded as "best" outcomes) are usually excluded from the stable set. Therefore, when players are not forced to delegate their decision, full cooperation may not be acceptable to the players. We also investigate the relationship between the strong equilibrium and the stable set in the model with delegations.

The rest of the paper is organized as follows. In Section 2, we briefly introduce the (abstract) system and the stable set. In Section 3, we define the social conflict game. We derive a system from the social conflict game and investigate the stable set in Section 4. In Section 5, we derive another system by incorporating delegations. We also state and prove the unique existence

and the characterization of the stable set for this system. In Section 6, we conclude with some remarks.

2 The (abstract) system and the stable set

This section briefly introduces the (abstract) system and the stable set for the system by Greenberg (1990).

A system is a pair (A, \succ) , where A is a nonempty set and \succ is a binary relation on A . The set A is called an outcome set, and the binary relation \succ is called the dominance relation. For each $x, y \in A$, $x \succ y$ means that x dominates y .

The stable set for the system is defined as follows, which is a direct application of the original one by von Neumann and Morgenstern (1944).

Definition 1 *Let (A, \succ) be a system. A subset K of A is a stable set for (A, \succ) iff K satisfies the following two stabilities.*

Internal stability: *For any $x, y \in K$, it is not true that $x \succ y$.*

External stability: *For any $x \in A \setminus K$, $y \succ x$ for some $y \in K$.*

3 The social conflict game

We consider a strategic game $\Gamma = (N, (X_i)_{i \in N}, (u_i)_{i \in N})$, where $N = \{1, \dots, n\}$ is the set of players, X_i is the set of strategies of $i \in N$, and u_i is the payoff function of $i \in N$. Let $\mathcal{N} = \{S \subset N \mid S \neq \emptyset\}$ be the set of coalitions. We denote $X_S = \times_{i \in S} X_i$ for all $S \in \mathcal{N}$ and $X = X_N$ for simplicity. A strategy profile $x \in X$ is (weakly Pareto) efficient iff there exists no $y \in X$ such that $u_i(y) > u_i(x)$ for all $i \in N$.

We say Γ is a (symmetric) social conflict game if $X_i = \{C, D\}$ for all $i \in N$ and there exists some $f : \{C, D\} \times \{0, \dots, n-1\} \rightarrow \mathbb{R}$ such that

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

for all $i \in N$ and $x \in X$, where $C(x) = \{i \in N \mid x_i = C\}$ and f satisfies Assumption 1 below.

- Assumption 1** (a) $f(C, h)$ and $f(D, h)$ are increasing in $h = 0, \dots, n-1$;
- (b) There exists some $\hat{h} = 0, \dots, n-1$ such that $f(C, h') < f(D, h')$ for all $h' \geq \hat{h}$ and $f(C, h'') > f(D, h'')$ for all $h'' < \hat{h}$;
- (c) $f(C, n-1) > f(D, 0)$.
- (d) for all $(a, h), (a', h') \in \{C, D\} \times \{0, \dots, n-1\}$, $(a, h) \neq (a', h')$ implies $f(a, h) \neq f(a', h')$.

In Assumption 1, (a) requires cooperation to benefit other players, (b) assumes the existence of a threshold, (c) assures the efficiency of full cooperation, and (d) is for simplicity. Note that (b) excludes the case where C is the dominant strategy for all players, which is less interesting.

In addition to \hat{h} , we define another threshold. Let $h^* = 0, \dots, n-1$ be the minimum integer such that $f(D, h^*) > f(C, n-1)$. Note that the existence and positivity of h^* are assured by Assumption 1(b) and (c). Note also that $\hat{h} \leq h^*$; otherwise, $f(C, h^*) > f(D, h^*) > f(C, n-1) > f(C, h^*)$, a contradiction.

We introduce some notations. Given $x \in X$, $j \in N$, and $S \in \mathcal{N}$, we denote $x_S = (x_i)_{i \in S}$, $x_{-S} = x_{N \setminus S}$, and $x_{-j} = (x_i)_{i \neq j}$. Given $S \in \mathcal{N}$, C_S denotes x_S such that $x_i = C$ for all $i \in S$, and D_S denotes x_S such that $x_i = D$ for all $i \in S$.

