Contributing or Free-Riding? A Theory of Endogenous Lobby Formation*

Taiji Furusawa† Hideo Konishi‡

July 9, 2007

Abstract

We consider a two-stage public good provision game: In the first stage, players simultaneously decide if they join a contribution group or not. In the second stage, players in the contribution group simultaneously offer contribution schemes in order to influence a third party agent's policy choice (say, the government chooses a level of public good provision). We use a communication-based self-enforcing equilibrium concept in a noncooperative two stage game, *perfectly coalition-proof Nash equilibrium* (Bernheim, Peleg and Whinston, 1987 JET). We show that, in public good economy, the outcome set of this equilibrium concept is equivalent to an "intuitive" hybrid solution concept *free-riding-proof core*, which always exists but does not necessarily achieve global efficiency. It is not necessarily true that the formed lobby group is the highest willingness-to-pay players, nor is a consecutive group with respect to their willingnesses-to-pay.

**Still Preliminary.* Konishi is grateful to the participants of the PGPPE Workshop 07 at CIRM in Marseille, Elena Paltseva and Sang-Seung Yi for discussions on the preliminary idea of this paper. He also thanks Kyoto Intitute of Economic Research for providing an excellent research environment. We are greateful to participants of seminars at various universities and the ones of the PET 2007 Conference for their comments. Our intellectual debt to Michel Le Breton is obvious.

†Department of Economics, Hitotsubashi University, Japan. (E-mail) `furusawa@econ.hit-u.ac.jp`

‡Corresponding author: Department of Economics, Boston College, USA. (E-mail) `hideo.konishi@bc.edu`

1 Introduction

In their seminar paper, Grossman and Helpman (1994) consider an endogenous trade policy formation problem when industries can influence the government's trade policy through lobbying activities by applying common agency game defined by Bernheim and Whinston (1986). A common agency game is a menu auction game in which there are multiple players/principals and an agent who can choose an action that affects all players' payoffs. Each player offers a contribution scheme to the agent promising how much money she will pay for each action. Observing contribution schemes, the agent chooses her action in order to maximize the total benefit she can obtain. Bernheim and Whinston (1986) analyze a communication-based equilibrium concept, coalition-proof Nash equilibrium (CPNE), in order to analyze common agency games. In Grossman and Helpman (1994), players/principals are lobbies who represents industries, and an agent is the government. The government cares about social welfare, while it also cares about flexible contribution money provided by lobby groups. Each lobby contributes money to the government in order to influence the government's trade policy for its favor. Each lobby represents one industry, and it prefers a high price for a commodity that is produced by the industry, while prefers low prices for all other commodities.[1] That is, in Grossman and Helpman (1994), there are conflicts of interests among lobbies. One of their main results is that in equilibrium lobby powers are cancelled out and that the government chooses a free trade (no tariff) policy, it can collect a big amount of contributions from conflicting industries.

Although the free trade outcome is an interesting result, it is based on special assumptions.[2] They assume that industry lobbies are preorganized, and that each lobby act as a single player. This implies that if there are multiple firms in an industry, each industry lobby has power to allocate contribution shares efficiently and forcefully among the member firms. However, in the real world, it is not necessarily the case that all firms in the same industry participate in a lobbying group. Since trade policies affect all firms

[1] This is because lobbies representing industries are ultimately consumers.

[2] Actually, their clean result is crucially based on their assumption that all lobbies are ultimately consumers who have identical utility functions. Bernheim and Whinston (1986) show that in equilibrium the agent chooses an action that maximizes the total surplus of the game. The result by Grossman and Helpman (1994) is a direct corollary of this under the representative consumer assumption.

in an industry in the same way, there are free-riding motives for firms.

Motivated by this, we consider a common agency game with players' endogenous participation decisions. In the first stage, players choose if they participate in lobbying or not.[3] In the second stage, among players who chose to participate in lobbying activities, a common agency game is played: contribution schemes are offered simultaneously, and the agent chooses an action. We use a dynamic extension of coalition-proof Nash equilibrium (CPNE), perfectly coalition-proof Nash equilibrium (PCPNE), as the solution concept. The reason that we need a solution concept other than Nash (or subgame perfect Nash) equilibrium is that we do not fix utility allocation rule for each possible lobby. Lobby formation and utility allocation within it are jointly determined. Unlike prefixed utility allocation rule, a coalitional deviation to form a new lobby needs to decide a utility allocation (or lobby cost sharing) that is available for the lobby. Thus, it is natural for us to use a communication based refinement of Nash equilibrium. This equilibrium concept also has a solid theoretical ground in a certain sense,[4] but characteristics of equilibria are not immediately clear due to its recursive definition. We use "guess and verify" method in order to characterize PCPNE: we define intuitive hybrid solution concepts for special classes of common agency games, and verify them with the PCPNE.[5]

We ask who participate in lobbying (and who free-ride others) in a special environment.[6] One is an environment without conflict of interests such as public good economies, in which all players have comonotonic preferences.[7]

[3]This is called open membership game (Yi, 1996).

[4]If only nested coalitional deviations are allowed, CPNE (and credible core in Ray, 1989) is a consistent solution concept in the sense that the original strategy profile and strategy profiles that are generated from deviations are treated in the same manner. However, it does not mean that CPNE is the only satisfactory solution concept. There can be many formulation of describing coalition formation process in noncooperative games. For example, Konishi and Ray (2003) and Gomes and Jehiel (2005) allow non-nested future deviations in defining consistent solutions.

[5]In various games, it is sometimes possible to show the equivalence between CPNE and intuitive solution concepts. See Bernheim and Whinston (1986), Thoron (1999), Conley and Konishi (2002) and Konishi and Ünver (2006).

[6]It is hard to characterize payoff structure of CPNEs in common agency games under general setup (see Laussel and Le Breton, 2001). Characterization of PCPNE is even harder. In fact, PCPNE may not exist in common agency games with endogenous participation decision.

[7]Preferences are **comonotonic** if for all pair of players i and j, and all pair of actions a and a', if i prefers a to a', then j also prefers a to a'.

This environment mimics an import competing industry case in which many firms decide lobbying or free-riding (Bombardini, 2005, and Paltseva, 2006), and pure public good provision problem. Although Bombardini (2005) provides some empirical evidence of free-rider firms in industries. Assuming symmetric firms and focusing on symmetric outcome among lobby participants in a common agency game, Paltseva (2006) consider Nash equilibrium to analyze free-riding incentives. The other is an environment in which there are two groups with pure conflicts of interests. This environment mimics the situation in which there are firms in import competing industries and exporting firms, and the government is deciding if it signs a free trade agreement with a foreign country.[8]

In such an environment, there is no rent for the government in PCPNE, while free-riding incentive is strong (equilibrium lobby participation is small). The equilibrium outcome is highly nonconvex, and the equilibrium lobby group may not be consecutive: i.e., weak firms may join the lobby together with strong firms, yet some medium firms may not. In contrast, if conflicts are present (export lobby and import lobby), the government gets a big rent, while lobby participation is strong.

This paper is organized as follows. In the next subsection, some related literature is discussed briefly. In Section 2, the common agency game is reviewed, then our game and the equilibrium concept, PCPNE, is introduced. In Section 3, we consider the environment without conflict of interests. We define an intuitive hybrid solution concept, free-riding-proof core, and prove the equivalence between PCPNE and the free-riding-proof core (Theorem). In Section 4, we provide an example that describes how free-riding-proof core looks like. Section 5 proves the key proposition for the proof of Theorem. Section 6 concludes.

1.1 Related Literature

Le Breton and Salaniè (2003) analyze a common agency problem with asymmetric information on agent's preferences. They show that equilibria can be inefficient even in the case that there is only one player in each interest group.[9] If there are multiple players in each interest group, then the failure

[8] A free trade agreement abandons trade barriers of the two countries. Exporting firms prefer a free trade agreement, while import competing firms prefer a protection policy.

[9] Laussel and Le Breton (1998) analyze public good case when the agent must sign a contract of participation when all contribution schemes are proposed before knowing

in internalizing the benefits of contributions within the group makes contributions even less. In this sense, Le Breton and Salaniè (2003) generate free-riding incentives under compulsory lobby participation. In contrast, we generate "free-riding" is a more obvious way by introducing participation decisions.

The environment without conflict of interests can be regarded as a public good provision problem. Groves and Ledyard (1977), Hurwicz (1979), and Walker (1981) showed that efficient public good provision can be achieved despite of Samuelson's pessimism (1954). However, they all assume that players must participate in the game. Saijo and Yamato (1999) considered voluntary participation game of public good provision by constructing a two stage game (participation, and public good provision). Negative results for efficiency due to free-riding incentives. Shinohara (2003) considers coalition-proof Nash equilibrium in the voluntary participation game by Saijo and Yamato (1999) with the Lindahl mechanism in the second stage. He shows that there can be multiple coalition-proof Nash equilibria with different sets of players participating in the mechanism in heterogenous player case. One of our results exhibits the same result but with a common agency game in the second stage (thus, payoffs are not fixed unlike in Shinohara, 2003). Such a voluntary public good provision problem can idealize the case of no conflict of interests. Maruta and Okada (2005) analyze a similar sort of heterogeneous agent binary public good provision game with evolutionary stability (see also Palfrey and Rosenthal, 1984).[10]

her cost type (then Nature plays and the agent chooses an agenda). They show that all equilibria are efficient, and there is no free-riding incentive.

[10]In contrast, Nishimura and Shinohara (2007) consider a multi-stage voluntary participation game in a discrete multi-unit public good problem, and show that Pareto-efficient allocations in subgame perfect Nash equilibrium through a mechanism that determines public good provision unit-by-unit. Their efficiency result crucially depends on the following assumption: a player who did not participate in the mechanism in early stages can participate in public good provision later on. This forgiving attitude allows a player not contributing at all until the time that all other players are no longer interested in contributing without her participation, and then contribute money just to bring one more unit of public good. Thus, we may say that their mechanism achieves Pareto efficiency by accommodating players' free-riding incentives.

2 A Noncooperative Game

We will consider a two stage game: in the first stage, each player decides if she join the lobby (contribution) group, or she stay outside (free-riding).[11] In the second stage, among the lobby group members, a common agency game by Bernheim and Whinston (1986). If a player choose to free-ride in the first stage, she cannot participate in the contribution game. Without free-riding incentive ($S = N$), Laussel and Le Breton (1998, 2001) extensively studied the equilibrium payoff structures of common agency games on general versions of this public good problem, and obtained many interesting and useful results. Our analysis will be built on theirs, but we consider possible free-riding: our focus is the conflict between contributing and free-riding.

We will focus on players' lobbying activities over government policies. We will consider a two stage game. In stage 1, players decide if they join a lobbying process or not (lobby formation stage).[12] In stage 2, the lobbying group lobby over government policies. In the next section, we analyze the second stage game.

2.1 Common Agency Game (the Second Stage)

There is a set of players, $N = \{1, ..., n\}$ and the government G. Suppose that $S \subseteq N$ is the contribution group, and $N \backslash S$ are passive free-riders. The government G can choose an agenda a from the set of agendas A. Each player i has utility function $v_i : A \to \mathbb{R}_+$, and similarly the government has utility function $v_G : A \to \mathbb{R}_+$. In public good provision problem, $v_G(a) = -C(a)$. Each player i offers a contribution scheme $\tau_i : A \to \mathbb{R}_+$. If the government chooses $a \in A$, then the government gets the payoff

$$u_G(a; (\tau_i(a))_{i \in S}) = \sum_{i \in S} \tau_i(a) + v_G(a),$$

[11] Note that in our game, there can be only one lobby that contributes to provision of public good. This does not seem a bad assumption given the nature of common agency game played in the second period. In contrast, Ray and Vohra (2001) assume that the second stage is a voluntary contribution game, thus it makes sense to assume that many groups can be formed in the first stage and they all provide public good simultaneously (see Ray and Vohra, 2001).

