Staff Paper Series ’07-07 October 2007

Coalition-proof Equilibria in a Voluntary Participation Game

Ryusuke Shinohara

Faculty of Economics
Shinshu University
Matsumoto 390-8621 Japan
Phone: +81-263-35-4600
Fax: +81-263-37-2344

Coalition-proof Equilibria in a Voluntary Participation Game *

Ryusuke Shinohara†

October 4, 2007

Abstract

We examine the coalition-proof equilibria of a participation game in a public good mechanism and study which Nash equilibria are achieved through the cooperative behavior of agents. The participation game may have multiple Nash equilibria, and various numbers of participants may be attained at the Nash equilibria. We provide sufficient conditions for the Nash equilibrium of the participation game to be a coalition-proof equilibrium and sufficient conditions under which the number of participants achieved at coalition-proof equilibria is unique. By applying these results, we more easily characterize the set of coalition-proof equilibria of the participation game with specific environments.

JEL classification: C72, D62, D71, H41.

* This paper is based on Chapters 3 and 4 of my doctoral dissertation submitted to Hitotsubashi University. I would like to thank Koichi Tadenuma for a number of helpful comments, suggestions, and constant support. I am also grateful to Tatsuro Ichiishi, Shin-ichi Takekuma, Akira Okada, Motohiro Sato, Tatsuyoshi Saijo, Takehiko Yamato, Hideo Konishi, Takashi Shimizu, and participants of the seminars at the autumn meeting of Japanese Economic Association at Meiji University and the Decentralization Conference in Japan. This research was supported by the Grand-in-Aid for Young Scientists (Start-up) from the Japan Society for the Promotion of Science. Of course, any remaining errors are my own.

† Faculty of Economics, Shinshu University, 3-1-1, Asahi, Matsumoto, Nagano, 390-8621, JAPAN. E-mail: ryusukes@shinshu-u.ac.jp Telephone : +81-263-37-2951 Fax : +81-263-37-2344

Key Words: Participation game; Coalition-proof equilibrium; Heterogeneous agents

1 Introduction

In this paper, we examine coalition-proof equilibria in a participation game in a public good mechanism. In the theory of implementation, many mechanisms have been constructed to solve the free-rider problem in economies with public goods. For example, Groves and Ledyard (1977) constructed a mechanism in which efficient allocations are attained at Nash equilibria, and Hurwicz (1979) and Walker (1981) constructed mechanisms that implement the Lindhal allocation rule in Nash equilibria. Thus, from the results of the theory of implementation, the construction of mechanisms is used to solve the free-rider problem.

However, the implementation theory supposes the participation of all agents, and the individual agents do not have the right to decide whether or not to participate in the mechanism. This may not be realistic in many situations. For example, in consideration of a case involving international environmental agreements, an environmental agreement is regarded as a mechanism that provides public goods or eliminates public bads in order to attain efficient allocations of resources. Each country has the right to decide whether or not to participate in the agreement.

Saijo and Yamato (1999) introduced a model of voluntary participation in a public good mechanism that provides a non-excludable public good. Their model consists of two stages. In the first stage, agents decide simultaneously whether or not to participate in the public good mechanism. In the second stage, the mechanism is played only by the participants, and, at equilibria, an allocation that is desirable with respect to only participants' preferences is attained. The participants bear the cost of the public good, but the non-participants can benefit from the public good provision at no cost because of the non-excludability of the public good. Therefore, agents have an incentive not to participate in the mechanism and enjoy the public good at no cost. On the other hand, agents may have an incentive to enter the mechanism in order to have their preferences taken into account. In an economy in which all the agents have the same Cobb-Douglas utility function and the same initial endowments of private goods, Saijo and Yamato (1999) showed that there are subgame perfect equilibria in which not all agents participate in the mechanism.

In this paper, we extend the analyses of the participation issues in two directions. First, we examined a coalition-proof equilibrium (Bernheim et al., 1987), which is a refinement of a Nash equilibrium that is stable against coalitional deviations. Second, we consider the case in which agents have heterogeneous preferences. It is likely that the participation game has multiple Nash equilibria when agents' preferences are heterogeneous and various sets of participants that consist of different numbers of participants are supported as the equilibria. However, few studies have been made in which Nash equilibria are more likely to occur. In this paper, considering the coalition-proof equilibrium, we investigate which Nash equilibria can be achieved through the cooperative behavior of agents and how many agents participate in the mechanism as a result of cooperative behavior.[*1]

Our results are as follows. First, we provide sufficient conditions under which a Nash equilibrium is a coalition-proof equilibrium in the participation game. We prove that a Nash equilibrium is coalition-proof if the condition of preservation of participation incentive holds at the Nash equilibrium and the Nash equilibrium is strict for non-participants. The condition of the preservation of participation incentives requires that, if agent i does not have an incentive to join in a set of participants P, then i does not have an incentive to join in a set of participants that produces more public goods than P. A Nash equilibrium is strict for non-participants if the non-participants at the equilibrium have a strict incentive to choose non-participation. Moreover, if these two conditions are satisfied at a Nash equilibrium, then the Nash equilibrium is not only a coalition-proof equilibrium but also a Pareto-superior Nash equilibrium.

Secondly, focusing on the maximal number of participants that can be achieved at Nash equilibria, we establish sufficient conditions under which the number of participants at coalition-proof equilibria is solely determined by the maximal number. We show that, if all participants have at least as high a marginal willingness to pay for the public good as the non-participants at a Nash equilibrium with the maximal number of participants, the preservation of the participation incentive condition holds at the Nash equilibrium, and the Nash equilibrium is strict for non-participants, then only the maximal number is attained at coalition-proof equilibria. Finally, we apply these

[*1] Cooperative behavior in the participation decision can be observed in the real world. For example, in the case of the Kyoto Protocol, ratification by Russia was essential to bring the protocol into force. The European Union, which is an environmentally conscious group, negotiated with Russia and tried to induce Russia to ratify the protocol.

results to a case in which agents' preferences are identical and to one in which agents have square-root benefit functions and the cost function of the public good is linear, and we identify the coalition-proof equilibria of the participation game.

The relationship between this work and other studies will be discussed. Several studies have investigated games that are similar to the participation game that is the focus of this paper, such as games of the ratification of international environmental agreements (e.g., Carraro and Siniscalco, 1993, 1998), cartel formation games (e.g., d'Aspremont et al., 1983; Thoron, 1998), and endogenous lobby formation games (e.g., Furusawa and Konishi, 2007). Most of the studies on ratification games have examined only the case of identical agents and clarified the characteristics of Nash equilibria, but, in this paper, we analyze the case of heterogeneous agents and consider coalition-proof equilibria. The same applies to games of cartel formation. Although Thoron (1998) characterized coalition-proof equilibria in cartel formation games, she also considered the case of identical firms. Moreover, Thoron (1998) used a condition that differs from ours. Furusawa and Konishi (2007) examined two-stage public good provision games. In the first stage, each agent decides simultaneously whether or not to join a contribution group. In the second stage, agents that join the contribution group play the common agency game (Berheim and Whinston, 1986) and supply the public good. Agents in the non-contribution group can free-ride the public good. These researchers showed that perfectly coalition-proof equilibria are equivalent to the free-ride proof core in this game. However, since Furusawa and Konishi (2007) considered the common agency game in the second stage, the cost of producing a public good is not always distributed in proportion to contributors' marginal willingness to pay, and the conditions in this paper do not necessarily hold in their model. Although these researchers and this paper analyzed the case in which agents ' preferences are quasi-linear, the results in this paper also hold in the case in which agents ' preferences are represented by Cobb-Douglas utility functions and all agents have the same initial endowments of private good.*2 Maruta and Okada (2005) also considered a two-stage public good provision game and studied the evolutional selection of the contribution group formation game. However, their model differs from ours.

It may be difficult to identify the coalition-proof equilibria because the equilibria are defined by using recursion with respect to the number of players. However, by ap-

*2 See Remark 1.

plying our results, it becomes easier to characterize the set of coalition-proof equilibria in a participation game with a specific environment. Our results would be useful to examine the coalition-proof equilibria of games that are similar to the participation game, such as the ratification game of international environmental treaties and cartel formation games.

2 The model

We consider the problem of providing a (pure) public good and distributing its cost. There are one private and one public good in the economy. The public good is perfectly divisible: the level of public good is in the set of non-negative real numbers $\mathbb{R}_+$. Let n be the number of agents. The set of agents is denoted by $N = \{1, \dots, n\}$. Each agent $i \in N$ has a preference relation that is represented by a quasi-linear utility function. If y and x_i designate the level of the public good and a monetary transfer from agent $i \in N$, respectively, then agent i's utility is $V_i(y, x_i) = \alpha_i v(y) - x_i$, in which $\alpha_i > 0$, $v(0) = 0$, $v' > 0$, and $v'' < 0$ are satisfied. For each level of public good y, the cost of producing y units of the public good is $c(y)$, in which $c(0) = 0$, $c' > 0$, and $c'' \geq 0$ are satisfied. We further assume that $v'(0) > c'(0)$.