4 The social conflict system

Let $\Gamma = (N, (X_i)_{i \in N}, (u_i)_{i \in N})$ be a social conflict game. The outcome set is defined by X , the set of strategy profiles. Given $x, y \in X$ 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 dominates x via S , denoted by $y \succ_S x$, iff $x \rightarrow_S y$ and $u_i(y) > u_i(x)$ for all $i \in S$. Given $x, y \in X$, we simply say y is inducible from x , denoted by $x \rightarrow y$, iff $x \rightarrow_S y$ for some $S \in \mathcal{N}$; we simply say y dominates x , denoted by $y \succ x$, iff $y \succ_S x$ for some $S \in \mathcal{N}$. We say (X, \succ) is a social conflict system (derived from Γ).

The Nash and strong equilibria were originally defined in strategic games. These concepts can be applied to the system (X, \succ) : $x \in X$ is a Nash equilibrium iff there exists no $y \in X$ and $i \in N$ such that $y \succ_{\{i\}} x$; $x \in X$ is a strong equilibrium iff there exists no $y \in X$ and $S \in \mathcal{N}$ such that $y \succ_S x$.

The social conflict system can be regarded as an open negotiation in the social conflict game allowing coalitional deviations. Indeed, it is essentially the same as the coalitional contingent threat situation in the theory of social situations (Greenberg, 1990, Chapter 7). Once an outcome (strategy profile) is proposed, a coalition may declare that the members will change their strategies if the players outside the coalition will stick to the proposed strategies. At the revised outcome, some coalition may declare its intention to change strategies if the players outside the coalition stick to the current strategies, and so on.

For later analysis, we confirm the following fundamental properties.

Lemma 1 (a) $x \in X$ is a Nash equilibrium if and only if $|C(x)| = \hat{h}$.

(b) A Nash equilibrium is a strong Nash equilibrium if and only if $\hat{h} = h^*$.

(c) $x \in X$ is efficient if and only if $|C(x)| \geq h^*$.

Proof. (a) Straightforward.

(b) Assume that $\hat{h} = h^*$. Fix an arbitrary Nash equilibrium x . By (a), $|C(x)| = \hat{h} = h^*$. Fix arbitrary $T \in \mathcal{N}$ and $y_T \in X_T$ with $y_i \neq x_i$ for all $i \in T$. First, assume that $y_i = D$ for all $i \in T$. Since x is a Nash equilibrium and by Assumption 1(a), $u_i(x) > u_i(D, x_{-i}) \geq u_i(y_T, x_{-T})$ for all $i \in T$. Thus, (y_T, x_{-T}) is false. Assume next that there exists some $j \in T$ with $y_j = C$. Note that $x_j = D$. Since x is a Nash equilibrium and $\hat{h} = h^*$,

$$u_j(x) = f(D, \hat{h}) > f(C, n-1) \geq f(C, |C(y_T, x_{-T})| - 1) = u_j(y_T, x_{-T}).$$

Thus, (y_T, x_{-T}) is again false. Hence x is a strong equilibrium.

Assume that $\hat{h} < h^*$. Let $z \in X$ be a Nash equilibrium. By Lemma 1(a), $|C(z)| = \hat{h}$. By $\hat{h} < h^*$ and Assumption 1(a), $u_i(C_N) > u_i(z)$ for all $i \in N$. Hence z is not a strong equilibrium.

(c) Let $x \in X$. First, assume that x is efficient. If $|C(x)| < h^*$, then $u_i(C_N) > u_i(x)$ for all $i \in N$ since $f(C, |C(x)| - 1) < f(C, n-1)$ by Assumption 1(a) and $f(D, |C(x)|) < f(C, n-1)$ by $|C(x)| < h^*$, contradicting the efficiency of x . Hence $f(D, |C(x)|) > f(C, n-1)$.

Next, assume that $|C(x)| \geq h^*$. Note that $|C(x)| \geq h^* \geq \hat{h}$. Fix an arbitrary $y \in X$. Assume that there exists some $j \in C(y) \setminus C(x)$. Then, $u_j(x) = f(D, |C(x)|) > f(C, n-1) \geq f(C, |C(y)| - 1) = u_j(y)$. Thus, $y \succ_N x$ is false. Therefore, assume $C(y) \subset C(x)$. Then, $u_i(y) = f(C, |C(y)| - 1) \leq f(C, |C(x)| - 1) = u_i(x)$ for all $i \in C(y)$. Thus, $y \succ_N x$ is again false. Hence, x is efficient. ■

In the social conflict system, the stable set fails to exist in general. Formally, the stable set exists if and only if a strong equilibrium exists.