[12] This is called open membership game (no exclusion is possible). See d'Aspremont et al. (1983), Yi (1996) and Thoron (1998). For excludable coalitions, see Hart and Kurz (1983), Yi (1996), and Ray and Vohra (2001). Bloch (1997) has a nice survey on the rules of coalition formation games.

and player i gets payoff

$$u_i(a; \tau_i(a)) = v_i(a) - \tau_i(a).$$

The government chooses a policy $a \in A$ that maximizes u_G:

$$a^*(S, \tau_S) \in \arg\max_{a \in A} u_G(a; (\tau_i(a))_{i \in S}).$$

In the game, the government is not a player: it is just a machine that maximizes its payoff given the contribution schemes.[13] A second-stage **common agency game** Γ is a list $\Gamma = (S, (\mathcal{T}_i, u_i)_{i \in S})$, where $\mathcal{T}_i$ is collection of all contribution schemes for i. Note that $N \backslash S$ are simply free-riders, and they do not affect game Γ. Thus, $N \backslash S$ can be regarded as irrelevant players in game Γ.

First consider joint payoff that can be achieved by each subgroup $T \subseteq S$. For each $T \subseteq S$, let

$$W_\Gamma(T) \equiv \max_{a \in A} \left[\sum_{i \in T} v_i(a) + v_G(a) \right],$$

and

$$W_\Gamma(\emptyset) \equiv \max_{a \in A} v_G(a),$$

The efficient public good provision for S (and G) is

$$a^*(S) \in \arg\max_{a \in A} \left(\sum_{i \in N} v_i(a) + v_G(a) \right).$$

Let

$$Z_\Gamma \equiv \left\{ \begin{array}{c} u \in \mathbb{R}^N_+ : \sum_{i \in T} u_i \leq W_\Gamma(S) - W_\Gamma(S \backslash T) \text{ for all } T \subset S \\ \text{and } u_j = v_j(a^*(S)) \text{ for all } j \notin S \end{array} \right\}.$$

The inequality that Z_Γ satisfies can be interpreted as what T can get since the complement set $S \backslash T$ can achieve total payoff $W_\Gamma(S \backslash T)$ by themselves (T

[13]Strictly speaking, since the government may have multiple optimal policy, we need to introduce a tie-breaking rule. However, it is easy to check the set of truthful equilibria (see below) would not depend on the choice of tie-breaking rules.

cannot ask more than $W_\Gamma(S) - W_\Gamma(S \backslash T)$). We take the Pareto-frontier of Z_Γ:[14]

$$\bar{Z}_\Gamma \equiv \{u \in Z_\Gamma : \nexists u' \in Z_\Gamma \text{ such that } u' > u\}.$$

Now, let us state the results in the literature. Bernheim and Whinston (1986) introduced a concept of truthful strategies, where τ_i is **truthful relative to** $\bar{a}$ if and only if for all $a \in A$ either $v_i(a) - \tau_i(a) = v_i(\bar{a}) - \tau_i(\bar{a})$, or $v_i(a) - \tau_i(a) < v_i(\bar{a}) - \tau_i(\bar{a})$ and $\tau_i(a) = 0$. A **truthful Nash equilibrium** (τ_S^*, a^*) is a Nash equilibrium such that τ_i^* is truthful relative to $a^* \in A$ for all $i \in S$. The first result by Bernheim and Whinston (1986) is the following:

Fact 1. (Bernheim and Whinston, 1986) Consider a common agency game Γ. In all truthful Nash equilibria G chooses an efficient action $a^*(S)$, and the vector of players' payoffs belongs to the Pareto frontier $\bar{Z}_\Gamma$. Moreover, every vector $u \in \bar{Z}_\Gamma$ can be supported by a truthful Nash equilibrium.

Bernheim and Whinston (1986) defined (strictly) coalition-proof Nash equilibrium. First define a reduced game. A **reduced game** of Γ is $\Gamma(T, \tau_{-T})$ that is a game with players in T by letting players in $S \backslash T$ passive players in Γ, who always play τ_{-T}. A **(strictly) coalition-proof Nash equilibrium (CPNE)** of common agency game Γ is defined as follows (Bernheim and Whinston, 1986; Bernheim, Peleg and Whinston, 1987):[15]

1. In a single player game Γ, (τ_1^*, a^*) is a CPNE of reduced game $\Gamma(\{i\}, \mathcal{T}_i, \bar{\tau}_{-i})$ if and only if it is a Nash equilibrium.

2. Let n be the number of players of the game. In a game $\Gamma(S, \mathcal{T}_S, \bar{\tau}_{-S})$ where $|S| = n$, $(\tau_S^*, a^*) = ((\tau_i^*)_{i \in S}, a^*)$ is a **(strictly) self-enforcing** strategy profile if for all $T \subsetneq S$, $(\tau_i^*)_{i \in T}$ is a CPNE of the reduced game $\Gamma(T, \mathcal{T}_T, \tau_{S \backslash T}^*, \bar{\tau}_{-S})$.

3. Let n be the number of players of the game. In a game $\Gamma(S, \mathcal{T}_S, \bar{\tau}_{-S})$ where $|S| = n$, $(\tau_S^*, a^*) = ((\tau_i^*)_{i \in S}, a^*)$ is a **CPNE** if it is self-enforcing and there is no other self-enforcing strategy profile τ_S' that yields at

[14] We follow the standard notational convention: $u' > u$ means (i) for all $i \in N$, $u_i' \geq u_i$, and (ii) $u' \neq u$.

[15] The definitions of CPNE in Bernheim and Whinston (1986) and Bernheim, Peleg and Whinston (1987) are different in defining coalitional deviations. The former uses weakly improving deviations, while the latter uses strictly improving deviations. On this issue, see Konishi, Le Breton and Weber (1999).

least as high a payoff to each player and a strictly higher payoff to at least one player in S.

The second result in Bernheim and Whinston (1986) is as follows:

Fact 2. (Bernheim and Whinston, 1986) Consider a common agency game Γ. In all CPNEs G chooses an efficient action $a^*(S)$, and the vector of players' payoffs belongs to the Pareto frontier $\bar{Z}_\Gamma$. Moreover, every truthful Nash equilibrium is coalition-proof, thus every vector $u \in \bar{Z}_\Gamma$ can be supported by a CPNE.

That is, there is essentially a one-to-one relationship between truthful Nash equilibria and CPNEs. Note that a truthful Nash equilibrium is a CPNE (this property will be used in the proof of Theorem 1). One of many results in Laussel and Le Breton (2001) provided a useful property, convex-game property, which applies to an interesting class of games. Consider a characteristic function $(W_\Gamma(T))_{T\subseteq S}$ generated from a common agency game Γ. We say that Γ has **convex-game property** if for all $T \subset T' \subset S$ with $i \in S\backslash T'$, $W_\Gamma(T\cup\{i\}) - W_\Gamma(T) \leq W_\Gamma(T'\cup\{i\}) - W_\Gamma(T')$ holds. Laussel and Le Breton (2001) shows the following:

Fact 3. (Laussel and Le Breton, 2001) Consider a common agency game Γ with convex-game property. Then, in all CPNEs G obtains $u_G = W_\Gamma(\emptyset)$ (no rent property), and the set of CPNE payoff vectors is equivalent to the core of characteristic function game $(W_\Gamma(T))_{T\subseteq S}$.[16]

This fact will be useful in analyzing public good case below, since public good economy satisfies the convex game property.[17]

2.2 Lobby Formation Game

In this section, we analyze an equilibrium lobby group and its allocation. Note that we are not only talking about coalition-proof Nash equilibrium

[16]Here, we normalize $W_\Gamma(\emptyset) = 0$ in order to make $(W_\Gamma(T))_{T\subseteq S}$ a characteristic function game. A payoff vector $u_S = (u_i)_{i\in S}$ is in *the core* iff $\sum_{i\in S} u_i = W_\Gamma(S)$, and $\sum_{i\in T} u_i \geq W_\Gamma(T)$ for all $T \subset S$.

[17]Actually, with the no-rent property (in all CPNE, $u_G = W_\Gamma(\emptyset)$ holds), the set of CPNEs is eqivalent to the set of strong equilibria (see Aumann, 1959, for the definition) in common agency games (see Konishi, Le Breton and Weber, 1999). Thus, with the convex-game property, the set of CPNE, the set of strong equilibria, and the core of $(W_\Gamma(T))_{T\subseteq S}$ are all equivalent. However, with participation stage, strong equilibrium tends to be empty.

allocation in the menu auction stage. We also require that the lobby group formation itself is coalition-proof as well. In order to do so, we first need to define the first stage lobby formation game in an appropriate manner, assuming that the outcome of each possible lobby S is a coalition-proof Nash equilibrium of a common agency game played by S. As an extension of CPNE in strategic form games to extensive form games, Bernheim, Peleg and Whinston (1987) provide a definition of coalition-proof Nash equilibrium for multi-stage games, *perfectly coalition-proof Nash equilibrium (PCPNE)*. The first stage **lobby formation game** is such that N is the set of players, and player i's action set is a list $A_i = \{0, 1\}$: i.e., player i announces if she wants to participate in the lobby. Once action profile $\sigma^1 = (\sigma_1^1, ..., \sigma_n^1) \in A = \Pi_{j \in N} A_j$ is determined, then in the second stage, lobbying game takes place with the set of active players $S(\sigma^1) = \{i \in N : \sigma_i^1 = 1\}$.

Next we extend the definition of CPNE for multi-stage games by following the definition by Bernheim, Peleg and Whinston (1987). In our game, there are only two stages $t = 1, 2$. Player i's strategy $\sigma_i = (\sigma_i^1, \sigma_i^2) \in \Sigma_i = \Sigma_i^1 \times \Sigma_i^2$ is such that $\sigma_i^1 \in \Sigma_i^1$ denotes i's lobby participation choice, and $\sigma_i^2 \in \Sigma_i^2$ is a function $\sigma_i^2 : \mathcal{S}(i) \to \mathcal{T}_i$ if $\sigma_i^1 = 1$, where $\mathcal{S}(i) = \{S \in 2^N : i \in S\}$ and $\mathcal{T}_i$ is the space of bid functions in the common agency game (if $\sigma_i^1 = 0$, then σ_i^2 is a trivial strategy).[18] When $\sigma_i^1 = 0$ (no participation in lobbying), the second stage strategy σ_i^2 is irrelevant. Each player's payoff function is $u_i : \Sigma \to \mathbb{R}$ that is the same payoff function of lobbying game when lobby group S is determined by $S(\sigma^1)$. For $T \subseteq N$, consider a **reduced game** $\Gamma(T, \sigma_{-T})$ that is a game with players in T by letting players in $N \backslash T$ passive players in Γ, who always play σ_{-T}. We also consider **subgames** for all $\sigma^1 \in \Sigma^1$, and **reduced subgames** $\Gamma(T, \sigma^1, \sigma_{-T}^2)$ in similar ways. A **perfectly coalition-proof Nash equilibrium** (**PCPNE**) $(\sigma^*, a^*) = ((\sigma_i^{1*}, \sigma_i^{2*})_{i \in N}, a^*)$ is recursively defined as follows:[19]

(a) In a single player, single stage subgame $\Gamma(\{i\}, \Sigma_i^2, \sigma^1, \sigma_{-\{i\}}^2)$, strategy $\sigma_i^{2*} \in \Sigma_i^2$ and the agenda chosen by the agent a^* is a **PCPNE** if σ_i^{2*} maximizes u_i via a^*.

[18] Thus, $\sigma_i^2(S) \in \mathcal{T}_i$ is $T_i : A \to \mathbb{R}_+$ in the last section.

[19] Note that in Bernheim, Peleg and Whinston (1987), the definition of PCPNE is based on strictly improving coalitional deviations. However, we adopt a definition based on weakly improving coalitional deviations, since the theorem on menu auction in Bernheim and Whinston (1986) uses CPNE based on weakly improving deviation. For details on these two definitions, see Konishi, Le Breton and Weber (1999).