In this paper, we consider a situation in which there is a mechanism to provide a public good and distribute the cost of the public good and each agent can simultaneously decide either participation or non-participation in the mechanism. A mechanism (or a game form) is a list of message spaces of all agents and an outcome function that associates an allocation with each profile of the messages. The following two-stage game is considered: in the first stage, each agent simultaneously decides whether or not to participate in the mechanism. In the second stage, knowing the participation decisions of other agents, the agents who choose participation in the first stage select their messages from their message spaces of the mechanism. Only the participants decide the quantity of the public good and the cost shares of each participant through the choice of messages. Let $(y^P, (x_j^P)_{j \in P})$ be the equilibrium outcome of the mechanism for a set of participants P. We assume that the *ratio allocation rule* introduced by Kaneko (1977a, 1977b) is achieved at the equilibrium of the mechanism. Thus, $y^{\emptyset} = 0$, and, for every non-empty subset P of N, $(y^P, (x_j^P)_{j \in P})$ satisfies the following conditions:

$$
\begin{aligned}
y^P &\in \arg\max_{y \in \mathbb{R}_+} \sum_{j \in P} \alpha_j v(y) - c(y) \text{ and} \\
x_i^P &= \frac{\alpha_i}{\sum_{j \in P} \alpha_j} c(y^P) \text{ for all } i \in P.
\end{aligned}
\tag{1}
$$

In this paper, we are not concerned with the implementation problem of the allocation rule. However, many researchers have constructed mechanisms that implement the ratio allocation rule and the Lindahl allocation rule under the various equilibrium concepts. Groves and Ledyard (1977) constructed a mechanism that achieves efficient allocations in Nash equilibria; Hurwicz (1979) and Walker (1981) designed mechanisms that implement Lindahl allocations in Nash equilibria. Besides these authors, Peleg (1996) and Tian (2000) constructed mechanisms that implement the Lindahl allocation rule in both strong and Nash equilibria; Corchon and Wilkie (1996) constructed mechanisms that implement the ratio equilibria in both strong and Nash equilibria.

In this paper, we assume that agents that selected non-participation can benefit from the public good at no cost because of the non-excludability of the public good.

Assumption 1 For every set of participants P and for every agent $i \notin P$, $x_i^P = 0$, and i consumes y^P.

Given the outcome of the second stage, the participation-decision stage can be reduced to the following simultaneous game. In the game, each agent i simultaneously chooses either $s_i = I$ (participation) or $s_i = O$ (non-participation), and then the set of participants is determined. Let P^s be the set of participants at an action profile $s = (s_1, \ldots, s_n)$. Then, each agent i obtains the utility $V_i(y^{P^s}, x_i^{P^s})$ at the action profile s. That is, the participants produce the public good, and they share the cost of the public good as above. Each non-participant can benefit from the public good at no cost. We call this reduced game a *participation game*, which is formally defined as follows.

Definition 1 (Participation game) A *participation game* is represented by $G = \left[N,\, S^n = \{I,\, O\}^n,\, (U_i)_{i \in N}\right]$, where U_i is the payoff function of i, which associates a real number $U_i(s)$ with each strategy profile $s \in S^n$: if P^s designates the set of participants at s, then $U_i(s) = V_i(y^{P^s}, x_i^{P^s})$ for all i.

We limit our attention to pure strategy profiles.

3 Equilibrium Concepts

Before introducing the notions of equilibria studied in this paper, we will introduce some notations. For all $D \subseteq N$, we denote the complement of D by $-D$. For all coalitions D, $s_D \in S^{\#D}$ denotes a strategy profile for D.[*3] We simply write $s_N = s$.

The first notion is very basic.

Definition 2 (Nash equilibrium) A strategy profile $s^* \in S^n$ is a *Nash equilibrium* if, for all $i \in N$ and for all $\widehat{s}_i \in S$, $U_i(s_i^*, s_{-i}^*) \geq U_i(\widehat{s}_i, s_{-i}^*)$.

Our second notion is the *coalition-proof equilibrium*. It was introduced by Bernheim, Peleg, and Whinston (1987) and is known as a refinement of Nash equilibria based on the stability against self-enforcing coalitional deviations. It is defined by using the notion of *restricted games*. A restricted game is a game in which a subset of agents play the game G, taking the strategy profiles of agents outside the subset as given. We formally define it as follows. Let $T \subsetneq N$ and $t = \#T$. Let $\bar{s}_{N \setminus T} \in S^{n-t}$. A *restricted game* $G|\bar{s}_{N \setminus T}$ is a game in which the set of agents is T, the set of strategy profiles is S^t, and the payoff function for each $i \in T$ is the function $U_i(\cdot, \bar{s}_{N \setminus T})$ that associates a real value $U_i(s_T,\, \bar{s}_{N \setminus T})$ with each element s_T in S^t such that: $U_i(s_T,\, \bar{s}_{N \setminus T}) = V_i(y, x_i)$, where $(y, (x_j)_{j \in N})$ is the allocation when agents play $(s_T,\, \bar{s}_{N \setminus T})$ in G.

Definition 3 A *coalition-proof equilibrium* $(s_1^*, \ldots, s_n^*)$ is defined inductively with respect to the number of agents t:

- When $t = 1$, for all $i \in N$, s_i^* is a coalition-proof equilibrium for $G|s_{N \setminus \{i\}}^*$ if $s_i^* \in \arg\max\ U_i(s_i, s_{N \setminus \{i\}}^*)$ s.t. $s_i \in S$.
- Let $T \subseteq N$ with $t = \#T \geq 2$. Assume that coalition-proof equilibria have been defined for all normal form games with fewer agents than t.
- Consider the restricted game $G|s_{N \setminus T}^*$ with t agents.
  - A strategy profile $s_T^* \in S^t$ is called *self-enforcing* if, for all $Q \subsetneq T$, s_Q^* is a coalition-proof equilibrium of $G|s_{N \setminus Q}^*$.

[*3] For every set $T \subseteq N$, $\#T$ means the cardinality of T.

- A strategy profile s_T^* is a coalition-proof equilibrium of $G|s_{N\backslash T}^*$ if it is a self-enforcing strategy profile and there is no other self-enforcing strategy profile $\widehat{s}_T \in S^t$ such that $U_i(\widehat{s}_T, s_{N\backslash T}^*) \geq U_i(s_T^*, s_{N\backslash T}^*)$ for all $i \in T$ and $U_i(\widehat{s}_T, s_{N\backslash T}^*) > U_i(s_T^*, s_{N\backslash T}^*)$ for some $i \in T$.

Coalition-proof equilibria are defined as the Pareto-efficient frontier within the set of self-enforcing strategy profiles. The self-enforcing strategy profiles are recursively defined with respect to the number of agents in coalitions. At a self-enforcing strategy profile of N, no proper coalition of N can coordinate its members' strategies in such a way that all members of the coalition are at least as well off and at least one of them is strictly better off and no proper subsets of the coalition further deviate in a self-enforcing way.

Example 1 indicates that the participation game may have multiple Nash equilibria at which various numbers of participants are achieved and the number of participants may be uniquely determined at coalition-proof equilibria.

Example 1 Consider an example in which $n = 3$, $v(y) = \sqrt{y}$, $\alpha = \alpha_1 = \alpha_2 = \alpha_3$, and $c(y) = y$. When the set of participants is P with $p = \#P$, the public good provision y^P maximizes

$$p\left(\alpha\sqrt{y} - \frac{y}{p}\right).$$

Hence,

$$y^P = \left(\frac{\alpha p}{2}\right)^2.$$

The payoff to every agent $i \in P$ is

$$\frac{\alpha^2 p}{2} - \frac{1}{p}\left(\frac{\alpha p}{2}\right)^2 = \frac{\alpha^2 p}{4},$$

and the payoff to every $i \in N\backslash P$ is

$$\frac{\alpha^2 p}{2}.$$

The payoff matrix is shown in Table 1, where agent 1 chooses rows, agent 2 chooses columns, and agent 3 chooses matrices. The first entry in each box is agent 1's payoff, the second is agent 2's, and the third is agent 3's. There are two types of Nash equilibria in this game. One is the Nash equilibrium with participation of one agent, and the other is the Nash equilibrium with participation of two agents. Clearly, every Nash equilibrium with two participants is coalition-proof, and only the Nash equilibria are coalition-proof.

	I	O
I	$\frac{3\alpha^2}{4}, \frac{3\alpha^2}{4}, \frac{3\alpha^2}{4}$	$\frac{\alpha^2}{2}, \alpha^2, \frac{\alpha^2}{2}$
O	$\alpha^2, \frac{\alpha^2}{2}, \frac{\alpha^2}{2}$	$\frac{\alpha^2}{2}, \frac{\alpha^2}{2}, \frac{\alpha^2}{4}$

I

	I	O
I	$\frac{\alpha^2}{2}, \frac{\alpha^2}{2}, \alpha^2$	$\frac{\alpha^2}{4}, \frac{\alpha^2}{2}, \frac{\alpha^2}{2}$
O	$\frac{\alpha^2}{2}, \frac{\alpha^2}{4}, \frac{\alpha^2}{2}$	0, 0, 0

O

Table. 1 Payoff matrix of Example 1

Our third notion is that of strong equilibria (Aumann, 1959).

Definition 4 (Strong equilibrium) A strategy profile $s^* \in S^n$ is a *strong equilibrium of* G if there exist no coalition $T \subseteq N$ and its strategy profile $\widetilde{s}_T \in S^{\#T}$ such that $U_i(\widetilde{s}_T, s^*_{-T}) \geq U_i(s^*)$ for all $i \in T$ with strict inequality for at least one $i \in T$.