Proposition 1 *In the social conflict system (X, \succ) , the stable set exists if and only if $\hat{h} = h^*$. If $\hat{h} = h^*$, then $\{x \in X \mid |C(x)| = \hat{h}\}$ is the unique stable*

set.

Proof. Let (X, \succ) be the social conflict system. First, assume that $\hat{h} = h^*$. Define $K = \{x \in X \mid |C(x)| = \hat{h}\}$. By Lemma 1(a) and (b), K is the set of strong equilibria. Therefore, K is internally stable.

Fix an arbitrary $\bar{x} \notin K$. Suppose that $|C(\bar{x})| > \hat{h}$. Let $x^* \in X$ be such that $C(x^*) \subset C(\bar{x})$ and $|C(x^*)| = \hat{h}$. By the choice of x^* , $\bar{x} \rightarrow_{C(\bar{x}) \setminus C(x^*)} x^*$. By $\hat{h} = h^*$, $u_i(x^*) = f(D, \hat{h}) > f(C, n-1) \geq f(C, |C(\bar{x})| - 1) = u_i(\bar{x})$ for all $i \in C(\bar{x}) \setminus C(x^*)$. Thus, $x^* \succ_{C(\bar{x}) \setminus C(x^*)} \bar{x}$.

Suppose that $|C(\bar{x})| < \hat{h}$. Let $y^* \in X$ be such that $C(\bar{x}) \subset C(y^*)$ and $|C(y^*)| = \hat{h}$. By the choice of y^* , $\bar{x} \rightarrow_{C(y^*) \setminus C(\bar{x})} y^*$. By the definition of \hat{h} , $u_i(y^*) = f(C, \hat{h} - 1) > f(D, \hat{h} - 1) \geq f(D, |C(\bar{x})|) = u_i(\bar{x})$ for all $i \in C(y^*) \setminus C(\bar{x})$. Thus, $y^* \succ_{C(y^*) \setminus C(\bar{x})} \bar{x}$. Hence, K is externally stable.

We turn to the uniqueness. Let $K' \subset X$ with $K' \neq K$. Assume that $K \setminus K' \neq \emptyset$. Let $x \in K \setminus K'$. Since x is a strong equilibrium by Lemma 1(b), there exists no $x' \in K'$ such that $x' \succ x$. Thus, K' is not externally stable. Therefore, assume that $K \setminus K' = \emptyset$. By $K' \neq K$, there exists some $y \in K'$ such that $y \notin K$. Therefore, there exists some $y' \in K$ such that $y' \succ y$ by the external stability of K . Since $K \subset K'$ by $K \setminus K' = \emptyset$, K' is not internally stable. Hence, there exists no stable set other than K .

Next, assume that $\hat{h} < h^*$. Suppose that there exists a stable set \bar{K} .

Claim 1 For any $x, y \in X$ with $C(x) \subsetneq C(y)$, either $x \succ y$ or $y \succ x$.

Proof of Claim 1. Let $x, y \in X$ such that $C(x) \subsetneq C(y)$. By $C(x) \subsetneq C(y)$, both $x \rightarrow_{C(y) \setminus C(x)} y$ and $y \rightarrow_{C(y) \setminus C(x)} x$. By Assumption 1(d), either $f(C, |C(y)| - 1) > f(D, |C(x)|)$ or $f(C, |C(y)| - 1) < f(D, |C(x)|)$. In the former case, $y \succ_{C(y) \setminus C(x)} x$. In the latter case, $x \succ_{C(y) \setminus C(x)} y$. \square

We prove that there exists no $x \in \bar{K}$ such that $|C(x)| > \hat{h}$ by a mathe-

matical induction.

Induction Base: $C_N \notin \bar{K}$.