(b-1) Let $(n, 2)$ be the numbers of players and stages of games. Pick any positive pair of integers $(m, r) \leq (n, 2)$ with $(m, r) \neq (n, 2)$.[20] For any $T \subseteq N$ with $|T| \leq m$, assume that PCPNE has been defined for all reduced games $\Gamma(T, \sigma_{-T})$ and their subgames $\Gamma(T, \sigma^1, \sigma^2_{-T})$ (if $r = 1$, then only for all reduced subgames $\Gamma(T, \sigma^1, \sigma^2_{-T})$). Then,

(i) for all reduced games $\Gamma(S, \sigma_{-S})$ and their subgames $\Gamma(S, \sigma^1, \sigma^2_{-S})$ with $|S| = n$, $(\sigma^*, a^*) \in \Sigma \times A$ is **perfectly self-enforcing** if for all $T \subset S$ we have (σ^*_T, a^*) is a PCPNE of reduced game $\Gamma(T, \sigma^*_{S \setminus T}, \sigma_{-S})$, and σ^{2*}_T is a PCPNE of reduced subgame $\Gamma(T, \sigma^1, \sigma^{2*}_{S \setminus T}, \sigma^2_{-S})$, and

(ii) for all $S \subseteq N$ with $|S| = n$, (σ^*_S, a^*) is a **PCPNE** of reduced game $\Gamma(S, \sigma_{-S})$ if (σ^*_S, a^*) is perfectly self-enforcing in reduced game $\Gamma(S, \sigma_{-S})$, and there is no other perfectly self-enforcing σ'_S such that $u_i(\sigma'_S, \sigma_{-S}) \geq u_i(\sigma^*_S, \sigma_{-S})$ for every $i \in S$ with at least one strict inequality.

(b-2) Let $(n, 1)$ be the numbers of players and stages of games. Pick any positive integer $m < n$. For any $T \subseteq N$ with $|T| \leq m$, assume that PCPNE has been defined for all reduced subgames $\Gamma(T, \sigma^1, \sigma^2_{-T})$. Then,

(i) for all reduced subgame $\Gamma(S, \sigma^1, \sigma^2_{-S})$ with $|S| = n$, $(\sigma^*, a^*) \in \Sigma \times A$ is **perfectly self-enforcing** if for all $T \subset S$ we have (σ^{2*}_T, a^*) is a PCPNE of reduced subgame $\Gamma(T, \sigma^1, \sigma^{2*}_{S \setminus T}, \sigma^2_{-S})$, and

(ii) for all $S \subseteq N$ with $|S| = n$, (σ^{2*}_S, a^*) is a **PCPNE** of reduced game $\Gamma(S, \sigma^1, \sigma_{-S})$ if (σ^{2*}_S, a^*) is perfectly self-enforcing in reduced subgame $\Gamma(S, \sigma^1, \sigma_{-S})$, and there is no other perfectly self-enforcing $\sigma^{2\prime}_S$ such that $u_i\left(\sigma^1, \sigma^{2\prime}_S, \sigma^2_{-S}\right) \geq u_i\left(\sigma^1, \sigma^{2*}_S, \sigma^2_{-S}\right)$ for every $i \in S$ with at least one strict inequality.

For any $T \subseteq N$ and any strategy profile σ, let $PCPNE(\Gamma(T, \sigma_{-T}))$ denote the set of PCPNE strategy profiles on T for the game $\Gamma(T, \sigma_{-T})$. For any strategy profile (σ, a), a strategic coalitional deviation (T, σ'_T, a') from (σ, a) is **credible** if $(\sigma'_T, a') \in PCPNE(\Gamma(T, \sigma_{-T}))$. A PCPNE is *a strategy profile that is immune to any credible coalitional deviation.*

[20] The numbers n and t represent the numbers of players and stages of a reduced (sub) game, respectively.

First note that PCPNE coincides with CPNE in the second stage. Thus, a CPNE needs to be assigned to each subgame. Second, if a coalition T wants to deviate in the first stage, within the reduced game $\Gamma(T, \sigma_{-T})$, it can orchestrate the whole plan of the deviation by assigning a new CPNE to each subgame so that the target allocation (by the deviation) would be attained as PCPNE of the reduced game $\Gamma(T, \sigma_{-T})$.

In general, it is hard to see the properties of PCPNE of lobby formation game with common agency including its existence of equilibrium. However, in public good provision problem, we can assure existence of equilibrium and provide a characterization of PCPNE in its equilibrium outcome set. Consider a PCPNE (σ^*, a^*). An **outcome allocation** for (σ^*, a^*) is a list $(S, a^*, u) \in 2^N \times A \times \mathbb{R}^N \times \mathbb{R}$, where $S = \{i \in N : \sigma_i^{1*} = 1\}$ and (u, u_G) is the resulting utility allocation for players and the agent such that for all $i \notin S$, $u_i = v_i(a^*)$. From the facts obtained in common agency game, (S, a^*, u) satisfies $a^* = a^*(S)$, and for any $T \subseteq S$, $\sum_{i \in T} u_i \leq W_\Gamma(S) - W_\Gamma(S \backslash T)$. The agent's payoff u_G is implicitly determined by $\sum_{i \in S} u_i + u_G = W_\Gamma(S)$.

3 A Public Good Provision Problem

In this section, we consider a case in which all players' interests are in the same direction, while the intensity of their interests can be heterogeneous. We will describe the game, and then propose a hybrid solution concept: free-riding-proof core.

A stylized public good model can be viewed as a special class of the above game. Agenda is a public good provision level, and is one-dimensional: $A = \mathbb{R}_+$, and the provision cost of public good is described by a C^2 cost function $C : A \to \mathbb{R}_+$ with $C(0) = 0$, $C'(a) > 0$ and $C''(a) > 0$ (for uniqueness: for simplicity). Player i's utility function is quasi linear in private good net consumption x and is written as $v_i(a) - x$, where $v_i : A \to \mathbb{R}_+$ is $v_i(0) = 0$, $v_i'(a) > 0$ and $v_i''(a) \leq 0$. In order to guarantee the existence of solution, we assume the Inada condition on the cost function: $\lim_{a \to 0} C'(a) = 0$ and $\lim_{a \to \infty} C'(a) = \infty$.

We will analyze PCPNE of our two stage game in this problem. First, we will define an intuitive but not well-grounded hybrid solution concept, *free-riding-proof core (FRP-core)*, which is the set of Foley-core allocations[21] that

[21] The Foley core of our public good economy is the standard core concept assuming that deviating coalitions have to provide public good by themselves. That is, it assumes

are immune to free-riding incentives and is Pareto-optimal in a constrained sense. The free-riding-proof core is always nonempty in the public good provision problem. Second, by an example, we investigate how the free-riding-proof core looks like. Finally, we prove that the set of outcomes of PCPNE is equivalent to the free-riding-proof core.

A public good provision problem determines two things: (i) which group provides public good and how much, and (ii) how to allocate the benefits from providing public good among the members of the group (or how to share the cost). Let $S \subseteq N$ with $S \neq \emptyset$. For $T \subseteq S$, let

$$V(S) \equiv \max_{a \in A} \left[\sum_{i \in S} v_i(a) - C(a) \right],$$

and

$$a^*(S) \equiv \arg\max_{a \in A} \left[\sum_{i \in S} v_i(a) - C(a) \right].$$

An **allocation for** S is $(S, a^*(S), u)$ such that $u \in \mathbb{R}_+^N$, $\sum_{i \in S} u_i \leq V(S)$,[22] and $u_j = v_j(a)$ for all $j \notin S$ (utility allocation). That is, $N \backslash S$ are passive free-riders, and they do not contribute at all. Given that S is the lobby group, a natural way to allocate utility among the members is to use the core (Foley, 1970). A **core allocation for** S, $(S, a^*(S), u)$, is an allocation for S such that $\sum_{i \in T} u_i \geq V(T)$ holds for all $T \subseteq S$.

However, a core allocation for S may not be immune to free-riding incentives by its members of S. So, we will define a hybrid solution concept of cooperative and noncooperative games. A **free-riding-proof core allocation for** S is a core allocation $(S, a^*(S), u)$ for S such that

$$u_i \geq v_i(a^*(S \backslash \{i\})) \text{ for all } i \in S.$$

A free-riding-proof core allocation is immune to unilateral deviations of the members of S. Note that, given the nature of public good provision problem, we can allow a coalitional deviation from S at no cost (since one person deviation is the most profitable). Let $Core^{FRP}(S)$ be the set of all free-riding-proof core allocations for S. Note that $Core^{FRP}(S)$ may be empty for large group S, while for small groups it is nonempty (especially, for singleton

that there is no spillover of public good across the groups.

[22]Note that we have $V(S) = W_\Gamma(S) - W_\Gamma(\emptyset)$ in our public good provision problem.

groups it is always nonempty). We collect free-riding-proof core allocations for all S, and take their Pareto frontiers: the set of **free-riding-proof core** is defined as

$$\begin{aligned} Core^{FRP} &= \left\{(S, a^*(S), u) \in \cup_{S' \in 2^N} Core^{FRP}(S') : \right. \\ &\quad \left. \forall T \in 2^N,\, \forall u' \in Core^{FRP}(T),\, \exists i \in N \text{ with } u_i > u'_i \right\}. \end{aligned}$$

That is, an element of $Core^{FRP}$ is a free-riding-proof core allocation for some S that is not weakly dominated by any other free-riding-proof core allocation for any T. *Note that $Core^{FRP}$ is **not** a subsolution of $Core(N)$: it only achieves constrained efficiency due to free-riding incentives, since we often have $Core^{FRP}(N) = \emptyset$.* Note that there always exists a free-riding-proof core allocation. since for all singleton set $S = \{i\}$, $Core^{FRP}(S)$ is nonempty.

Proposition 1. $Core^{FRP} \neq \emptyset$.

Now, we will characterize PCPNE with free-riding-proof core. In the public good provision problem, Fact 3 (Laussel and Le Breton, 2001) is useful. Note that the core of $(W_\Gamma(T))_{T \subseteq S}$ is equivalent to $Core(S)$ in our game. It can be seen as follows. Since in a public good provision problem preferences are **comonotonic**, i.e., $v_i(a) \geq v_i(a')$ if and only if $v_j(a) \geq v_j(a')$ for all $i, j \in S$ and all $a, a' \in A$, $(W_\Gamma(T))_{T \subseteq S}$ is a convex game (Laussel and Le Breton, 2001). Thus, no rent property $u_G = W_\Gamma(\emptyset)$ holds, and $W_\Gamma(\emptyset) = 0$ in public good game. This implies

$$\sum_{i \in S} u_i = W_\Gamma(S) - W_\Gamma(\emptyset) = W_\Gamma(S).$$

This further implies,

$$\sum_{i \in T} u_i \geq W_\Gamma(T),$$

holds for all $T \subset S$, since for $S \backslash T$, the complement of T, we have

$$\begin{aligned} \sum_{i \in S \backslash T} u_i &\leq W_\Gamma(S) - W_\Gamma(T) \\ &= \sum_{i \in S} u_i - W_\Gamma(T). \end{aligned}$$

Since $W_\Gamma(T) = V(T)$ in the public good economy, the second stage CPNE outcomes coincide $Core(S)$ of a characteristic function form game $(V(T))_{T\subseteq S}$ with $V(S) = \max_{a\in A}\left(\sum_{i\in T} v_i(a) - C(a)\right)$.[23] This is nothing but Foley's core in a public good economy (Foley, 1970). This gives us some insight in our two-stage noncooperative game. Given the setup of our lobby formation game in the first stage, if a CPNE outcome u in a subgame S can realize as the equilibrium outcome (on-equilibrium path), it is *necessary* to have $u \in Core^{FRP}(S)$, since otherwise, some member of S would deviate in the first stage and obtain a secured free-riding payoff. This observation is useful in our analysis in the equivalence theorem. With some constructions, we can show the following:

Proposition 2. In public good provision problem, if an allocation $(S, a^*(S), u)$ is in the FRP-core, then there is a PCPNE σ of which outcome is $(S, a^*(S), u)$.

We postpone the proof of this proposition to Section 5, since it is somewhat involved. Once this direction is proved then the other direction is trivial. Notice that being PCPNE requires free-riding-proofness. Every PCPNE must be a free-riding-proof core allocation for some S. Since $Core^{FRP}$ is the Pareto-frontier of $\cup_{S\subseteq N} Core(S)$, Proposition 2 actually proves that all Pareto-dominated free-riding-proof core allocation for S can be defeated by a free-riding-proof core allocation.

Theorem. *In public good provision problem, an allocation $(S, a^*(S), u)$ is in the FRP-core, if and only if there is a PCPNE σ of which outcome is $(S, a^*(S), u)$.*

Proof of Theorem. We will show the other direction of Proposition 2: every PCPNE σ generates a free-riding-proof core allocation as its outcome. It is easy to see that the outcome $(S, a^*(S), u)$ of a PCPNE σ is a free-riding-proof core allocation **for** S, since otherwise the resulting allocation will not be a subgame perfect Nash equilibrium: the resulting allocation must be in the core, and needs to be free-riding-proof: if not, a player who has free-riding incentive certainly unilaterally (so credibly) deviates in the first stage. Thus,

[23] Actually, with no rent property, CPNE and strong Nash equilibrium (Aumann, 1959, but with weakly improving deviations) are equivalent in common agency game. See Konishi, Le Breton and Weber (1999).

