Implementation of Kernel Correspondence

Yusuke Samejima

Abstract

This paper proposes a game form that fully implements the kernel correspondence in subgame-perfect equilibrium. This finite-stage extensive game form incorporates objections and counter-objections in the definition of the kernel proposed by Osborne and Rubinstein [1994]. In the game form, agents do not have to report preferences, which are difficult to communicate in reality. Furthermore, the game form achieves implementation for any class of allocation problems where preferences are quasi-linear and the associated coalitional form game is superadditive.

1. Introduction

This paper proposes a game form that implements the kernel correspondence in subgame-perfect equilibrium.

The kernel is a solution concept for coalitional form games with transferable utility and was first introduced by Davis and Maschler [1965]. Osborne and Rubinstein [1994] proposed another equivalent definition of the kernel in terms of objections and counter-objections in negotiations. The kernel is the set of payoff divisions such that for every objection, there is a counter-objection. Imagine that there is a payoff division, which we call the status quo, on the table of negotiations. Suppose that some agent k makes an objection against agent ℓ to the status quo by naming a coalition S to which k belongs that excludes ℓ, and saying S's dissatisfaction is very large; S's dissatisfaction is a difference between S's payoff achievable on its own and S's payoff at the status quo. However, if the status quo is in the kernel, ℓ can make a counter-objection against k by naming another coalition T to which ℓ belongs that excludes k, and saying T's dissatisfaction is not less than S's. Such a proper counter-objection can make the objection ineffective.

By its definition, the kernel is stable against objections. In other words, a payoff division in the kernel indicates a way of sharing the society's benefits, which is desirable in the sense that it is immune

to any complaint from any society member.

Implementation theory aims to develop a tool for the uninformed social planner who wishes to realize certain desirable allocations. When the social planner, or the society, attempts to realize desirable allocations, he must collect information on the preferences of members in the society. However, it is often the case that the social planner has difficulty in collecting such information while the concerned members share much information on each other. For such circumstances, a game form can be used as a tool for the social planner. The game form itself is defined independently of the preferences of members in the society. As the literature on implementation theory has proposed, properly designed game forms can realize desirable allocations in equilibrium of the games even if the social planner is given an insufficient amount of information.

This paper presents a game form that realizes the set of kernel allocations in subgame-perfect equilibrium. We incorporate objections and counter-objections in Osborne and Rubinstein's definition of the kernel into a finite-stage extensive game form. In the first stage of the game form, each agent is given an opportunity for an objection. If there is no proper counter-objection to the objection, the objector can obtain a higher payoff than at the status quo no matter what counter-objection is tried to make. However, if there exists a proper counter-objection, it turns out that the game form chooses the status quo. With the game form, every kernel allocation is attainable in subgame-perfect equilibrium. Moreover, every equilibrium allocation achieves a payoff division in the kernel. That is, the game form *fully* implements the kernel correspondence in subgame-perfect equilibrium.

Our investigation of implementation of the kernel correspondence might be worthwhile in the following three aspects.

First, our attempt to design a game form that incorporates the objection and counter-objection phase to implement the kernel correspondence in allocation problems appears new in the literature. Important solutions for coalitional form games include the Shapley value, the bargaining set, the nucleolus, and the kernel.[1] There are many studies on implementation of the Shapley value.[2] Studies on implementation of the bargaining set, which is also defined in terms of objections and counter-objections as the kernel, include Serrano and Vohra [2002], Einy and Wettstein [1999], and Perez-Castrillo and Wettstein [2000] among others. These papers present game forms that employ the objection and counter-objection phase, which has inspired the idea of the present paper. Implementation of the nucleolus in allocation problems as in the present paper is studied by Samejima [2004]. However,

---

[1] For the definitions of these solutions, readers are referred to Osborne and Rubinstein [1994].

[2] The literature on implementation and non-cooperative foundations of the Shapley value includes Bag and Winter [1999], Gul [1989], Hart and Mas-Colell [1996], Perez-Castrillo and Wettstein [2001], and others.

studies on implementation of the kernel in allocation problems appear to have been missing.

Second, our game form possesses various desirable properties compared to the canonical game form developed by Moore and Repullo [1988]. Moore and Repullo present a sufficient condition, called the condition C^+, and the canonical game form for subgame-perfect implementation in the general environment with three or more agents. The condition C^+ involves the existence of a finite sequence of allocations satisfying certain preference relations, but such a sequence is difficult to find. Furthermore, agents have to report their preferences in the canonical game form while agents do not have to do so in our game form. This is a desirable property of our game form because preferences are difficult to communicate in reality. Our game form is always balanced in the sense that there are neither deficits nor surpluses even on off-equilibrium paths. The game form also satisfies individual feasibility, which is an important requirement for natural mechanisms (Saijo, Tatamitani, and Yamato [1996]). Furthermore, the game form achieves implementation for any class of allocation problems where preferences are quasi-linear and the associated coalitional form game is superadditive.

Third, our approach to the kernel is what Bergin and Duggan [1999] call an implementation-theoretic approach to non-cooperative foundations of cooperative solutions. Studies on non-cooperative foundations aim to provide non-cooperative models of negotiations whose equilibrium outcomes agree with cooperative solutions of coalitional form games. Bergin and Duggan discuss the advantage of the implementation-theoretic approach over the coalitional-function approach to non-cooperative foundations. The implementation-theoretic approach advocates explicitly modelling the physical environment and individual preferences and constructing a game form independent of preferences to implement a certain cooperative solution. On the other hand, the coalitional-function approach models a process of negotiations directly on the coalitional function that gives payoffs or worths of coalitions without specifying details of the environment and preferences, which are critical for determining feasible payoffs. Following Bergin and Duggan's spirit, the present paper takes the implementation-theoretic approach to non-cooperative foundations of the kernel, while Serrano [1997] takes the coalitional-function approach. Besides its approach, Serrano [1997] differs from the present paper in that Serrano considers sequential bilateral matching of agents in the infinite time horizon model.

The remaining part of this paper is organized as follows. Section 2 defines allocation problems and introduces the notions of the kernel and implementation. Section 3 proposes a game form that implements the kernel correspondence in subgame-perfect equilibrium. Section 4 provides some concluding remarks.

2. Preliminaries

2–1. Preferences

Let $N = \{1, 2, \ldots, n\}$ be a finite set of agents with at least three members ($n \geq 3$). A *coalition* is a non-empty subset of N, and it is typically denoted by S or T. Each agent $i \in N$ has the consumption space denoted by $X_i \times \mathbb{R}$ where X_i is a non-empty set and $\mathbb{R}$ is the set of real numbers. A consumption bundle in $X_i \times \mathbb{R}$ is denoted by (x_i, m_i) where $m_i \in \mathbb{R}$ is the amount of monetary transfer that agent i receives and $x_i \in X_i$ is the remaining part of the consumption bundle. Agent i's endowment bundle is $a_i^\omega \equiv (\omega_i, 0) \in X_i \times \mathbb{R}$.

We assume that each agent has a quasi-linear preference relation since the definition of the kernel requires interpersonal comparisons of payoffs.

P1. Each agent $i \in N$ has a preference relation represented by a real-valued quasi-linear utility function $u_i(x_i) + m_i$ defined over $X_i \times \mathbb{R}$.

We call $u_i(\cdot)$ agent i's *valuation function*. The set of all admissible valuation functions of agent i is denoted by $\mathcal{U}_i$. We call $u \equiv (u_1, u_2, \ldots, u_n)$ a *valuation profile*. The set of all admissible valuation profiles is denoted by $\mathcal{U} \equiv \mathcal{U}_1 \times \mathcal{U}_2 \times \cdots \times \mathcal{U}_n$.

For notational convenience, given $a = (x_i, m_i)_{i \in S}$, we define

$$u_S(x) \equiv \sum_{i \in S} u_i(x_i),$$

$$m_S \equiv \sum_{i \in S} m_i,$$

$$U_i(a) \equiv u_i(x_i) + m_i, \quad \text{and}$$

$$U_S(a) \equiv u_S(x) + m_S.$$

2–2. Allocations

The set of allocations that a coalition S can achieve on its own is predetermined exogenously and is denoted by $A_S \subseteq \prod_{i \in S}(X_i \times \mathbb{R})$. Elements in A_S are referred to as *S-feasible allocations*, on which we assume the following.

A1. For each coalition S, each S-feasible allocation $(x_i, m_i)_{i \in S} \in A_S$, and each transfer vector $(m_i')_{i \in S} \in \mathbb{R}^S$ such that $m_S' = 0$, we have $(x_i, m_i + m_i')_{i \in S} \in A_S$.

A2. For each pair of coalitions $S, T \subsetneq N$ such that $S \cap T = \emptyset$, each S-feasible allocation $(x_i, m_i)_{i \in S} \in A_S$, and each T-feasible allocation $(x_i, m_i)_{i \in T} \in A_T$, we have $(x_i, m_i)_{i \in S \cup T} \in A_{S \cup T}$.

A1 says that any transfer among agents in a coalition is feasible as long as it is balanced within the coalition. A2 says that for any two disjoint coalitions and two allocations feasible for the coalitions, the combined allocation is feasible for the combined coalition. Throughout the paper, we fix allocation feasibilities $(A_S)_{\emptyset \neq S \subseteq N}$.

We define S-efficient allocations as follows. An S-feasible allocation $a \in A_S$ is S-efficient if $U_S(a) \geq U_S(a')$ for all $a' \in A_S$. Note that for each S-efficient allocation $(x_i, m_i)_{i \in S}$ and each transfer vector $m' \in \mathbb{R}^S$ such that $m'_S = 0$, another allocation $(x_i, m_i + m'_i)_{i \in S}$ is also S-efficient.