Proof of Induction Base. Suppose that $C_N \in \bar{K}$. Fix an arbitrary $i \in N$. By $n > \hat{h}$, $(D, C_{-i}) \succ_{\{i\}} C_N$. Then, there exists some $y \in X$ such that $y \succ (D, C_{-i})$ by the external stability of \bar{K} . By Claim 1 and $C(y) \subsetneq N$, either $C_N \succ y$ or $y \succ C_N$. This contradicts the internal stability of \bar{K} . Hence, $C_N \notin \bar{K}$.

Induction Step $n - \ell + 1$ ($\ell > \hat{h}$): Let $\ell > \hat{h}$. Assume that $x' \notin \bar{K}$ for all $x' \in X$ with $|C(x')| > \ell$. Then, $x \notin \bar{K}$ for all $x \in X$ with $|C(x)| = \ell$.

Proof of Induction Step ℓ . Let $\ell > \hat{h}$. Suppose that there exists some $x \in \bar{K}$ with $|C(x)| = \ell$. By $|C(x)| > \hat{h}$, $(D, x_{-i}) \succ_{\{i\}} x$ for all $i \in S$. Fix an arbitrary $j \in S$. By the internal stability of \bar{K} , $(D, x_{-j}) \notin \bar{K}$. By the external stability of \bar{K} , there exists some $y \in \bar{K}$ such that $y \succ (D, x_{-j})$. By the induction hypothesis, $|C(y)| \leq \ell$. By Claim 1 and the internal stability of \bar{K} , $C(y) \setminus C(x) \neq \emptyset$. Thus, $C(y) \setminus (C(x) \setminus \{j\}) \neq \emptyset$ and $(D, x_{-j}) \rightarrow_Q y$ implies $C(y) \setminus (C(x) \setminus \{j\}) \subset Q$. By $\ell > \hat{h}$, $u_i(y) = f(C, |C(y)| - 1) \leq f(C, \ell - 1) < f(D, \ell - 1) = u_i(D, x_{-j})$ for all $i \in C(y) \setminus (C(x) \setminus \{j\})$. This contradicts $y \succ (D, x_{-j})$. Hence, $x \notin \bar{K}$ for all $x \in X$ with $|C(x)| = \ell$. \square

Thus, $\bar{K} \subset \{x \in X \mid |C(x)| \leq \hat{h}\}$. Then, $x \succ C_N$ is false for all $x \in \bar{K}$ by $\hat{h} < h^*$. This contradicts the external stability of \bar{K} . Hence no stable set exists. \blacksquare

Note that the stable set never exists in the prisoners' dilemma case where $\hat{h} = 0 < h^*$.

5 The social conflict system with delegations

We derive another system from the social conflict game by incorporating an agent and allowing the players to delegate their decisions to the agent. We give a formal definition first.

Let Γ be a social conflict game. Define

$$X^* = \bigcup_{S \subset N} (\{S\} \times \{C_S, D_S\} \times X_{N \setminus S}).$$

Define the dominance relation \succ^* on X^* as follows. Let $(S, x), (T, y) \in X^*$ and $Q \in \mathcal{N}$. We say (T, y) is inducible from (S, x) via Q , denoted by $(S, x) \rightarrow_Q^* (T, y)$, iff either the following **D1** or **D2** is satisfied.

D1 $S \subset T \subset Q$ and $x_{N \setminus Q} = y_{N \setminus Q}$;

D2 $S = T$, $Q \cap S = \emptyset$, and $x_{N \setminus Q} = y_{N \setminus Q}$.

We simply say (T, y) is inducible from (S, x) , denoted by $(S, x) \rightarrow^* (T, y)$, iff $(S, x) \rightarrow_Q^* (T, y)$ for some $Q \in \mathcal{N}$. We say (T, y) dominates (S, x) via Q , denoted by $(T, y) \succ_Q^* (S, x)$ iff $(S, x) \rightarrow_Q^* (T, y)$ and $u_i(x) > u_i(y)$ for all $i \in Q$. We simply say (T, y) dominates (S, x) , denoted by $(T, y) \succ^* (S, x)$ iff $(T, y) \succ_Q^* (S, x)$ for some $Q \in \mathcal{N}$. We say (X^*, \succ^*) is a social conflict system with delegations (derived from the social conflict game Γ).

In the social conflict system with delegations, a responsible agent is prepared. Given $(S, x) \in X^*$, S denotes the set of players who have delegated their decisions to the agent. Therefore, the agent has the right to change the strategies of S at (S, x) . The agent behaves on behalf of all delegating players. More formally, the agent changes the strategies when it benefits all the delegating players.