$(S, a^*(S), u) \in Core^{FRP}(S)$. Suppose to the contrary that $u \notin Core^{FRP}$. Then, there is an free-riding-proof core allocation $(S', a^*(S'), u') \in Core^{FRP}$ with $u' > u$. Consider a coalitional deviation with a grand coalition N by preparing a PCPNE σ' that achieves u'. There is such a σ' by Proposition 2. This implies that there is a credible coalitional deviation from σ. This is a contradiction. Thus, every PCPNE achieves a free-riding-proof core allocation.□

This Theorem together with Proposition 1 guarantees nonemptiness of PCPNE. However, it may still not clear how free-riding-proof core looks like. In the next section, we will use a simple example to illustrate the properties of free-riding-proof core allocations, thus the outcome of PCPNE.

4 A Linear-Utility and Quadratic-Cost Public Good Example

Let $v_i(a) = \theta_i a$ for all $i \in N$ and $C(a) = \frac{1}{2}a^2$, where $\theta_i > 0$ is a parameter.[24] With this setup, for group S, the optimal public good provision is determined by the first order condition, $\sum_{i \in S} \theta_i - a = 0$: i.e.,

$$a^*(S) = \sum_{i \in S} \theta_i.$$

Thus, the value of S is written as

$$\begin{aligned} V(S) &= \sum_{i \in S} \theta_i \left(\sum_{i \in S} \theta_i \right) - \frac{1}{2} \left(\sum_{i \in S} \theta_i \right)^2 \\ &= \frac{\left(\sum_{i \in S} \theta_i \right)^2}{2}. \end{aligned}$$

For an outsider $j \in N \backslash S$, the payoff is

$$v_j(a^*(S)) = \theta_j \left(\sum_{i \in S} \theta_i \right).$$

[24] Coefficient 1/2 of $C(a)$ function is just matter of normalization. For any $k > 0$ with $C(a) = ka^2$, we get isomorphic results.

Consider the following example.

Example 1. Let $N = \{1, 3, 5, 11\}$ with $\theta_i = i$ for each $i \in N$.

Suppose first that a simultaneous voluntary contribution of public goods are done, instead of our two stage public good provision process. Then, only $i = 11$ contributes, and all others free-ride. The public good provision level is $a = 11$.

Now, let us move to our problem. First suppose that $S = N$. Then, we have $a^*(N) = \sum_{i \in N} i = 20$, and $V(N) = \frac{20^2}{2} = 200$. However, in order to have free-riding-proofness, we need to give each player the following payoff at the very least:

$$\begin{aligned} v_{11}(a^*(N\backslash\{11\})) &= (20 - 11) \times 11 = 99, \\ v_5(a^*(N\backslash\{5\})) &= (20 - 5) \times 5 = 75, \\ v_3(a^*(N\backslash\{3\})) &= (20 - 3) \times 3 = 51, \\ v_1(a^*(N\backslash\{1\})) &= (20 - 1) \times 1 = 19. \end{aligned}$$

The sum of all the above values exceeds the value of the grand coalition $V(N)$. As a result, we can conclude $Core^{FRP}(N) = \emptyset$. Next, consider $S = \{11, 5\}$. Then, $a^*(S) = 16$, and $V(S) = 128$. In order to check if the free-riding-proof core for S is nonempty, first again check the free-riding-incentives.

$$\begin{aligned} v(a^*(S\backslash\{11\})) &= (16 - 11) \times 11 = 55, \\ v(a^*(S\backslash\{5\})) &= (16 - 5) \times 5 = 55. \end{aligned}$$

Thus, if there is a free-riding-proof core allocation $u = (u_{11}, u_5)$ for S, u must satisfy

$$\begin{aligned} u_{11} + u_5 &= 128, \\ u_{11} &\geq 55, \\ u_5 &\geq 55, \\ u_{11} &\geq \frac{11 \times 11}{2} = 60.5, \\ u_5 &\geq \frac{5 \times 5}{2} = 12.5. \end{aligned}$$

The last two conditions are obtained by the core requirement. Thus, we have[25]

$$Core(\{11,5\}) = \left\{ \begin{array}{c} \tilde{u} \in \mathbb{R}^5_+ : u_{11} + u_5 = 128,\ u_{11} \geq 60.5,\ u_5 \geq 12.5, \\ \tilde{u}_3 = 48,\ \tilde{u}_2 = 32,\ \tilde{u}_1 = 16 \end{array} \right\},$$

and

$$Core^{FRP}(\{11,5\}) = \left\{ \begin{array}{c} \tilde{u} \in \mathbb{R}^5_+ : u_{11} + u_5 = 128,\ u_{11} \geq 60.5,\ u_5 \geq 55, \\ \tilde{u}_3 = 48,\ \tilde{u}_2 = 32,\ \tilde{u}_1 = 16 \end{array} \right\}.$$

As is easily seen, $Core^{FRP}(\{11,5\}) \neq \emptyset$, but it is a smaller set than $Core(\{11,5\})$. That is, the first observation is obvious:

- "*Free-riding-proof constraints may narrow the set of attainable core allocations.*"

Now, let us consider a simultaneous move voluntary public good provision game by Bergstrom, Blume and Varian (1986). Each player i chooses her monetary contribution $m_i \geq 0$ to provide public good. The public good provision level is determined by $a(m) = \sqrt{2\sum_{i\in N} m_i}$ reflecting the cost function of public good production. Consider player i. Given that others are contributing M_{-i} together, player i maximizes $\theta_i\sqrt{2\left(m_i + M_{-i}\right)} - m_i$. Thus, the best response for player i is $m_i^* = \max\left\{\frac{i^2}{2} - M_{-i}, 0\right\}$. This implies that only player 11 contributes, and the public good provision level is 11. Thus, by forming a contribution group in the first stage, it is possible to increase the public good provision level in equilibrium.[26]

Now, let us characterize the free-riding-proof core, the FRP-core. Since the FRP-core requires Pareto-efficiency on the union of free-riding-proof cores

[25] For notational simplicity, without confusion, we abuse notations by dropping irrelevant arguments of allocations. Thus, in this subsection, allocations are utility allocations.

[26] In relation to this, the readers may wonder about the Lindahl equilibrium allocation for $S = \{11,5\}$. Unfortunately, this example is not very useful since utility function is quasi-linear. The result would totally dependent on how the profits are distributed as is seen below. The Lindahl prices are $p_{11} = 11$ and $p_5 = 5$ given $\theta_{11} = 11$ and $\theta_5 = 5$, since $a^*(\{11,5\}) = 16$ means marginal cost is $16 (= 11 + 5)$. Since there are pure profits in producing public goods (cost function is strictly convex), we need to specify the way to allocate the profits 128. If they are distributed equally, then both get 64 each as profit share, and this is the only source of their utilities. If they are distributed according to players' willingnesses-to-pay, then players get 88 and 40. In the former case, the free-riding-proof conditions are satisfied, but in the latter case, they are not satisfied.

for all subsets of the players, we need to find free-riding-proof core for each S, in order to find the set of free-riding-proof core. The following lemma helps us to do the task.

Lemma 1. In the linear-utility-quadratic-cost public good problem, the free-riding-proof core for S is nonempty if and only if S satisfies $\Phi(S) \equiv \sum_{i\in S}\theta_i a^*(S) - \frac{1}{2}(a^*(S))^2 - \sum_{i\in S}\theta_i a^*(S\backslash\{i\}) \geq 0$ (aggregated "no free riding condition").

Even in this simple setup, we can make a few interesting observations.

Example 1. (continued) The free-riding-proof core allocations are attained by groups $\{11,5,1\}$, $\{11,3,1\}$, $\{11,5\}$, $\{11,3\}$, and $\{5,3\}$.

First by Lemma1, we can easily check for which S, $Core^{FRP}(S) \neq \emptyset$ holds. There are 12 such contribution groups: $\{11,5,1\}$, $\{11,3,1\}$, $\{11,5\}$, $\{11,3\}$, $\{11,1\}$, $\{5,3\}$, $\{5,1\}$, $\{3,1\}$, $\{11\}$, $\{5\}$, $\{3\}$, and $\{1\}$.

Note that $S = \{11,5,3\}$ does not have nonempty free-riding-proof core for S. Let $S = \{11,5,3\}$. Then, $a^*(S) = 19$ and $W(S) = 180.5$. Now, $11v(a^*(S\backslash\{11\})) = 88$, $5v(a^*(S\backslash\{5\}) = 70$ and $3v(a^*(S\backslash\{3\})) = 48$. Since $88+70+48 > 180.5$, there is no free-riding-proof core allocation for $S = \{11,5,3\}$. Thus, $\{11,5,1\}$ is the group that achieves the highest level of public good provision, and has nonempty free-riding-proof core.[27] This analysis gives an interesting observation:[28]

- *(Even the largest) group that achieves a free-riding-proof core allocation may not be consecutive.*

The intuition of this result is simple. Suppose $\Phi(S)$ is positive (say, $S = \{11,5\}$). Then by Lemma 1, there is an internally stable allocation for S. Now, we may try to find $S' \supset S$ that still keeps $\Phi(S') \geq 0$. If the value of $\Phi(S)$ is positive yet the value is not so large, then adding high θ player (say, player 3) may make $\Phi(S') < 0$, since adding such a player may increase $a^*(S')$ a lot, making free-riding problem severer. However, if low θ player

[27] As is seen below, group $\{11,5,1\}$ supports some allocations in $Core^{FRP}$.

[28] Although the context and approach are very different, in political science and sociology, formation of such non-consecutive coalitions is of a tremendous interest. For a game theoretical treatment of this line of literature (known and "Gamson's law"), see Le Breton et al. (2007).

(say, player 1) is added, the free-rider problem does not become too severe, and $\Phi(S') \geq 0$ may be satisfied relatively easily.

Among the above 12 groups, it is easy to see that groups $\{5,1\}$, $\{3,1\}$, $\{11\}$, $\{5\}$, $\{3\}$, and $\{1\}$ do not survive the test of Pareto-domination by free-riding-proof core allocations for other groups. For example, consider $S = \{11,5\}$ and $u' = (73, 55, 48, 32, 16) \in Core^{FRP}(\{11,5\})$.[29] Since the payoff of 11 by free-riding is $v_{11}(a) = 11a$, every allocation for the above groups are dominated by the above u'. On the other hand, $\{5,3\}$ is not dominated, since player 11 gets 88 by free-riding, respectively. Thus, player 11 would not join a deviation (11 can obtain maximum 73 in a free-riding-proof core allocation for $S \ni 11$). Without player 11's cooperation, there is no free-riding core allocation that dominates those of $\{5,3\}$.

By the same reasons, free-riding-proof core allocations for $S = \{11,1\}$ are dominated by the one for $S' = \{11,5\}$. Under $S = \{11,1\}$, player 5 gets 60, but S' can attain $u' = (63, 65, 48, 32, 16)$.[30] However, free-riding-proof core allocations for $S = \{11,3,1\}$ and $\{11,3\}$ cannot be beaten by the ones for $S' = \{11,5\}$, since player 5 gets 70 even under $\{11,3\}$.[31]

Finally, $S = \{11,5\}$, $\{11,3\}$. The free-riding-proof core allocations for $S = \{11,5\}$ is characterized by $u_{11} + u_5 = 128$, $u_{11} \geq 60.5$ and $u_5 \geq 55$, with $u_3 = 48$, $u_2 = 32$ and $u_1 = 16$. Now, consider $S' = \{11,5,1\}$. The free-riding-proof core allocations for S' is characterized by $u'_{11} + u'_5 + u'_1 = 144.5$, $u'_1 \geq 66$, $u'_5 \geq 60$ and $u'_1 \geq 16$, with $u'_3 \geq 51$ and $u'_2 \geq 34$. Thus, S' can attain $u'_{11} + u'_5 = 144.5 - 16 = 128.5$ as long as $u'_{11} \geq 66$ and $u'_5 \geq 60$. Thus, if $u \in Core^{FRP}(\{11,5\})$ satisfies $u_{11} + u_5 = 128$, $60.5 \leq u_{11} \leq 68.5$, and $55 \leq u_5 \leq 62.5$, then u is improved upon by an allocation in $Core^{FRP}(\{11,5,1\})$. However, if $u \in Core^{FRP}(\{11,5\})$ satisfies $u_{11} + u_5 = 128$, $u_{11} > 68.5$, or $u_5 > 62.5$, then u cannot be improved upon by forming group $\{11,5,1\}$. Free-riding-proof core allocations for $S = \{11,3\}$ has a similar property with possible deviations by group $S' = \{11,3,1\}$. This phenomenon illustrates another interesting observation:

- *An expansion of group definitely increases the total value of the group,*

[29] The best allocation for player 11 in $Core^{FRP}(\{11,5\})$. See the characterization of $Core^{FRP}(\{11,5\})$ in Example 1. Other players are free-riders, and their payoffs are directly generated from $a^*(\{11,5\}) = 16$.