A strong equilibrium is a strategy profile in which no subset of agents, taking the strategies of others as given, can jointly deviate in such a way that all members are at least as well off and at least one of its members is strictly better off. It is clear that both coalition-proof and strong equilibria are Nash equilibria. The set of strong equilibria is included in that of coalition-proof equilibria because the coalition-proof equilibria are required to be stable only against *self-enforcing* coalitional deviations, while the strong equilibria are defined to be stable against *all possible* coalitional deviations. The converse inclusion relation does not always hold.

Example 2 Consider the same situation as Example 1 except for $n = 3$. In this example, let $n = 5$. There are two types of Nash equilibria as in Example 1. The set of Nash equilibria with two participants coincides with that of coalition-proof equilibria in this example. However, there is no strong equilibrium. In every Nash equilibrium with two participants, non-participants obtain the payoff α^2. When the number of participants is five, the payoffs of all the agents are $5\alpha^2/4$. Hence, the three non-participants in the Nash equilibrium can gain higher payoffs if all of the non-participants jointly deviate from O to I. Therefore, none of the Nash equilibria with two participants is a strong equilibrium.

4 Basic properties of the participation game

4.1 Properties for payoff functions

We first introduce the payoff function that associates a real number with each set of participants.

Definition 5 A payoff function of $i \in N$, $u_i : 2^N \to \mathbb{R}_+$, is defined as follows:

$$\text{For every set of participants } P,\ u_i(P) = \begin{cases} \alpha_i v(y^P) - \frac{\alpha_i}{\sum_{j \in P} \alpha_j} c(y^P) & \text{if } i \in P, \\ \alpha_i v(y^P) & \text{otherwise.} \end{cases}$$

Lemma 1 proves that the level of the public good gets higher as the sum of the marginal willingness to pay of participants increases.

Lemma 1 For all sets of participants P, $Q \subseteq N$, if $\sum_{i \in P} \alpha_i > \sum_{i \in Q} \alpha_i$, then $y^P > y^Q$.

Proof. Let $P,\ Q \subseteq N$ be such that $\sum_{j \in P} \alpha_j > \sum_{j \in Q} \alpha_j$. Let y^P and y^Q be levels of the public good when the set of participants are P and Q, respectively. The public good provision y^P and y^Q satisfy the following conditions

$$v'(y^P) = \frac{c'(y^P)}{\sum_{j \in P} \alpha_j} \text{ and } v'(y^Q) = \frac{c'(y^Q)}{\sum_{j \in Q} \alpha_j}.$$

Suppose, on the contrary, that $y^Q \geq y^P$. Then, the following inequalities are satisfied:

$$v'(y^P) = \frac{c'(y^P)}{\sum_{j \in P} \alpha_j} < \frac{c'(y^P)}{\sum_{j \in Q} \alpha_j} \leq \frac{c'(y^Q)}{\sum_{j \in Q} \alpha_j} = v'(y^Q).$$

Hence, we have $v'(y^P) < v'(y^Q)$. Since v' is strictly decreasing, we have $y^P > y^Q$. This is a contradiction. ■

Lemma 2 For all sets of participants P, $Q \subseteq N$, if $\sum_{i \in P} \alpha_i > \sum_{i \in Q} \alpha_i$, then conditions (2) and (3) are satisfied:

$$u_i(P) > u_i(Q) \text{ for all } i \notin P \cup Q, \quad \text{and} \tag{2}$$

$$u_j(P) > u_j(Q) \text{ for all } j \in P \cap Q. \tag{3}$$

Proof. It is immediate from Lemma 1 that (2) holds. We show (3). Let P, $Q \subseteq N$ be such that $\sum_{j\in P}\alpha_j > \sum_{j\in Q}\alpha_j$, and let $i \in P \cap Q$. Since y^P maximizes the sum of the utilities of agents in P,

$$\sum_{j\in P} u_j(P) = v(y^P)\sum_{j\in P}\alpha_j - c(y^P) \geq v(y^Q)\sum_{j\in P}\alpha_j - c(y^Q). \tag{4}$$

Multiplying the both sides of (4) by $\alpha_i/\sum_{j\in P}\alpha_j$, together with $\sum_{j\in P}\alpha_j > \sum_{j\in Q}\alpha_j$, yields

$$\begin{aligned}\alpha_i\, v(y^P) - \frac{\alpha_i}{\sum_{j\in P}\alpha_j}c(y^P) &\geq \alpha_i\, v(y^Q) - \frac{\alpha_i}{\sum_{j\in P}\alpha_j}c(y^Q)\\ &> \alpha_i\, v(y^Q) - \frac{\alpha_i}{\sum_{j\in Q}\alpha_j}c(y^Q).\end{aligned}$$

Hence, we obtain $u_i(P) > u_i(Q)$. ■

Lemma 2 is a basic property for the payoff functions. From Lemma 2, the payoffs to participants and those to non-participants increase with respect to the sum of the marginal willingness to pay for the public good of participants. This property will play an important role in showing the main results.

4.2 Nash equilibria and Pareto domination

The following is the definition of Pareto domination of strategy profiles.

Definition 6 A strategy profile $s \in S^n$ is Pareto-dominated by a strategy profile $\widetilde{s}$ if $U_i(\tilde{s}) \geq U_i(s)$ for all $i \in N$ and $U_i(\tilde{s}) > U_i(s)$ for some $i \in N$.

Let $s \in S^n$ be a strategy profile and let P be the set of participants at s. Define $R(s)$ as the set of strategy profiles that can be reached from s by deviations of agents in P. Set $R(s)$ is formally defined as follows.

Definition 7 Let s be a profile of strategies and let P be the set of agents that choose I. The subset of strategy profile $R(s)$ is defined as

$$\left\{\widehat{s} \in S^n | \text{there exists } D \in 2^P\backslash\{\emptyset\} \text{ such that } \widehat{s}_i = O \text{ for all } i \in D \text{ and } \widehat{s}_i = s_i \text{ for all } i \notin D\right\}.$$

For example, $R((I,\dots,I))$ is equivalent to the set $S^n\backslash\{(I,\dots,I)\}$ and $R((O,\dots,O))$ is empty.

Lemma 3 Let $s \in S^n$ be a Nash equilibrium of the participation game. Then, s Pareto-dominates all the strategy profiles $\widehat{s} \in R(s)$.

Proof. Let s be a Nash equilibrium of G in which P is the set of participants, and let $\widehat{s} \in R(s)$ be a strategy profile in which $\widehat{P}$ is a set of participants. Note that $\widehat{P} \subsetneq P$ by the definition of $R(s)$. Thus, it follows that $\sum_{i\in P}\alpha_i > \sum_{i\in\widehat{P}}\alpha_i$. From $\sum_{i\in P}\alpha_i > \sum_{i\in\widehat{P}}\alpha_i$, (2), (3), and the definition of Nash equilibrium, we have the following three conditions:

$$u_i(P) > u_i(\widehat{P}) \text{ for all } i \in \widehat{P}, \tag{5}$$

$$u_i(P) > u_i(\widehat{P}) \text{ for all } i \in N\backslash P, \text{ and} \tag{6}$$

$$u_i(P) \geq u_i(P\backslash\{i\}) \geq u_i(\widehat{P}) \text{ for all } i \in P\backslash\widehat{P}. \tag{7}$$

Conditions (5) and (6) are immediate from $\sum_{i\in P}\alpha_i > \sum_{i\in\widehat{P}}\alpha_i$, (2) and (3). The first inequality of (7) follows from the definition of Nash equilibrium, and the second follows from (2) and holds with equality if $\widehat{P} = P\backslash\{i\}$. By (5), (6) and (7), $\widehat{s}$ is Pareto-dominated by s. ■

5 Main results

5.1 A sufficient condition for a coalition-proof equilibrium

We provide a sufficient condition for a Nash equilibrium to be a coalition-proof equilibrium in the participation game. The first condition is related to the participation incentives.

Condition 1 (Preservation of the participation incentive with respect to $P \subseteq N$) Let P denote a set of participants. The payoff function satisfies the condition of *the preservation of the participation incentive with respect to* $P \subseteq N$ if the following condition is satisfied: for all $Q \subseteq N$ with $\sum_{j\in Q}\alpha_j > \sum_{j\in P}\alpha_j$ and for all $i \in P \cap Q$, if $u_i(P \setminus \{i\}) - u_i(P) \geq 0$, then $u_i(Q \setminus \{i\}) - u_i(Q) > 0$.

For every P, Q with $\sum_{j\in Q}\alpha_j > \sum_{j\in P}\alpha_j$ and $i \in P \cap Q$, Condition 1 holds if $u_i(Q \setminus \{i\}) - u_i(Q) > u_i(P \setminus \{i\}) - u_i(P)$. This inequality is satisfied if $v(y(\alpha))$ is a concave function with respect to α and $\frac{\beta}{\alpha+\beta}c(y(\alpha+\beta))$ is an increasing function in both α and β, in which $y(\alpha) := y^P$ such that $\alpha = \sum_{j\in P}\alpha_j$. For example, consider a case in which $v(y) = y^{\frac{1}{a}}$ $(a \geq 2)$ and $c(y) = y$. If P designates a set of participants, $\alpha_P := \sum_{j\in P}\alpha_j$, and $\alpha_{P\setminus\{i\}} := \alpha_P - \alpha_i$ for every $i \in P$, then

$$y(\alpha_P) = \left(\frac{\alpha_P}{a}\right)^{\frac{a}{a-1}},$$

$$v(y(\alpha_P)) = \left(\frac{\alpha_P}{a}\right)^{\frac{1}{a-1}}, \text{ and}$$

$$\frac{\alpha_i}{\alpha_i + \alpha_{P\setminus\{i\}}}c(y(\alpha_i + \alpha_{P\setminus\{i\}})) = \alpha_i\left(\frac{\alpha_i + \alpha_{P\setminus\{i\}}}{a^a}\right)^{\frac{1}{a-1}}.$$

Thus, $v(y(\alpha_P))$ is concave in α_P, and $\frac{\alpha_i}{\alpha_i+\alpha_{P\setminus\{i\}}}c(y(\alpha_i + \alpha_{P\setminus\{i\}}))$ increases in both α_i and $\alpha_{P\setminus\{i\}}$. Later, we prove that Condition 1 with respect to some P is satisfied if all agents have identical preferences.