This system also describes an open negotiation. Given a strategy profile, a coalition of players, which may include the agent, openly declare that it

will change its strategies if the others stick to the current strategies. This declaration can involve some players newly delegating their decisions to the agent, and the agent accept the delegations when the agent is in the coalition **(D1)**. Of course, no additional delegation occurs when the the agent is not in the coalition **(D2)**. For simplicity, the agent is assumed to choose identical strategies for all delegating players.

We also assume that delegation is irreversible. An imaginary situation is that each player leaves the negotiation after he delegates his decision. Therefore, a player cannot withdraw his delegation. With this irreversibility, we emphasize that players are never forced to delegate their decisions to the agent.

The strong equilibrium can be extended to the case in the social conflict system with delegations (X^*, \succ^*) : $(S, x) \in X^*$ is a strong equilibrium (in (X^*, \succ^*)) iff there exists no $(T, y) \in X^*$ such that $(T, y) \succ^* (S, x)$. In contrast with the model without delegations, there is a strong equilibrium in the social conflict system with delegations even if $\hat{h} < h^*$. For example, (N, C_N) is always a strong equilibrium by Assumption 1(c) (Lemma 4 below). On the other hand, (\emptyset, x) where x is efficient is not a strong equilibrium when $\hat{h} < h^*$. In this case, any player taking C at x can be made better off by changing his strategy from C to D . In the following theorem, we show the unique existence and the efficiency of the stable set in the social conflict system with delegations. It is also shown that some efficient outcomes where no delegation takes place are also included in the stable set.

Theorem 1 *In the social conflict system with delegations (X^*, \succ^*) ,*

$$K^* = \{(S, x) \in X^* \mid |C(x)| = h^*\} \cup \{(S, x) \in X^* \mid C(x) = S, |S| > h^*\}$$

is the unique stable set. Note that x is efficient for all $(S, x) \in K^$.*

Before starting the proof of Theorem 1, we prove two lemmas.

Lemma 2 *Let $(S, x), (T, y) \in X^*$ such that x is efficient. If $(T, y) \succ^* (S, x)$, then $C(y) \subset C(x)$.*

Proof. Let $(S, x), (T, y) \in X^*$ such that x is efficient. Assume $(T, y) \succ^* (S, x)$. Suppose that $C(y) \setminus C(x) \neq \emptyset$. Let $z \in X$ such that $C(z) = C(x) \cup C(y)$ and $Q = \{i \in N \mid x_i \neq y_i\}$. We have $|C(z)| > |C(x)|$ and $Q \subset C(z)$. By $(T, y) \succ^* (S, x)$, $u_i(y) > u_i(x)$ for all $i \in Q$. Thus, $u_i(z) = f(C, |C(z)| - 1) \geq f(C, |C(y)| - 1) = u_i(y) > u_i(x)$ for all $i \in Q$. For all $i \notin Q$, $u_i(z) > u_i(x)$ by Assumption 1(a) and $|C(z)| > |C(x)|$. This contradicts the efficiency of x . Hence, $C(y) \setminus C(x) = \emptyset$. ■

Lemma 3 *$(N, C_N) \succ_N^* (S, x)$ for all $S \subset N$ and inefficient $x \in X$.*

Proof. Let $(S, x) \in X^*$ where x is inefficient. By Lemma 1(c), $|C(x)| < h^*$. Obviously, $(S, x) \rightarrow_N^* (N, C_N)$. By the definition of h^* , $u_i(C_N) = f(C, n - 1) > f(D, |C(x)|) = u_i(x)$ for all $i \in N \setminus C(x)$. By Assumption 1(a), $u_i(C_N) = f(C, n - 1) > f(C, |C(x)| - 1) = u_i(x)$ for all $i \in C(x)$. Hence, $(N, C_N) \succ_N^* (S, x)$. ■

Lemma 4 *(N, C_N) is a strong equilibrium.*

Proof. Suppose that (N, C_N) is not a strong equilibrium. Then, there exist some $Q \in \mathcal{N}$ and $(S, z) \in X^*$ such that $(S, z) \succ_Q^* (N, C_N)$. By the definition of \rightarrow_Q^* , $S = Q = N$. By the definition of X^* , $(N, z) \in X^*$ implies $z \in \{C_N, D_N\}$. Then, $(N, z) \succ_N^* (N, C_N)$ is impossible by Assumption 1(c), a contradiction. Hence, (N, C_N) is a strong equilibrium. ■

Proof of Theorem 1. Let (X^*, \succ^*) be a social conflict system with delegations. The internal stability of K^* follows from Lemma 2.