[30] Under $S = \{11,2\}$, player 11 can get at most 62.5 in order to satisfy the free-riding-proofness for player 2 ($v_2(\{11\}) = 22$).

[31] Since $V(\{11,5\}) = 128$, and player 5 demands at least 70, player 11 can get at most 58. However, $V(\{11\}) = 60.5$. Thus, involving player 5 is not feasible.

while it gives less flexibility in allocating it since free-riding incentives are strengthened by having more public good. As a result, some unequal free-riding-proof core allocations for the original group may not be improved upon by expanding the group.

In summary, the free-riding-proof core is *union* of the following sets of allocations attained by five different groups.

1. $S = \{11, 5, 1\}$, then $a^*(S) = 17$ and all free-riding-proof core allocations for S are attained:

$$Core^{FRP}(\{11,5,1\}) = \left\{ \begin{array}{c} \tilde{u} \in \mathbb{R}^5_+ : \tilde{u}_{11} + \tilde{u}_5 + \tilde{u}_1 = 144.5,\ \tilde{u}_3 = 51,\ \tilde{u}_2 = 34, \\ 66 \leq \tilde{u}_{11},\ 60 \leq \tilde{u}_5,\ 16 \leq \tilde{u}_1 \end{array} \right\}$$

2. $S = \{11, 3, 1\}$, then $a^*(S) = 15$ and all free-riding-proof core allocations for S are attained:

$$Core^{FRP}(\{11,3,1\}) = \left\{ \begin{array}{c} \tilde{u} \in \mathbb{R}^5_+ : \tilde{u}_{11} + \tilde{u}_3 + \tilde{u}_1 = 112.5,\ \tilde{u}_5 = 75,\ \tilde{u}_2 = 30, \\ 60.5 \leq \tilde{u}_{11},\ 36 \leq \tilde{u}_3,\ 14 \leq \tilde{u}_1 \end{array} \right\}$$

3. $S = \{11, 5\}$, then $a^*(S) = 16$ and only subset of free-riding-proof core allocations for S can be attained:

$$\begin{array}{cl} & \left\{ \tilde{u} \in Core^{FRP}(\{11,5\}) : \tilde{u}_{11} > 68.5,\ \text{or}\ \tilde{u}_5 > 62.5 \right\} \\ = & \left\{ \begin{array}{c} \tilde{u} \in \mathbb{R}^5_+ : \tilde{u}_{11} + \tilde{u}_5 = 128,\ \tilde{u}_3 = 48,\ \tilde{u}_2 = 32,\ \tilde{u}_1 = 16, \\ [68.5 < \tilde{u}_{11} \leq 73 \text{ and } 55 \leq \tilde{u}_5 < 59.5] \\ \text{or } [62.5 < \tilde{u}_5 \leq 67.5 \text{ and } 60.5 \leq \tilde{u}_{11} < 65.5] \end{array} \right\} \end{array}$$

4. $S = \{11, 3\}$, then $a^*(S) = 14$ and only subset of free-riding-proof core allocations for S can be attained:

$$\begin{array}{cl} & \left\{ \tilde{u} \in Core^{FRP}(\{11,3\}) : \tilde{u}_{11} > 62.5 \right\} \\ = & \left\{ \begin{array}{c} \tilde{u} \in \mathbb{R}^5_+ : \tilde{u}_{11} + \tilde{u}_3 = 98,\ \tilde{u}_5 = 70,\ \tilde{u}_2 = 28,\ \tilde{u}_1 = 14, \\ [62.5 < \tilde{u}_{11} \leq 65 \text{ and } 33 \leq \tilde{u}_3 < 35.5] \end{array} \right\} \end{array}$$

5. $S = \{5, 3\}$, then $a^*(S) = 8$ and all free-riding-proof core allocations for S are attained:

$$Core^{FRP}(\{5,3\}) = \left\{ \begin{array}{c} \tilde{u} \in \mathbb{R}^5_+ : \tilde{u}_5 + \tilde{u}_3 = 32,\ \tilde{u}_{11} = 88,\ \tilde{u}_2 = 16,\ \tilde{u}_1 = 8, \\ 15 \leq \tilde{u}_5,\ 15 \leq \tilde{u}_3 \end{array} \right\}$$

We can observe that in the last two groups, the levels of public good provision are less than the Nash equilibrium provision level of the standard voluntary contribution game (recall that $a = 11$ by player 11's contribution only is the unique Nash equilibrium.):

- *There may be free-riding-proof core allocations that achieve less public good provision than Nash equilibrium one of a simple voluntary contribution game by Bergstrom, Blume and Varian (1986).*

This occurs since in our setup, player 11 can commit to being an outsider in the first stage. In a simultaneous move voluntary contribution game, this cannot happen. Finally, needless to say, we have:

- *The free-riding-proof core may be a highly nonconvex set.*

5 Summary

This paper added players' participation decisions to common agency games. The solution concept we used is a natural extension of coalition-proof Nash equilibrium to a dynamic game, perfectly coalition-proof Nash equilibrium (PCPNE). We considered a special class of common agency games: an environment without conflict of interests (comonotonic proferences) such as public good economies. In this case, we show that PCPNE is equivalent to an intuitive hybrid solution in transferrable utility case, the *free-riding-proof core*, which is the Pareto-frontier of a union of all core allocations for subset of players that are immune to unilateral free-riding incentives. With a simple example, we found that the equilibrium lobby group may not be consecutive (with respect to willingness-to-pay), and public good can be underprovided.

6 Proof of Proposition 2.

First, we construct a strategy profile σ below, which will be shown to support $(S^*, a^*(S^*), u^*)$ as a PCPNE. By definition, we have $u^* \in Core^{FRP}(S^*)$. In defining σ, we need to assign a CPNE utility profile to every subgame S' (although this does not happen in the equilibrium, it matters when deviations are considered). Then, we show that there is no credible and profitable deviation from σ by way of contradiction.

The second stage strategy profile is described by utility allocations assigned in each subgame (we utilize truthful strategies). Steps 1 and 4 are trivial, and step 2's allocation can be anything as long as it is a core allocation. Step 3 is the key, since we need to consider coalitional deviations from σ later. A credible deviation requires both free-riding-proofness and profitability. Thus, for player i to join a coalitional deviation that achieves S', utility level $\bar{u}_i = \max\{u_i^*, v_i(S' \backslash \{i\})\}$ plays important roles. We construct a core allocation (CPNE) for subgame S' by utilizing utility vector $\bar{u}$ in step 3.

1. We assign $(S^*, a^*(S^*), u^*) \in Core^{FRP}$ to the on-equilibrium subgame S^*.

2. For any S' with $S' \cap S^* = \emptyset$, we assign a CPNE that achieves an extreme point of the core for S' of a convex game (just to assign a concrete core allocation). For an arbitrarily selected order ω over S', we assign payoff vector $u_{\omega(1)} = V(\{\omega(1)\}) - V(\emptyset)$, $u_{\omega(2)} = V(\{\omega(1), \omega(2)\}) - V(\{\omega(1)\})$,... etc. following Shapley (1971). Call the allocation $\hat{u}_{S'} \in Core(S')$.[32]

3. For any S' with $S' \cap S^* \neq \emptyset$, we will assign a CPNE (or a core allocation). It requires a few steps. First, we deal with the outsiders. Let $\omega : |S' \backslash S^*| \to S' \backslash S^*$ be an arbitrary bijection, and let $u_{\omega(1)} = V(\{\omega(1)\})$, $u_{\omega(2)} = V(\{\omega(1), \omega(2)\}) - V(\{\omega(1)\})$,..., and $u_{\omega(|S' \backslash S^*|)} = V(S' \backslash S^*) - V(S' \backslash S^* \backslash \{\omega(|S' \backslash S^*|)\})$. Such a core allocation suppresses the total payoffs of $S' \backslash S^*$ the most (Shapley, 1971). The rest $V(S') - V(S' \backslash S^*)$ goes to $S' \cap S^*$. Consider a characteristic function form game $(\tilde{V}(Q; S' \backslash S^*))_{Q \subseteq S' \cap S^*}$ such that $\tilde{V}(Q; S' \backslash S^*) = V(Q \cup (S' \backslash S^*)) - V(S' \backslash S^*)$. For each $i \in S' \cap S^*$, let $\bar{u}_i = \max\{u_i^*, v_i(S' \backslash \{i\})\}$. With a vector $\bar{u}_{S' \cap S^*} = (\bar{u}_i)_{i \in S' \cap S^*}$ and a game $\tilde{V} = (\tilde{V}(Q; S' \backslash S^*))_{Q \subseteq S' \cap S^*}$, we will construct a core allocation from $\bar{u}$ in the following manner. Pick

[32] Let the ordering be a bijection $\omega : \{1, 2, ..., |S'|\} \to S'$. Let $u_{\omega(i)} = W(\{\omega(1), ..., \omega(i)) - W(\{1, ..., \omega(i-1)\})$. The allocation $(u_{\omega(i)})_{i=1}^{|S'|}$ is in the core for S', since the game W is convex.

$u = (u_i)_{i \in S' \cap S^*}$. Let

$$\begin{aligned}
\mathcal{Q}^+(u_{S'\cap S^*}) &= \{Q \in 2^{S'\cap S^*} : \sum_{j\in Q} u_j > \tilde{V}(Q; S'\backslash S^*)\}, \\
\mathcal{Q}^0(u_{S'\cap S^*}) &= \{Q \in 2^{S'\cap S^*} : \sum_{j\in Q} u_j = \tilde{V}(Q; S'\backslash S^*)\}, \\
\mathcal{Q}^-(u_{S'\cap S^*}) &= \{Q \in 2^{S'\cap S^*} : \sum_{j\in Q} u_j < \tilde{V}(Q; S'\backslash S^*)\}.
\end{aligned}$$

Since the game $\tilde{V}$ satisfies the convex game property, for all $Q \subseteq S' \cap S^*$, all $Q' \subsetneqq Q$, if we have $Q' \in \mathcal{Q}^0(u_{S'\cap S^*}) \cup \mathcal{Q}^+(u_{S'\cap S^*})$ with at least one $Q' \in \mathcal{Q}^0(u_{S'\cap S^*})$, then $Q \in \mathcal{Q}^0(u_{S'\cap S^*}) \cup \mathcal{Q}^-(u_{S'\cap S^*})$ holds.[33] To achieve a core allocation $\hat{u}_{S'\cap S^*}$, we need both $\mathcal{Q}^-(\hat{u}_{S'\cap S^*}) = \emptyset$ and $\sum_{i\in S'\cap S^*} \hat{u}_i = \tilde{V}(S' \cap S^*; S'\backslash S^*)$.

(a) Suppose $\mathcal{Q}^-(\bar{u}_{S'\cap S^*}) = \emptyset$. Starting with $u_{S'\cap S^*}(0) = \bar{u}_{S'\cap S^*}$, we modify utility vector $u_{S'\cap S^*}(t)$ continuously until it reaches at a core allocation, where t represents a stage of the modification process. Then, $2^{S'\cap S^*} = \mathcal{Q}^0(\bar{u}_{S'\cap S^*}) \cup \mathcal{Q}^+(\bar{u}_{S'\cap S^*})$. Reduce u_is by the same amount simultaneously and continuously for $i \in (S' \cap S^*)\backslash(\cup_{Q\in\mathcal{Q}^0(u_{S'\cap S^*})} Q)$.[34] Since all elements in $\mathrm{Q}^0(u_{S'\cap S^*})$ stay in $\mathrm{Q}^0(u_{S'\cap S^*})$ as the process continues, while some of elements of $\mathcal{Q}^+(u_{S'\cap S^*}(t))$ start switching to $\mathcal{Q}^0(u_{S'\cap S^*}(t))$, $\mathcal{Q}^0(u_{S'\cap S^*}(t))$ monotonically expands in the process. At some stage $t = \hat{t}$, $S' \cap S^* \in \mathcal{Q}^0(u_{S'\cap S^*}(\hat{t}))$ occurs. Then we immediately stop the process. This terminates the process, and the final outcome is $\hat{u}_{S'\cap S^*} = u_{S'\cap S^*}(\hat{t})$.