Example. *Excludable public good economies.* The following model is one of the examples of allocation problems to which our analysis can be applied. A public good is *excludable* if it is possible to prevent any agent from the benefit of the public good. For simplicity, we assume a linear production technology, that is, one unit of the public good is produced with one unit of monetary contribution from agents. An excludable public good economy consists of

(1) a finite set of agents N,

(2) for each $i \in N$, the consumption space $\mathbb{R}_+ \times \mathbb{R}$ with its typical element (x_i, m_i) where $x_i \geq 0$ is the level of the public good,

(3) for each $i \in N$, the endowment bundle $(0, 0) \in \mathbb{R}_+ \times \mathbb{R}$,

(4) for each $i \in N$, a quasi-linear utility function $u_i(x_i) + m_i$ defined over the consumption space where u_i is a differentiable, increasing, and concave valuation function such that $\frac{d}{dx_i} u_i(x_i) < 1$ for sufficiently large $x_i \in \mathbb{R}_+$,

(5) for each coalition $S \subseteq N$, the set of S-feasible allocations

$$A_S = \{(x_i, m_i)_{i \in S} \in \prod_{i \in S} (\mathbb{R}_+ \times \mathbb{R}) : \text{For each partition } \{S_h\}_{h=1}^H \text{ of } S,$$

$$[\, i, j \in S_h \Rightarrow x(S_h) \equiv x_i = x_j \,] \quad \text{and} \quad \sum_{h=1}^H x(S_h) = -\sum_{i \in S} m_i \}.$$

Item (5) describes the excludability of the public good. If two agents i and j belong to the same sub-coalition S_h, they consume the same level $x(S_h)$ of the public good. But if they belong to different sub-coalitions, they may consume different levels. This model of an excludable public good is studied by Bag and Winter [1999], who investigate implementation of the core and the Shapley value. A1 and A2 are clearly satisfied. The existence of S-efficient allocations is ensured by the assumption that for each $i \in S$, $\frac{d}{dx_i} u_i(x_i) < 1$ for large x_i.

2–3. The kernel correspondence

Given an S-efficient allocation a, we can define the *worth* of a coalition S by $v_S \equiv U_S(a)$. Such v_S is uniquely determined independently of the choice of an S-efficient allocation. We denote $v_{\{i\}}$ by v_i.

The list $(N, (v_S)_{\emptyset \neq S \subseteq N})$, simply denoted by (N, v), constitutes a *coalitional form game* with transferable utility. A payoff division among agents is represented by a vector $q \equiv (q_1, q_2, \ldots, q_n)$. For notational convenience, let $q_S \equiv \sum_{i \in S} q_i$. A payoff vector q is *N-efficient* if $q_N = v_N$. A payoff vector q is *individually rational* if $q_i \geq v_i$ for all $i \in N$. An *imputation* of (N, v) is an N-efficient and individually rational payoff vector. By A2, the game (N, v) is *superadditive* in the sense that for coalitions $S, T \subsetneq N$ such that $S \cap T = \emptyset$, we have $v_{S \cup T} \geq v_S + v_T$. Furthermore, the set of imputations is non-empty by the superadditivity.

The kernel is a solution concept for coalitional form games; it is a collection of payoff divisions that are stable in a certain sense. The kernel was first introduced by Davis and Maschler [1965], and Osborne and Rubinstein [1994] presented another equivalent definition of the kernel. To describe the definition due to Osborne and Rubinstein, we define *objections* and *counter-objections* as follows.

K1. An *objection* of $k \in N$ against $\ell \in N$ to an imputation q is a coalition S such that $k \in S$, $\ell \notin S$, and $q_\ell > v_\ell$.

K2. A *counter-objection* to the objection S of k against ℓ is a coalition T such that $\ell \in T$, $k \notin T$, and $v_S - q_S \leq v_T - q_T$.

Definition (Osborne and Rubinstein, 1994)**.** The *kernel* of (N, v) is the set of imputations q such that for every objection S of any agent k against any other agent ℓ to q, there is a counter-objection to S.

The number $(v_S - q_S)$ is called the *excess* of S at q, and $(v_T - q_T)$ is the excess of T at q. The excess represents the amount of dissatisfaction of a coalition. An objection represents a complaint by $k \in S$ who is dissatisfied with the payoff given to his coalition S that excludes ℓ for whom $q_\ell > v_\ell$. A counter-objection represents a counter-argument by $\ell \in T$ who claims that the amount of dissatisfaction of S does not exceed that of his coalition T that excludes k. A payoff division in the kernel is stable in the sense that every objection can be made ineffective by some counter-objection.

Fact. The kernel of (N, v) is non-empty if the set of imputations for (N, v) is non-empty.

A proof of this fact is found in Osborne and Rubinstein [1994]. Since a game (N, v) induced by our allocation problem is superadditive and hence the set of imputations is non-empty, the kernel of (N, v)

is non-empty.

Given a valuation profile $u \in \mathcal{U}$, we obtain a coalitional form game (N, v) associated with it. A *kernel allocation* $a \in A_N$ is an allocation that realizes a payoff division q in the kernel of the associated coalitional form game. That is, for a kernel allocation a, we have $U_i(a) = q_i$ for all $i \in N$ where q is in the kernel. Let $\varphi(u)$ be the set of all kernel allocations for a valuation profile $u \in \mathcal{U}$. This multi-valued function $\varphi \colon \mathcal{U} \twoheadrightarrow A_N$ is called the *kernel correspondence*.

2–4. Implementation

We consider a finite-stage extensive game form Γ with perfect information. The *game form* Γ consists of a game tree with the set of choices available to agents at each decision node, and an outcome function O. A list of agent i's choice at each decision node is called agent i's *strategy* s_i. A strategy profile $(s_i)_{i \in N}$ is typically denoted by s, while another strategy profile $(s'_k, (s_i)_{i \in N \setminus \{k\}})$ is denoted by (s'_k, s_{-k}). With each strategy profile s, an outcome function O associates an N-feasible allocation $O(s) \in A_N$.

Given $u \in \mathcal{U}$, a pair (Γ, u) constitutes an extensive form *game*. A strategy profile s is called a *subgame-perfect equilibrium* of (Γ, u) if the choices specified in the strategy profile constitute a Nash equilibrium in every subgame of (Γ, u). It is well known that a pure-strategy subgame-perfect equilibrium can be obtained by backward induction. The set of outcome allocations corresponding to pure-strategy subgame-perfect equilibria of (Γ, u) is denoted by $SPE(\Gamma, u)$.

We say that a game form Γ *fully implements the kernel correspondence* φ *in subgame-perfect equilibrium* if $\varphi(u) = SPE(\Gamma, u)$ for all $u \in \mathcal{U}$. The full implementation requires that every kernel allocation can be realized as an equilibrium allocation as well as every equilibrium allocation is in fact a kernel allocation.

3. Result

3–1. The game form

This section presents a game form that implements the kernel correspondence in subgame-perfect equilibrium. The game form is depicted in Figure 1. Formally, the game form Γ is defined as follows.

The game form Γ.

The choices available to agents and the outcome function O are defined in the following description.

We note that a list of agent i's choice at each decision node is called agent i's strategy s_i, and a strategy profile $s = (s_i)_{i \in N}$ determines one terminal node at which the game stops. For each terminal node, the game form Γ describes an outcome allocation. So, for each strategy profile s, the game form Γ chooses one outcome

allocation, which implicitly defines the outcome function O that associates an N-feasible allocation $O(s) \in A_N$ with each strategy profile s.

In Stages 2A, 4A, and 4B of the game form Γ, agents move sequentially in order of index numbers.

Stage 1. Each agent $i \in N$ *simultaneously* chooses $(a^i, \hat{a}_i, \ell^i, S^i, e^i, z^i)$ where

$\quad\quad\quad a^i \in A_N$ is an N-feasible allocation,

$\quad\quad\quad \hat{a}_i \in A_{\{i\}}$ is an $\{i\}$-feasible allocation,

$\quad\quad\quad \ell^i \in N \setminus \{i\}$ is any other agent,

$\quad\quad\quad S^i \subsetneq N$ is a non-empty proper subset of N such that $i \in S^i$ and $\ell^i \notin S^i$, and

$\quad\quad\quad e^i$ and z^i are real numbers.

Rule 1A. If $a^i = a$ for all $i \in N$, then the outcome is a. STOP.

Rule 1B. If $a^i = a$ for all $i \in N \setminus \{k\}$ and $a^k \neq a$ for some $k \in N$, then apply the following Rule 1B1, 1B2, or 1B3.

$\quad$*Rule 1B1.* If $z^k < 0$, then go to Stage 2A.

$\quad$*Rule 1B2.* If $z^k > 0$, then define $\ell \equiv \ell^k$, $S \equiv S^k$, $e \equiv e^k$, $z \equiv z^k$, and $j \equiv \max(N \setminus \{k, \ell\})$ [3] and go to Stage 2B.

$\quad$*Rule 1B3.* If $z^k = 0$, then the outcome is such that k is allocated $\hat{a}_k$ and each $i \in N \setminus \{k\}$ is allocated his endowment bundle a_i^ω. STOP.

Rule 1C. If none of the above applies, the outcome is a^i with $i = \max(\arg\max_{h \in N} z^h)$.[4] STOP.

Stage 2A. Each agent in $N \setminus \{k\}$ *sequentially* chooses to either agree or disagree with a^k.

Rule 2A1. If all agree, then the outcome is a^k. STOP.

Rule 2A2. Otherwise, the outcome is a. STOP.

Stage 2B. Agent ℓ chooses one of the following rules.

Rule 2B1. Go to Stage 3A.