We show the external stability. Fix an arbitrary $(S, x) \in X^* \setminus K^*$. If x is not efficient, then $(N, C_N) \succ^* (S, x)$ by Lemma 3, where $(N, C_N) \in K^*$. Therefore, assume that x is efficient. By $(S, x) \notin K^*$, the efficiency of x , and Lemma 1(c), $|C(x)| > h^*$ and $C(x) \neq S$. Let $x^* \in X$ such that (i) $C(x^*) = S$ if $|S| > h^*$ and $S \subset C(x)$, (ii) $|C(x^*)| = h^*$ and $S \subset C(x^*) \subset C(x)$ if $|S| \leq h^*$ and $S \subset C(x)$, and (iii) $|C(x^*)| = h^*$ and $C(x^*) \subset C(x)$ if $C(x) \subset N \setminus S$. In any case, $C(x^*) \subset C(x)$ and $C(x) \setminus C(x^*) \subset N \setminus S$. Thus, $(S, x) \rightarrow_{C(x) \setminus C(x^*)}^* (S, x^*)$. In any case of (i)-(iii), $(S, x^*) \in K^*$ and $|C(x^*)| \geq h^*$. Then, $u_i(x^*) = f(D, |C(x^*)|) > f(C, n - 1) \geq f(C, |C(x)| - 1) = u_i(x)$ for all $i \in C(x) \setminus C(x^*)$. Thus, $(S, x^*) \succ_{C(x) \setminus C(x^*)}^* (S, x)$. Hence, K^* is externally stable.

We show the uniqueness. Let $K \subset X^*$ be a stable set for (X^*, \succ^*) . By Lemma 4, $(N, C_N) \in K$. By $(N, C_N) \in K$, Lemma 3, and the internal stability of K , $(S, x) \notin K$ for all $(S, x) \in X^*$ with Pareto inefficient x . Thus, $K \subset \{(S, x) \in X^* \mid |C(x)| \geq h^*\}$. By Lemma 2, for any $(S, x) \in K^*$, there exists no $(T, y) \in K$ such that $(T, y) \succ^* (S, x)$. Thus, $K^* \subset K$ by the external stability of K . Then, $K = K^*$ follows from the internal stability of K and the external stability of K^* .

The efficiency immediately follows from Lemma 1(c). ■

In the social conflict system (without delegation), the stable set consists of the strong equilibria if it exists (Proposition 1). Now, we turn to the problem of the stable set coinciding with the set of strong equilibria in the social conflict system with delegations. We show that $\hat{h} = h^*$, the same condition as Proposition 1, is a necessary and sufficient condition for the coincidence of the stable set and the set of strong equilibria.

Proposition 2 *Let K^* be the stable set for the social conflict system with delegations (X^*, \succ^*) defined in Theorem 1. Then, K^* coincides with the set*

of strong equilibria in (X^*, \succ^*) if and only if $\hat{h} = h^*$.

Proof. Let K^* be the stable set defined in Theorem 1 and Z be the set of strong equilibria in (X^*, \succ^*) . It is straightforward that $Z \subset K^*$ from the external stability of K^* independent of the relationship between \hat{h} and \hat{h}^* . Therefore, it suffices to show that $K^* \subset Z$ if and only if $\hat{h} = h^*$.

Assume that $\hat{h} = h^*$. Fix an arbitrary $(S, x) \in K^*$. We distinguish the proof in two cases. The first case is $|C(x)| > h^*$. Then, $C(x) = S$. Suppose that there exists some $(T, y) \in X^*$ such that $(T, y) \succ^* (S, x)$. By the efficiency of x , Lemma 2, and the definition of \rightarrow^* , $y = D_N$. By $\hat{h} = h^*$ and Assumption 1(c), $\hat{h} > 0$. Then, $u_i(D_N) = f(D, 0) < f(C, 0) \leq f(C, |C(x)| - 1) = u_i(x)$ for all $i \in S$. This contradicts $(T, y) \succ^* (S, x)$ and $y_i \neq x_i$ for all $i \in S$. Hence, (S, x) is a strong equilibrium.