(b) Suppose $\mathcal{Q}^-(\bar{u}_{S'\cap S^*}) \neq \emptyset$. Start with $u_{S'\cap S^*}(0) = \bar{u}_{S'\cap S^*}$. Let

$$\mathcal{Q}^-_{\min}(u_{S'\cap S^*}) = \{Q \in \mathcal{Q}^-(u_{S'\cap S^*}) : \nexists Q' \in \mathcal{Q}^-(u_{S'\cap S^*}) \text{ with } Q \supsetneqq Q'\}.$$

In phase 1 ($t \in [0, \tilde{t}]$), increase u_is by the same amount continuously and simultaneously for all $i \in \cup_{Q\in\mathcal{Q}^-_{\min}(u_{S'\cap S^*}(t))} Q$. All other

[33]Here, we are assigning a core allocation to each subgame, since otherwise, a CPNE is not played in an off-equilibrium subgame, and the equilibrium cannot be a PCPNE. We propose an algorithm that achives a core allocation from a convex game and an arbitrary utility vector (here it is $\bar{u}_{S'\cap S^*}$).

[34]This means that if $S' \cap S^* \in \mathcal{Q}^0(u)$, then the process terminates.

players' utilities are kept intact. If some $Q \in \mathcal{Q}^{-}_{\min}(u_{S' \cap S^*}(t))$ falls into $\mathcal{Q}^0(u_{S' \cap S^*}(t))$ at some stage, then the utility levels are kept constant afterwards for the members of such Q. Continue this process until $\mathcal{Q}^{-}_{\min}(u_{S' \cap S^*}(t))$ (thus $\mathcal{Q}^{-}(u_{S' \cap S^*}(t))$) becomes empty. Call the end stage of phase 1 $t = \tilde{t}$.[35] In phase 2 ($t \in (\tilde{t}, \hat{t}]$), there are only $\mathcal{Q}^0(u_{S' \cap S^*}(t))$ and $\mathcal{Q}^{+}(u_{S' \cap S^*}(t))$. Now, we can repeat the procedure in (a), and we reach at a final outcome $\hat{u}_{S' \cap S^*} = u_{S' \cap S^*}(\hat{t})$ when $S' \cap S^* \in \mathcal{Q}^0(u_{S' \cap S^*}(\hat{t}))$ occurs.

4. Let $\sigma^1_i = 1$ for $i \in S^*$, and $\sigma^1_i = 0$ for $j \notin S^*$. Let $\sigma^2_i(S^*)$ be a truthful strategy relative to $a^*(S^*)$ with $\tau_i(a^*(S^*)) = v_i(a^*(S^*)) - u^*_i$ for all $i \in S^*$. And let $\sigma^2_i(S')$ be a truthful strategy relative to $a^*(S')$ with $\tau_i(a^*(S')) = v_i(a^*(S')) - \hat{u}_i(S')$ for all $i \in S'$.

Here, for step 3, we provide a useful observation for the later purpose.

Lemma 2. Suppose that game V is convex. Then, in cases (a) and (b) of step 3, we have the following results, respectively:

(a) We have $\hat{u}_{S' \cap S^*} \in Core(S' \cap S^*, \tilde{V})$, and

$$\hat{u}(S') = ((\hat{u}_i)_{i \in S' \cap S^*}, (u_{\omega(k)})_{k \in S' \setminus S^*}) \in Core(S').$$

For all $i \in S' \cap S^*$, $\hat{u}_i \leq \bar{u}_i$ holds with equalities only for $i \in Q \in \mathcal{Q}^0(\bar{u}_{S' \cap S^*})$. In particular, we have $\hat{u}_i = \bar{u}_i$ for all $i \in S' \cap S^*$ only when phase 1 does not start ($S' \cap S^* \in \mathcal{Q}^0(\bar{u}_{S' \cap S^*})$).

(b) We have $\hat{u}_{S' \cap S^*} \in Core(S' \cap S^*, \tilde{V})$, and

$$\hat{u}(S') = ((\hat{u}_i)_{i \in S' \cap S^*}, (u_{\omega(k)})_{k \in S' \setminus S^*}, (v_j(a^*(S')))_{j \in N \setminus S'}) \in Core(S').$$

For all $i \in W$, $\hat{u}_i > \bar{u}_i$ holds and there is $Q \ni i$ such that $\sum_{j \in Q} \hat{u}_j = \tilde{V}(Q)$ and $Q \subseteq W$. While, for all $j \in L$, $\hat{u}_j \leq \bar{u}_j$ holds with equalities only for $i \in Q \in \{Q' \in \mathcal{Q}^0(u_{S' \cap S^*}(\tilde{t})) : Q' \subseteq L\}$, where

$$W = (S' \cap S^*) \cap \left(\cup_{t \in [0,\tilde{t}]} \left(\cup_{Q \in \mathcal{Q}^{-}_{\min}(u_{S' \cap S^*}(t))} Q \right) \right),$$

[35] This process guarantees that a player i who belonged to some $Q \in \mathcal{Q}^{-}_{\min}(u_{S' \cap S^*}(t))$ (at some moment $t \in [0, \tilde{t}]$) must belongs to some $Q' \in \mathcal{Q}^0(u_{S' \cap S^*}(\tilde{t}))$ at the end of phase 1.

and

$$L = (S' \cap S^*) \backslash W.$$

In particular,

(α) when $L = \emptyset$, phase 2 does not take place ($\hat{t} = \tilde{t}$), and $\hat{u}_i \geq \bar{u}_i$ holds for all $i \in S' \cap S^*$.

(β) when $L \neq \emptyset$, for all $i \in W$, there is $Q \in \mathcal{Q}^0(u_{S' \cap S^*}(\tilde{t}))$ such that $\sum_{i \in Q} \hat{u}_i = \hat{V}(Q)$, and for all $j \in L$, $\hat{u}_j \leq \bar{u}_j$ holds with at least one strict inequality.

Proof of Lemma 2. First, we show that algorithm in case (a) terminates with $\sum_{i \in S' \cap S^*} u_i(\hat{t}) = \tilde{V}(S' \cap S^*)$. This is shown if $(S' \cap S^*) \backslash (\cup_{Q \in \mathcal{Q}^0(u_{S' \cap S^*})} Q) \neq \emptyset$ holds whenever $\sum_{i \in S' \cap S^*} u_i > \tilde{V}(S' \cap S^*)$ is the case (otherwise, $u_{S' \cap S^*}$ is infeasible while the algorithm stops). Suppose that $\sum_{i \in S' \cap S^*} u_i > \tilde{V}(S' \cap S^*)$, while $(S' \cap S^*) \backslash (\cup_{Q \in \mathcal{Q}^0(u_{S' \cap S^*})} Q) = \emptyset$ in case (a). Then, for all $i \in S' \cap S^*$, there exists $Q \in \mathcal{Q}^0(u_{S' \cap S^*})$ with $i \in Q$. Then, we can construct a balanced family $\mathcal{B}$ by collecting these Qs (see, say, Ichiishi, 1986). Then, with balanced weight $\{\lambda_Q\}_{Q \in \mathcal{B}}$ such that $\sum_{Q \ni i, Q \in \mathcal{B}} \lambda_Q = 1$ for all $i \in S' \cap S^*$. This implies

$$\sum_{Q \ni i, Q \in \mathcal{B}} \lambda_Q u_i = u_i.$$

Since for all $Q \in \mathcal{B}$, $\sum_{j \in Q} u_j = \tilde{V}(Q)$ by definition, we have

$$\sum_{Q \in \mathcal{B}} \lambda_Q \tilde{V}(Q) = \sum_{i \in S' \cap S^*} u_i.$$

By assumption, have $\sum_{i \in S' \cap S^*} u_i > \tilde{V}(S' \cap S^*)$, and we can conclude

$$\sum_{Q \in \mathcal{B}} \lambda_Q \tilde{V}(Q) > \tilde{V}(S' \cap S^*).$$

This means that the game $\tilde{V}$ is not balanced. However, convex games are balanced. This is a contradiction. Thus, in case (a), the algorithm terminates at a feasible allocation. The same argument applies to phase 2 of case (b).

Once the above is shown, it is easy to see the final outcome of the algorithm is in the core. By definition, $u_{S' \cap S^*} \in Core(S' \cap S^*, \tilde{V})$ if and only if $S' \cap S^* \in \mathcal{Q}^0(u_{S' \cap S^*})$ and $Q \in S' \cap S^* \in \mathcal{Q}^+(u_{S' \cap S^*}) \cap \mathcal{Q}^0(u_{S' \cap S^*})$ for all $Q \in 2^{S' \cap S^*}$. Thus, $\hat{u}_{S' \cap S^*} \in Core(S' \cap S^*, \tilde{V})$ is obvious for both cases. By the properties of convex games (Shapley, 1971), it is also easy to see $\hat{u}(S') \in Core(S')$ for both cases. We start with case (a). If $Q \in \mathcal{Q}^0(\bar{u}_{S' \cap S^*}) = \mathcal{Q}^0(u_{S' \cap S^*}(0))$, then utility is not adjusted for its members in the entire process. Thus, for all i who belong to some $Q \in \mathcal{Q}^0(\bar{u}_{S' \cap S^*})$, $\hat{u}_i = \bar{u}_i$ holds. For all others, it is obvious from the algorithm that $\hat{u}_i < \bar{u}_i$ holds. It is easy to see that we have $\hat{u}_i = \bar{u}_i$ for all $i \in S' \cap S^*$ only when phase 1 does not start ($S' \cap S^* \in \mathcal{Q}^0(\bar{u}_{S' \cap S^*})$).

Now, we check case (b). In phase 1, if player i belonged to some $Q \in \mathcal{Q}^-_{\min}(u_{S' \cap S^*}(t))$ at some moment $t \in [0, \tilde{t}]$, $u_i(\tilde{t}) > \bar{u}_i$ holds, since nobody gets utility reduction in phase 1. Moreover, such a player i belongs to some $Q \in \mathcal{Q}^0(u_{S' \cap S^*}(\tilde{t}))$ at the end of phase 1. This implies two things: $\sum_{j \in Q} \hat{u}_i = \tilde{V}(Q)$ for some $Q \in \mathcal{Q}^0(u_{S' \cap S^*}(\tilde{t}))$, and by the algorithm in case (a), such i's utility is intact in phase 2 (thus, $Q \subseteq W$). This implies that $\hat{u}_i > \bar{u}_i$ holds for all $i \in W$. All $j \in L$ was not affected in phase 1, and j belongs to either $Q \in \mathcal{Q}^0(u_{S' \cap S^*}(\tilde{t})) \cap L$ or $Q \in \mathcal{Q}^+(u_{S' \cap S^*}(\tilde{t})) \cap L$. The rest is a repetition of case (a). Convex game property (phase 1) assures that $\hat{u}_{S' \cap S^*} \in Core(S' \cap S^*, \tilde{V})$ holds, thus we have $\hat{u}(S') \in Core(S')$ (again, convex game). Statements of (α) and (β) are easy to show.$\square$

Since every subgame has a core allocation with truthful strategies, it is a CPNE. Thus, if there is a deviation from σ, then it must happen in the first stage. The rest of the proof of σ being a PCPNE is by way of a contradiction. Suppose to the contrary that coalition T profitably and credibly deviates from the equilibrium σ. Note that in the reduced game by T, it must be a PCPNE deviation given σ_{-T} fixed. In the original equilibrium, S^* is the lobby group. This implies that all $i \in (N \backslash S^*) \backslash T$ play $\sigma^1_i = 0$ in the first stage and they free-ride, while all $i \in S^* \backslash T$ play $\sigma^1_i = 1$ in the first stage and they play the same strategies ($\sigma^2_i(S')$ a menu contingent to formed lobby S') in the second stage. Note that all $i \in T \backslash S^*$ play $\sigma^{1\prime}_i = 1$ in the first period after the deviation (by definition), while $i \in T \cap S^*$ may or may not play $\sigma^{1\prime}_i = 1$. Some may choose to free-ride by switching to 0, while others stay in the lobby with adjustment of their strategies in the second stage.