Rule 2B2. Choose $T \subsetneq N$ that is a non-empty proper subset of N such that $\ell \in T$ and $k \notin T$ and go to Stage 3B.

Rule 2B3. Choose an $\{\ell\}$-feasible allocation $\hat{a}_\ell \in A_{\{\ell\}}$. The outcome is such that ℓ is allocated $\hat{a}_\ell$ and each $i \in N \setminus \{\ell\}$ is allocated his endowment bundle a_i^ω. STOP.

---

[3]In words, agent j is the agent with the largest index number among all agents except for k and ℓ.

[4]In words, agent i in Rule 1C is the agent with the largest index number among those who have chosen the largest real number $z^i = \max_{h \in N} z^h$.

Stage 3A. Agent k chooses to either quit or propose $a^S \in A_S$.

Rule 3A1. If k quits, then the outcome is a. STOP.

Rule 3A2. If k proposes a^S and $S \neq \{k\}$, then go to Stage 4A.

Rule 3A3. If k proposes a^S and $S = \{k\}$, the outcome is constructed as in Rule 4A1. STOP.

Stage 4A. Each agent in $S \setminus \{k\}$ *sequentially* chooses to either agree or disagree with a^S.

Rule 4A1. If all agree, then the outcome is constructed as follows.

      First, allocate a^S to agents in S and a_i^ω to each $i \in N \setminus S$, respectively.

      Second, if $e \geq 0$ and $j \notin S$, then k pays the amount e to j.

      Third, if $e < 0$, then ℓ pays the amount $(-e)$ to k. STOP.

Rule 4A2. Otherwise, the outcome is constructed as follows.

      First, allocate a to agents in N.

      Second, if $e \geq 0$ and $j \in S$, then k pays the amount e to j. STOP.

Stage 3B. Agent ℓ chooses to either quit or propose $a^T \in A_T$.

Rule 3B1. If ℓ quits, then the outcome is constructed as follows.

      First, allocate a to agents in N.

      Second, ℓ pays the amount z to k. STOP.

Rule 3B2. If ℓ proposes a^T and $T \neq \{\ell\}$, then go to Stage 4B.

Rule 3B3. If ℓ proposes a^T and $T = \{\ell\}$, the outcome is constructed as in Rule 4B1. STOP.

Stage 4B. Each agent in $T \setminus \{\ell\}$ *sequentially* chooses to either agree or disagree with a^T.

Rule 4B1. If all agree, then the outcome is constructed as follows.

      First, allocate a^T to agents in T and a_i^ω to each $i \in N \setminus T$, respectively.

      Second, if $e + z \geq 0$ and $j \notin T$, then ℓ pays the amount $(e + z)$ to j.

      Third, if $e + z < 0$, then k pays the amount $(-e - z)$ to ℓ. STOP.

Rule 4B2. Otherwise, the outcome is constructed as follows.
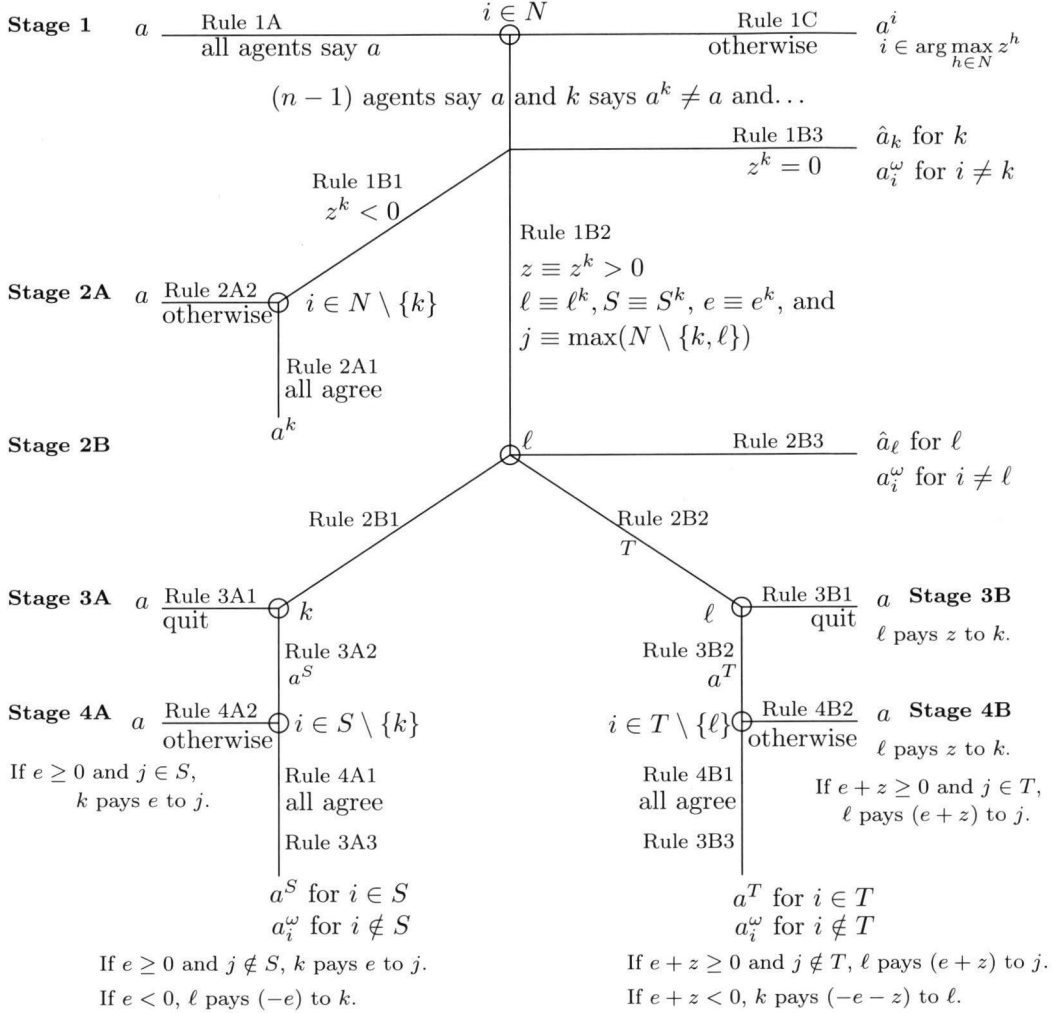
      First, allocate a to agents in N.

      Second, ℓ pays the amount z to k.

      Third, if $e + z \geq 0$ and $j \in T$, then ℓ pays the amount $(e + z)$ to j. STOP.

Note that the game form Γ is defined independently of a valuation profile u. The result of the present paper is the following.

Theorem. *The game form Γ fully implements the kernel correspondence φ in subgame-perfect equilibrium. That is, for all $u \in \mathcal{U}$, we have $\varphi(u) = SPE(\Gamma, u)$.*

Figure 1: The game form Γ

Stage 1 a $\dfrac{\text{Rule 1A}}{\text{all agents say } a}$ $i \in N$ $\dfrac{\text{Rule 1C}}{\text{otherwise}}$ a^i $i \in \arg\max\limits_{h \in N} z^h$

$(n-1)$ agents say a and k says $a^k \neq a$ and...

$\dfrac{\text{Rule 1B3}}{z^k = 0}$ $\hat{a}_k$ for k a_i^ω for $i \neq k$

Rule 1B1
$z^k < 0$

Rule 1B2
$z \equiv z^k > 0$
$\ell \equiv \ell^k, S \equiv S^k, e \equiv e^k$, and
$j \equiv \max(N \setminus \{k, \ell\})$

Stage 2A a $\dfrac{\text{Rule 2A2}}{\text{otherwise}}$ $i \in N \setminus \{k\}$

Rule 2A1
all agree

a^k

Stage 2B ℓ $\dfrac{\text{Rule 2B3}}{}$ $\hat{a}_\ell$ for ℓ a_i^ω for $i \neq \ell$

Rule 2B1

Rule 2B2
T

Stage 3A a $\dfrac{\text{Rule 3A1}}{\text{quit}}$ k ℓ $\dfrac{\text{Rule 3B1}}{\text{quit}}$ a **Stage 3B**

ℓ pays z to k.

Rule 3A2
a^S

Rule 3B2
a^T

Stage 4A a $\dfrac{\text{Rule 4A2}}{\text{otherwise}}$ $i \in S \setminus \{k\}$ $i \in T \setminus \{\ell\}$ $\dfrac{\text{Rule 4B2}}{\text{otherwise}}$ a **Stage 4B**

If $e \geq 0$ and $j \in S$,
   k pays e to j.

Rule 4A1
all agree

Rule 4B1
all agree

ℓ pays z to k.

If $e + z \geq 0$ and $j \in T$,
   ℓ pays $(e+z)$ to j.

Rule 3A3

Rule 3B3

a^S for $i \in S$
a_i^ω for $i \notin S$

a^T for $i \in T$
a_i^ω for $i \notin T$

If $e \geq 0$ and $j \notin S$, k pays e to j.

If $e < 0$, ℓ pays $(-e)$ to k.

If $e + z \geq 0$ and $j \notin T$, ℓ pays $(e+z)$ to j.

If $e + z < 0$, k pays $(-e - z)$ to ℓ.

- -

In Stage 1, each $i \in N$ simultaneously chooses

   $a^i \in A_N$,

   $\hat{a}_i \in A_{\{i\}}$,

   $\ell^i \in N \setminus \{i\}$,

   $S^i \subsetneq N$ such that $i \in S^i$ and $\ell^i \notin S^i$, and

   e^i and z^i are real numbers.

k is a deviator in Stage 1 and chooses ℓ.

i's endowment bundle is denotd by a_i^ω.