Condition 2 (Strictness for non-participants) A strategy profile $s \in S^*$ is *strict for non-participants* if $u_i(P^s) > u_i(P^s \cup \{i\})$ for all $i \notin P^s$, where $P^s := \{j \in N |\, s_j = I\}$.

Proposition 1 Let $s^* \in S^n$ be a Nash equilibrium of the participation game, and let $P^{s*} := \{j \in N |\, s^*_j = I\}$. If Condition 1 with respect to P^{s*} and s^* is strict for non-participants, then s^* is a coalition-proof equilibrium, and it is also a Nash equilibrium that is not Pareto-dominated by any other Nash equilibrium.

Proof. Let s^* be a Nash equilibrium that is strict for non-participants, and let $P^{s*} := \{j \in N |\, s^*_j = I\}$. Suppose that Condition 1 with respect to P^{s*} holds. Suppose, to the contrary, that s^* is not coalition-proof in the participation game. Then, there are coalitions $D \subset N$ and $\widetilde{s_D} \in S^{\#D}$, such that $\widetilde{s_D}$ is self-enforcing in $G|s^*_{-D}$ and no members of D are worse off and at least one member of D is better off. Let $\widetilde{P}$ be a set of participants at $(\widetilde{s}_D, s^*_{-D})$.

Claim 1 If $\sum_{j\in P^{s*}}\alpha_j \geq \sum_{j\in\widetilde{P}}\alpha_j$, then there is a member of D that is worse off after the deviation. Therefore, the deviation by D with $\sum_{j\in P^{s*}}\alpha_j \geq \sum_{j\in\widetilde{P}}\alpha_j$ is not profitable.

Proof of Claim 1 If $\widetilde{P} \subseteq P^{s*}$, then the deviation by D is not profitable by Lemma 3. We consider the case in which $i \in \widetilde{P}$ exists such that $i \notin P^{s*}$: i switches from O to I by the deviation. Payoff to i before the deviation is $u_i(P^{s*}) = \alpha_i v(y^{P^{s*}})$ and payoff to i after the deviation is $u_i(\widetilde{P}) = \alpha_i v(y^{\widetilde{P}}) - \frac{\alpha_i}{\sum_{j\in\widetilde{P}}} c(y^{\widetilde{P}})$. By Lemma 1, we have $y^{P^{s*}} \geq y^{\widetilde{P}}$ and $\frac{\alpha_i}{\sum_{j\in\widetilde{P}}} c(y^{\widetilde{P}}) > 0$. Therefore, payoff to i after the deviation is fewer than that before deviation and i is worse off by the deviation. **(End of Proof of Claim 1)**

Claim 2 If $\sum_{j\in P^{s*}} \alpha_j < \sum_{j\in\widetilde{P}} \alpha_j$, then $\widetilde{s}_D$ is not self-enforcing in $G|s^*_{-D}$.

Proof of Claim 2 It is noteworthy that $i \in \widetilde{P}$ exists such that $i \notin P^{s*}$: otherwise, $i \in P^{s*}$ for all $i \in \widetilde{P}$ and $\sum_{j\in P^{s*}} \alpha_j \geq \sum_{j\in\widetilde{P}} \alpha_j$. Since the deviation is profitable, we have $u_i(\widetilde{P}) \geq u_i(P^{s*})$ for the agent i. From strictness for non-participants of s^*, we have $u_i(P^{s*}) > u_i(P^{s*} \cup \{i\})$. Therefore, we have

$$u_i(\widetilde{P}) - u_i(P^{s*} \cup \{i\}) > 0. \tag{8}$$

It follows from (8) that $\sum_{j\in\widetilde{P}} \alpha_j > \sum_{j\in P^{s*}} \alpha_j + \alpha_i$: otherwise, we obtain $u_i(\widetilde{P}) \leq u_i(P^{s*} \cup \{i\})$ from (3) of Lemma 2. It is clear that Condition 1 with respect to P^{s*} implies Condition 1 with respect to $P^{s*} \cup \{i\}$. From Condition 1 with respect to $P^{s*} \cup \{i\}$ and the strictness of s^*,

$$0 < u_i(\widetilde{P} \setminus \{i\}) - u_i(\widetilde{P}).$$

Thus, we have $u_i(\widetilde{P} \setminus \{i\}) > u_i(\widetilde{P})$. We confirm from this inequality that $\widetilde{s}_D$ is not a Nash equilibrium of $G|s^*_{-D}$, which indicates that $\widetilde{s}_D$ is not self-enforcing in $G|s^*_{-D}$. **(End of Proof of Claim 2)**

It follows from Claim 1 and Claim 2 that no group of agents deviates from s^* profitably in a self-enforcing way. Therefore, s^* is a coalition-proof equilibrium of the participation game.

By substituting N for D in the proof above, we can show that s^* is a Nash equilibrium that is not Pareto-dominated by any other Nash equilibrium. ■

A Nash equilibrium satisfying Conditions 1 and 2 also belongs to the Pareto-efficient frontier within the set of Nash equilibria. Hence, if, in the participation game, there are multiple Nash equilibria and one of the Nash equilibria satisfies these conditions, a Nash equilibrium in the Pareto-efficient frontier can be achieved by a coalition-proof equilibrium; the negotiation among agents achieves Pareto-superior Nash equilibria. It is clear from the definition of coalition-proof equilibria that the set of coalition-proof equilibria coincides with the Pareto-efficient frontier of the set of Nash equilibria in every two-player game. However, the two sets do not necessarily coincide in games with more than two players. Bernheim et al. (1987) provided an example of a three-player game in which there are two Nash equilibria; one of them is coalition-proof, and the other is not, and the former is dominated by the latter. Thus, we can say that the relationship between the Pareto-superior Nash equilibria and the coalition-proof equilibria in the participation game is a notable feature of our model.

5.2 Coalition-proof equilibria and the number of participants

Although the coalition-proof equilibrium is a refinement of the Nash equilibrium, the number of participants at coalition-proof equilibria is not always unique. The following is such an example.

Example 3 (The multiple numbers of participants in coalition-proof equilibria) Consider a game with three agents. Let α_1, α_2, and α_3 be such that $\alpha_1 > \alpha_2 = \alpha_3$ and $\alpha_1 < \alpha_2 + \alpha_3$, for example $\alpha_1 = 5$, and $\alpha_2 = \alpha_3 = 3$. In this example, we assume that one unit of the private good yields one unit of the public good. The payoff matrix of this game appears in Table 2. In this example, there are two coalition-proof equilibria. One is $s = (s_1, s_2, s_3) = (O, I, I)$ and the other is $s' = (s'_1, s'_2, s'_3) = (I, O, O)$. Two agents participate in the mechanism at s, while one agent enters the mechanism at s'. Thus, the number of participants attained at coalition-proof equilibria is not unique in this example.

	I	O
I	13.75, 8.25, 8.25	10, 12, 6
O	15, 4.5, 4.5	7.5, 4.5, 2.25

I

	I	O
I	10, 6, 12	6.25, 7.5, 7.5
O	7.5, 2.25, 4.5	0, 0, 0

O

Table. 2 Payoff matrix of Example 3

As Example 3 shows, the number of participants at coalition-proof equilibria is not necessarily unique. In the following, we focus on a Nash equilibrium at which the number of participants is maximal within the set of Nash equilibria and establish a sufficient condition under which only the maximal number of participants can be supported as a coalition-proof equilibrium.

Let p^{max} be the maximal number of participants that is attained at Nash equilibria of G. Let $s^{max} \in S^n$ be a Nash equilibrium at which p^{max} agents choose I. Let us denote the set of participants at s^{max} by P^{max}.

Condition 3 $\alpha_i \geq \alpha_j$ for all $i \in P^{max}$ and all $j \in N \setminus P^{max}$.

Condition 3 means that all agents in P^{max} have at least as high marginal willingness to pay for the public good as the agents in $N \setminus P^{max}$. It is noteworthy that the condition is not satisfied in Example 3. The following proposition proves that p^{max} agents choose I in every coalition-proof equilibrium under Condition 3.

Proposition 2 Let p^{max} denote the maximal number of participants attained in the set of Nash equilibria of G. Let s^{max} be a Nash equilibrium at which $P^{max} := \{i \in N | s_i^{max} = I\}$ and $p^{max} = \#P^{max}$. If $s^{\max}$ is strict for non-participants, Conditon 1 with respect to P^{max} holds, and P^{max} satisfies Condition 3, then p^{max} is the unique number of participants that is achieved in the set of coalition-proof equilibria.

Before proving Proposition 2, we show the following lemma.

Lemma 4 Let s^{max} be a Nash equilibrium at which P^{max} is the set of participants and p^{max} agents choose I. Let us suppose that P^{max} satisfies Condition 3. Then, (i) no profiles of strategies with participation of p^{max} agents Pareto-dominate s^{max}, and (ii) s^{max} Pareto-dominates every strategy profile with the participation of fewer than p^{max} agents.