Let S' be the lobby formed by T's deviation: $S' = S(\sigma^1_{-T}, \sigma^{1\prime}_T)$. Then, there are five groups of players (see Figure 1):

(i) the members of $S^* \backslash S' \subset T$ free-ride after the deviation,

(ii) the members of $S'\backslash S^* \subset T$ join the lobby,

(iii) the members of $(S^* \cap S')\backslash T \subset S'$ do not change their strategies in any stage (participate in lobbying, while keep the same menu in the second stage),

(iv) the members of $(S^* \cap S') \cap T \subset S'$ change strategies in the second stage,

(v) the members of $N\backslash(S' \cup S^*)$ are outsiders before or after the deviation.

Let the resulting allocation be $(S', a^*(S'), u')$. Since T is a profitable and credible deviation, the members in (i), (ii) and (iv) are better-off after T deviates. That is,

$$\begin{aligned} v_i(a^*(S')) &\geq u_i^* \text{ for all } i \in S^*\backslash S', \\ u_i' &\geq \bar{u}_i \text{ for all } i \in S'\backslash S^*, \\ u_i' &\geq \bar{u}_i \text{ for all } i \in (S^* \cap S') \cap T, \end{aligned}$$

must hold, where $\bar{u}_i = \max\{u_i^*, v_i^*(a^*(S'\backslash\{i\}))\}$.

Given our supposition, we will provide a sequence of claims below.

First note that since members of (ii) are better off, we have $a^*(S') > a^*(S^*)$. It is because (ii) is nonempty, since otherwise, $S' \subset S^*$ holds, and a coalitional deviation cannot be profitable.

Claim 1. $a^*(S') > a^*(S^*)$.

Since in σ, all players use truthful strategies, even after T's deviation, the members in (iii) (outsiders of T) get the same payoff vector $\hat{u}_{(S^*\cap S')\cap T}(S')$ as in the original subgame CPNE for S'. It is because in subgame S' (even after deviation), $a^*(S')$ must be provided since CPNE (core) must be assigned to the subgame. Thus, we have the following for group (iii).

Claim 2. After deviation by T, all $i \in (S^* \cap S')\backslash T \subset S'$ receives exactly $u_i' = \hat{u}_i(S')$.

Note that, since u' needs to be is a CPNE in the second stage of the reduced game by T, we have $\sum_{i\in S'\backslash S^*} u_i' \geq V(S'\backslash S^*)$, (to be in $Core(S')$). By construction of $\hat{u}(S')$, we have $\sum_{i\in S'\backslash S^*} \hat{u}_i = V(S'\backslash S^*)$. Thus, we have the following for group (ii).

Claim 3. $\sum_{i\in S'\backslash S^*} u_i' \geq \sum_{i\in S'\backslash S^*} \hat{u}_i = V(S'\backslash S^*)$.

Now, we consider group (iv). By Claims 2 and 3, the members of (iv) together can get at most

$$\sum_{i \in S' \cap S^* \cap T} u'_i \leq \sum_{i \in S' \cap S^* \cap T} \hat{u}_i,$$

since group (iv) cannot get transfers from groups (ii). Since group (iv) is better off and free-riding-proofness is satisfied for them after the deviation (PCPNE deviation), $u'_i \geq \bar{u}_i = \max\{u^*_i, v_i(a^*(S' \backslash \{i\}))\}$ must be satisfied for all $i \in S' \cap S^* \cap T$.

Claim 4. Suppose that for group $S' \cap S^* \cap T$, either case (a) or case (b) with $L \neq \emptyset$ holds. Then, $u'_i = \hat{u}_i$ for all $i \in S' \cap S^* \cap T$.

Proof of Claim 4. In case (a), Lemma 1 says that for all $i \in S' \cap S^*$, we have $\hat{u}_i \leq \bar{u}_i$. Claims 2 and 3 requires $\sum_{i \in S' \cap S^* \cap T} u'_i \leq \sum_{i \in S' \cap S^* \cap T} \hat{u}_i$. However, we need $u'_i \geq \bar{u}_i$ for all $i \in S' \cap S^* \cap T$. Thus, we have $u'_i = \hat{u}_i$ for all $i \in S' \cap S^* \cap T$.

In case (b) with $L \neq \emptyset$, by Lemma 1, some of the members of W must belong to (iv). However, for all $i \in W$, the winner group, there is $Q \in \mathcal{Q}^0(\hat{u}_{S' \cap S^*})$ with $i \in Q$, thus $\tilde{V}(Q) = \sum_{j \in Q'} \hat{u}_j(S')$. Now consider a reduced characteristic function form game played by $S' \cap S^* \cap T$, after group (iii) takes $\hat{u}_j$s (Claim 2). A reduced characteristic function form game $(\hat{V}(Q))_{Q \subseteq S' \cap S^* \cap T}$, where for all $Q \subseteq T \cap (S' \cap S^*)$, $\hat{V}(Q) = \tilde{V}(Q \cup ((S' \cap S^*) \backslash T)) - \sum_{j \in (S' \cap S^*) \backslash T} \hat{u}_j$. That is, we have $\hat{V}(Q \cap T) = \tilde{V}(Q) - \sum_{j \in Q \backslash T} \hat{u}_j = \sum_{i \in Q \cap T} \hat{u}_i$ for $Q \in \mathcal{Q}^0(\hat{u}_{S' \cap S^*})$. Therefore, we conclude that for all $i \in W$, there is $Q \in \mathcal{Q}^0(\hat{u}_{S' \cap S^*})$ with $i \in Q$ such that $\hat{V}(Q \cap T) = \sum_{j \in Q \cap T} \hat{u}_j(S')$. Since Claims 2 and 3 requires $\sum_{i \in S' \cap S^* \cap T} u'_i \leq \sum_{i \in S' \cap S^* \cap T} \hat{u}_i$, for $u'_{S' \cap S^* \cap T}$ to be in the core of $\hat{V}$ (so, CPNE of the reduced game), we need $u'_{S' \cap S^* \cap T} = \hat{u}_{S' \cap S^* \cap T}$.□

Claims 2, 3 and 4 immediately imply the following for group (ii).

Claim 5. Suppose that for group $S' \cap S^* \cap T$, either case (a) or case (b) with $L \neq \emptyset$ holds. Then, we have

$$\sum_{i \in S' \backslash S^*} u'_i = V(S' \backslash S^*) = \sum_{i \in S' \backslash S^*} v_i(a^*(S' \backslash S^*)) - C(a^*(S' \backslash S^*)).$$

Thus, we have shown that if $L \neq \emptyset$, then group (ii) can deviate profitably and credibly (together with group (iv)) achieve $u'_{S' \cap S^*}$ with a limited resource $V(S' \backslash S^*)$. Due to profitability, $V(S' \backslash S^*) \geq \sum_{i \in S' \backslash S^*} v_i(a^*(S^*))$, we have $a^*(S' \backslash S^*) > a^*(S^*)$. We consider a new allocation that is achieved only by group (ii).

Claim 6. Suppose that for group $S' \cap S^* \cap T$, either case (a) or case (b) with $L \neq \emptyset$ holds. Then, an allocation $(S' \backslash S^*, a^*(S' \backslash S^*), (u'_i)_{i \in S' \backslash S^*}, (v_j(a^*(S' \backslash S^*))_{j \notin S' \backslash S^*})$ can be achieved only by $S' \backslash S^*$ (u'_T is the deviators' allocation by T), and this allocation Pareto-dominates $(S^*, a^*(S^*), u^*)$.

Proof of Claim 6. First, groups (i) and (v) are better off, since $a^*(S' \backslash S^*) > a^*(S^*)$. By assumption, group (ii) are better off and have no free-riding incentives. Thus, the only groups which need investigation are groups (iii) and (iv). We check if there can be $i \in S' \cap S^*$ with $u^*_i > v_i(a^*(S' \backslash S^*))$ despite of $a^*(S' \backslash S^*) > a^*(S^*)$. Since $u^*_i \in Core(S^*)$, and the game V is convex, $u^*_i \leq V(S^*) - V(S^* \backslash \{i\})$ (Shapley, 1971). Since

$$
\begin{aligned}
& V(S^*) - V(S^* \backslash \{i\}) \\
= \; & \sum_{j \in S^*} v_j(a^*(S^*)) - C(a^*(S^*)) - \left(\sum_{j \in S^* \backslash \{i\}} v_j(a^*(S^* \backslash \{i\})) - C(a^*(S^* \backslash \{i\})) \right) \\
< \; & v_i(a^*(S' \backslash S^*)) \\
& + \sum_{j \in S^* \backslash \{i\}} v_j(a^*(S^*)) - C(a^*(S^*)) - \left(\sum_{j \in S^* \backslash \{i\}} v_j(a^*(S^* \backslash \{i\})) - C(a^*(S^* \backslash \{i\})) \right) \\
< \; & v_i(a^*(S' \backslash S^*)).
\end{aligned}
$$

The last inequality holds since $\sum_{j \in S^* \backslash \{i\}} v_j(a) - C(a)$ is maximized at $a = a^*(S^* \backslash \{i\})$. This proves that all members of (iii) and (iv) are better off in $(S' \backslash S^*, a^*(S' \backslash S^*), (u'_i)_{i \in S' \backslash S^*}, (v_j(a^*(S' \backslash S^*))_{j \notin S' \backslash S^*})$. Hence, we conclude that $(S^*, a^*(S^*), u^*) \in Core^{FRP}$ is Pareto-dominated by $(S' \backslash S^*, a^*(S' \backslash S^*), (u'_i)_{i \in S' \backslash S^*}, (v_j(a^*(S' \backslash S^*))_{j \notin S' \backslash S^*}) \in Core^{FRP}(S' \backslash S^*)$, since the members of (ii), $S' \backslash S^*$, have no free-riding incentive.$\square$

The last part of proof of Proposition 2. The statement of Claim 6 is an apparent contradiction to $(S^*, a^*(S^*), u^*) \in Core^{FRP}$, since $u'_i \geq v_i(a^*(S' \backslash \{i\}))$ for all $i \in S' \backslash S^*$ implies $u'_i \geq v_i(a^*((S' \backslash S^*) \backslash \{i\}))$ for all $i \in S' \backslash S^*$ (pure public good). Thus, we conclude that case (b) occurs and $L = \emptyset$ holds.

However, under case (b) with $L = \emptyset$, we have $\hat{u}_i > \bar{u}_i = \max\{u_i^*, v_i(a^*(S' \backslash \{i\}))\}$ for all $i \in S' \cap S^*$. Thus, members of group (iii) are better off and have no free-riding incentive. Players in groups (i), (ii) and (iv) deviate credibly and profitably by T, they are better-off and have no free-riding incentive for groups (ii) and (iv). Group (v) is better-off by Claim 1. This means that $(S', a^*(S'), (u_i')_{i \in S' \cap T}, (\hat{u}_i)_{i \in (S' \cap S^*) \backslash T}, (v_j(a^*(S')))_{j \in N \backslash S'}) \in Core^{FRP}(S')$, and Pareto-dominates $(S^*, a^*(S^*), u^*) \in Core^{FRP}$. This is a contradiction. Hence, $(S^*, a^*(S^*), u^*)$ is supportable with a PCPNE σ.□

Appendix

Lemma 1. In the linear utility- quadratic cost public good problem, the free-riding-proof core for S is nonempty if and only if S satisfies $\sum_{i \in S} \theta_i a^*(S) - \frac{1}{2}(a^*(S))^2 \geq \sum_{i \in S} \theta_i a^*(S \backslash \{i\})$ (aggregated "no free riding conditions").