In Stage 2B, ℓ chooses one of the following:

   1) Rule 2B1,

   2) Rule 2B2 with $T \subsetneq N$ such that

       $\ell \in T$ and $k \notin T$, or

   3) Rule 2B3 with $\hat{a}_\ell \in A_{\{\ell\}}$.

In Stage 3A, k quits or proposes $a^S \in A_S$.

In Stage 3B, ℓ quits or proposes $a^T \in A_T$.

In Stages 2A, 4A, and 4B, agents move sequentially and choose to agree or disagree.

3–2. Proof

3–2–1. $\varphi(u) \subset SPE(\Gamma, u)$

Let us fix a valuation profile $u \in \mathcal{U}$. Take any kernel allocation $a \in \varphi(u)$. We define a strategy s_i for each agent $i \in N$. We will prove that the strategy profile $s = (s_i)_{i \in N}$ is a subgame-perfect equilibrium of the game (Γ, u).

Agent i's strategy s_i.

Stage 1.

(*Choice 1*) Agent i chooses $(a^i, \hat{a}_i, \ell^i, S^i, e^i, z^i)$ such that

$$a^i = a \text{ (the kernel allocation chosen in the previous paragraph)},$$
$$\hat{a}_i = a_i^\omega,$$
$$\ell^i = i(\mod n) + 1,$$
$$S^i = \{i\}, \text{ and}$$
$$e^i = z^i = 0.$$

Stage 2A. In case i moves in Stage 2A,

(*Choice 2A*) if $U_i(a^k) \geq U_i(a)$ then i agrees, and if $U_i(a^k) < U_i(a)$ then i disagrees.

Stage 2B. In case i moves as ℓ in Stage 2B, his choice is the following. Choose any

$$\hat{T} \in \arg\max(v_T - U_T(a)) \text{ subject to } T \subsetneq N, \ell \in T, \text{ and } k \notin T.$$

(*Choice 2B1*) Consider the case where $e \geq \max\{v_S - U_S(a), v_{\hat{T}} - U_{\hat{T}}(a)\}$.

　　If $v_\ell < U_\ell(a)$, then ℓ chooses Rule 2B1 to go to Stage 3A.

　　If $v_\ell \geq U_\ell(a)$, then ℓ chooses Rule 2B3 and an $\{\ell\}$-efficient allocation $\hat{a}_\ell$.

(*Choice 2B2*) Consider the case where $e < \min\{v_S - U_S(a), v_{\hat{T}} - U_{\hat{T}}(a)\}$.

　　If $v_\ell < v_{\hat{T}} - U_{\hat{T}}(a) + U_\ell(a) - e - z$, then ℓ chooses Rule 2B2 and $T = \hat{T}$ to go to Stage 3B.

　　If $v_\ell \geq v_{\hat{T}} - U_{\hat{T}}(a) + U_\ell(a) - e - z$, then ℓ chooses Rule 2B3 and an $\{\ell\}$-efficient allocation $\hat{a}_\ell$.

(*Choice 2B3*) Consider the case where $v_{\hat{T}} - U_{\hat{T}}(a) \leq e < v_S - U_S(a)$.

　　If $v_\ell < U_\ell(a) - z$, then ℓ chooses Rule 2B2 and $T = \hat{T}$ to go to Stage 3B.

　　If $v_\ell \geq U_\ell(a) - z$, then ℓ chooses Rule 2B3 and an $\{\ell\}$-efficient allocation $\hat{a}_\ell$.

(*Choice 2B4*) Consider the case where $v_S - U_S(a) \leq e < v_{\hat{T}} - U_{\hat{T}}(a)$.

　　If $U_\ell(a) > v_\ell$ and $U_\ell(a) \geq v_{\hat{T}} - U_{\hat{T}}(a) + U_\ell(a) - e - z$, then ℓ chooses Rule 2B1 to go to Stage 3A.

　　If $v_{\hat{T}} - U_{\hat{T}}(a) + U_\ell(a) - e - z > \max\{v_\ell, U_\ell(a)\}$, then ℓ chooses Rule 2B2 and $T = \hat{T}$ to go to Stage 3B.

　　If $v_\ell \geq \max\{U_\ell(a), v_{\hat{T}} - U_{\hat{T}}(a) + U_\ell(a) - e - z\}$, then ℓ chooses Rule 2B3 and an $\{\ell\}$-efficient $\hat{a}_\ell$.

Stage 3A. In case i moves as k in Stage 3A,

(*Choice 3A1*) if $e \geq v_S - U_S(a)$ then k quits.

(*Choice 3A2*) If $e < v_S - U_S(a)$ then k propose a^S that is S-efficient and satisfies the following:

$U_j(a^S) = U_j(a) + \max\{0, e\}$ if $j \in S$ (recall that $j \equiv \max(N \setminus \{k, \ell\})$), and

$U_h(a^S) = U_h(a)$ for all $h \in S \setminus \{j, k\}$.

Stage 4A. In case i moves in Stage 4A, i either agrees or disagrees according to the following.

(*Choice 4A1*) Consider the case where $e < 0$ or $i \neq j$.

If $U_i(a^S) \geq U_i(a)$ then i agrees, and if $U_i(a^S) < U_i(a)$ then i disagrees.

(*Choice 4A2*) Consider the case where $e \geq 0$ and $i = j$.

If $U_j(a^S) \geq U_j(a) + e$ then j agrees, and if $U_j(a^S) < U_j(a) + e$ then j disagrees.

Stage 3B. In case i moves as ℓ in Stage 3B,

(*Choice 3B1*) if $e \geq v_T - U_T(a)$ then ℓ quits.

(*Choice 3B2*) If $e < v_T - U_T(a)$ then ℓ proposes a^T that is T-efficient and satisfies the following:

$U_j(a^T) = U_j(a) + \max\{0, e + z\}$ if $j \in T$, and

$U_h(a^T) = U_h(a)$ for all $h \in T \setminus \{j, \ell\}$.

Stage 4B. In case i moves in Stage 4B, i either agrees or disagrees according to the following.

(*Choice 4B1*) Consider the case where $e + z < 0$ or $i \neq j$.

If $U_i(a^T) \geq U_i(a)$ then i agrees, and if $U_i(a^T) < U_i(a)$ then i disagrees.

(*Choice 4B2*) Consider the case where $e + z \geq 0$ and $i = j$.

If $U_j(a^T) \geq U_j(a) + e + z$ then j agrees, and if $U_j(a^T) < U_j(a) + e + z$ then j disagrees.

To apply backward induction arguments, our analysis starts from the stages closest to terminal nodes.

Lemma 1. *For any subgame that starts in Stage 2A, the strategy profile induced by s is a subgame-perfect equilibrium.*

Proof. We first note that agents in $N \setminus \{k\}$ move sequentially in Stage 2A, so the set of subgames that start in Stage 2A includes many subgames with various lengths. For any such subgame, the following backward induction arguments apply.

Consider a decision node that is one step before a terminal node, and let us call the mover of the decision node agent i. By Rules 2A1 and 2A2 of the game form, if i agrees then the outcome is either a^k or a, depending on the history of moves. If i disagrees then the outcome is a. So, if $U_i(a^k) \geq U_i(a)$ then

agreeing is weakly preferred to disagreeing, and if $U_i(a^k) < U_i(a)$ then disagreeing is weakly preferred to agreeing. Therefore, Choice 2A induced by s_i is an optimal choice for i.

Next, consider a decision node that is two steps before a terminal node. If the agent of the node agrees then the outcome is either a^k or a, depending on the history of moves and the choices of subsequent movers. If the agent disagrees then the outcome is a. So, Choice 2A is an optimal choice for him.

Repeating the above arguments, we conclude that Choice 2A is an optimal choice for each agent at every decision node of the subgame of our concern. Therefore, the strategy profile induced by s is a subgame-perfect equilibrium of this subgame. $\qquad\square$

Lemma 2. *For any subgame that starts in Stage 4A, the strategy profile induced by s is a subgame-perfect equilibrium.*

Proof. Similar arguments as in the proof of Lemma 1 can show that Choice 4A1 induced by s_i is an optimal choice for all $i \in S \setminus \{k\}$ when $e < 0$ or $i \neq j \equiv \max(N \setminus \{k, \ell\})$.

A remark on Choice 4A2 follows. Choice 4A2 is valid for the case $e \geq 0$ and for agent j when he is in S. For such agent j, if j disagrees then his utility is $(U_j(a) + e)$ because he receives the amount e by Rule 4A2. If j agrees then his utility is either $(U_j(a) + e)$ or $U_j(a^S)$, depending on the history of moves. Therefore, Choice 4A2 is an optimal choice for him. $\qquad\square$

Lemma 3. *For any subgame that starts in Stage 4B, the strategy profile induced by s is a subgame-perfect equilibrium.*

Proof. Similar arguments as in the proof of Lemma 2 apply. $\qquad\square$

Lemma 4. *For any subgame that starts in Stage 3A, the strategy profile induced by s is a subgame-perfect equilibrium.*

Proof. It is sufficient to show that Choices 3A1 and 3A2 are optimal for agent k at the initial node of this subgame. This is because by Lemma 2 the strategy profile induced by s is a subgame-perfect equilibrium of any subgame that starts in Stage 4A that is reached after Stage 3A. Recall that $j \equiv \max(N \setminus \{k, \ell\})$.

Case 1: $e < 0$ or $j \notin S$.

For agent k, feasible utilities under the strategy profile induced by s are the following: $U_k(a)$ by Rule 3A1, $(U_k(a^S) - e)$ by Rule 3A3 or Rule 4A1, and $U_k(a)$ by Rule 4A2.