Proof of Lemma 4. Let $\widehat{s} \in S^n$ be a profile of strategies. Let $\widehat{P}$ be a set of participants that is attained in $\widehat{s}$, and let $\widehat{p}$ be the number of agents in $\widehat{P}$. We consider the following two cases: one is the case of $p^{max} = \widehat{p}$ and the other is the case of $p^{max} > \widehat{p}$.

First, we consider $p^{max} = \widehat{p}$. If $\widehat{P} = P^{max}$, then $\widehat{s}$ does not Pareto-dominate s^{max} trivially. Let us consider the case of $\widehat{P} \neq P^{max}$. Since $p^{max} = \widehat{p}$ and $\widehat{P} \neq P^{max}$, we have $\#[P^{max} \setminus \widehat{P}] = \#[\widehat{P} \setminus P^{max}] > 0$. For every agent $i \in \widehat{P} \setminus P^{max}$, we have

$$u_i(P^{max}) \geq u_i(P^{max} \cup \{i\}) > u_i(\widehat{P}). \tag{9}$$

The first inequality follows from the definition of Nash equilibrium, and the second inequality holds since

$$\begin{aligned} \sum_{j \in P^{max} \cup \{i\}} \alpha_j &= \sum_{j \in P^{max} \cap \widehat{P}} \alpha_j + \sum_{j \in P^{max} \setminus \widehat{P}} \alpha_j + \alpha_i \\ &\geq \sum_{j \in P^{max} \cap \widehat{P}} \alpha_j + \sum_{j \in \widehat{P} \setminus P^{max}} \alpha_j + \alpha_i \\ &> \sum_{j \in P^{max} \cap \widehat{P}} \alpha_j + \sum_{j \in \widehat{P} \setminus P^{max}} \alpha_j \\ &= \sum_{j \in \widehat{P}} \alpha_j. \end{aligned} \tag{10}$$

Note that $\alpha_k \geq \alpha_l$ for all $k \in P^{max} \setminus \widehat{P}$ and all $l \in \widehat{P} \setminus P^{max}$ from Condition 3. Hence, we obtain $\sum_{j \in P^{max} \setminus \widehat{P}} \alpha_j \geq \sum_{j \in \widehat{P} \setminus P^{max}} \alpha_j$. The second inequality of (10) follows from this, and the third inequality of (10) follows from $\alpha_i > 0$. It follows from (9) that every agent $i \in \widehat{P} \setminus P^{max}$ is worse off by switching from s^{max} to $\widehat{s}$. This completes the proof of (i).

Secondly, we consider the case of $p^{max} > \widehat{p}$. Note that $\#[P^{max} \setminus \widehat{P}] > \#[\widehat{P} \setminus P^{max}]$ must be satisfied in this case. If $\widehat{s} \in R(s^{max})$, then s^{max} Pareto-dominates $\widehat{s}$ by Lemma 3. If $\widehat{s} \notin R(s^{max})$, then the following claim is satisfied.

Claim 3 It follows that $\#[P^{max} \setminus \widehat{P}] > \#[\widehat{P} \setminus P^{max}] \geq 1$.

Proof of Claim 3. Since $\widehat{s} \notin R(s^{max})$, $\widehat{P} \setminus P^{max}$ is non-empty. Thus, $\#[\widehat{P} \setminus P^{max}] \geq 1$. We obtain $\#[P^{max} \setminus \widehat{P}] > \#[\widehat{P} \setminus P^{max}]$ because $p^{max} > \widehat{p}$. **(End of Proof of Claim 3)**

Claim 4 Every agent i with $s_i^{max} = \widehat{s}_i$ is worse off:

$$\begin{aligned} &u_i(P^{max}) > u_i(\widehat{P}) \text{ for all } i \in P^{max} \cap \widehat{P}, \quad \text{and} \\ &u_i(P^{max}) > u_i(\widehat{P}) \text{ for all } i \in N \setminus (P^{max} \cup \widehat{P}). \end{aligned} \tag{11}$$

Proof of Claim 4. We first show $\sum_{j \in P^{max}} \alpha_j > \sum_{j \in \widehat{P}} \alpha_j$. Note that

$$\sum_{j \in P^{max}} \alpha_j = \sum_{j \in P^{max} \cap \widehat{P}} \alpha_j + \sum_{j \in P^{max} \setminus \widehat{P}} \alpha_j. \tag{12}$$

It follows from Condition 3 that $\alpha_k \geq \alpha_l$ for all $k \in P^{max} \setminus \widehat{P}$ and all $l \in \widehat{P} \setminus P^{max}$. From this condition and Claim 3, we have $\sum_{j \in P^{max} \setminus \widehat{P}} \alpha_j > \sum_{j \in \widehat{P} \setminus P^{max}} \alpha_j$. Hence,

$$\begin{aligned}(12) &> \sum_{j \in P^{max} \cap \widehat{P}} \alpha_j + \sum_{j \in \widehat{P} \setminus P^{max}} \alpha_j \\ &= \sum_{j \in \widehat{P}} \alpha_j.\end{aligned}$$

Therefore, we obtain $\sum_{j \in P^{max}} \alpha_j > \sum_{j \in \widehat{P}} \alpha_j$. It follows from Lemma 2 and $\sum_{j \in P^{max}} \alpha_j > \sum_{j \in \widehat{P}} \alpha_j$ that (11) holds. **(End of Proof of Claim 4)**

Claim 5 Every agent i with $s^{max} = I$ and $\widehat{s}_i = O$ is worse off: $u_i(P^{max}) \geq u_i(\widehat{P})$ for every $i \in P^{max} \setminus \widehat{P}$.

Proof of Claim 5. Let $i \in P^{max} \setminus \widehat{P}$. By the definition of Nash equilibrium, $u_i(P^{max}) \geq u_i(P^{max} \setminus \{i\})$. We first obtain the following equations:

$$\begin{aligned}\sum_{j \in P^{max} \setminus \{i\}} \alpha_j &= \sum_{j \in P^{max}} \alpha_j - \alpha_i \\ &= \sum_{j \in P^{max} \cap \widehat{P}} \alpha_j + \sum_{j \in (P^{max} \setminus \widehat{P}) \setminus \{i\}} \alpha_j.\end{aligned} \tag{13}$$

By Condition 3 and $\#[P^{max} \setminus \widehat{P}] - 1 \geq \#[\widehat{P} \setminus P^{max}]$, we obtain $\sum_{j \in (P^{max} \setminus \widehat{P}) \setminus \{i\}} \alpha_j \geq \sum_{j \in \widehat{P} \setminus P^{max}} \alpha_j$. Therefore,

$$\begin{aligned}(13) &\geq \sum_{j \in P^{max} \cap \widehat{P}} \alpha_j + \sum_{j \in \widehat{P} \setminus P^{max}} \alpha_j \\ &= \sum_{j \in \widehat{P}} \alpha_j.\end{aligned}$$

It is straightforward from Lemma 2 to show $u_i(P^{max}) \geq u_i(\widehat{P})$. **(End of Proof of Claim 5)**

Claim 6 Every agent i with $s_i^{max} = O$ and $\widehat{s}_i = I$ is worse off: $u_i(P^{max}) > u_i(\widehat{P})$ for every $i \in \widehat{P} \setminus P^{max}$.

Proof of Claim 6. Let $i \in \widehat{P} \setminus P^{max}$. By the definition of Nash equilibrium, $u_i(P^{max}) \geq u_i(P^{max} \cup \{i\})$. We show $\sum_{j \in P^{max} \cup \{i\}} \alpha_j > \sum_{j \in \widehat{P}} \alpha_j$ to prove $u_i(P^{max} \cup \{i\}) > u_i(\widehat{P})$. Note that the following conditions are satisfied:

$$\sum_{j\in P^{max}\cup\{i\}} \alpha_j = \alpha_i + \sum_{j\in P^{max}\cap\widehat{P}} \alpha_j + \sum_{j\in P^{max}\setminus\widehat{P}} \alpha_j. \tag{14}$$

It follows from Condition 3 and Claim 3 that $\sum_{j\in P^{max}\setminus\widehat{P}} \alpha_j > \sum_{j\in\widehat{P}\setminus P^{max}} \alpha_j$. Thus,

$$\begin{aligned}(14) &> \sum_{j\in P^{max}\cap\widehat{P}} \alpha_j + \sum_{j\in\widehat{P}\setminus P^{max}} \alpha_j + \alpha_i \\ &> \sum_{j\in P^{max}\cap\widehat{P}} \alpha_j + \sum_{j\in\widehat{P}\setminus P^{max}} \alpha_j \\ &= \sum_{j\in\widehat{P}} \alpha_j.\end{aligned}$$

Therefore, we have $\sum_{j\in P^{max}\cup\{i\}} \alpha_j > \sum_{j\in\widehat{P}} \alpha_j$, which indicates $u_i(P^{max}) \geq u_i(P^{max}\cup\{i\}) > u_i(\widehat{P})$. **(End of Proof of Claim 6)**

By Claim 4, 5, and 6, s^{max} Pareto-dominates every strategy profile with participation of fewer than p^{max} agents. Therefore, (ii) holds. **(End of Proof of Lemma 4)**