The second case is $|C(x)| = h^*$. By the definitions of \rightarrow and \rightarrow^* , for all $(S, x), (T, y) \in X^*$, $(S, x) \rightarrow^* (T, y)$ implies $x \rightarrow y$. Thus, it follows that (S, x) is a strong equilibrium from Lemma 1(a) and (b). Hence, $K^* \subset Z$.

Then, assume that $\hat{h} < h^*$. Let $x' \in X$ be such that $|C(x')| = h^*$. We have $(\emptyset, x') \in K^*$. Fix an arbitrary $j \in C(x')$. By $\hat{h} < h^*$, $u_j(x') = f(C, |C(x')| - 1) < f(D, |C(x')| - 1) = u_j(D, x'_{-j})$. By the definition of \rightarrow^* , $(\emptyset, x') \rightarrow_{\{j\}}^* (\emptyset, (x'_{-j}, D))$. Thus, $(\emptyset; D, x'_{-j}) \succ_{\{j\}}^* (\emptyset, x')$, and (\emptyset, x') is not a strong equilibrium. Hence, $K^* \setminus Z \neq \emptyset$. ■

6 Concluding remarks

This paper studies the stable sets in two systems derived from a social conflict game. We show that the stable set uniquely exists and consists of certain efficient outcomes in the social conflict system with delegations, while the stable set fails to exist in the social conflict system (without delegations)

when the Nash equilibrium is inefficient. Note that the stable set contains an efficient outcome without any delegation. We also characterize the stable sets by strong equilibria. We conclude with two remarks.

First, the agent and the possibility of delegations are exogenously arranged institutions. We may consider the possibility of endogenously arranged institutions as in Okada (1993) and investigate the stable set for such models. Second, the irreversibility of the delegations plays an important role. A similar analysis was given by Hirai (2012), in which payoff transfers are allowed and full cooperation is achieved. We may also consider what kinds of institutions lead to full cooperation without payoff transfers.

References

- Aumann, R.J. (1959) "Acceptable points in general cooperative n -person games," *Annals of Mathematical Studies* **40**, 287-324
- Bernheim, B.D., Peleg, B., Whinston, M.D. (1987) "Coalition-proof equilibria I. Concepts," *Journal of Economic Theory* **42**, 1-12
- Chwe, M.S.-Y. (1994) "Farsighted coalitional stability," *Journal of Economic Theory* **63**, 299-325
- Greenberg, J. (1990) "*The Theory of Social Situations*," Cambridge University Press
- Hirai, T. (2012) "Renegotiations and the diffusion of the technology with positive externalities," *Working Paper, University of Toyama* **No. 277**
- Nakanishi, N. (2001) "On the existence and the efficiency of the von Neumann-Morgenstern stable set in a n -player prisoners' dilemma," *International Journal of Game Theory* **30**, 291-307
- Nakanishi, N. (2009) "Noncooperative farsighted stable set in an n -player prisoners' dilemma," *International Journal of Game Theory* **38**, 249-261

- Okada, A. (1993) "The possibility of cooperation in an n -person prisoners' dilemma with institutional arrangements," *Public Choice* **77**, 629-656
- Saijo, T. and Yamato, T. (1999) "A voluntary participation game with a non-excludable public good," *Journal of Economic Theory* **84**, 227-242
- Schelling, T. (1960) "*The Strategy of Conflict*," Harvard University Press
- Shinohara, R. (2010) "Coalition-proof equilibria in a voluntary participation game," *International Journal of Game Theory* **39**, 603-615
- Suzuki, A. and Muto, S. (2005) "Farsighted stability in an n -person prisoner's dilemma," *International Journal of Game Theory* **33**, 431-445
- Taylor, M. and Ward, H. (1982) "Chickens, whales, and lumpy goods: Alternative models of public goods provision," *Political Studies* **30**, 350-370
- von Neumann, J. and Morgenstern, O. (1944) "*Theory of Games and Economic Behavior*," Princeton University Press