First, we investigate what $a^S \in A_S$ maximizes $(U_k(a^S) - e)$. When $S = \{k\}$, the solution is a^S that is S-efficient, and the maximized value is $(v_k - e)$. When $S \neq \{k\}$, Rule 4A1 applies only when all agree in Stage 4A, which occurs when $U_i(a^S) \geq U_i(a)$ for each $i \in S \setminus \{k\}$ by Choice 4A1. Note that $U_k(a^S) \equiv U_S(a^S) - \sum_{i \in S \setminus \{k\}} U_i(a^S)$. So, we want to consider a problem

$$\max_{a^S \in A_S} U_S(a^S) - \sum_{i \in S \setminus \{k\}} U_i(a^S) - e \quad \text{subject to} \quad U_i(a^S) \geq U_i(a) \text{ for each } i \in S \setminus \{k\}.$$

It is clear that the solution to the problem is a^S that is S-efficient and $U_i(a^S) = U_i(a)$ for all $i \in S \setminus \{k\}$, and the maximized value is $(v_S - \sum_{i \in S \setminus \{k\}} U_i(a) - e)$.

Second, we investigate the condition under which the maximized $(U_k(a^S) - e)$ exceeds $U_k(a)$, that is, agent k prefers proposing a^S to quitting in Stage 3A. When $S = \{k\}$, the condition is such that $v_k - e > U_k(a)$ which is equivalent to the condition $e < v_S - U_S(a)$. When $S \neq \{k\}$, the condition is such that

$$v_S - \sum_{i \in S \setminus \{k\}} U_i(a) - e > U_k(a)$$

which is equivalent to the condition $e < v_S - U_S(a)$.

By the above arguments, Choices 3A1 and 3A2 are optimal for agent k in Case 1.

Case 2: $e \geq 0$ and $j \in S$.

For agent k, feasible utilities under the strategy profile induced by s are the following: $U_k(a)$ by Rule 3A1, $U_k(a^S)$ by Rule 4A1, and $(U_k(a) - e)$ by Rule 4A2. Agent k weakly prefers Rule 3A1 to Rule 4A2.

First, we investigate what $a^S \in A_S$ maximizes $U_k(a^S)$. Rule 4A1 applies only when all agree in Stage 4A, which occurs when $U_j(a^S) \geq U_j(a) + e$ for j by Choice 4A2 and $U_i(a^S) \geq U_i(a)$ for each $i \in S \setminus \{j, k\}$ by Choice 4A1. Therefore, $U_k(a^S)$ is maximized by a^S that is S-efficient and $U_j(a^S) = U_j(a) + e$ for j and $U_i(a^S) = U_i(a)$ for each $i \in S \setminus \{j, k\}$, and the maximized value is $(v_S - \sum_{i \in S \setminus \{k\}} U_i(a) - e)$.

Second, we investigate the condition under which the maximized $U_k(a^S)$ exceeds $U_k(a)$, that is, agent k prefers proposing a^S to quitting in Stage 3A. The condition is such that

$$v_S - \sum_{i \in S \setminus \{k\}} U_i(a) - e > U_k(a)$$

which is equivalent to the condition $e < v_S - U_S(a)$.

By the above arguments, Choices 3A1 and 3A2 are optimal for agent k in Case 2. □

Lemma 5. *For any subgame that starts in Stage 3B, the strategy profile induced by s is a subgame-perfect equilibrium.*

Proof. We apply similar arguments as in the proof of Lemma 4. It is sufficient to show that Choices 3B1 and 3B2 are optimal for agent ℓ at the initial node of this subgame. This is because by Lemma 3 the strategy profile induced by s is a subgame-perfect equilibrium of any subgame that starts in Stage 4B that is reached after Stage 3B. Recall that $j \equiv \max(N \setminus \{k, \ell\})$.

Case 1: $e + z < 0$ or $j \notin T$.

For agent ℓ, feasible utilities under the strategy profile induced by s are the following: $(U_\ell(a) - z)$ by Rule 3B1, $(U_\ell(a^T) - e - z)$ by Rule 3B3 or Rule 4B1, and $(U_\ell(a) - z)$ by Rule 4B2.

First, we investigate what $a^T \in A_T$ maximizes $(U_\ell(a^T) - e - z)$. When $T = \{\ell\}$, the solution is a^T that is T-efficient, and the maximized value $(v_\ell - e - z)$. When $T \neq \{\ell\}$, Rule 4B1 applies only when all agree in Stage 4B, which occurs when $U_i(a^T) \geq U_i(a)$ for each $i \in T \setminus \{\ell\}$ by Choice 4B1. Note that $U_\ell(a^T) \equiv U_T(a^T) - \sum_{i \in T \setminus \{\ell\}} U_i(a^T)$. So, we want to consider a problem

$$\max_{a^T \in A_T} U_T(a^T) - \sum_{i \in T \setminus \{\ell\}} U_i(a^T) - e - z \quad \text{subject to} \quad U_i(a^T) \geq U_i(a) \text{ for each } i \in T \setminus \{\ell\}.$$

It is clear that the solution to the problem is a^T that is T-efficient and $U_i(a^T) = U_i(a)$ for all $i \in T \setminus \{\ell\}$, and the maximized value is $(v_T - \sum_{i \in T \setminus \{\ell\}} U_i(a) - e - z)$.

Second, we investigate the condition under which the maximized $(U_\ell(a^T) - e - z)$ exceeds $(U_\ell(a) - z)$, that is, agent ℓ prefers proposing a^T to quitting in Stage 3B. When $T = \{\ell\}$, the condition is such that $v_\ell - e - z > U_\ell(a) - z$ which is equivalent to the condition $e < v_T - U_T(a)$. When $T \neq \{\ell\}$, the condition is such that

$$v_T - \sum_{i \in T \setminus \{\ell\}} U_i(a) - e - z > U_\ell(a) - z$$

which is equivalent to the condition $e < v_T - U_T(a)$.

By the above arguments, Choices 3B1 and 3B2 are optimal for agent ℓ in Case 1.

Case 2: $e + z \geq 0$ and $j \in T$.

For agent ℓ, feasible utilities under the strategy profile induced by s are the following: $(U_\ell(a) - z)$ by Rule 3B1, $U_\ell(a^T)$ by Rule 4B1, and $(U_\ell(a) - z - (e + z))$ by Rule 4B2. Agent ℓ weakly prefers Rule 3B1 to Rule 4B2.

First, we investigate what $a^T \in A_T$ maximizes $U_\ell(a^T)$. Rule 4B1 applies only when all agree in Stage 4B, which occurs when $U_j(a^T) \geq U_j(a) + e + z$ for j by Choice 4B2 and $U_i(a^T) \geq U_i(a)$ for each $i \in T \setminus \{j, \ell\}$ by Choice 4B1. Therefore, $U_\ell(a^T)$ is maximized by a^T that is T-efficient and $U_j(a^T) = U_j(a) + e + z$ for j and $U_i(a^T) = U_i(a)$ for each $i \in T \setminus \{j, \ell\}$, and the maximized value is $(v_T - \sum_{i \in T \setminus \{\ell\}} U_i(a) - e - z)$.

Second, we investigate the condition under which the maximized $U_\ell(a^T)$ exceeds $(U_\ell(a) - z)$, that is, agent ℓ prefers proposing a^T to quitting in Stage 3B. The condition is such that

$$v_T - \sum_{i \in T \setminus \{\ell\}} U_i(a) - e - z > U_\ell(a) - z$$

which is equivalent to the condition $e < v_T - U_T(a)$.

By the above arguments, Choices 3B1 and 3B2 are optimal for agent ℓ in Case 2. $\qquad\square$

Lemma 6. *For any subgame that starts in Stage 2B, the strategy profile induced by s is a subgame-perfect equilibrium.*

Proof. It is sufficient to show that Choices 2B1, 2B2, 2B3, and 2B4 are optimal for agent ℓ at the initial node of this subgame. This is because by Lemmas 4 and 5 the strategy profile induced by s is a subgame-perfect equilibrium of any subgame that starts in Stages 3A and 3B that are reached after Stage 2B. We need two claims before proving the lemma.

Claim 1. *When ℓ optimally chooses Rule 2B3, ℓ should choose $\hat{a}_\ell \in A_{\{\ell\}}$ that is $\{\ell\}$-efficient, and the maximized value of ℓ's utility is v_ℓ.*

Proof of Claim 1. The claim holds because ℓ's utility under Rule 2B3 is $U_\ell(\hat{a}_\ell)$. *Q.E.D.*

Claim 2. *When ℓ optimally chooses Rule 2B2 in order to obtain the utility under Rule 3B3 or 4B1, ℓ should choose $T = \hat{T} \in \arg\max(v_T - U_T(a))$ subject to $T \subsetneq N, \ell \in T$, and $k \notin T$.*

Proof of Claim 2. Since the above choice of $\hat{T}$ maximizes $(v_T - U_T(a))$, this $\hat{T}$ also maximizes $(v_T - U_T(a) + U_\ell(a) - e - z)$ because the last three terms are independent of T. Note that

$$\text{if } T = \{\ell\} \text{ then } v_T - U_T(a) + U_\ell(a) - e - z = v_\ell - e - z, \text{ and}$$

$$\text{if } T \neq \{\ell\} \text{ then } v_T - U_T(a) + U_\ell(a) - e - z = v_T - \sum_{i \in T \setminus \{\ell\}} U_i(a) - e - z.$$

These values, $(v_\ell - e - z)$ and $(v_T - \sum_{i \in T \setminus \{\ell\}} U_i(a) - e - z)$, are ℓ's utilities under Rule 3B3 or 4B1 when the strategy profile induced by s is taken, as we have seen in the proof of Lemma 5. Therefore, $\hat{T}$ maximizes ℓ's utility under Rule 3B3 or 4B1. *Q.E.D.*

To show that Choices 2B1, 2B2, 2B3, and 2B4 are optimal for agent ℓ, we now investigate agent ℓ's utility realized by a choice of each of Rules 2B1, 2B2, and 2B3, in the following exhaustive cases. We note that $z > 0$, by Rule 1B2, in these cases.