Proof of Proposition 2. Let p^{max} denote the maximal number of participants that is supportable as Nash equilibria and let s^{max} be a Nash equilibrium with the participation of p^{max} agents. From Proposition 1, s^{max} is a coalition-proof equilibrium of G. Since p^{max} is the maximal number of participants within the set of Nash equilibria, the participation of more than p^{max} agents is not achieved at coalition-proof equilibria. By Lemma 4, no strategy profiles with participation of fewer than p^{max} agents are coalition-proof, because such strategy profiles are Pareto-dominated by the coalition-proof equilibrium s^{max}. ■

Remark 1 It is noteworthy that Propositions 1 and 2 hold if Lemmas 1 and 2 are satisfied. Thus, the assumption that all agents have a quasi-linear preference does not matter. The same results hold when agents have the preferences that are represented by the Cobb-Douglas utility function and the same initial endowments of the private good. In the economy, Conditions 1 and Lemma 2 are satisfied. Thus, if a Nash equilibrium is strict for non-participants, then the equilibrium is coalition-proof. If there is a Nash equilibrium that satisfies Condition 3, then the number of participants is solely determined by the maximal number of participants attained at Nash equilibria.[*4]

Remark 2 Yi (1999) investigated the equivalence between the set of coalition-proof equilibria and the Pareto-efficient frontier of the set of Nash equilibria. Yi (1999) considered a game in which the strategy space of each player is a subset of the real line and he showed that, if a game satisfies *anonymity*, *monotone externality*[*5], and *strategic substitutability*[*6], then the set of coalition-proof equilibria and the Pareto-efficient frontier of the set of Nash equilibria coincide. If strategies 1 and 0 designate participation and non-participation in the mechanism, respectively, then the anonymity condition is satisfied in the participation game only when agents are identical. Since our model allows for heterogeneous agents, Yi (1999)'s results cannot be applied to our model in order to characterize the set of coalition-proof equilibria.

Furthermore, we mention that the coalition-proof equilibrium of our model is based on *weak domination*, while that of Yi (1999) is based on *strict domination*. These two coalition-proof equilibria are not necessarily related by inclusion. (See Konishi et al. (1999) and Shinohara (2005)) Hence, it is not trivial that the set of coalition-proof equilibria under weak domination coincides with the strictly Pareto-efficient frontier of the set of Nash equilibria under Yi (1999)'s conditions.

Remark 3 Thoron (1998) examined the coalition-proof equilibria of a cartel formation game, which is similar to the participation game in a public good mechanism. In the cartel formation game, each firm decides whether or not to join the cartel. Only the firms that join the cartel follow its agreements, and the other firms behave independently. However, Thoron (1998) considered a case in which every firm is identical and the condition that is satisfied in Thoron (1998)'s model differs from ours. Hence, we can not use the results in Thoron (1998) to clarify the properties of coalition-proof equilibria in the participation game.[*7]

[*4] See the Appendix for a detailed discussion.

[*5] A game satisfies monotone externality if, for all $i \in N$, all $s_i \in S$, and all s_{-i} and $\widehat{s}_{-i} \in S^{n-1}$, if $\sum_{j \neq i} s_j > \sum_{j \neq i} \widehat{s}_j$, then either $U_i(s_i, s_{-i}) \geq U_i(s_i, \widehat{s}_{-i})$ or $U_i(s_i, s_{-i}) \leq U_i(s_i, \widehat{s}_{-i})$ holds. If the former holds, the condition means *positive* externalities, and it represents *negative* externalities if the latter is satisfied.

[*6] A game satisfies strategic substitutability if, for all $i \in N$, all s_i, $\widehat{s}_i \in S$, and all s_{-i}, $\widehat{s}_{-i} \in S^{n-1}$, if $s_i > \widehat{s}_i$ and $\sum_{j \neq i} s_j > \sum_{j \neq i} \widehat{s}_j$, then $U_i(s_i, s_{-i}) - U_i(\widehat{s}_i, s_{-i}) < U_i(s_i, \widehat{s}_{-i}) - U_i(\widehat{s}_i, \widehat{s}_{-i})$.

[*7] See Remark 4 for a detailed discussion.

6 Applications of the main results

6.1 Application to the case of identical agents

We consider the case in which all agents have preference relations that is represented by the same quasi-utility function: $\alpha_1 = \cdots = \alpha_n$. We normalize $\alpha_1 = \cdots = \alpha_n = 1$.[*8] Note that the payoff to participants and the payoff to non-participants depend on the number of participants in the case of identical agents. We will introduce the following notation for convenience.

Definition 8 Let $u_i : \{0, 1, \ldots, n-1\} \times \{I, O\} \to \mathbb{R}_+$ denote a payoff function of agent i that depends on the number of agents other than i and i's participation decision. If $p \in \{0, 1, \ldots, n-1\}$ designates the number of participants other than i and $s_i \in \{I, O\}$ designates i's participation decision, then i receives the payoff $u_i(p, s_i)$.

Let y^p be the level of the public good when p agents choose I for every $p \in \{0, \ldots, n\}$. Note that, for every $i \in N$ and for every p, $u_i(p, I) = v(y^{p+1}) - \frac{c(y^{p+1})}{p+1}$ and $u_i(p, O) = v(y^p)$. Since agents are identical, we have $u_i(p, I) = u_j(p, I)$ for all $i,\ j \in N$ and for all $p \leq n-1$ and $u_i(p, O) = u_j(p, O)$ for all $i,\ j \in N$ and for all $p \leq n-1$. Therefore, we can hereafter omit agents' indices of the payoff functions.

From the direct application of Lemma 1 and Lemma 2, the following properties are satisfied.

Lemma 5 For all numbers of participants $p,\ q \in \{0, \ldots, n\}$, if $p > q$, then $y^p > y^q$.

Lemma 6 The payoff function of participants and that of non-participants are increasing functions with respect to the number of participants other than them: (i) for all $p,\ q \in \{0, \ldots, n-1\}$, if $p > q$, then $u(p, O) > u(q, O)$ and (ii) for all $p,\ q \in \{0, \ldots, n-1\}$, if $p > q$, then $u(p, I) > u(q, I)$.

Using Propositions 1 and 2, we show that the number of participants in coalition-proof equilibria is uniquely determined by the largest number of participants in the set of Nash equilibria.

[*8] The results in this subsection also hold in the case of $\alpha_1 = \cdots = \alpha_n \neq 1$.

Proposition 3 Suppose that agents' preferences are represented by the same quasi-linear utility function. Let $p^{max} \leq n$ be the maximal number of participants in the set of Nash equilibria: $p^{max} = \max_{s \in NE(G)} \#\{i \in N \mid s_i = I\}$, in which $NE(G)$ is the set of Nash equilibria of G. Then, the set of coalition-proof equilibria coincides with the set of Nash equilibria at which p^{max} agents choose I in the participation game.

Let $s^{max} \in S^n$ be a Nash equilibrium at which p^{max} agents select participation. As preparations for proving Proposition 3, we show the following lemma.

Lemma 7 If $p^{max} < n$, then $u(t, O) - u(t, I) > 0$ for all $t \in \{p^{max}, \ldots, n-1\}$.

Proof of Lemma 7. Suppose that $p^{max} < n$ and that there exist $t \in \{p^{max}, \ldots, n-1\}$ such that $u(t, O) - u(t, I) \leq 0$. Hence, $u(t, O) \leq u(t, I)$.

If $t = n-1$, then $u(n-1, I) \geq u(n-1, O)$, which indicates that the participation of n agents is supported as a Nash equilibrium. This contradicts $p^{max} < n$. Consider a case in which $t < n-1$. Since $t+1 > p^{max}$, we must have $u(t+1, O) < u(t+1, I)$. Otherwise, condition $u(t+1, O) \geq u(t+1, I)$, together with $u(t, O) \leq u(t, I)$, implies that the participation of $t+1$ agents can be supported as a Nash equilibrium and p^{max} is not the maximal number of participants in the set of Nash equilibria, which is a contradiction. By the same way, the following inequalities are obtained:

$$\begin{aligned} u(t+2, O) &< u(t+2, I), \\ u(t+3, O) &< u(t+3, I), \\ &\vdots \\ \text{and } u(n-1, O) &< u(n-1, I). \end{aligned} \tag{15}$$

From (15), participation of n agents is a Nash equilibrium, which is a contradiction. Therefore, we have $u(t, O) - u(t, I) > 0$ for all $t \in \{p^{max}, \ldots, n-1\}$. ■

Proof of Proposition 3. Note that Condition 3 holds at s^{max}in the case of identical agents, trivially. If $p^{max} = n$, no coalitional deviations can improve their members' payoffs by Lemma 4. Then, s^{max} is a strong equilibrium of G. Hence, it is a coalition-proof equilibrium. Let us consider a case in which $p^{max} \leq n-1$. It is immediate from Lemma 7 that s^{max} is strict for non-participants. Since $P^{max} := \{i \in N | s_i^{max} = I\}$ is supported as a Nash equilibrium, we have $u(O, p^{max} - 1) - u(I, p^{max} - 1) \leq 0$. If

$u(O, p^{max} - 1) - u(I, p^{max} - 1) < 0$, then Condition 1 with respect to P^{max} holds, vacuously. If $u(O, p^{max} - 1) - u(I, p^{max} - 1) = 0$, then Condition 1 with respect to P^{max} holds from Lemma 7. Therefore, s^{max} is a coalition-proof equilibrium in G from Proposition 1. From Proposition 2, the number of participants at coalition-proof equilibria is equal to p^{max}. Therefore, all coalition-proof equilibria are Nash equilibria with the participation of p^{max} agents. ■

Corollary 1 A coalition-proof equilibrium exists in the participation game with identical agents .