Proof. If the above condition is violated, there is no allocation that satisfies no free riding for S. Thus, we only need to show that if the above condition is satisfied then we can find a core allocation that satisfies $\sum_{i \in T} u_i \geq V(T) = \sum_{i \in T} \theta_i a^*(T) - \frac{1}{2}(a^*(T))^2$. To be instructive, we will not explicitly solve $a^*(T)$ for a while. The strategy we take is to construct an allocation, and verify that it is in the core. Let $u_S \in \mathbb{R}_+^S$ be such that for all $i \in S$

$$u_i = \theta_i a^*(S \backslash \{i\}) + \frac{\theta_i}{\sum_{j \in S} \theta_j} \left(\sum_{i \in S} \theta_i a^*(S) - k(a^*(S))^2 - \sum_{j \in S} \theta_j a^*(S \backslash \{j\}) \right).$$

Notice that the contents of the parenthesis is the aggregated "no free riding" surplus: given the no free riding conditions, the most surplus the lobby group S can distribute for their members. The above formula distribute this surplus proportionally according to members' willingnesses-to-pay θs. Obviously, we have $\sum_{i \in S} u_i = V(S) = \sum_{i \in S} \theta_i a^*(S) - k(a^*(S))^2$, and $u_i \geq \theta_i a^*(S \backslash \{i\})$.

Thus, we only need to check condition 2. For a coalition $T \subsetneq S$, we have

$$\begin{aligned}
& \sum_{i \in T} u_i - V(T) \\
= \; & \sum_{i \in T} \theta_i a^*(S \backslash \{i\}) + \frac{\sum_{i \in T} \theta_i}{\sum_{j \in S} \theta_j} \left(\sum_{j \in S} \theta_j a^*(S) - \frac{1}{2}(a^*(S))^2 - \sum_{j \in S} \theta_j a^*(S \backslash \{j\}) \right) \\
& - \left(\sum_{i \in T} \theta_i a^*(T) - \frac{1}{2}(a^*(T))^2 \right) \\
= \; & \frac{\sum_{i \in T} \theta_i}{\sum_{j \in S} \theta_j} \left(\sum_{j \in S} \theta_j a^*(S) - \frac{1}{2}(a^*(S))^2 \right) - \left(\sum_{i \in T} \theta_i a^*(T) - \frac{1}{2}(a^*(T))^2 \right) \\
& + \sum_{i \in T} \theta_i a^*(S \backslash \{i\}) - \frac{\sum_{i \in T} \theta_i}{\sum_{j \in S} \theta_j} \sum_{j \in S} \theta_j a^*(S \backslash \{j\}).
\end{aligned}$$

We want this to be nonnegative for all $T \subset S$. Now, we use quadratic cost and linear utility. The first order condition for optimal public good provision is

$$a^*(S) = \sum_{i \in S} \theta_i.$$

Thus, we have

$$\sum_{i \in S} \theta_i a^*(S) - k(a^*(S))^2 = \frac{\left(\sum_{i \in S} \theta_i\right)^2}{2},$$

and

$$\theta_i a^*(S \backslash \{i\}) = \theta_i \left(\sum_{j \in S} \theta_j - \theta_i \right).$$

Thus, we have

$$
\begin{aligned}
&\sum_{i\in T} u_i - V(T) \\
= \;& \frac{\sum_{i\in T}\theta_i}{2\sum_{j\in S}\theta_j}\left(\sum_{j\in S}\theta_j\right)^2 - \frac{1}{2}\left(\sum_{i\in T}\theta_i\right)^2 + \sum_{i\in T}\theta_i \sum_{j\neq i, j\in S}\theta_j - \frac{\sum_{i\in T}\theta_i}{\sum_{i\in S}\theta_i}\sum_{i\in S}\theta_i \sum_{j\neq i, j\in S}\theta_j \\
= \;& \frac{1}{2}\left(\sum_{i\in T}\theta_i\right)\left(\sum_{j\in S}\theta_j\right) + \sum_{i\in T}\theta_i\left(\sum_{j\in S}\theta_j - \theta_i\right) - \frac{\sum_{i\in T}\theta_i}{\sum_{i\in S}\theta_i}\sum_{i\in S}\theta_i\left(\sum_{j\in S}\theta_j - \theta_i\right) \\
= \;& \frac{1}{2}\left(\sum_{i\in T}\theta_i\right)\left(\sum_{j\in S}\theta_j\right) + \sum_{i\in T}\theta_i\left(\sum_{j\in S}\theta_j\right) - \sum_{i\in T}\theta_i^2 - \sum_{i\in T}\theta_i\left(\sum_{j\in S}\theta_j\right) + \frac{\sum_{i\in T}\theta_i}{\sum_{i\in S}\theta_i}\sum_{i\in S}\theta_i^2 \\
= \;& \frac{1}{2}\left(\sum_{i\in T}\theta_i\right)\left(\sum_{j\in S}\theta_j\right) - \sum_{i\in T}\theta_i^2 + \frac{\sum_{i\in T}\theta_i}{\sum_{i\in S}\theta_i}\sum_{i\in S}\theta_i^2 \\
= \;& \left(\sum_{i\in T}\theta_i\right)\left[\frac{\sum_{j\in S}\theta_j}{2} - \frac{\sum_{i\in T}\theta_i^2}{\sum_{i\in T}\theta_i} + \frac{\sum_{i\in S}\theta_i^2}{\sum_{i\in S}\theta_i}\right] \\
= \;& \left(\sum_{i\in T}\theta_i\right)\left[\frac{\sum_{j\in S}\theta_j}{2} - \sum_{j\in T}\frac{\theta_j}{\sum_{i\in T}\theta_i}\times\theta_j + \sum_{j\in S}\frac{\theta_j}{\sum_{i\in S}\theta_i}\times\theta_j\right].
\end{aligned}
$$

The second term is the only negative term, and it takes maximum absolute value when T is composed by the players with the highest values of θ_j. Let us call such value $\theta_{\max}$. Suppose that $\sum_{i\in S} u_i - V(T) < 0$. Then, by focusing the first two terms, we know $\theta_{\max} > \frac{1}{2}\sum_{i\in S}\theta_i$. However, if it is the case, we have

$$
\begin{aligned}
& \frac{\sum_{j\in S}\theta_j}{2} - \sum_{j\in T}\frac{\theta_j}{\sum_{i\in T}\theta_i}\times\theta_j + \sum_{j\in S}\frac{\theta_j}{\sum_{i\in S}\theta_i}\times\theta_j \\
\geq \;& \frac{\sum_{j\in S}\theta_j}{2} - \theta_{\max} + \sum_{j\in S}\frac{\theta_j}{\sum_{i\in S}\theta_i}\times\theta_j \\
\geq \;& \frac{\theta_{\max}}{2} - \theta_{\max} + \frac{\theta_{\max}}{\sum_{i\in S}\theta_i}\times\theta_{\max} \\
> \;& \frac{\theta_{\max}}{2} - \theta_{\max} + \frac{1}{2}\times\theta_{\max} = 0.
\end{aligned}
$$

This is a contradiction. Therefore, u is in $Core^{FRP}(S)$.■

References

[1] Aumann, R., Acceptable Points in General Cooperative N-Person Games, in "Contributions to the Theory of Games IV," H.W. Kuhn and R.D. Luce eds. (1959), Princeton University Press (Princeton), 287-324.

[2] Bergstrom, T., L. Blume, and H. Varian, 1986, On the Private Provision of Public Goods, Journal of Public Economics 29, 25-49.

[3] Bernheim, B.D., B. Peleg, and M.D. Whinston, 1987, Coalition-Proof Nash Equilibria I: Concepts, Journal of Economic Theory 42, 1-12.

[4] Bernheim, B.D., and M.D. Whinston, 1986, Menu Auctions, Resource Allocation, and Economic Influence, Quarterly Journal of Economics 101, 1-31.

[5] Bloch, F., Non-Cooperative Models of Coalition Formation in Games with Spillovers, in "New Directions in the Economic Theory of the Environment," C. Carraro and D. Siniscalco eds. (1997), Cambridge University Press (Cambridge, U.K.), 311-352.

[6] Bombardini, M., Firm Heterogeneity and Lobby Participation, UBC Working Paper, University of British Columbia

[7] Conley, J.P., and H. Konishi, 2002, Migration-Proof Tiebout Equilibrium: Existence and Asymptotic Efficiency, Journal of Public Economics 86, 241-260.

[8] d'Aspremont, C., A. Jacquemin, J.J. Gabszevicz, and J.A. Weymark, 1983, On the Stability of Collusive Price Leadership, Canadian Journal of Economics 16, 17-25.

[9] Foley, D.K., 1970, Lindahl's Solution and the Core of an Economy with Public Goods, Econometrica 38, 66-72.

[10] Gomes, A., and P. Jehiel, 2005, Dynamic Processes of Social and Economic Interactions: On the Persistence of Inefficiencies, Journal of Political Economy 113, 626-667.

[11] Grossman, G., and E. Helpman, 1994, Protection for Sale, American Economic Review 84, 833-850.

[12] Groves T., and J. Ledyard, 1977, Optimal Allocation of Public Goods: a Solution to the “Free Rider” Problem, Econometrica 45, 783-811.

[13] Hart, S., and M. Kurz, 1983, Endogeneous Formation of Coalitions, Econometrica 51, 1047-1064.

[14] Hurwicz, L., 1979, Outcome Functions Yielding Walrasian and Lindahl Allocations at Nash Equilibrium Points, Review of Economic Studies 46, 217-225.

[15] Ichishi, T., 1983, *Game Theory for Economic Analysis*, (Academic Press, New York).

[16] Konishi, H., M. Le Breton and S. Weber, 1999, On Coalition-Proof Nash Equilibria in Common Agency Games, Journal of Economic Theory 85, 122-139.

[17] Konishi, H., and D. Ray, 2003, Coalition Formation as a Dynamic Process, Journal of Economic Theory 110, 1-41.

[18] Konishi, H., and M.U. Ünver, 2006, Credible Group-Stability in Many-to-Many Matching Problems, Journal of Economic Theory 129, 57-80 (2006).

[19] Laussel, D., and M. Le Breton, 1998, Efficient Private Provision of Public Goods under Common Agency, Games and Economic Behavior 25, 194-218.

[20] Laussel, D., and M. Le Breton, 2001, Conflict and Cooperation: the Structure of Equilibrium Payoffs in Common Agency, Journal of Economic Theory 100, 93-128.

[21] Le Breton, M., and F. Salaniè, 2003, Lobbying under Political Uncertainty, Journal of Public Economics 87, 2589-2610.

[22] Mitra, D., 1999, Endogenous Lobby Formation and Endogenous Protection: A Long-Run Model of Trade Policy Determination, American Economic Review 89, 1116-1134.

[23] Maruta, T., and A. Okada, 2005, Group Formation and Heterogeneity in Collective Action Games, Hitotsubashi University Discussion Paper #2005-7.

[24] Nishimura, Y., and R. Shinohara, 2006, A Voluntary Participation Game through a Unit-by-Unit Cost Share Mechanism of a Non-Excludable Public Good, Working Paper, Shinshu University.

[25] Palfrey, R.T., and H. Rosenthal, 1984, Participation and the Provision of Discrete Public Goods: a Strategic Analysis, Journal of Public Economics 24, 171-193.

[26] Paltseva, E., 2006, Protection for Sale to Oligopolists, IIES Working Paper, Stockholm University.

[27] Ray, D., 1989, Credible Coalitions and the Core, International Journal of Game Theory 18, 185-187.

[28] Ray, D., and R. Vohra, 2001, Coalitional Power and Public Goods, Journal of Political Economy 109, 1355–1384.

[29] Saijo, T., and T. Yamato, 1999, A Voluntary Participation Game with a Non-Excludable Public Good, Journal of Economic Theory 84, 227-242.

[30] Samuelson, P.A., 1954, The Pure Theory of Public Expenditure, Review of Economics and Statistics 36, 387-389.

[31] Shapley, L., 1971, Core of Convex Games, International Journal of Game Theory 1, 11-26.

[32] Shinohara, R., 2003, Coalition-proof Equilibria in a Voluntary Participation Game, Working Paper, Hitotsubashi University.

[33] Thoron, S., 1998, A Coalition-Proof Stable Cartel, Canadian Journal of Economics 31, 63-76.

[34] Walker, M., 1981, A Simple Incentive Compatible Scheme for Attaining Lindahl Allocations, Econometrica 49, 65-71.

[35] Yi., S.-S., 1996, Endogenous Formation of Customs Unions under Imperfect Competition: Open Regionalism Is Good, Journal of International Economics 41, 153-177.