Case 1: $e \geq \max\{v_S - U_S(a), v_{\hat{T}} - U_{\hat{T}}(a)\}$.

For agent ℓ, feasible utilities under the strategy profile induced by s are the following.

Rule 2B1 leads to $U_\ell(a)$ since Rule 3A1 applies due to Choice 3A1.

Rule 2B2 leads to $(U_\ell(a) - z)$ since Rule 3B1 applies due to Choice 3B1 for any T.

Rule 2B3 leads to v_ℓ due to Claim 1.

Therefore, Choice 2B1 is optimal for agent ℓ in Case 1.

Case 2: $e < \min\{v_S - U_S(a), v_{\hat{T}} - U_{\hat{T}}(a)\}$.

For agent ℓ, feasible utilities under the strategy profile induced by s are the following.

Rule 2B1 leads to $U_\ell(a_\ell^\omega)$ or $(U_\ell(a_\ell^\omega) - |e|)$ since Rule 3A3 or 4A1 applies because Choice 3A2 is taken and all agree in Stage 4A. Note that $U_\ell(a_\ell^\omega) \leq v_\ell$ by the definition of v_ℓ.

Rule 2B2, with an optimal choice of $T = \hat{T}$ due to Claim 2, leads to $(v_{\hat{T}} - U_{\hat{T}}(a) + U_\ell(a) - e - z)$ since Rule 3B3 or 4B1 applies because Choice 3B2 is taken and all agree in Stage 4B.

Rule 2B3 leads to v_ℓ due to Claim 1.

Therefore, Choice 2B2 is optimal for agent ℓ in Case 2.

Case 3: $v_{\hat{T}} - U_{\hat{T}}(a) \leq e < v_S - U_S(a)$.

For agent ℓ, feasible utilities under the strategy profile induced by s are the following.

Rule 2B1 leads to $U_\ell(a_\ell^\omega)$ or $(U_\ell(a_\ell^\omega) - |e|)$ since Rule 3A3 or 4A1 applies because Choice 3A2 is taken and all agree in Stage 4A. Note that $U_\ell(a_\ell^\omega) \leq v_\ell$ by the definition of v_ℓ.

Rule 2B2 leads to $(U_\ell(a) - z)$ since Rule 3B1 applies due to Choice 3B1 for any T.

Rule 2B3 leads to v_ℓ due to Claim 1.

Therefore, Choice 2B3 is optimal for agent ℓ in Case 3.

Case 4: $v_S - U_S(a) \leq e < v_{\hat{T}} - U_{\hat{T}}(a)$.

For agent ℓ, feasible utilities under the strategy profile induced by s are the following.

Rule 2B1 leads to $U_\ell(a)$ since Rule 3A1 applies due to Choice 3A1.

Rule 2B2, with an optimal choice of $T = \hat{T}$ due to Claim 2, leads to $(v_{\hat{T}} - U_{\hat{T}}(a) + U_\ell(a) - e - z)$ since Rule 3B3 or 4B1 applies because Choice 3B2 is taken and all agree in Stage 4B.

Rule 2B3 leads to v_ℓ due to Claim 1.

Therefore, Choice 2B4 is optimal for agent ℓ in Case 4. $\qquad\square$

Lemma 7. *For the game (Γ, u), the strategy profile s is a subgame-perfect equilibrium.*

Proof. Take any agent $k \in N$. Under the strategy profile s, Rule 1A of the game form applies and the outcome $O(s)$ is a, which is the kernel allocation chosen at the beginning of Section 3–2–1. Hence agent k's utility under s is $U_k(a)$. We show that agent k's utility cannot be strictly higher than $U_k(a)$ by any deviation from s_k to s_k'' when the other agents choose s_{-k}; that is, $U_k(O(s)) \geq U_k(O(s_k'', s_{-k}))$ for any s_k''. Denote by $(a^k, \hat{a}_k, \ell, S, e, z)$ agent k's choice in Stage 1 under s_k''. If $a^k = a$, then $O(s_k'', s_{-k}) = a$, and agent k's utility remains the same after the deviation. So, we assume that $a^k \neq a$ henceforth.

Let s_k' be agent k's strategy such that his choice is the same as s_k'' in Stage 1 and his choices are the same as s_k in the other stages. By Lemmas 1 and 6, in any subgame starting after Stage 1, agent k's strategy induced by s_k is a best response to the strategy profile induced by s_{-k}, and hence $U_k(O(s_k', s_{-k})) \geq U_k(O(s_k'', s_{-k}))$. So, it is sufficient to show that $U_k(a) \geq U_k(O(s_k', s_{-k}))$. Under the strategy profile (s_k', s_{-k}), either one of the following rules applies in Stage 1: Rule 1B1, 1B2, or 1B3. We consider the following exhaustive cases.

Case 1: $z = 0$ and Rule 1B3 applies.

In this case, $O(s_k', s_{-k}) = \hat{a}_k$ and agent k's utility is $U_k(\hat{a}_k)$, which is maximized when $\hat{a}_k \in A_{\{k\}}$ is $\{k\}$-efficient. So, agent k's maximum feasible utility is v_k.

On the other hand, since any payoff vector in the kernel is individually rational by its definition and a is a kernel allocation, we must have $U_k(a) \geq v_k$.

Hence, $U_k(a) \geq U_k(O(s_k', s_{-k}))$ in this case.

Case 2: $z < 0$ and Rule 1B1 applies.

In this case, the game proceeds to Stage 2A, where either Rule 2A1 or 2A2 applies.

When Rule 2A2 applies, $O(s_k', s_{-k}) = a$ and agent k's utility is $U_k(a)$.

When Rule 2A1 applies, $O(s_k', s_{-k}) = a^k$ and agent k's utility is $U_k(a^k)$. Since this rule applies when all agents in $N \setminus \{k\}$ agree in Stage 2A, it must be the case that $U_i(a^k) \geq U_i(a)$ according to Choice 2A described in s_i for each $i \in N \setminus \{k\}$. Suppose, by contradiction, that $U_k(a^k) > U_k(a)$. Then we must have $U_N(a^k) > U_N(a)$, which means that the kernel allocation a is not N-efficient. This is a contradiction because payoff vectors in the kernel are N-efficient.

Hence, $U_k(a) \geq U_k(O(s_k', s_{-k}))$ in this case.

Case 3: $z > 0$ and Rule 1B2 applies.

In this case, the game proceeds to Stage 2B. For the subgame starting in Stage 2B, the strategy

profile induced by (s'_k, s_{-k}) is the same as the one induced by s. Choose any

$$\bar{T} \in \arg\max(v_T - U_T(a)) \text{ subject to } T \subsetneq N, \ell \in T, \text{ and } k \notin T.$$

This $\bar{T}$ may be different from $\hat{T}$ that is to be chosen by agent ℓ in Stage 2B. However, by their definitions, we must have $v_{\bar{T}} - U_{\bar{T}} = v_{\hat{T}} - U_{\hat{T}}$. We consider the following four subcases.

Subcase 1: $e \geq \max\{v_S - U_S(a), v_{\bar{T}} - U_{\bar{T}}(a)\}$.

In this subcase, agent ℓ in Stage 2B takes Choice 2B1 induced by s_ℓ. For agent k, feasible utilities under the strategy profile induced by (s'_k, s_{-k}) are the following: $U_k(a)$ by Rule 3A1, and $U_k(a_k^\omega)$ by Rule 2B3. As we have argued in Case 1, $U_k(a) \geq v_k$ by the individual rationality of the kernel. Moreover, we have $v_k \geq U_k(a_k^\omega)$ by the definition of v_k. Therefore, we have $U_k(a) \geq U_k(O(s'_k, s_{-k}))$ in this subcase.

Subcase 2: $e < \min\{v_S - U_S(a), v_{\bar{T}} - U_{\bar{T}}(a)\}$.

In this subcase, agent ℓ in Stage 2B takes Choice 2B2 induced by s_ℓ. For agent k, feasible utilities under the strategy profile induced by (s'_k, s_{-k}) are the following: $U_k(a_k^\omega)$ or $(U_k(a_k^\omega) - |e+z|)$ by Rule 3B3 or 4B1, and $U_k(a_k^\omega)$ by Rule 2B3. As we have argued in Subcase 1, we have $U_k(a) \geq v_k \geq U_k(a_k^\omega)$. Hence, $U_k(a) \geq U_k(O(s'_k, s_{-k}))$ in this subcase.

Subcase 3: $v_{\bar{T}} - U_{\bar{T}}(a) \leq e < v_S - U_S(a)$.

In this subcase, agent ℓ in Stage 2B takes Choice 2B3 induced by s_ℓ. As we will show, the only feasible utility for agent k under the strategy profile induced by (s'_k, s_{-k}) is $U_k(a_k^\omega)$ by Rule 2B3, for which we have $U_k(a) \geq v_k \geq U_k(a_k^\omega) = U_k(O(s'_k, s_{-k}))$. For the rest of this subcase, we show that $v_\ell \geq U_\ell(a) - z$ in this subcase and hence Rule 2B3 is always chosen by agent ℓ under Choice 2B3. The fact that a is a kernel allocation plays a critical role in the following argument.