Proof. It is clear from Proposition 3 that the existence of a Nash equilibrium implies that of a coalition-proof equilibrium in the participation game. The Nash equilibrium is shown to exist in the same way as in d'Aspremont et al. (1983). ■

Remark 4 Thoron (1998) obtained a similar result to Proposition 3 in the cartel formation problem. However, Thoron (1998) used different conditions from ours. She used the following two conditions: (i) $u(p, O) > u(q, O)$ for all $p,\ q \in \{0, \ldots, n-1\}$ such that $p > q$, and (ii) $u(p, O) > u(p-1, I)$ for all $p \in \{1, \ldots, n-1\}$. Although conditions (i) and (ii) are satisfied in our model, we do not use condition (ii) in the proof; we use the condition $u(p, I) > u(q, I)$ for all $p,\ q \in \{0, \ldots, n-1\}$ such that $p > q$, instead, and this condition is not satisfied in the cartel formation game of the Cournot competition.

Remark 5 Saijo and Yamato (1999) studied the *symmetric Cobb-Douglas economy*, in which all agents have the preference that is represented by the same Cobb-Douglas utility function and all agents are assumed to have the same initial endowments of the private good. From discussion in Remark 1, Lemma 1, Lemma 2, and Condition 1 hold. If p^{max} designates the greatest number of participants supported as a Nash equilibrium and s^{max} designates a Nash equilibrium with the maximal number of participants, then s^{max} is shown to be strict for non-participants in a similar way to Lemma 7 and Condition 3 is satisfied at s^{max} in the symmetric Cobb-Douglas economy. Hence, the participation game in the symmetric Cobb-Douglas economy has a coalition-proof equilibrium, and only the participation of p^{max} agents is supportable as coalition-proof equilibria.

We provide another characterization of the set of coalition-proof equilibria in the participation game.

Corollary 2 In the participation game with identical agents, a strategy profile is a coalition-proof equilibrium if and only if it is a Nash equilibrium that is not Pareto-dominated by any other Nash equilibrium.

Proof. From Proposition 3, the set of coalition-proof equilibria coincides with that of Nash equilibria with the participation of p^{max} agents. It follows from Lemma 4 that every Nash equilibrium that is not Pareto-dominated by any other Nash equilibrium is that with p^{max} participants and no Nash equilibria Pareto-dominate Nash equilibria with the participation of p^{max} agents. ■

6.2 Application to the case of a square-root benefit function and a linear cost function

In the case of $v(y) = \sqrt{y}$ and $c(y) = y$, there may be multiple Nash equilibria in the participation game, as Example 3 shows. However, Example 3 indicates that the number of participants at coalition-proof equilibria is not always unique in this case. In this subsection, applying Propositions 1 and 2, we investigate the existence of coalition-proof equilibria and identify a necessary and sufficient condition under which the number of participants is unique at coalition-proof equilibria.

We consider the case of $v(y) = \sqrt{y}$ and $c(y) = y$. Let P be a set of participants. Then, y^P maximizes $\sum_{j\in P} \alpha_j \sqrt{y} - y$; hence, we obtain $y^P = \left(\frac{\sum_{j\in P}\alpha_j}{2}\right)^2$. The payoff functions to the participants and the non-participants are as follows:

$$u_i(P) = \begin{cases} \frac{\alpha_i(\sum_{j\in P}\alpha_j)}{4} & \text{if } i \in P, \text{ and} \\ \frac{\alpha_i(\sum_{j\in P}\alpha_j)}{2} & \text{if } i \notin P. \end{cases} \tag{16}$$

It follows from (16) that

$$u_i(P \setminus \{i\}) - u_i(P) = \frac{\alpha_i(\sum_{j\in P\setminus\{i\}}\alpha_j - \alpha_i)}{4} \lesseqgtr 0 \text{ if } \sum_{j\in P\setminus\{i\}} \alpha_j \lesseqgtr \alpha_i \tag{17}$$

for all $i \in P$. Thus, i chooses O if the sum of the marginal willingness to pay of participants other than i is greater than i's willingness to pay, and he chooses

I if the sum of the marginal willingness to pay of participants other than i is less than i's willingness to pay. From (17), the Nash-equilibrium sets of participants are characterized as follows.

Proposition 4 A set of participants $P \subseteq N$ is a Nash-equilibrium set of participants in the participation game if and only if P satisfies (i) $\sum_{j \in P \setminus \{i\}} \alpha_j \leq \alpha_i$ for all $i \in P$ and (ii) $\sum_{j \in P} \alpha_j \geq \alpha_i$ for all $i \notin P$.

Proposition 5 Let $n \geq 2$. There is no Nash equilibrium at which more than two agents participate in the mechanism.

Proof. Suppose that there is a Nash equilibrium at which P with $\#P \geq 3$ is the set of participants. Then, we have $\alpha_i \geq \sum_{j \in P \setminus \{i\}} \alpha_j$ for all $i \in P$. Summing up this inequality for all $i \in P$ yields $\sum_{i \in P} \alpha_i \geq \sum_{i \in P} \sum_{j \in P \setminus \{i\}} \alpha_j = (\#P - 1) \sum_{i \in P} \alpha_i$. However, $\sum_{i \in P} \alpha_i < (\#P - 1) \sum_{i \in P} \alpha_i$ must be satisfied because $\#P \geq 3$. This is a contradiction. ■

Proposition 6 In the participation game with $n \geq 3$, there is a Nash equilibrium that is strict for non-participants.

Proof. If agents' utility functions are identical ($\alpha_1 = \cdots = \alpha_n$), then the Nash equilibria with the maximal number of participants within the set of the Nash-equilibrium number of participants are strict for non-participants from the analysis of the previous subsection. Consider the case in which $\alpha_i \neq \alpha_j$ for some $i,\ j \in N$. Let $i^* \in \arg\max_{i \in N} \alpha_i$. If $\{i^*\} = \arg\max_{i \in N} \alpha_i$, then $\{i^*\}$ is supported as a Nash equilibrium that is strict for non-participants from Proposition 4. If $\arg\max_{i \in N} \alpha_i$ consists of more than one agent, then $\{i^*, j^*\} \subseteq \arg\max_{i \in N} \alpha_i$ is attained at a Nash equilibrium that is strict for non-participants. Therefore, there is a Nash equilibrium that is strict for non-participants. ■

Let s denote a Nash equilibrium that is strict for non-participants and let $P^s := \{j \in N | s_j = I\}$. Let $Q \subseteq N$ be a set of participants such that $\sum_{j \in Q} \alpha_j > \sum_{j \in P} \alpha_j$ and $i \in P \cap Q$. Since $\sum_{j \in Q \setminus \{i\}} \alpha_j > \sum_{j \in P \setminus \{i\}} \alpha_j$, Condition 1 with respect to P^s is satisfied from (17). Hence, it follows from Proposition 1 that the participation game has a coalition-proof equilibrium.

Corollary 3 There is a coalition-proof equilibrium in the participation game.

As Example 4 shows, the number of participants at coalition-proof equilibria may be multiple. The following corollary provides a necessary and sufficient condition under which two agents enter the mechanism at every coalition-proof equilibrium.

Corollary 4 The number of participants at coalition-proof equilibria is solely determined by two if and only if $\arg\max_{i \in N} \alpha_i$ is not a singleton set.

Proof. Suppose that $\arg\max_{i \in N} \alpha_i$ is not a singleton set. Let $\{i^*, j^*\} \subseteq \arg\max_{i \in N} \alpha_i$. Then, we have $\alpha_k \geq \alpha_l$ for all $k \in \{i^*, j^*\}$ and all $l \in N \setminus \{i^*, j^*\}$; hence, Condition 3 with respect to $\{i^*, j^*\}$ is satisfied. It follows from Proposition 2 that two agents choose I at every coalition-proof equilibrium. Conversely, suppose that the number of participants at coalition-proof equilibria is solely determined by two. If $\arg\max_{i \in N} \alpha_i$ is a singleton set and $\arg\max_{i \in N} \alpha_i = \{i^*\}$, then $\{i^*\}$ is supportable as a Nash equilibrium that is strict for non-participants. Therefore, $\{i^*\}$ is also attained at a coalition-proof equilibrium, which is a contradiction. ■

Finally, using the results in this subsection, we characterize the set of coalition-proof equilibria in the following example.

Example 4 Let $N = \{1, 2, 3, 4\}$, $v(y) = \sqrt{y}$, $\alpha_1 = \alpha_2 = 3$, $\alpha_3 = \alpha_4 = 2$, and $c(y) = y$. We confirm from Proposition 4 that $\{1\}$, $\{2\}$, $\{1, 2\}$, and $\{3, 4\}$ are supported as Nash equilibria. Note that $\{1, 2\}$ and $\{3, 4\}$ are supported as Nash equilibria that are strict for non-participants. Therefore, $\{1, 2\}$ and $\{3, 4\}$ are sets of participants attained at coalition-proof equilibria from Proposition 1. Since $\{1, 2\}$ satisfies Condition 3, two agents choose participation at every coalition-proof equilibria, which implies that only $\{1, 2\}$ and $\{3, 4\}$ are the sets of participants that are achieved at coalition-proof equilibria.

In the case of $n = 4$, we need to consider one-agent games, two-agent games, three-agent games, and the whole game, in that order. Since there are four one-agent coalitions, six-agent coalitions, and four three-agent coalitions in this case, it is time-consuming to identify the set of coalition-proof equilibria according to the definition of coalition-proof equilibria. However, from the results in this paper, we can characterize the set of coalition-proof equilibria by just checking Conditions 1, 2, and

3. Therefore, by applying these results, we more easily characterize the equilibrium set of the participation game.

7 Concluding Remarks

In this paper, we have investigated the coalition-proof equilibria of the participation game in a public good mechanism. First, we provided sufficient conditions for a Nash equilibrium to be a coalition-proof equilibrium. We showed that a Nash equilibrium is coalition-proof if the condition of the preservation of the participation incentive holds at the Nash equilibrium and the Nash equilibrium is strict for non-participants. Secondly, we examined the number of participants. Focusing on the maximal number of participants, we established sufficient conditions under which the number of participants at coalition-proof equilibria is solely determined by the maximal number. Finally, we applied these results to a case in which agents have an identical preference and one in which agents have square-root benefit functions and the cost function of the public good is linear, and we identified the coalition-proof equilibria of the participation game.

The participation game may have multiple Nash equilibria. However, few studies have focused on which Nash equilibria are more likely to occur. We considered the possibility that agents coordinate the participation decision, and we clarified the characteristics of coalition-proof equilibria. Our results could be used to discuss that one type of Nash equilibrium is more 'focal' than another. Our results would also be useful to characterize the coalition-proof equilibria in cartel formation games and participation games in international environmental agreements as well as the participation game in a public good mechanism.

Appendix: Cobb-Douglas Economy

It is noteworthy that the results presented in this paper hold in every economic domain in which Lemmas 1 and 2 are satisfied. For example, our results are satisfied in the following Cobb-Douglas economy.

Let us denote agent i's utility function by $V_i(x_i, y) = (1 - \alpha_i) \ln x_i + \alpha_i \ln y$, where $\alpha_i \in (0, 1)$, x_i is a consumption of private good i and y is the level of the public good. Every agent is assumed to have the same initial endowment of private good $\omega > 0$.

Let us suppose that y units of the private good are needed for the provision of y units of the public good.

Let P be a set of participants, and let $\beta_i^P \in (0,\, 1)$ be a cost-share rate of the public good: $\sum_{j \in P} \beta_j = 1$. Then, in the ratio allocation, $(y^P, (x_i^P)_{i \in P})$ is as follows:

$$\begin{aligned} \beta_i^P &= \frac{\alpha_i}{\sum_{j \in P} \alpha_j} \text{ for every } i \in P. \\ y^P &= \sum_{j \in P} \alpha_j \, \omega. \\ x_i^P &= \omega - \beta_i^P y^P = \omega(1 - \alpha_i) \text{ for every } i \in P. \end{aligned}$$

We then have $u_i(P) = (1 - \alpha_i) \ln[\omega(1 - \alpha_i)] + \alpha_i \ln[\sum_{j \in P} \alpha_j \omega]$ for every $i \in P$, and $u_i(P) = (1 - \alpha_i) \ln \omega + \alpha_i \ln[\sum_{j \in P} \alpha_j \omega]$ for every $i \notin P$. Therefore, the payoff functions satisfy the following properties, which are the same as those in Lemma 2:

(1) For all $i \in N$ and for all $P,\ \widetilde{P} \subseteq N$, if $i \in P \cap \widetilde{P}$ and $\sum_{j \in P} \alpha_j > \sum_{j \in \widetilde{P}} \alpha_j$, then $u_i(P) > u_i(\widetilde{P})$.

(2) For all $i \in N$ and for all $P,\ \widetilde{P} \subseteq N$, if $i \notin P \cup \widetilde{P}$ and $\sum_{j \in P} \alpha_j > \sum_{j \in \widetilde{P}} \alpha_j$, then $u_i(P) > u_i(\widetilde{P})$.

We obtain

$$\begin{aligned} & u_i(P \setminus \{i\}) - u_i(P) \\ &= (1 - \alpha_i) \left(\ln \omega - \ln[\omega(1 - \alpha_i)] \right) + \alpha_i \left(\ln[\sum_{j \in P \setminus \{i\}} \alpha_j \omega] - \ln[\sum_{j \in P} \alpha_j \omega] \right) \\ &= (1 - \alpha_i) \ln \frac{1}{1 - \alpha_i} + \alpha_i \ln \frac{\sum_{j \in P \setminus \{i\}} \alpha_j}{\alpha_i + \sum_{j \in P \setminus \{i\}} \alpha_j}. \end{aligned} \tag{18}$$

From (18), $u_i(P \setminus \{i\}) - u_i(P)$ decreases with respect to $\sum_{j \in P \setminus \{i\}} \alpha_j$. Therefore, for all $P \subseteq N$, for all $Q \subseteq N$ with $\sum_{j \in Q} \alpha_j > \sum_{j \in P} \alpha_j$, and for all $i \in P \cap Q$, we have $u_i(P \setminus \{i\}) - u_i(P) < u_i(Q \setminus \{i\}) - u_i(Q)$, which indicates that Condition 1 with respect to every P is satisfied.

Therefore, it follows from Proposition 1 and Proposition 2 that the following statements hold in the Cobb-Douglas economy.

Corollary 5 Let $s^* \in S^n$ be a Nash equilibrium of the participation game. If s^* is strict for non-participants, then s^* is a coalition-proof equilibrium and it is also a Nash equilibrium that is not Pareto-dominated by any other Nash equilibrium.

Corollary 6 Let p^{max} denote the maximal number of participants attained in the set of Nash equilibria of G. Let s^{max} be a Nash equilibrium at which $P^{max} := \{i \in N | s_i^{max} = I\}$ and $p^{max} = \#P^{max}$. If $s^{\max}$ is strict for non-participants and P^{max} satisfies Condition 3, then p^{max} is the unique number of participants that is achieved in the set of coalition-proof equilibria.

References

[1] d'Aspremont, C., Jaskold Gabszewics, A. Jacquemin, and J. A. Weymark (1983) "On the Stability of Collusive Price Leadership," *Canadian Journal of Economics*, vol.16, 17-25.

[2] Aumann, R. (1959) "Acceptable Points in General Cooperative n-person Games," in *Contributions to the theory of games IV* by H. W. Kuhn and R. D. Luce, Eds., Princeton University Press: Princeton, 287-324.

[3] Bernheim, D., and M. Whinston (1986) "Menu Auctions, Resource Allocation, and Economic Influence," *Quarterly Journal of Economics*, vol.101,1-31.

[4] Bernheim, D., B. Peleg and M. Whinston (1987) "Coalition-proof Nash Equilibria I. Concepts," *Journal of Economic Theory*, vol.42, 1-12.

[5] Carraro C., and D. Siniscalco (1993) "Strategies for the International Protection of the Environment," *Journal of Public Economics*, vol.52, 309-328.

[6] Carraro, C., and D. Siniscalco (1998) "International Environmental Agreements: Incentives and Political Economy," *European Economic Review*, vol.42, 561-572.

[7] Corchon, L. and S. Wilkie (1996) "Double Implementation of the Ratio Correspondence by a Market Mechanism," *Economic Design*, vol.2, 325-337.

[8] Furusawa, T. and H. Konishi (2007) "Contributing or Free-riding?: A Theory of Endogeneous Lobby Formation," mimeo, Boston College.

[9] Groves, T. and J. Ledyard (1977) "Optimal Allocation of Public Goods: A Solution to the Free Rider Problem," *Econometrica*, vol.45, 783-811.

[10] Hurwicz, L. (1979) "Outcome Functions Yielding Walrasian and Lindahl Allocations at Nash Equilibrium Points," *Review of Economic Studies*, vol.46, 217-225.

[11] Kaneko, M. (1977a) "The Ratio Equilibrium and a Voting Game in a Public Goods Economy," *Journal of Economic Theory*, vol.16, 123-136.

[12] Kaneko, M. (1977b) "The Ratio Equilibria and the Core of the Voting Game

$G(N, W)$ in a Public Goods Economy," *Econometrica*, vol.45, 1589-1594.

[13] Konishi, H., M. Le Breton and S. Weber (1999) "On Coalition-proof Nash Equilibria in Common Agency Games," *Journal of Economic Theory*, vol.85, 122-139.

[14] Maruta, T. and A. Okada (2005) "Group Formation and Heterogeneity in Collective Action Games," Disucussion Paper Series #2005-7, Graduate School of Economics, Hitotsubashi University.

[15] Peleg, B. (1996) "Double Implementation of the Lindahl Allocation by a Continuous Mechanism," *Economic Design*, vol.2, 311-324.

[16] Saijo, T. and T. Yamato (1999) "A Voluntary Participation Game with a Non-excludable Public Good," *Journal of Economic Theory*, vol.84, 227-242.

[17] Shinohara, R. (2005) "Coalition-proofness and Dominance Relations," *Economics Letters*, vol.89, 174-179.

[18] Thoron, S. (1998) "Formation of a Coalition-proof Stable Cartel," *Canadian Journal of Economics*, vol.31, 63-76.

[19] Tian, G. (2000) "Double Implementation of Lindahl Allocations by a Pure Mechanism," *Social Choice and Welfare*, vol.17, 125-141.

[20] Walker, M. (1981) "A Simple Incentive Compatible Scheme for Attaining Lindahl Allocations," *Econometrica*, vol. 49, 65-71.

[21] Yi, S. (1999) "On Coalition-proofness of the Pareto Frontier of the Set of Nash Equilibria," *Games and Economic Behavior*, vol.26, 353-364.