Suppose, by contradiction, that $v_\ell < U_\ell(a) - z$, which implies that $U_\ell(a) > v_\ell$ since $z > 0$. Since $k \in S$, $\ell \notin S$, and $U_\ell(a) > v_\ell$, the coalition S is an objection of k against ℓ as described in K1 in Section 2–3. Because the payoff vector $(U_i(a))_{i \in N}$ is in the kernel, there exists a counter-objection T such that $\ell \in T$, $k \notin T$, and $v_S - U_S(a) \leq v_T - U_T(a)$ as described in K2 in Section 2–3. However, by the definitions of $\bar{T}$ and this subcase, we must have the following: $v_T - U_T(a) \leq v_{\bar{T}} - U_{\bar{T}}(a) < v_S - U_S(a)$. A contradiction obtains.

Subcase 4: $v_S - U_S(a) \leq e < v_{\bar{T}} - U_{\bar{T}}(a)$.

In this subcase, agent ℓ in Stage 2B takes Choice 2B4 induced by s_ℓ. For agent k, feasible utilities

under the strategy profile induced by (s'_k, s_{-k}) are the following: $U_k(a)$ by Rule 3A1, $U_k(a^\omega_k)$ or $(U_k(a^\omega_k) - |e + z|)$ by Rule 3B3 or 4B1, and $U_k(a^\omega_k)$ by Rule 2B3. As we have argued in Subcase 1, we have $U_k(a) \geq v_k \geq U_k(a^\omega_k)$. Hence, $U_k(a) \geq U_k(O(s'_k, s_{-k}))$ in this subcase.

By the above arguments, for any $k \in N$, we have $U_k(O(s)) \geq U_k(O(s'_k, s_{-k})) \geq U_k(O(s''_k, s_{-k}))$ for any s''_k. So, the strategy profile s is a Nash equilibrium of the game (Γ, u). Together with Lemmas 1 and 6, we conclude that s is a subgame-perfect equilibrium of the game (Γ, u). $\square$

Up to this point, we have shown the following. For an arbitrary valuation profile $u \in \mathcal{U}$, we have chosen a kernel allocation $a \in \varphi(u)$. We have constructed a pure strategy s_i for each agent i, and have proved that the strategy profile s is a subgame-perfect equilibrium of the game (Γ, u). Since the outcome allocation corresponding to this equilibrium is a, we have $a \in SPE(\Gamma, u)$. Therefore, the following proposition holds.

Proposition 1. *For any valuation profile $u \in \mathcal{U}$, we have $\varphi(u) \subset SPE(\Gamma, u)$.*

3–2–2. $SPE(\Gamma, u) \subset \varphi(u)$

Let us fix a valuation profile $u \in \mathcal{U}$. Take any equilibrium allocation $a \in SPE(\Gamma, u)$ and let $s = (s_i)_{i \in N}$ be its associated subgame-perfect equilibrium of the game (Γ, u), that is, $O(s) = a$. We show that a is a kernel allocation in $\varphi(u)$. Denote agent i's choice in Stage 1 under s_i by $(a^i, \hat{a}_i, \ell^i, S^i, e^i, z^i)$ for each $i \in N$.

Lemma 8. *Under the strategy profile s, Rule 1A applies, and hence $a^i = a$ for all $i \in N$.*

Proof. If Rule 1B applies, then $a^i = a'$ for all $i \in N \setminus \{k\}$ and $a^k \neq a'$ for some $k \in N$, where a' is some N-feasible allocation. In this case, agent $h \in N \setminus \{k\}$ can gain by a deviation that induces Rule 1C. He only has to choose an N-feasible allocation $a'' \notin \{a', a^k\}$ such that $U_h(a'') > U_h(a)$ [5] together with a real number z'' such that $z'' > \max(z^i)_{i \in N}$ in Stage 1. If Rule 1C applies, then any agent can gain by a deviation that is similar to the one in the previous case. Since the strategy profile s is an equilibrium, Rule 1A must apply. $\square$

Lemma 9. *The equilibrium allocation a is an N-efficient allocation.*

Proof. Suppose not. Then, there exists another allocation a' such that $U_i(a') > U_i(a)$ for all $i \in N$. In this case, any agent $k \in N$ can gain by a deviation that induces Rule 1B1 as follows. Let s'_k be

---

[5] Such $a'' = (x''_i, m''_i)_{i \in N}$ can be constructed by making m''_h sufficiently large.

agent k's strategy where he chooses $a^k = a'$ together with $z' < 0$ in Stage 1. Under the strategy profile (s'_k, s_{-k}), Rule 1B1 applies and the game proceeds to Stage 2A. Backward induction arguments as in the proof of Lemma 1 can show that the unique equilibrium path in the subgame starting in Stage 2A is the "all agree" path because $U_i(a') > U_i(a)$ holds with strict inequality for all $i \in N \setminus \{k\}$. Therefore, $O(s'_k, s_{-k}) = a'$, for which we have $U_k(O(s'_k, s_{-k})) = U_k(a') > U_k(a) = U_k(O(s))$. This is a contradiction to the fact that s is an equilibrium associated with a. $\square$

Lemma 10. *The payoff vector for a is individually rational; $U_i(a) \geq v_i$ for all $i \in N$.*

Proof. Suppose not. Then, $U_k(a) < v_k$ for some agent k. We show that he can gain by a deviation that induces Rule 1B3. Let s'_k be agent k's strategy where he chooses $a^k \neq a$, $\hat{a}_k$ that is $\{k\}$-efficient, and $z' = 0$ in Stage 1. Under (s'_k, s_{-k}), Rule 1B3 applies and the outcome is such that agent k is allocated $\hat{a}_k$, for which we have $U_k(O(s'_k, s_{-k})) = v_k > U_k(a) = U_k(O(s))$. This is a contradiction to the fact that s is an equilibrium associated with a. $\square$

Lemma 11. *The equilibrium allocation a is a kernel allocation in $\varphi(u)$.*

Proof. By Lemmas 9 and 10, the payoff vector $(U_i(a))_{i \in N}$ is an imputation. Suppose, by contradiction, that $(U_i(a))_{i \in N}$ is not in the kernel. Then, there exists an objection to the imputation $(U_i(a))_{i \in N}$ to which no counter-objection exists. That is, there exists (S, k, ℓ) such that $S \subsetneq N$, $k \in S$, $\ell \notin S$, $U_\ell(a) > v_\ell$, and $v_{\bar{T}} - U_{\bar{T}}(a) < v_S - U_S(a)$ where

$$\bar{T} \in \arg\max(v_T - U_T(a)) \text{ subject to } T \subsetneq N, \ell \in T, \text{ and } k \notin T.$$

We show that agent k can gain by a deviation. Let s'_k be agent k's strategy where he chooses $(a^k, \hat{a}_k, \ell^k, S^k, e^k, z^k)$ in Stage 1 as follows:

$$a^k \neq a,$$
$$\ell^k = \ell,$$
$$S^k = S,$$
$$e^k = e \equiv (v_S - U_S(a) + v_{\bar{T}} - U_{\bar{T}}(a))/2, \text{ and}$$
$$z^k = z \equiv (U_\ell(a) - v_\ell)/2.$$

Choices not specified above are arbitrary for s'_k. Note that e and z satisfy the following conditions: $v_{\bar{T}} - U_{\bar{T}}(a) < e < v_S - U_S(a)$ and $U_\ell(a) - v_\ell > z > 0$ and $U_\ell(a) - z > v_\ell$.

Claim 1. *Under the strategy profile (s'_k, s_{-k}), the game proceeds to Stage 2B.*

Proof of Claim 1. The claim holds because $z > 0$ and Rule 1B2 applies. Q.E.D.

Henceforth, we restrict our attention to subgames that are contained in the subgame starting in Stage 2B under the strategy profile (s'_k, s_{-k}).

Claim 2. *Consider a subgame that starts in Stage 3A, where agent k moves given the condition $e < v_S - U_S(a)$. In this subgame, the strategy profile induced by s leads to the outcome under Rule 3A3 or 4A1.*

Proof of Claim 2. In this subgame, the strategy profile induced by s is a subgame-perfect equilibrium. It is sufficient to show that Rules 3A1 and 4A2 never apply in this equilibrium.

Note that agent k's utility for the outcome under Rules 3A1 and 4A2 is $U_k(a)$ or less. We show that agent k's utility for the outcome under Rule 3A3 or 4A1 can be strictly higher than $U_k(a)$ given the condition $e < v_S - U_S(a)$. This fact will complete the proof.

Consider agent k's choice in Stage 3A, where he proposes a^S such that

$U_S(a^S) = v_S$ (that is, a^S is S-efficient),

$U_j(a^S) = U_j(a) + \max\{0, e\} + \varepsilon$ if $j \in S$ (recall that $j \equiv \max(N \setminus \{k, \ell\})$), and

$U_i(a^S) = U_i(a) + \varepsilon$ for all $i \in S \setminus \{j, k\}$,

where $\varepsilon = (v_S - U_S(a) - e)/|S|$ and $|S|$ is the cardinality of S.

Note that $\varepsilon > 0$ because $e < v_S - U_S(a)$.

When $S = \{k\}$, agent k's utility for the outcome under Rule 3A3 by the above choice is $(v_k - e)$, which is higher than $U_k(a)$ by the condition $e < v_S - U_S(a)$.

When $S \neq \{k\}$, all agents in $S \setminus \{k\}$ prefer the outcome under Rule 4A1 to the one under Rule 4A2 due to the above choice of a^S. Backward induction arguments as in the proof of Lemma 1 can show that the unique equilibrium path in the subgame starting in Stage 4A is the "all agree" path. Here, agent k's utility for the outcome under Rule 4A1 is the following.

If $e < 0$ or $j \notin S$, then agent k's utility under Rule 4A1 is

$$U_k(a^S) - e = U_S(a^S) - \sum_{i \in S \setminus \{k\}} U_i(a^S) - e = v_S - \sum_{i \in S \setminus \{k\}} U_i(a) - (|S| - 1)\varepsilon - e$$

$$= U_k(a) + (v_S - U_S(a) - e) - (|S| - 1)\varepsilon = U_k(a) + \varepsilon > U_k(a).$$

If $e \geq 0$ and $j \in S$, then agent k's utility under Rule 4A1 is

$$U_k(a^S) = U_S(a^S) - \sum_{i \in S \setminus \{k\}} U_i(a^S) = v_S - \sum_{i \in S \setminus \{k\}} U_i(a) - (|S| - 1)\varepsilon - e$$

$$= U_k(a) + (v_S - U_S(a) - e) - (|S| - 1)\varepsilon = U_k(a) + \varepsilon > U_k(a).$$

Therefore, for the above choice of a^S, agent k's utility for the outcome under Rule 3A3 or 4A1 is strictly higher than $U_k(a)$ under Rules 3A1 and 4A2 given the condition $e < v_S - U_S(a)$. Since the strategy profile s is an equilibrium of the subgame of our concern, agent k's best response induced by s_k must guarantee him the utility at least as high as the utility for the above choice of a^S. Therefore, the equilibrium outcome must be the one under Rule 3A3 or 4A1. *Q.E.D.*

Claim 3. *Consider a subgame that starts in Stage 3B, where agent ℓ moves given the condition $e > v_{\bar{T}} - U_{\bar{T}}(a)$. In this subgame, the strategy profile induced by s leads to the outcome under Rule 3B1 or 4B2, where agent ℓ's utility is $(U_\ell(a) - z)$.*

Proof of Claim 3. In this subgame, the strategy profile induced by s is a subgame-perfect equilibrium. Since $e > v_{\bar{T}} - U_{\bar{T}}(a) \geq v_T - U_T(a)$ for any T chosen in Stage 2B, similar arguments as in the proof of Lemma 5 can show that agent ℓ in Stage 3B prefers the outcome under Rule 3B1 to the one under Rule 3B3 or 4B1. Furthermore, ℓ's utility under Rule3B1 is obviously at least as good as the one under Rule 4B2. Therefore, ℓ's utility for the equilibrium outcome is $(U_\ell(a) - z)$ under Rule 3B1 or 4B2. A more detailed argument follows.

Rule 3B3 applies when $T = \{\ell\}$ and agent ℓ in Stage 3B proposes a^T. In this case, the condition $e > v_T - U_T(a)$ is equivalent to the condition $U_\ell(a) - z > v_\ell - e - z$, which means that agent ℓ prefers the outcome under Rule 3B1 to the one under Rule 3B3 even if a^T is optimally chosen.

Rule 4B1 applies when $T \neq \{\ell\}$ and agent ℓ in Stage 3B proposes a^T. In this case, agent ℓ must propose a^T to which all agents in $T \setminus \{\ell\}$ agree. To make them all agree in Stage 4B, the outcome under Rule 4B1 must guarantee them an outcome at least as good as the one under Rule 4B2. As argued in Cases 1 and 2 in the proof of Lemma 5, agent ℓ's maximal achievable utility for the outcome under Rule 4B1 is the left hand side of the following inequality:

$$v_T - \sum_{i \in T \setminus \{\ell\}} U_i(a) - e - z < U_\ell(a) - z.$$

The right hand side is agent ℓ's utility for the outcome under Rule 3B1. This inequality, which means that agent ℓ prefers the outcome under Rule 3B1 to the one under Rule 4B1 even if a^T is optimally chosen, holds because it is equivalent to the condition $e > v_T - U_T(a)$. *Q.E.D.*

Claim 4. *According to s_ℓ, agent ℓ in Stage 2B chooses Rule 2B2 to go to Stage 3B given the conditions $v_{\bar{T}} - U_{\bar{T}}(a) < e < v_S - U_S(a)$ and $U_\ell(a) - z > v_\ell$.*

Proof of Claim 4. In the subgame that starts in Stage 2B, the strategy profile induced by s is

a subgame-perfect equilibrium. We investigate feasible utilities for agent ℓ given $s_{-\ell}$. If ℓ chooses Rule 2B1 to go to Stage 3A, then, by Claim 2, his utility is $U_\ell(a_\ell^\omega)$ or less under Rule 3A3 or 4A1. If ℓ chooses Rule 2B2 to go to Stage 3B, then his utility is $(U_\ell(a) - z)$ by Claim 3. If ℓ chooses Rule 2B3, then his utility is at most v_ℓ. Since $U_\ell(a) - z > v_\ell \geq U_\ell(a_\ell^\omega)$, Rule 2B2 should give him the highest utility. Therefore, agent ℓ chooses Rule 2B2 according to s_ℓ given the conditions of our concern. *Q.E.D.*

Claim 5. $U_k(O(s_k', s_{-k})) = U_k(a) + z > U_k(a) = U_k(O(s))$, *that is, agent k can gain by a deviation from s_k to s_k', given s_{-k}.*

Proof of Claim 5. Under the strategy profile (s_k', s_{-k}), the game proceeds to Stage 2B by Claim 1. Since $v_{\bar{T}} - U_{\bar{T}}(a) < e < v_S - U_S(a)$ and $U_\ell(a) - z > v_\ell$ according to s_k', agent ℓ in Stage 2B chooses Rule 2B2 to go to Stage 3B according to s_ℓ by Claim 4. Since $e > v_{\bar{T}} - U_{\bar{T}}(a)$, the outcome $O(s_k', s_{-k})$ is the one constructed under Rule 3B1 or 4B2 by Claim 3, for which agent k's utility is $U_k(a) + z$. The strict inequality $U_k(O(s_k', s_{-k})) > U_k(O(s))$ holds because $z > 0$. *Q.E.D.*

By Claim 5, agent k can gain by a deviation from s_k to s_k', given s_{-k}. This is a contradiction to the fact that s is an equilibrium associated with a. Therefore, the equilibrium allocation a is a kernel allocation in $\varphi(u)$. $\square$

Up to this point, we have shown the following. For an arbitrary valuation profile $u \in \mathcal{U}$, we have chosen an equilibrium allocation $a \in SPE(\Gamma, u)$ and proved that a is a kernel allocation $a \in \varphi(u)$. Therefore, the following proposition holds, which completes the proof of the theorem.

Proposition 2. *For any valuation profile $u \in \mathcal{U}$, we have $SPE(\Gamma, u) \subset \varphi(u)$.*

4. Conclusion

This paper has presented a game form that fully implements the kernel correspondence in subgame-perfect equilibrium. Our finite-stage extensive game form incorporates objections and counter-objections in the definition of the kernel due to Osborne and Rubinstein [1994], and achieves implementation for any class of allocation problems where preferences are quasi-linear and the associated coalitional form game is superadditive. Our game form possesses several desirable properties. One of them is the fact that agents do not have to report preferences, which are difficult to communicate in reality.

The present paper is an attempt to pursue non-cooperative foundations of the kernel. Following the spirit of Bergin and Duggan [1999], we have taken an implementation-theoretic approach and explicitly modelled the physical environment and individual preferences. We have succeeded in designing a game

form that achieves full implementation, but the constructed game form is not simple. The reason why the game form has turned out to be complicated is probably because the definition of the kernel requires interpersonal and inter-coalitional comparisons of payoffs and we have to incorporate Stage 2B and its subsequent stages to achieve the comparisons. It would be a significant contribution if one could design a simpler game form that achieves full implementation of the kernel correspondence. Such a simple game form would enhance non-cooperative foundations and relevance of the kernel of coalitional form games.

Acknowledgement

References

Bag, P. K. and E. Winter [1999], "Simple subscription mechanisms for excludable public goods," *Journal of Economic Theory* Vol.87, pp.72–94.

Bergin, J. and J. Duggan [1999], "An implementation-theoretic approach to non-cooperative foundations," *Journal of Economic Theory* Vol.86, pp.50–76.

Davis, M. and M. Maschler [1965], "The kernel of a cooperative game," *Naval Research Logistics Quarterly* Vol.12, pp.223–259.

Einy, E. and D. Wettstein [1999], "A non-cooperative interpretation of bargaining sets," *Review of Economic Design* Vol.4, pp.219–230.

Gul, F. [1989], "Bargaining foundations of Shapley value," *Econometrica* Vol.57, pp.81–95.

Hart, S. and A. Mas-Colell [1996], "Bargaining and value," *Econometrica* Vol.64, pp.357–380.

Moore, J. and R. Repullo [1988], "Subgame perfect implementation," *Econometrica* Vol.56, pp.1191–1220.

Osborne, M. and A. Rubinstein [1994], *A Course in Game Theory,* Cambridge: MIT Press.

Perez-Castrillo, D. and D. Wettstein [2000], "Implementation of bargaining sets via simple mechanisms," *Games and Economic Behavior* Vol.31, pp.106–120.

Perez-Castrillo, D. and D. Wettstein [2001], "Bidding for the surplus: a non-cooperative approach to the Shapley value," *Journal of Economic Theory* Vol.100, pp.274–294.

Saijo, T., Y. Tatamitani, and T. Yamato [1996], "Toward natural implementation," *International Economic Review* Vol.37, pp.949–980.

Samejima, Y. [2004], "Implementation of nucleolus correspondence," *Essays on Social Choice and Mechanism Design,* Doctoral Dissertation, Columbia University.

Serrano, R. [1997], "Reinterpreting the kernel," *Journal of Economic Theory* Vol.77, pp.58–80.

Serrano, R. and R. Vohra [2002], "Bargaining and bargaining sets," *Games and Economic Behavior* Vol.39, pp.292–308.